The typology and diachrony of nominal classification
Passer, M.B.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Matthias Benjamin Passer

The Typology and Diachrony of Nominal Classification

There are two ways for a language to classify its nouns: either by means of classifiers, which specify the semantics of the classified noun, or by means of grammatical gender, which groups all nouns of a language into formal classes.

This thesis investigates the common assumption that classifier systems may develop into grammatical gender systems. Because this diachronic phenomenon has not yet been documented for any language, the likeliness that such a development would occur is examined by means of a typological study of synchronic systems.

In analyzing the data, this study adopts a new perspective on the development of nominal classification by separating how the means of formal expression develops from the development of those components that have to do with a system’s semantic transparency.

This twofold account for the data from a variety sample of 40 languages shows that there is indeed a number of systems that lie at the intersection of classifiers and gender systems, but that a direct shift from classifier to gender is not likely to occur.
The Typology and Diachrony of Nominal Classification
The Typology and Diachrony of Nominal Classification

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op donderdag 15 december 2016, te 12.00 uur

door

Matthias Benjamin Passer

geboren te Burghausen, Duitsland
Promotiecommissie:

Promotores: 
- prof. dr. P.C. Hengeveld 
  Universiteit van Amsterdam
- prof. dr. H.H. Zeijlstra 
  Georg-August-Universität Göttingen

Overige leden: 
- prof. dr. E.O. Aboh 
  Universiteit van Amsterdam
- dr. J. Audring 
  Universiteit Leiden
- prof. dr. W. Bisang 
  Johannes Gutenberg-University
- dr. F.C. Seifart 
  Universiteit van Amsterdam
- prof. dr. F.P. Weerman 
  Universiteit van Amsterdam

Faculteit der Geesteswetenschappen
omnis labor finem habet premivm eivs non habet fnem

[matthias] scripsit istam partem

do gratias qvod ego perfeci opvs mevm

freely adapted from Madalfrid
In: Johannes Chrysostomus: In Epistolam ad Hebraeos.
Freising (approx. 840), p. 246.
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Grammaticalising Systems</td>
<td>119</td>
</tr>
<tr>
<td>5.1</td>
<td>Overview</td>
<td>120</td>
</tr>
<tr>
<td>5.2</td>
<td>Classifier Systems</td>
<td>126</td>
</tr>
<tr>
<td>5.3</td>
<td>Concordial Class Systems</td>
<td>139</td>
</tr>
<tr>
<td>5.4</td>
<td>Hybrid Systems</td>
<td>143</td>
</tr>
<tr>
<td>5.5</td>
<td>The Grey Area</td>
<td>157</td>
</tr>
<tr>
<td>5.6</td>
<td>Discussion</td>
<td>164</td>
</tr>
<tr>
<td>6</td>
<td>Grammaticalising Properties</td>
<td>175</td>
</tr>
<tr>
<td>6.1</td>
<td>Inventory Size</td>
<td>178</td>
</tr>
<tr>
<td>6.2</td>
<td>Obligatoriness</td>
<td>187</td>
</tr>
<tr>
<td>6.3</td>
<td>Multiple vs. Single Marking</td>
<td>194</td>
</tr>
<tr>
<td>6.4</td>
<td>Locus Operandi</td>
<td>205</td>
</tr>
<tr>
<td>6.5</td>
<td>Flexibility</td>
<td>210</td>
</tr>
<tr>
<td>6.6</td>
<td>Assignment</td>
<td>216</td>
</tr>
<tr>
<td>6.7</td>
<td>Semantic Contribution</td>
<td>222</td>
</tr>
<tr>
<td>6.8</td>
<td>Correlations Between Properties</td>
<td>227</td>
</tr>
<tr>
<td>6.9</td>
<td>Discussion</td>
<td>236</td>
</tr>
<tr>
<td>7</td>
<td>Mapping Nominal Classification</td>
<td>251</td>
</tr>
<tr>
<td>7.1</td>
<td>Developmental Pathways in the Sample</td>
<td>252</td>
</tr>
<tr>
<td>7.2</td>
<td>Theoretical Considerations</td>
<td>262</td>
</tr>
<tr>
<td>7.2.1</td>
<td>The Dimension of Form</td>
<td>265</td>
</tr>
<tr>
<td>7.2.2</td>
<td>The Dimension of Transparency</td>
<td>272</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Graphic Representation of Form and Transparency</td>
<td>278</td>
</tr>
<tr>
<td>7.3</td>
<td>Application of the Mapping Instrument</td>
<td>281</td>
</tr>
<tr>
<td>7.3.1</td>
<td>The Formalisation Value</td>
<td>281</td>
</tr>
<tr>
<td>7.3.2</td>
<td>The Transparency Value</td>
<td>290</td>
</tr>
<tr>
<td>7.3.3</td>
<td>Combining the Values</td>
<td>297</td>
</tr>
<tr>
<td>8</td>
<td>Diachronic Study</td>
<td>307</td>
</tr>
<tr>
<td>8.1</td>
<td>Grammaticalisation in Mandarin</td>
<td>309</td>
</tr>
<tr>
<td>8.2</td>
<td>Dyirbal Crossing the Transparency Threshold</td>
<td>319</td>
</tr>
<tr>
<td>8.3</td>
<td>Loss and Gain in Transparency in West Germanic</td>
<td>325</td>
</tr>
<tr>
<td>8.4</td>
<td>Discussion</td>
<td>333</td>
</tr>
<tr>
<td>9</td>
<td>Conclusions</td>
<td>343</td>
</tr>
<tr>
<td>9.1</td>
<td>The Grammaticalisation Hypothesis</td>
<td>343</td>
</tr>
<tr>
<td>9.2</td>
<td>Nominal Classification</td>
<td>347</td>
</tr>
<tr>
<td>9.3</td>
<td>Outlook</td>
<td>350</td>
</tr>
</tbody>
</table>
Appendix

1. Bora ........................................... 355
2. Boumaa Fijian .............................. 364
3. Burushaski ................................. 373
4. Chichewa ................................. 381
5. Chukchi ................................. 388
6. Dongo-Ko ................................. 394
7. Dyirbal ................................. 408
8. German .................................. 415
9. Great Andamanese ...................... 423
10. Hadza .................................... 433
11. Hausa .................................... 441
12. Hmong Daw ............................... 449
13. Hungarian ............................... 457
14. Ingush .................................. 463
15. Jakaltek .................................. 469
16. Ket ....................................... 478
17. Kiowa .................................... 486
18. Korean .................................... 493
19. Maasai ................................... 502
20. Mandarin .................................. 509
21. Masri ...................................... 516
22. Moro ....................................... 522
23. Munduruku ............................... 529
24. Nsioi ..................................... 536
25. Ngan gityemerrji ......................... 543
27. Nung ....................................... 557
28. Sumerian .................................. 564
29. Taimne ..................................... 571
30. Tamil ....................................... 579
31. Tlingit .................................... 586
32. Tohono O'odham ......................... 592

Bibliography ................................................. 599
Index ................................................. 619

English Summary .......................................... 629
Samenvatting in het Nederlands .................. 636
Curriculum Vitae ........................................... 645
Acknowledgements

Honour your father as a youth,
And receive his teaching in your soul.
As a man, then, add to scientific truth,
So your son can achieve a higher goal.

J. W. von Goethe, Faust I

Like Madalfrid, I am grateful to finish my work, and able to look back on my days in Amsterdam with both joy and nostalgia. The last four years have posed me the greatest of personal challenges, and my studies were the one constant and retreat from an often troublesome world. The more I consider myself fortunate at this point, where life has found its way and another chapter waits to be written. Before I engage in this enterprise, it is in order to give my sincere thanks to those who have supported me work- and otherwise.

First of all, I would like to thank Kees Hengeveld, who gave advice and support to this project from the first draft versions of the project proposal to the final version of the present thesis’ manuscript. Despite the great amount of PhD projects that he supervises, I never felt as but one of the many, but could always rely on an open ear and frank and constructive feedback. Thank you, Kees, for your trust, support, and encouragement!

It would be futile to even think of mentioning all the things that I am grateful for when it comes to Hedde Zeijlstra. Not only is he the one who taught me to read and write, to think and reason, but also made me the scholar that I am today. Hedde, I do not see a way in which I could have carried out this project without your care and supervision, your understanding, hospitality and friendship. You are an inspiration and a model for me, both as an academic and a human being. Thank you for everything that I was allowed to learn from you.

Besides my Doktorväter Kees and Hedde, I owe thanks to a great number of people here at the ACLC for their friendship and company, advice when I was clueless, feedback on work in progress, and support when it came to proof-
reading, finalising manuscripts, or mere procrastination. Among these, I would like to mention by name Roland Pfau, Marlou van Rijn, Caitlin Meyer, Margot Kraaikamp, Brechje van Osch, Margreet van Koert, Maja Ćurčić, Camille Welie, Jing Lin, Hernán Labbé Grünberg, Jelke Bloem, Tessa Spätgens, Bibi Janssen, Rosalinde Stadt, Patrick Schetters, and Sanne Berends — in no particular order, of course. Moreover, I give my utter thanks to Seid Tvica, who I deem not only my much valued brother in PhD and my office buddy, but a true friend and most inspiring antagonist in endless discussions about everything under the sun.

During the years of my project, I had the great joy and luck to meet many fellow researchers of nominal classification, who shared their much appreciated views and opinions and are a true source of motivation and insight for me. Most of all, I owe cordial thanks to Grev Corbett, Sebastian Fedden, Timothy Feist, Gunter Senft, Jenny Audring, and Mike Franjieh. Thank you for organising a wonderful series of workshops on nominal classification at Guildford and Nijmegen, for challenging discussions not only on nominal classification, and for consistently fuelling my curiosity and interest in that specific matter. Furthermore, I would like to thank Marcin Kilarski for our collaboration and for inviting me to spend one month as a guest researcher at Uniwersytet im. Adama Mickiewicza at Poznań, Poland. Beyond those mentioned, there are many more colleagues that shared their expertise and knowledge with me in personal and written communication. They helped improving the quality of this thesis, and I would like to express my cordial thanks for their support. Furthermore, I would like to thank my dear colleagues Marta Castella, Kamil Kaźmierski, and Sophie Kuiper for their true companionship and personal as well as moral support during the past years.

Speaking of personal and moral support, one name must not remain unmentioned: Michael Peter Vereno, we met on the very first day of our study in Salzburg. Since then, you have proven to be the most loyal friend one can find. You have helped me from a distance through the most difficult times, and to master them *viribus unitis*. Thank you for being who you are.

Moreover, I would like to thank my parents Bernd and Monika. I could always rely on your trust and loving support, even when I decided to move far away from you. It has not always been easy for you to accept or understand my decisions and actions, which makes me the more grateful for you always being there for me. I hope to have shown that your efforts and care have not been entirely in vain.

Finally, I owe a great debt of gratitude to you, Sonja; for your love, willpower, support, and believing in both me and us. Everything about my life has improved since you entered it, and I cannot emphasise enough how much I look forward to a bright and joint future in love — under will.

Amsterdam, on the Feast of Saint Lydia, A.D. 2016
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Selected scholars of nominal classification</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>The semantic and discourse properties of nominal classification devices</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>The reduced set of semantic and discourse properties of nominal classification</td>
<td>18</td>
</tr>
<tr>
<td>2.3</td>
<td>Lexical origins of the Persian numeral classifiers</td>
<td>29</td>
</tr>
<tr>
<td>2.4</td>
<td>‘Possessive pronoun’ series in Kilivila</td>
<td>33</td>
</tr>
<tr>
<td>2.5</td>
<td>Bound pronominals in Kilivila: The possessive category IV and emphatic pronouns</td>
<td>33</td>
</tr>
<tr>
<td>2.6</td>
<td>The inventory the verbal classifier system of Manam and Diegueno</td>
<td>37</td>
</tr>
<tr>
<td>3.1</td>
<td>Lexical origins of the Persian numeral classifiers</td>
<td>75</td>
</tr>
<tr>
<td>3.2</td>
<td>The inventory of Korean numeral classifiers (selection)</td>
<td>75</td>
</tr>
<tr>
<td>3.3</td>
<td>Properties of nouns and prototypical devices of nominal classification</td>
<td>79</td>
</tr>
<tr>
<td>4.1</td>
<td>The development of the vegetable food class in Australian</td>
<td>99</td>
</tr>
<tr>
<td>4.2</td>
<td>Functional and hybrid properties of nominal classification devices</td>
<td>104</td>
</tr>
<tr>
<td>4.3</td>
<td>Formal and hybrid properties of nominal classification devices</td>
<td>104</td>
</tr>
<tr>
<td>4.4</td>
<td>Number of languages from the basic phyla in a forty-language sample</td>
<td>108</td>
</tr>
<tr>
<td>4.5</td>
<td>The language sample</td>
<td>110</td>
</tr>
<tr>
<td>4.6</td>
<td>The functional properties of the language sample</td>
<td>113</td>
</tr>
<tr>
<td>4.7</td>
<td>Overview: Structure of the second subsection of section three of the individual language dossiers</td>
<td>116</td>
</tr>
<tr>
<td>5.1</td>
<td>The form-based properties of the language sample</td>
<td>120</td>
</tr>
<tr>
<td>5.2</td>
<td>The ‘hybrid island’</td>
<td>123</td>
</tr>
<tr>
<td>5.3</td>
<td>The ‘mechanised island’</td>
<td>124</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>5.4 The systemic distribution within the language sample.</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>5.5 The classifier systems from the sample.</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>5.6 The form-based properties of the classifier systems from the sample.</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>5.7 The semantics of the Great Andamanese classifiers in different</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>morphosyntactic contexts.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.8 The semantic domains of bodily and possessive reference of the</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Great Andamanese classifiers.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.9 The semantic domains of spatial reference of the Great Andamanese</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>classifiers.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.10 The class inventories of Tohono O’odham, Boumaa Fijian, and North</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Ambrym.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.11 The property composition of the systemic classifier types.</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>5.12 The property composition of ‘single host’ and multiple classifier systems.</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>5.13 The form-based properties of the concordial class systems from</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>the sample.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.14 The form-based properties of the ‘truly’ intermediate systems</td>
<td>144</td>
<td></td>
</tr>
<tr>
<td>from the sample.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.15 The form-based properties of the ‘truly’ intermediate systems</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>from the sample.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.16 A selection of Nasioi class markers.</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>5.17 Selected polysyllabic classifiers of Bora.</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>5.18 Selected classificatory markers of Mundurukú.</td>
<td>150</td>
<td></td>
</tr>
<tr>
<td>5.19 Class marking in Ngan gityemerrt.</td>
<td>153</td>
<td></td>
</tr>
<tr>
<td>5.20 The grey area in its widest definition.</td>
<td>157</td>
<td></td>
</tr>
<tr>
<td>5.21 Class markers in Awtuw.</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>5.22 The form-based properties of the language sample.</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>5.23 The grey area.</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>5.24 Definite and indefinite articles in German.</td>
<td>165</td>
<td></td>
</tr>
<tr>
<td>5.25 The number values of the Kiowa concordial classes.</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>5.26 The sample’s revised concordial class systems.</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>5.27 The sample’s revised classifier systems.</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>6.1 The form-based properties of the revised language sample.</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>6.2 The systemic inventory size in the language sample.</td>
<td>178</td>
<td></td>
</tr>
<tr>
<td>6.3 The sample systems’ form-based properties: inventory size.</td>
<td>179</td>
<td></td>
</tr>
<tr>
<td>6.4 Individual languages’ inventory size.</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>6.5 A selection of Korean measure nouns.</td>
<td>183</td>
<td></td>
</tr>
<tr>
<td>6.6 Selection of general and specific class markers in the hybrid sys-</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>tems of Nasioi, Bora, and Mundurukú.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.7 The semantic domains of the smaller classifier inventories of the</td>
<td>186</td>
<td></td>
</tr>
<tr>
<td>sample.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.8 Obligatory classification in the language sample.</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>6.9 The Jakaltek classificatory categories.</td>
<td>188</td>
<td></td>
</tr>
</tbody>
</table>
6.10 Obligatoriness in the sample revised. 
6.11 Multiple marking in the language sample. 
6.12 Multiple marking revised: Agreement vs. referential marking in the sample. 
6.13 The revised sample's form-based properties. 
6.14 The locus operandi in the language sample. 
6.15 Flexibility in the language sample. 
6.16 Assignment in the language sample. 
6.17 The inventory of the Great Andamanese possessive classifiers. 
6.18 The inventory of North Ambrym classifiers. 
6.19 Semantic contribution to the NP in the language sample. 
6.20 The Nùng and Hmong Daw classifiers. 
6.21 Correlation between multiple marking and locus operandi. 
6.22 Correlation between assignment and flexibility. 
6.23 Correlation between obligatoriness, agreement, and the locus operandi. 
6.24 The revised form-based properties of the language sample. 
6.25 The revised grey area. 
6.26 Inventory size in the revised sample. 

7.1 Morphosyntactic hosts of class markers in the sample's grammaticalising classifier systems. 
7.2 The properties of prototypical classifier systems, the sample's formalising classifier systems, and prototypical concordial class systems. 
7.3 The semantics of the Great Andamanese classifiers on verbs and adverbs. 
7.4 Quantificational categories in Carrier. 
7.5 Selected languages' inventory size. 
7.6 Formalisation measures: Inventory size. 
7.7 Formalisation measures: Host number. 
7.8 Formalisation measures: Locus operandi. 
7.9 Formalisation measures: Obligatoriness. 
7.10 Formalisation measures: Boundedness. 
7.11 Formalisation measures: Single vs. multiple marking. 
7.12 Formalisation measures: Exhaustivity. 
7.13 Formal criteria: Summary. 
7.14 Formalisation measures: Semantic class structure. 
7.15 Formalisation measures: Rule types and productivity. 
7.16 Formalisation measures: Rule number. 
7.17 Formalisation measures: Categorial independence. 
7.18 Formalisation measures: Discreteness of the markers. 
7.19 Formalisation measures: Redundancy. 
7.20 Formalisation measures: Flexibility. 
7.21 Transparency criteria: Summary.
7.22 The formalisation value of Hungarian .................................................. 282
7.23 The formalisation value of Mandarin .................................................. 283
7.24 The formalisation value of Burushaski .............................................. 284
7.25 The formalisation value of German ................................................... 285
7.26 The formalisation value of Bora ......................................................... 286
7.27 Overview over the representative systems' formalisation criteria and value .................................................. 288
7.28 Overview over the sample systems' formalisation criteria and value .... 289
7.29 The transparency value of Hungarian ................................................ 290
7.30 The transparency value of Mandarin .................................................. 291
7.31 The transparency value of Burushaski .............................................. 292
7.32 The third person forms of the Burushaski auxiliary b(a)- .................. 293
7.33 The transparency value of German ................................................... 293
7.34 The transparency value of Bora ......................................................... 294
7.35 Overview over the representative systems' transparency criteria and value .................................................. 295
7.36 Overview over the sample systems' transparency criteria and value .... 296
7.38 The mapped sample systems and their correspondences to systemic areas .................................................. 300
7.39 The grey area ..................................................................................... 303
8.1 The formalisation value in the diachrony of the Mandarin system ......... 311
8.2 The transparency value in the diachrony of the Chinese system .......... 314
8.3 The core classifiers of Mandarin ......................................................... 315
8.4 The formalisation value in the diachrony of the linguistic registers of Mandarin .................................................. 316
8.5 The transparency value of the linguistic registers of Mandarin ......... 318
8.6 The inventory of the Dyirbal concordial class system ......................... 320
8.7 Class membership of selected nouns in Traditional and Younger Dyirbal .................................................. 321
8.8 The formalisation value in the diachrony of the Dyirbal system .......... 322
8.9 The transparency value in the diachrony of the Dyirbal system .......... 324
8.10 The formalisation value in the diachrony of the West Germanic system .................................................. 329
8.11 The transparency value in the diachrony of the Germanic system ...... 332
8.12 Frequency stabilising irregular grammatical phenomena: to be and to exist .................................................. 335
8.13 The Nùng classifiers .......................................................................... 335
9.1 The revised form-based properties of the language sample ............ 348
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Preliminary taxonomy of classifier systems.</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>Reanalysis of the possessive construction in Kilivila.</td>
<td>34</td>
</tr>
<tr>
<td>2.3</td>
<td>Taxonomy of classifier systems.</td>
<td>42</td>
</tr>
<tr>
<td>3.1</td>
<td>Pathway of grammaticalisation.</td>
<td>55</td>
</tr>
<tr>
<td>3.2</td>
<td>Grammaticalisation: (possibly) involved processes.</td>
<td>56</td>
</tr>
<tr>
<td>3.3</td>
<td>Sigmoid curve.</td>
<td>66</td>
</tr>
</tbody>
</table>
| 3.4    | The functional continuum of nominal classification and its relation to gramma
ticalisation.                                                        | 68   |
| 3.5    | Position of classifier systems on a lexico-grammatical axis.                | 69   |
| 3.6    | Grammaticalisation: (possibly) involved processes.                          | 70   |
| 3.7    | Position of classifier systems on a lexico-grammatical axis.                | 85   |
| 3.8    | The basic properties of systems of nominal classification.                  | 89   |
| 4.1    | The basic properties of systems of nominal classification.                  | 103  |
| 4.2    | Taxonomy of the systems of nominal classification.                          | 103  |
| 4.3    | Formal and functional properties of nominal classification and their use in
the sample study.                                                        | 106  |
| 4.4    | Sample coverage of the terminal subtypes of nominal classification systems. | 112  |
| 4.5    | Overview: Structure of the first section of the individual language dossiers.| 115  |
| 4.6    | Overview: Structure of the second section of the individual language dossiers.| 115  |
| 4.7    | Overview: Structure of the first subsection of section three of the individual language dossiers. | 116  |
| 4.8    | Overview: Structure of the third subsection of section three of the individual language dossiers. | 118  |
| 5.1    | Revised taxonomy of the systems of nominal classification.                  | 171  |
6.1 The inventory size of the sample languages. . . . . . . . . . . . 180
6.2 Obligatory classification and the positioning of the nominal class-
sification systems of Hungarian, Persian, and Ngan'gityemerri on
a lexico-grammatical axis. . . . . . . . . . . . . . . . . . . . . . 192
6.3 Order of grammaticalising properties in the sample. . . . . . . 235
6.4 Final taxonomy of systems of nominal classification. . . . . . 239
6.5 Possible developmental pathways of systems of nominal classifi-
cation. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 249

7.1 The sample systems at their respective developmental state. . . 253
7.2 The development of the Great Andamanese classifiers. . . . . . 260
7.3 Grammaticalisation in nominal classification devices: Formal ex-
pression vs. transparency. . . . . . . . . . . . . . . . . . . . . . 263
7.4 The basic mapping instrument. . . . . . . . . . . . . . . . . . . . 279
7.5 Corresponding mapping areas and systems. . . . . . . . . . . . 279
7.6 Class assignment in Burushaski. . . . . . . . . . . . . . . . . . . 292
7.7 Comparative mapping of the full language sample. . . . . . . . 297
7.8 Comparative mapping of the individual sample systems. . . . . 298
7.9 Corresponding areas and systems in the model. . . . . . . . . . 299
7.10 Comparative mapping of the sample systems: Distribution of
systemic types. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 301
7.11 Systemic areas and developmental pathways in the sample. . . 303
7.12 The grey area: Deviations and prototypes. . . . . . . . . . . . . 304

8.1 Systemic areas and developmental pathways in the sample. . . 308
8.2 Developments in the Mandarin system. . . . . . . . . . . . . . . 314
8.3 Developments in the Mandarin linguistic registers. . . . . . . . 318
8.4 Developments in the Dyirbal system. . . . . . . . . . . . . . . . 325
8.5 Developments in the English system. . . . . . . . . . . . . . . . 333
8.6 Developmental shifts in the history of the English system. . . . 337
8.7 Comparative mapping of the sample systems: Distribution of
systemic types. . . . . . . . . . . . . . . . . . . . . . . . . . . . 339
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>first person</td>
</tr>
<tr>
<td>1D</td>
<td>one-dimensional</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
</tr>
<tr>
<td>2D</td>
<td>two-dimensional</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
</tr>
<tr>
<td>3D</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>A</td>
<td>agent</td>
</tr>
<tr>
<td>ABL</td>
<td>ablative</td>
</tr>
<tr>
<td>ABS</td>
<td>absolutive</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ADJ</td>
<td>adjective</td>
</tr>
<tr>
<td>ADV</td>
<td>adverb</td>
</tr>
<tr>
<td>AGR</td>
<td>agreement</td>
</tr>
<tr>
<td>ALI</td>
<td>alienable</td>
</tr>
<tr>
<td>AOR</td>
<td>aorist</td>
</tr>
<tr>
<td>ART</td>
<td>article</td>
</tr>
<tr>
<td>ASS</td>
<td>associative</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary</td>
</tr>
<tr>
<td>C</td>
<td>common</td>
</tr>
<tr>
<td>CC</td>
<td>concordial class</td>
</tr>
<tr>
<td>CCS</td>
<td>concordial class system</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
</tr>
<tr>
<td>CLS</td>
<td>classifier system</td>
</tr>
<tr>
<td>COLL</td>
<td>collective</td>
</tr>
<tr>
<td>COMPL</td>
<td>completive</td>
</tr>
<tr>
<td>CON</td>
<td>sentence connector pronoun</td>
</tr>
<tr>
<td>CONJ</td>
<td>subjunctive</td>
</tr>
<tr>
<td>CONT</td>
<td>continuous</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>CP</td>
<td>complementiser phrase</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>DEICT</td>
<td>deictic</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
</tr>
<tr>
<td>DERIV</td>
<td>derivational</td>
</tr>
<tr>
<td>DIST</td>
<td>distal</td>
</tr>
<tr>
<td>DL</td>
<td>dual</td>
</tr>
<tr>
<td>DP</td>
<td>determiner phrase</td>
</tr>
<tr>
<td>DUR</td>
<td>durative</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EXCL</td>
<td>exclusive</td>
</tr>
<tr>
<td>F/F.</td>
<td>feminine</td>
</tr>
<tr>
<td>FOC</td>
<td>focal</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>GS</td>
<td>gender system</td>
</tr>
<tr>
<td>H</td>
<td>hypothesis</td>
</tr>
<tr>
<td>HAB</td>
<td>habitual</td>
</tr>
<tr>
<td>HORIZ</td>
<td>horizontal</td>
</tr>
<tr>
<td>HS</td>
<td>hybrid system</td>
</tr>
<tr>
<td>IMAG</td>
<td>imaginative</td>
</tr>
<tr>
<td>INAL</td>
<td>inalienable</td>
</tr>
<tr>
<td>INCL</td>
<td>inclusive</td>
</tr>
<tr>
<td>IND</td>
<td>indicative</td>
</tr>
<tr>
<td>INDEF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INT</td>
<td>intensifier</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>INV</td>
<td>inverse</td>
</tr>
<tr>
<td>INSTR</td>
<td>instrumental</td>
</tr>
<tr>
<td>IPF</td>
<td>preterite</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>M/m.</td>
<td>masculine</td>
</tr>
<tr>
<td>mCLS</td>
<td>multiple classifier system</td>
</tr>
<tr>
<td>MOD</td>
<td>modal</td>
</tr>
<tr>
<td>MULTI</td>
<td>multiple</td>
</tr>
<tr>
<td>N/n.</td>
<td>neuter, noun</td>
</tr>
<tr>
<td>nCLS</td>
<td>noun classifier system</td>
</tr>
<tr>
<td>NCS</td>
<td>noun class system</td>
</tr>
<tr>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>NUM</td>
<td>numeral</td>
</tr>
<tr>
<td>O</td>
<td>object</td>
</tr>
<tr>
<td>OBJ</td>
<td>object</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>P</td>
<td>prediction</td>
</tr>
<tr>
<td>PART</td>
<td>participle</td>
</tr>
<tr>
<td>pCLS</td>
<td>possessive classifier system</td>
</tr>
<tr>
<td>PF</td>
<td>perfect</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>PRD</td>
<td>predicative</td>
</tr>
<tr>
<td>PROGR</td>
<td>progressive</td>
</tr>
<tr>
<td>PRS</td>
<td>present</td>
</tr>
<tr>
<td>PRT</td>
<td>participle</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>Q</td>
<td>research question</td>
</tr>
<tr>
<td>qCLS</td>
<td>numeral classifier system</td>
</tr>
<tr>
<td>R/r.</td>
<td>residue</td>
</tr>
<tr>
<td>REC</td>
<td>recent</td>
</tr>
<tr>
<td>REF</td>
<td>referential</td>
</tr>
<tr>
<td>REL</td>
<td>relative</td>
</tr>
<tr>
<td>RM</td>
<td>remote</td>
</tr>
<tr>
<td>S</td>
<td>subject</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SR_SEQ</td>
<td>same referent sequence</td>
</tr>
<tr>
<td>SUB</td>
<td>subordinate clause</td>
</tr>
<tr>
<td>SUBJ</td>
<td>subject</td>
</tr>
<tr>
<td>s.v.</td>
<td>sub voce</td>
</tr>
<tr>
<td>TR</td>
<td>transitive</td>
</tr>
<tr>
<td>V</td>
<td>verb</td>
</tr>
<tr>
<td>VEG</td>
<td>vegetable</td>
</tr>
<tr>
<td>VERT</td>
<td>vertical</td>
</tr>
<tr>
<td>VOC</td>
<td>vocative</td>
</tr>
<tr>
<td>VP</td>
<td>verb phrase</td>
</tr>
</tbody>
</table>
Nominal classification is a grammatical device that groups a given language’s nouns into semantic and/or morphological groups. These groups are morphosyntactically marked either by classifiers or gender/noun class markers. Cf. 1.1, where a semantic categorisation of yabuyugu “girl” and gangu:l “wallaby” is indicated by the classifier for humans (bama:l) and for edible flesh (miña) respectively, and 1.2 and 1.3, where formal nominal classes are differentiated by agreement markers.

(1.1) Yidiny (cf. Dixon 1977: 480)

<table>
<thead>
<tr>
<th>bama:l</th>
<th>yabuyugu</th>
<th>miña</th>
<th>gangu:l</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL:PERSON.ERG</td>
<td>girl.ERG</td>
<td>CL:EDIBLE.FLESH.ABS</td>
<td>wallaby.ABS</td>
</tr>
</tbody>
</table>

wawa:l
see.PAST
“The girl saw the wallaby.”

(1.2) Ancient Greek

A ha ge̥r̥-dɔ farmer-NOM.M.SG short.sighted-NOM.M.SG
“the short-sighted farmer”

B ha mukr-ɔ small-NOM.F.SG ant.hill-NOM.F.SG
“the small ant hill”
The examples [1.1] [1.2] and [1.3] display evident formal and semantic differences between classifier and gender/noun class systems: Consider for instance the transparent semantics of the formally independent classifiers in [1.1] as opposed to [1.2] where the feminine gender of Ancient Greek *marmêkia* "ant hill" or the neuter gender of Latin *corpus* "body" is expressed by affixed markers and cannot be motivated by an ant hill’s or a body’s real life properties. This entails the central question in nominal classification research: Considering that both are means of nominal classification, why are they so different with respect to their form and function? A common answer to this question brings in grammaticalisation as the link between the two types of systems: The semantically more productive and formally more independent classifier systems may grammaticalise and consequently lose their semanticity and formal independence, which ultimately results in a semantically vain system of formal classes of nouns. This latter state is generally assumed to be prototypical for gender/noun class systems. This assumed diachronic shift is henceforth referred to as the “grammaticalisation hypothesis” and can be found throughout the literature of the past decades. It constitutes the focus of this study. The grammaticalisation hypothesis presupposes that all systemic types of nominal classification devices are related. It is theoretically sound and constitutes a well-imaginable scenario, since it takes into account all regular processes involved in grammaticalisation and relates them to the typical properties found in classifier systems and gender/noun class systems. However, besides a handful of descriptions of intermediate systems that might represent ‘hinge systems’ between classifier and gender/noun class systems, there is no documented case of such a shift. Since the complete lack of historical evidence does not justify such a widespread, this study was carried out.

The research history of nominal classification is related to a number of challenges to the grammaticalisation hypothesis and common conceptualisations of nominal classification in general. It provides a historical basis for a number of challenges towards nominal classification theory, which are addressed in this study. Nominal classification is object of study since the earliest days of linguistic and philological research and discussion; descriptions of nominal classification in the form of gender/noun class systems date back as far as to Pâñini (Faddegon 1963: 56) and Protagoras. Ever since it has puzzled and challenged researchers due to its seeming arbitrariness and appealing lack of transparent functionality and convenience for the speaker. This even led to the widespread
Introduction

belief that nominal classification in general was a useless, inherited, and arbitrary baggage of earlier grammatical stages. Though earlier descriptions of nominal classification systems are “often dismissed as prescientific”, “they not only reflect prevailing views but also reveal recurrent motifs that have shaped our understanding of this grammatical category” and despite their lack of linguistic theory frequently come to “equally insightful and accurate conclusions” (Kilarski 2013: 59). Kilarski (2013) divides the research history in three bigger sections; this division is followed here: “From Protagoras to the philosophical grammars” (4th century BC – 18th century AC; 59–115), “From the 'Romantics’ to the Neogrammarians” (18th and 19th century, 117–157), and studies of the 20th and 21st century (199–319).

The two basic mechanisms behind gender/noun class systems, analogy and anomaly, are known since antiquity. Analogy constitutes the conventionalised, systematic, and formal imitation of a given linguistic unit’s behaviour (e.g. *sing, sang, sung* vs. *ring, rang, rung*); “anomaly” in this specific context refers to seemingly unsystematic phenomena that occur within nominal classification (cf. the ‘arbitrary’ gender assignment in German: *der Löffel* m. “the spoon”, *die Gabel* f. “the fork”, *das Messer* n. “the knife”) and that are supposed to mirror ‘natural’ extralinguistic values. From the earliest studies of gender/noun class systems on, the approaches to nominal classification can be loosely divided into two groups that are related to the mechanisms of analogy and anomaly, the ‘semantics first’ approach, which emphasises the meaningfulness of nominal classification systems, and the ‘form first’ approach, which emphasises the arbitrariness and semantic redundancy of nominal classification. These approaches toward nominal classification can also be referred to as the ‘naturalist’ approach, which considers nominal classification a semantic reflection of real world objects’ qualities, and the ‘conventionalist’ approach, which considers concordial class correlations a matter of communicative convention and a more or less arbitrary phenomenon. Across the research history, individual grammarians usually tend to largely follow either the conventionalists’ or the naturalists’ approach; scholars that do not emphasise either opinion and/or acknowledge both mechanisms involved in nominal classification are rather the exception (cf. table 1.1).

Table 1.1: Selected scholars of nominal classification (cf. Kilarski 2013).

<table>
<thead>
<tr>
<th>naturalists</th>
<th>conventionalists</th>
<th>independent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>antiquity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stoics</td>
<td>Socrates, Aristotle</td>
<td>Protagoras</td>
</tr>
<tr>
<td>Apollonius Dyscolus</td>
<td>Sextus Empiricus</td>
<td>Dionysius Thrax</td>
</tr>
<tr>
<td>Priscian</td>
<td>Marcus Terentius Varro</td>
<td></td>
</tr>
<tr>
<td>artes-grammarians</td>
<td>regular-grammarians</td>
<td></td>
</tr>
<tr>
<td><strong>middle ages</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ælfric</td>
<td>St Boniface</td>
<td>Modistæ</td>
</tr>
<tr>
<td>John of Genoa</td>
<td>William of Ockham</td>
<td>Smaragdus</td>
</tr>
<tr>
<td></td>
<td>critics of Modistæ</td>
<td>Peter Helias</td>
</tr>
<tr>
<td></td>
<td>Peter Abelard</td>
<td></td>
</tr>
<tr>
<td><strong>15th–18th century</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>James Harris</td>
<td>Nicolas Beauzée</td>
<td>John Patrægave</td>
</tr>
<tr>
<td>Port Royal grammarians</td>
<td></td>
<td>T. Campanella</td>
</tr>
<tr>
<td>Charles Pinot Duclos</td>
<td></td>
<td>Paul LeJeune</td>
</tr>
</tbody>
</table>
Chapter 1

The naturalist and structuralist factions shaped the discussion, study, conceptualisation, and theory of nominal classification, which virtually only involved grammatical gender until the second half of the 20th century. The semantic motivation was often drawn from earliest and earlier stages of language and language research and thus fairly overlaps with diachronic research on concordial class systems, while the structuralist motivations of concordial class systems mostly are investigations of synchronic language stages, where a given system’s semantics are often considerably eroded. Due to these different approaches, discussions of the matter at times led to polemic arguments between the two ‘camps’.

It was only since the 1970s that classifier systems also found a raised interest in nominal classification studies. Since then, a greater number of descriptions of classifier systems and attempts to account for nominal classification in the form of both classifier and gender/noun class systems in general have been published (cf. e.g. Aikhenvald 2000, Allan 1977, Craig 1986c, Dixon 1986). However, due to the comparatively late consideration of classifier systems, our understanding, definitions, and conceptualisation of classifier systems and their formal and functional nature yet lags behind the one of gender/noun class systems, which for instance led to a number of phenomena to be considered classifiers (e.g. ‘relational classifiers’ and ‘verb classifiers’), while their functional and/or formal behaviour suggests that these are no means of nominal classification. This issue is addressed more in detail in chapter 2. Today’s research community assumes nominal classification to constitute a functional continuum (cf. e.g. Aikhenvald 2000: 13), which is characterised by a semantics-based pole (corresponding to the fairly semantically rich and inter-

---

As an example, take the following quote from Michels (1891: 125, cited and translated in Kilarski 2013: 139), which dates from the heated 19th century discussion of Grimm’s analysis of grammatical gender: “[T]he most obvious assumption would be that it [personification and sexualisation; MP] was at its strongest at the beginnings of human linguistic and cultural history. Consequently, we should however hold it up to be even stronger in animals, for example in those apes similar to humans, of which until now animal psychologists have suspected nothing. But putting oneself in the mind of an ape does have its difficulties.”

Another example of the highly polemic manner of the discussion of Grimm’s analysis is the ‘pathological state’ of the human soul that is presupposed by Brugmann when it comes to the personification and sexualisation of the world (cf. Kilarski 2013: 134).

The probably most extreme view on nominal classification stems from Jan Ignacy Nieceşław Badouin de Courtenay, who thought of gender as a blessing and a curse in that it positively contributed to artistic, literary, and scientific creativity, but also led “to such disorders as an excessive presence of sex in thought, nightmares, pathological behaviour, chaotic associations, unhealthy moods, erotic and religious delusions and sadism” (Kilarski 2013: 188).
active classifier systems) and a morphosyntax-based pole (corresponding to the fairly automatised and system-sided gender/noun class systems).

This assumption of a functional continuum of nominal classification is crucial for the grammaticalisation hypothesis, and yet faces a theoretical problem: Since modern research has been focusing on the more graspable and obvious formal properties, the functionality of nominal classification systems is underresearched until today. Most available literature is descriptive and not analytic, because data on many systems is either not sufficiently available or insufficient in terms of quality. In addition, a given system’s functional characteristics often appear to be individual rather than comparable. Both these factors led to an unsatisfactory understanding of these systems in general; for instance, it led to the widespread assumption that there is some main function of a given type of nominal classification system (e.g. reference tracking for gender/noun class systems or individuation for numeral classifier systems), which optionally may be complemented by additional functions an individual system may take. While this basic conception is not wrong, it is oversimplifying the situation. There is evidence for given systemic types serving more than one ‘main’ function and additional functions a system may employ are likely not to be arbitrary.

Having outlined this study’s main subject, the grammaticalisation hypothesis, and the research history of nominal classification, it is in order to address the organisation of this thesis. Chapter 2 provides a general theoretical and typological background on nominal classification devices. It presents an overview on their functions (section 2.1) and commonly acknowledged properties and this study’s definition of classifier systems (section 2.2) and gender/noun class systems (section 2.4). Section 2.3 addresses those classifier types that are not included in this study.

Chapter 3 provides a theoretical background on grammaticalisation, and how it applies to the study of nominal classification. Section 3.1 discusses the principles and properties of grammaticalisation as well as the involved processes; it also distinguishes grammaticalisation from other phenomena of grammatical change such as reanalysis and analogy or lexicalisation. Section 3.2 relates the properties of nominal classification devices to the distinct subprocesses of grammaticalisation and outlines the expected change phenomena. Section 3.3 presents these theoretical considerations in a more general context and addresses the consequences of grammaticalisation for nominal classification systems with respect to their interactivity, semantics, nature of classification, and formal expression.

Chapter 4 introduces the focus and method of this study: Section 4.1 lists

---

3Descriptions of functional relations between individual systems and other grammatical categories (cf. e.g. Bisang 1999 and 2002, a more comprehensive study of functionality in languages of South East Asia and in general) or specific functions of individual systems exist, but a general comparative and typological investigation has not yet been carried out. Contini-Morava and Kilarski (2013) propose a sevenfold discrimination of four semantic and three discourse functions, which apply to both classifier and gender/noun class systems and are introduced in chapter 2.
its research questions, from which hypotheses and predictions are derived. Section 4.2 outlines the theoretical motivation and feasibility of a synchronic and typological approach to a diachronic hypothesis, which lacks documented proof, and further presents arguments for the grammaticalisation in form of potential ‘hinge systems’ which are not included in this study (Olgolo, Zande) as well as reconstructional attempts (Australian, Proto-Indo-European). Section 4.3 introduces the sampling method that is used in this study, and presents the language sample. Beyond this, the functional typology presented in section 2.1 is applied to the language sample. Section 4.4 gives an overview on the data that this study is based on: For each sample language, a dossier is compiled; this section addresses the data sources and presents the structure of the dossiers, which are attached to this study in the appendix (cf. page 353).

Chapter 5 presents the sample study’s outcomes from the perspective of the systemic types of classification systems. Section 5.1 provides an overview on the sample’s properties, where a clear grammaticalisation cline is observable. The sections 5.2–5.4 present the properties of classifier systems, gender/noun class systems, and hybrid systems respectively. Section 5.5 defines three different types of deviations from the prototypical systems defined in the sections 2.2 and 2.4; these deviations together form the maximum extension of the ‘grey area’ that connects classifier and gender/noun class systems.

Chapter 6 presents the sample study’s outcomes from the perspective of the individual properties of nominal classification and their distribution in the sample. The sections 6.1–6.7 discuss the individual properties introduced in section 3.2 and 3.3. Section 6.8 addresses correlations between properties, and the prevailing order of grammaticalising properties. Section 6.9 evaluates the reliability and validity of the individual properties; it concludes the chapter by distinguishing two observed dimensions of grammaticalisation in nominal classification, the grammaticalisation of form (“formalisation”) and transparency (“conventionalisation”).

Chapter 7 builds upon the outcomes of the chapters 5 and 6 and introduces a separate account for both formalisation and conventionalisation. Section 7.1 recapitulates the pathways of development that are observable in the sample and provides the theoretical considerations that are important for approaching both dimensions separately. Section 7.2 describes the account in detail: A package of seven properties defines each dimension; the measuring points of these properties can be translated into an approximative numerical value. The properties, their measure points, and the correspondence to numerical values are introduced in section 7.2.1 for the formal dimension, and in section 7.2.2 for the dimension of transparency. The combination of the “formalisation value” and “transparency value” allows a graphic representation of nominal classification systems on a two-dimensional graph; this mapping instrument is presented in section 7.2.3 together with the possible types of systems and pathways of development. Section 7.3 applies the account for both dimensions and the mapping instrument to the language sample and discusses the outcomes.

Chapter 8 comprises diachronic case studies of the systemic developments
in the systems of Mandarin (section 8.1), Dyirbal (section 8.2), and West Germanic (section 8.3). The insights gained from the detailed investigation of innersystemic changes are summarised and put in a more general theoretical context in section 8.4.

Chapter 9 recapitulates the outcomes of the ‘unitary’ grammaticalisation account of the chapters 5 and 6 and the ‘split’ grammaticalisation account of the chapters 7 and 8. This study’s observations are projected onto the expectations from the grammaticalisation hypothesis as well as the general context of nominal classification theory.
CHAPTER 2

Nominal Classification

This chapter is devoted to the introduction of the theoretical considerations and definitions of the present study of nominal classification devices. It introduces an in-depth description and definition of nominal classification devices. Since the better researched properties of nominal classification devices are the formal ones and since these mainly serve the distinction of classifier and gender/noun class systems, these are described in separate sections (2.2 and 2.4), while the commonalities of nominal classification devices are described in section 2.1. A number of classifier types (‘relational’ and ‘verb’ classifiers as well as ‘minor classifier types’) are not considered in this study; section 2.3 discusses these types and the reasons for their exclusion.

2.1 Functionality

“The most obvious connection between these systems of nominal classification is their function” (Senft 2007: 686). Function complements the form-based criteria which are provided in section 2.2 and 2.4. The functionality of nominal classification in general, i.e. the functionality of both classifier and gender/noun class systems, is a fairly new object of study, and consequently no broader comparative studies have been conducted so far. I sketch out the common understanding of the functionality of nominal classification in the following in order to provide a basic understanding of their raison d’être prior to the in-depth definition and description of classifier and concordial class systems.

The only commonly established functional criteria are the semantic contribution (or lack thereof) of the classificatory marker to its noun phrase and the
contribution of a system of nominal classification to the parsing of discourse referents. This latter function is usually referred to by the term “reference tracking”, which denotes the usage of classificatory markers as tracking device for discourse referents (cf. 2.1 and 2.2, where the respective classificatory marker enables the reader to identify the intended referent).

(2.1) Hmong Daw (cf. Mottin 1978: 29)


“What house? That one.”

(2.2) German

Sonja und Michael können nicht kommen.
Sonja:NOM.F.SG and Michael:NOM.M.SG can not come
Er ist krank.
Er:NOM.M is sick

“Sonja and Michael are not able to come. He is sick.”

Reference tracking is assumed to play a role in virtually any system of nominal classification. Next to it, the most basic assumed function of systems of nominal classification concerns their (not further specified) semantic contribution to the noun phrase: Does the classificatory marker add or manipulate the semantics of the noun phrase, as it is commonly assumed for classifier systems, or is the classificatory contribution semantically virtually empty, as is commonly assumed for gender/noun class systems? Both the semantic contribution and the lack thereof can be accounted for by two linguistic concepts, referentiality and agreement respectively. Referential markers “provide independent reference” and “are capable of conveying semantic information”. Agreement markers copy information from their source of agreement (cf. van Rijn 2016: 12). The source of agreement is the head noun in the case of nominal classification systems. The grammaticalisation of a system’s semantics thus corresponds to a decline in referential potential, with the extremes being full semantic contribution to the noun phrase (i.e. the marker’s semantic contribution is necessary to encode the intended content; cf. 2.3) and no semantic contribution at all (i.e. the agreement marker merely mirrors the head noun’s referential class; cf. 2.4). In addition, the instances of markers that grammatically agree do not fulfil a semantic function, since they only copy already available information, whereas non-agreeing markers do.

\footnote{It is important to note that in principle both functions, reference and grammatical agreement, can be achieved independently from the markers’ formal realisation. Despite the preference of agreement markers to occur in the form of affixes and the preference of referential markers to occur as independent elements or clitics, the grammaticalisation of a system’s semantics is not depending on the degree of its formal grammaticalisation. This means that a given system does not have to wait until the grammaticalising marker has lost its formal independence before it may drop semantic content (cf. van Rijn 2016: 7f.).}
Nominal Classification

(2.3) Mákú (cf. Farmer and Bruil 2015: 4)

\[ kíù-gaba, \quad kíù-naka \]

metal-CL:RING  metal-CL:POINTED

“chain link, nail”

(2.4) German

\[
\begin{align*}
\text{der} & \quad \text{Stuhl,} \\
\text{die} & \quad \text{Tasse,} \\
\text{das} & \quad \text{Kissen}
\end{align*}
\]

\text{ART}  \quad \text{NOM.M.SG}  \\
\text{ART}  \quad \text{NOM.F.SG}  \\
\text{ART}  \quad \text{NOM.N.SG}

“the chair, the cup, the pillow”

Nominal classification systems often employ functions beyond reference tracking and the semantic contribution to the noun phrase. Contini-Morava and Kilarski (2013) propose a sevenfold functional typology, which is applicable to all types of nominal classification devices.\(^2\) The criteria proposed in Contini-Morava and Kilarski (2013) could however not yet be tested on a more general scale, as especially rich discourse data is both necessary and lacking for its application. Also, these functions are not exclusively employed in nominal classification, but are all observed in other grammatical domains such as the category of number. Nevertheless, the criteria of Contini-Morava and Kilarski (2013) are focussing on the function of nominal classification devices rather than their formal behaviour and allow for a broad and cross-linguistic comparison of these systems. Therefore, I adopt the criteria of Contini-Morava and Kilarski (2013), which are listed in table 2.1 and explained below.\(^3\)

Table 2.1: The semantic and discourse properties of nominal classification devices (Contini-Morava and Kilarski 2013).

<table>
<thead>
<tr>
<th>Semantics (S)</th>
<th>Discourse (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S(_1) expansion of lexicon</td>
<td>D(_1) identification (anaphora, deixis, disambiguation, . . .)</td>
</tr>
<tr>
<td>S(_2) differentiating referents</td>
<td>D(_2) management (definiteness/specificity, topic/focus, . . .)</td>
</tr>
<tr>
<td>S(_3) individuation</td>
<td>D(_3) re-presentation of referents (recategorisation)</td>
</tr>
<tr>
<td>S(_4) attribution of properties</td>
<td></td>
</tr>
</tbody>
</table>

\(^2\) This typology is largely based on Seiler’s (e.g. 1986), approach to nominal classification within the framework of apprehension. In this framework, two underlying principles of predicativity and indicativity stand in competition; these can be loosely translated into the competition of linguistic expressivity and economy. The principles are encoded to a varying degree in individual languages. The scope of Seiler (1986) goes beyond the realm of nominal classification. Therefore, the typology of Contini-Morava and Kilarski (2013) is adopted for this study, as it applies Seiler’s (1986) findings in a way that allows accounting for nominal classification devices only.

\(^3\) The interested reader may refer to Contini-Morava and Kilarski (2013) for a more comprehensive elaboration of this typology. Note that the sparse documentation of the individual sample languages did not always allow for an assessment of all seven functions, which is why a reduced version of the functional typology was applied to this study. The discussion for the exclusion or fusion of functions is embedded in the respective function’s description.
2.1. Functionality

**Ad S₁:** The expansion of the lexicon’s referential power is accomplished by productively using classification markers in a derivational function, i.e. by a reclassification of the root (as shown in 2.5 and 2.6 where a range of classificatory markers fuses with the stem in order to create a new nominal lexeme).

(2.5) **Máih̄kî** (cf. Farmer and Bruil 2015: 4)

\[
\begin{align*}
&\text{kíù-gaba,} & &\text{metal-CL:RING} \\
&\text{kíù-naka,} & &\text{metal-CL:POINTED} \\
&\text{kíò-ró,} & &\text{metal-CL:CONCAVE} \\
&\text{kíù-tòtó} & &\text{metal-CL:FLAT.RIGID}
\end{align*}
\]

“chain link, nail, metal pot, sheet of metal”

(2.6) **Bora** (cf. Seifart 2005: 114)

\[
\begin{align*}
&\text{wáhi-20,} & &\text{banana-CL:3D.LONG.OBJECT} \\
&\text{wáhi-kọ,} & &\text{banana-CL:1D.POINTED} \\
&\text{wáhi-pájkọ} & &\text{banana-CL:LIQUID}
\end{align*}
\]

“banana fruit, banana plant, banana drink”

Homophonous roots which occur in different classes and display distinct, but related meanings are also included in this function (e.g. German *der Schild* m. “the (battle-)shield” vs. *das Schild* n. “the sign”; cf. also Contini-Morava and Kilarski 2013: 268f.).

**Ad S₂:** The function of referent differentiation denotes the provision of more detailed information on a lexical item which is otherwise underspecified for the classificational markers distinction. By specifying the meaning of a noun (phrase) by means of a classificational marker, its existing interpretation is narrowed down in a certain aspect. This function occurs in both gender/noun class systems (e.g. in cases of epicene class, where a common class noun is further specified by reassignment to the masculine or feminine class; cf. 2.7) and in classifier systems (cf. 2.8).

(2.7) **Hausa** (cf. Newman 2000: 201)

\[
\begin{align*}
&wáfàwà, & &\text{dwarf:M/F.SG} \\
&wáwà, & &\text{fool:M/F.SG} \\
&zàbyà, & &\text{albino:M/F.SG} \\
&kùmàmà & &\text{feeble.person:M/F.SG}
\end{align*}
\]

“dwarf, fool, albino, feeble person”

(2.8) **Mandarin** (cf. Contini-Morava and Kilarski 2013: 274)

A \[
\begin{align*}
&\text{yì} & &\text{tóng} & &\text{diàn-huà} \\
&\text{one} & &\text{CL:COMMUNICATION} & &\text{telephone}
\end{align*}
\]

“one phone call”

B \[
\begin{align*}
&\text{yì} & &\text{ge} & &\text{diàn-huà} \\
&\text{one} & &\text{CL:GENERAL} & &\text{telephone}
\end{align*}
\]

“one phone call/telephone (context-sensitive)”
Nominal Classification

Note the seeming similarity of examples such as 2.9 and 2.10 to the derivational function of the classificatory markers in 2.5 and 2.6 which fuse with the stem in order to coin a new lexical item, whereas in 2.9 and 2.10 the referent differentiation occurs formally independent from the noun and outside of the lexicon.

(2.9) Yucatec Maya (Lucy 1992: 74, cited in Lucy 2000: 329)

A `un-tz’íit há’as
one-CL:1D banana
“one banana fruit”

B `un-wáal há’as
one-CL:2D banana
“one banana leaf”

C `un-kúul há’as
one-CL:PLANTED banana
“one banana tree”

D `un-kúuch há’as
one-CL:LOAD banana
“one bunch of bananas”

E `um-p’íit há’as
one-CL:BIT banana
“one bit of a banana”


A myi? to ya?
river one CL:PLACE
“one river (e.g. destination for a picnic)”

B myi? to tan
river one CL:LINE
“one river (e.g. on a map)”

C myi? to hmwa
river one CL:SECTION
“one river (e.g. a fishing area)”

D myi? to ’sin
river one CL:DISTINCT.ARC
“one river (e.g. a path on the sea)”

E myi? to ’kwin
river one CL:CONNECTION
“one river (e.g. tying two villages)”

F myi? to ’pa
river one CL:SACRED.OBJECT
“one river (e.g. in mythology)”

G myi? to khü
river one CL:CONCEPTUAL.UNIT
“one river (e.g. in a general discussion on rivers)”
2.1. Functionality

H myi? to myi?
river one CL:RIVER
“one river (unmarked)”

“The distinction between subcategorizing undifferentiated lexical items and creating new lexical items is not always easy to draw” (Contini-Morava and Kilarski 2013: 273): Cf. again 2.5, which is a reported example for the derivation of new lexical items by classificatory markers. Example 2.11 is a reported case of classificatory variation. In terms of form, there is no visible difference between the derivational and the classificational process. Therefore, S₁ (expansion of the nominal lexicon by productive derivation) and S₂ (referent differentiation by variable classification) are fused in this study.

(2.11) Great Andamanese (cf. Abbi 2011: 764)

A ot=cala, er=cala, og=cala
   CL:4=scar CL:2=scar CL:3=scar
“scar left by arrow-head, scar on the head, scar on the limbs”

B e=tei, ot=tei, og=tei
   CL:5=blood CL:4=blood CL:3=blood
“blood inside the body, blood outside the body, blood on/from finger”

In fact, the available data of many languages even does not allow at all to draw a reliable conclusion on whether their systems employed derivational or classificational reference, whereas others seem to employ both. Since systems of a certain language type are not even able to serve S₁ (i.e. systems of isolating languages or systems that employ formwise independent classificatory markers), the two criteria are fused into a single one, which applies to all kinds of systems and language types found in the sample. The fused criterion S₁/₂ thus indicates whether a system serves the function of referent differentiation by either derivation or variable classification.

Ad S₃: Individuation is found in numeral classifier systems as well as in other kinds of classifier systems and in gender/noun class systems. In classifier systems, an underspecified lexical item is turned into a countable unit and thus ‘individuated’ by the classifier (cf. 2.9), which itself is obligatory in quantificational constructions (cf. 90).


A sb tug neeg
   one CL:ANIMATE man/person
   “one man/person”

B *sb neeg
   one man/person
   “one man/person”
Unlike earlier claims, individuation can relate to both gender/noun class and number. It is frequently involved in count/mass distinctions, where a neuter noun often refers to unbounded and uncountable referents, whereas a non-neuter noun refers to specific and countable referents (cf. 2.13 and Contini-Morava and Kilarski 2013: 276).

(2.13) Dutch (cf. Kraaikamp 2012: 216)

“the stone, the diamond, the cork (mass)”

B] de ARTDEF:C.SG steen, de ARTDEF:C.SG diamant, de ARTDEF:C.SG kerk
“the stone, the diamond, the cork (single object)”

Ad $S_4$: Nominal classification systems may allow the speaker to ascribe perceived or real properties to a referent or to highlight them in order to express his or her attitude (affection, up- or downgrade of a referent’s social status, anthropomorphism) or focus on an unusual e.g. physical property of the referent (‘cross-addressing’; cf. Contini-Morava and Kilarski 2013: 277f.). It is found in both gender/noun class systems (cf. 2.14, where the Dyirbal noun for “man” is assigned to the feminine class in order to refer to a hermaphrodite) and classifier systems (cf. 2.15, where the social status of a discourse referent varies with respect to the speaker’s intention, and 2.16, where the mountain is anthropomorphised).

(2.14) Dyirbal (cf. Dixon 1968: 111)

$balan$ here:NOM.F $yara$ man
“this hermaphrodite” (literally: “this effeminate man”)


drel, chubs-drel
“mule (plain form), mule (honorific form)”

(2.16) Jakaltek (cf. Craig 1986b: 263)

$naj$ CL:MAN witz
“the spirit of the mountain”
2.1. Functionality

**Ad D₁:** Referent identification denotes the commonly acknowledged function of reference tracking. It is assumed to be the main (and only common) function of gender/noun class systems (Corbett 1991: 322). It occurs in form of deixis, anaphora or kataphora, and disambiguation (Contini-Morava and Kilarski 2013: 279) and embodies the ability to track a discourse referent over stretches of discourse by the marking of its gender/noun class on nominal modifiers without a cooccurring noun (cf. 2.17 where the pronoun is marked for its referent noun’s class and allows a doubtless identification of the intended referent) or by the anaphoric usage of a classifier without a cooccurring noun (cf. 2.18 and 2.19 where a classifier-noun pair introduces a new referent, and the classifier serves as a carrier of anaphora in the later discourse).

(2.17) German

```
(2.17) German
Ein
Auto
stieß
mit
einem
ART:NOM.N.SG
car:NOM.N.SG
push:IPF3SG
with
ART:DAT.M.SG
Zug
zusammen.
Er
trug
erheblichen
train:DAT.M.SG
together
3SG.M
carry:IPF3SG
considerable:ACC.M.SG
Schatدن
away.
damage:ACC.M.SG

“A car collided with a train. It [the train] took considerable damage.”
```

(2.18) Hmong Daw (cf. Mottin 1978: 29)

```
(2.18) Hmong Daw (cf. Mottin 1978: 29)
Lab
tsev
tug?
Lab
CL:GENERAL
house
which
CL:GENERAL
DEM

“Which house? That one.”
```

**Ad D₂:** A noun’s classification may be related to a referent’s status in discourse. For instance, Mandarin allows the speaker to use a specific classifier for introducing new discourse referents, which are in a later context classified by the general classifier ge (Contini-Morava and Kilarski 2013: 284). Bora follows a different strategy in that a discourse referent may be introduced by a lexical noun, which is categorised by a general classifier (cf. 2.19 A). The general classifier may further be used for referent identification (cf. 2.19). Towards the end of the discourse, the general classificatory marker is changed for a specific classificatory marker (cf. 2.19 C).


```
A  zhú-pé
   te-ne
   tög-kj-hìn
yesterday-PST
3SG-CL:INANIMATE
blowgun-CL:TUBE
o
pakpixngjkhun-kš
[...]
1SG.SUBORDINATE
rasp-PURPOSE

“Yesterday I rasped the [inanimate] blowgun, [...].”

B  [...]
   a-ke
   ó
   nthu-stún-ti
   übtí-kó?
[...]
CON-CL:INANIMATE
1SG
try:out-PRD
good

“[...] and I tried it [inanimate] out: good [...].”
```
Nominal Classification

While especially classifier systems seem to be often involved in discourse management, this aspect is yet underresearched. In many cases, their involvement in discourse management is merely mentioned and complemented by the comment that this aspect of a given system demands yet more research. Many descriptions of nominal classification systems allow no conclusion on whether or not they employ $D_2$ (reference management), since their documentation is too sparse. Others allow only for an intuitive valuation; since both cannot constitute a reliable basis for a comparative study from which can be generalised, $D_2$ was not adopted for this study.

**Ad $D_3$:** By re-presenting the referent, the speaker changes its perspective and adds information to it. This is achieved by a noun’s association with a given classificational marker; at a later point, the noun is recategorised. This is the case e.g. in grammatical vs. semantic agreement in gender/noun class systems (Contini-Morava 2013: 287; cf. 2.20).


\[
\text{Ich suchte ein kleines Auto, aber elegant und schick sol} \text{te es sein [\ldots] Da stand er [\ldots]}\]

“I was looking for a small car – but it had to be elegant and chic [\ldots] There he stood [\ldots]”

$D_3$ is ignored here, since its value is mainly determined by an evaluation of $S_{1/2}$, $S_3$, and $S_4$. Language documentations often do not provide information on whether a system that allows for a flexible classification actually makes use of the re-presentation of a nominal referent. The reduced set of functional properties of nominal classification devices that was applied to this study is summarised in table 2.2. Note that the individual functional properties are relabeled for this study and henceforth are referred to by $C_1$ (“first criterion”) – $C_4$ (“fourth criterion”).
Table 2.2: The reduced set of semantic and discourse properties of nominal classification.

<table>
<thead>
<tr>
<th>semantics</th>
<th>discourse</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1/2 = C_1$ referent differentiation (derivation, variable classification)</td>
<td>$D_1 = C_4$ reference identification (anaphora, deixis, disambiguation,...)</td>
</tr>
<tr>
<td>$S_3 = C_2$ individuation (interaction with number)</td>
<td></td>
</tr>
<tr>
<td>$S_4 = C_3$ attribution of properties (speaker-sided)</td>
<td></td>
</tr>
</tbody>
</table>

Recall that all of the functions that apply to nominal classification can be expressed by linguistic means that are different from nominal classification. Example [2.21] provides instances of languages that employ nominal classification devices, but express the individual functions from table [2.2] by e.g. lexical means or the categories person and number: Cf. [2.21] A (\(\sim C_1\)), where referent differentiation is achieved by the use of adjectives, which are no means of nominal classification, but a lexical device. In [2.21] B (\(\sim C_2\)), the unbounded referent of the noun \(\text{Salz} \) “salt” is specified by the measure noun \(\text{Prise} \) “pinch”, which is not a classifier; this is recognisable by e.g. the plural marking it takes, which is not an option available for nominal classifiers. In [2.21] C (\(\sim C_3\)), a judgemental attribution of pejorative values is achieved by the insertion of a lexical noun that refers to the opposite sex of the addressed ones; this is a common strategy in German, which does not allow its lexical nouns to switch in gender/noun class according to the speaker’s discoursal needs. Cf. again [2.14] for an example from Dyirbal, a gender/noun class system which allows for such a shift in gender/noun class in order to achieve the same effect. In [2.22] (\(\sim C_4\)), it is the inflectional affix of the verb, which allows for reference tracking due to its singular value.

(2.21) German

**A**  
Für welche Pille entscheidest Du Dich? Die *rote* oder die *blaue*?  
“Which pill are you choosing for? The red one or the blue one?”

**B**  
eine Prise Salz, zwei Prisen Salz  
“one pinch:SG salt two pinch:PL salt”

**C**  
Wenn ihr *Ladies* meine Insel verlasst, [...] *seid ihr* Priester des Todes und betet um Krieg.  
“If you ladies (addressee: male recruits) leave my island, [...] you will be a minister of death praying for war.”
Nominal Classification

(2.22) Latin

Cesar:NOM.M.SG et legiones:NOM.F.PL Rubicon:ACC.M.SG
transierunt. Sic creauit:PF-3SG proverbum:ACC.N.SG
cognitum.
popular:ACC.N.SG

“Cesar and the legions transgressed the Rubicon. This way, he created the well-known proverb.”

The same effect can be achieved by means of nominal classification, if a pronoun is inserted, the gender value of allows to identify Cesar (and not his legions) as the intended discourse referent (cf. [2.23]).

(2.23) Latin

Cesar:NOM.M.SG et legiones:NOM.F.PL Rubicon:ACC.M.SG
transierunt. Ill-e sic creauit:PF-3SG proverbum:ACC.N.SG
cognitum.
popular:ACC.N.SG

“Cesar and the legions transgressed the Rubicon. He created the well-known proverb this way.”

This functional overlap of nominal classification and other grammatical means demands a very careful analysis of supposed nominal classification systems: There are linguistic phenomena different from nominal classification, which formally resemble e.g. classifier systems, employ an underlying cognitive classification, and may even serve functions which are typically attributed to nominal classification devices, e.g. reference tracking. This has led to these phenomena being mingled together with nominal classification devices, though they do not serve nominal classification, as I show in the course of this chapter. The functional typology of Contini-Morava and Kilarski (2013) serves as a filter for candidate systems for this study’s sample: All systems that are included employ at least one of the presented functions. Systems that do not employ at least one of the presented functions are discarded as not falling under the scope of nominal classification. The functional setup of the sample languages is addressed more in detail in section [4.3]. The next section presents a theoretical and typological background of classifier systems.

2.2 Classifier Systems

This section is dedicated to the introduction of a general and theoretical background on classifier systems (CLSs), which have found a broader interest of
study in the second half of the 20th century (Kilarski 2013: 180). This section introduces various aspects of classifier systems in the following order: the frequency and location of classifier systems, their various definitions in the literature, and their formal and functional properties. Beyond that, the nature of their classification is discussed and their different systemic types introduced, followed by an introduction of their common diachronic sources, and a concluding definition that serves as the working definition of this study.

**Distribution** The distribution of classifier systems is related to both the genetic setup of a given language — isolating languages tend to employ classifier systems rather than gender/noun class systems — and the character of the linguistic area where the language is spoken (Dixon 1986: 109); classifier systems are thus frequently found in (South) East Asia, the Americas, and also in Africa and Australia.

**Earlier definitions** An early and basic definition of classifier systems stems from Allan (1977: 285), where classifiers are rendered as morphemes which occur in surface structures under specifiable conditions (i.e., adjacent to a specific and predictable morphosyntactic host) and that carry semantics, i.e. the classifier renders a salient perceived, inherent, or imputed property of the nominal referent. Dixon (1986) offers a first typological distinction between classifier and gender/noun class languages, which takes the form of characteristic properties of the respective systems. Dixon’s package of properties of classifier systems includes the following aspects (cf. Dixon 1986, Grinevald 2000: 62, and Rijkhoff 2002a: 74): They are independent linguistic elements, which are not fused with other categories, and their number is typically largish. They occur once, not multiply, and are either morphosyntactically independent or clitics. Classifiers do not classify all nouns in all contexts; their scope may be limited to parts of the nominal lexicon (e.g. alienable nouns) or certain constructional types (e.g. quantificational ones). Beyond that, they mainly operate in the domain of semantics and pragmatics, and their use may trigger changes in the linguistic register. Finally, classifiers are assigned to a noun variably at the speaker’s will or discoursal demands.

Aikhenvald (2000: 13) represents a more recent approach towards systems of nominal classification than Dixon. She defines classifier constructions as morphosyntactic units that are characterised by the requirement of a specific morphemic type (the classifier), whose choice is determined by semantic characteristics of the NP referent. Lucy (2000: 330) adds with respect to the semanticity of classifiers that they are semantically potent in the sense that they carry semantic features beyond the ones of the noun that is categorised (cf. 2.2.4 where the classifier adds specific semantic information to the general semantics of the noun).
Nominal Classification

(2.24) Bora (cf. Weber 2002: 3)

A müttsʰtsʰʰía -pá
pear.apple CL:FRUIT
“pear apple fruit”
B müttsʰtsʰʰía -ʔë
pear.apple CL:TREE
“pear apple tree”
C müttsʰtsʰʰía -pâhì
peear.apple CL:GROVE
“pear apple grove”

More recently, Contini-Morava and Kilarski (2013: 266) define classifiers as all means of nominal classification that are not gender/noun class systems; they lack agreement, do occur on the classified noun itself, are assigned semantically or lexically, constitute a largish or even open set, and may serve the variable presentation of the referent by highlighting different aspects about it.

Seifart (2010: 721) also notes that classifier systems employ large sets of classifiers, lack agreement, and are assigned semantically. These various definitions are largely consistent with each other and serve as a basic understanding for the following description of formal and functional features of classifiers, before the working definition for this study is phrased at the end of this section.

Formal properties  Classifier systems are fairly simple and consistent with respect to the formal aspects they have in common across their diverse subtypes. The classifier occurs as an independent or clitic element overtly within an NP and occurs only once. It is the head of their host NP which is classified (cf. 2.25 2.29).


ma- pëna- si- na yena
DEM- CL:ROTTEN- PL- DEM fish
“these rotten fish”


ʔon ʔai khwan thiat phoon lmy
I have child four CL:PERSON
“I have four children.”

---

4 The individual systems may show a sophisticated formal behaviour or function; what is referred to as ‘simplistic’, are those formal aspects that they do have in common.

5 The only exception from NP-internal occurrence of the classifier are the supposed ‘verb classifier’ systems, which attach the classifier to the verb. These systems are introduced and described after the general description of form, function, and classificatory nature of classifiers. Note that I argue ‘verb classifier’ systems not to exist (cf. Passer 2015).
2.2. Classifier Systems

(2.27) Panare (Mattei-Muller 1974, cited in Grinevald 2000: 66)

\[ \text{yu} -\text{kon} \text{ ka’ka} \]

\[ \text{GEN/1} \quad \text{CL:WEAPON} \quad \text{bow} \]

“my bow”

(2.28) Tariana (Aikhenvald 2000: 94)

\[ \text{kuda} \quad -\text{ma} \quad -\text{maka} \]

\[ \text{garment} \quad \text{CL:F} \quad \text{CL:CLOTH.LIKE} \]

“woman’s garment”


\[ d\text{âw} \quad \text{tuúm} \]

\[ \text{CL:HUMAN} \quad \text{eye} \]

“human eye”

The occurrence of a classifier is predictable in terms of the specific construction where it occurs: The main distinctive criterion for the different subtypes of classifier systems is the morphosyntactic host or neighbour of the classifier (cf. again (2.25)–(2.29) and e.g. Seifart 2010: 721). For instance, classifier systems that occur with quantifiers or numerals as their main host are referred to as numeral classifier systems (2.26), systems employing the noun itself as the main host are called noun classifier systems (2.29), and possessives as the main host of classifiers are generally labeled possessive classifier systems (2.27). Besides for those prototypical systems that employ their classifier set in just one morphosyntactic environment (thus with only one type of morphosyntactic host such as possessives or nouns or quantifiers), it is also not unusual for a classifier system to expand the range of morphosyntactic classifier hosts. These systems are referred to as “multiple classifier systems” (147). Illustrates this phenomenon in Mandarin, where classifiers can occur with a quantifier, the bare noun, and demonstratives.

(2.30) Mandarin (cf. Sun 2006: 159f.)

A nà yí-pí hēi mă

\[ \text{DEM:REMOTE} \quad \text{one-CL:HORSE} \quad \text{black horse} \]

“That black horse.”

B na pí hēi mă

\[ \text{DEM:REMOTE} \quad \text{CL:HORSE} \quad \text{black horse} \]

“That black horse.”

C măi pí hēi mă

\[ \text{buy} \quad \text{CL:HORSE} \quad \text{black horse} \]

“Buy one black horse.”
Functional properties  Classifiers diverge with respect to their functionality. Their primary functional mechanism is the categorisation of their nominal referent, which is the referent of the noun that is categorised. The categorisation mainly occurs in terms of inherent properties that the speaker perceives as inherent, salient, or characteristic for the discourse context at hand. This includes a categorisation by generic categories such as “person” or “animal” (cf. 2.31 A), but also the common categorisation by shape, size, consistency, function, animacy, or structure (cf. 2.31 B–D and e.g. Aikhenvald 2000: 98 & 115). In classifier languages with a (semi-)open classifier inventory, it is also possible to employ lexical nouns as their own classifiers, which are then called “repeaters” (cf. 2.32).


A  yi  zhi  laohu,  yi  zhi  niao,  yi  zhi  pangxie
one  CL:ANIMAL  tiger  one  CL:ANIMAL  bird  one  CL:ANIMAL  crab
“one tiger, one bird, one crab”

B  yi  tiao  she,  yi  tiao  kuzi,  yi  tiao  he
one  CL:LONGISH  snake  one  CL:LONGISH  pants  one  CL:LONGISH  river
“a snake, a pair of pants, a river”

C  yi  jian  chenshan,  yi  jian  diyi
one  CL:CLOTHING  shirt  one  CL:CLOTHING  overcoat
“a shirt, an overcoat”

D  yi  zhi  lanzi,  yi  zhi  qiqiu,  *yi  zhi  boluo
one  CL:CONTAINER  basket  one  CL:CONTAINER  balloon  one  CL:CONTAINER  pineapple
“one basket, one balloon, one pineapple”

(2.32) Thai (Hundius and Kölever 1983: 190, cited in Aikhenvald 2000: 103)

prathet  sāam  prathet
land  three  CL:LAND
“three countries”

The semanticity of classifiers allows a threefold functionality (Contini-Morava/Kilarski 2013: 268–278): They serve as a means to expand the nominal lexicon (cf. 2.33), they may differentiate a distinct referent of an otherwise semantically neutral noun and ‘individuate’ nouns, i.e. mark them as countable objects (cf. 2.9 A and B = 2.34), and they may attribute speaker-sidedly perceived properties to a referent (cf. 2.10 C and F = 2.35 where the speaker chooses for a certain aspect about the nominal referent, which he or she then highlights by the insertion of a specific classifier which expresses these aspects).
2.2. Classifier Systems

(2.33) Jakaltek (cf. Craig 1986a: 265 and 273)

A  *na’*  
  *CL:ANIMAL*  *sasim*
  “animal liver”

B  *ca-b’*  
  *two-PL:INANIMATE*  *na’*  *xila*
  “two saddles” (inanimate object made of animal product)

(2.34) Yucatec Maya (Lucy 1992: 74, cited in Lucy 2000: 329)

A  *’un-ts’ít*  *há’as*
  *one-CL:1D*  *banana*
  “one banana fruit”

B  *’un-wáá*  *há’as*
  *one-CL:2D*  *banana*
  “one banana leaf”


A  *myi*  *ta*  *hmwa*
  *river*  *one*  *CL:SECTION*
  “one river (e.g. a fishing area)”

B  *myi*  *ta*  *’pa*
  *river*  *one*  *CL:SACRED.OBJECT*
  “one river (e.g. in mythology)”

This core function of classifiers may bleach with the increasing grammaticalisation of the system. The combination of a given classifier and noun may conventionalise and lexicalise, until a flexibility of the assignment is not anymore given. This may affect only a part of the system, e.g. the most frequently used combinations, while others are not (yet) conventionalised. This is the case in Mandarin, where certain classifier-noun pairs in (2.36) have conventionalised and lexicalised and thus cannot be varied.

(2.36) Mandarin (Zhang 2007: 48 and Lin, p.c.)

A  *yu*  *wan*  *yueliang*
  *one*  *CL:CURVED*  *moon*
  “a crescent moon”

B  *yi*  *lun*  *(zīn)*  *yuè*
  *one*  *CL:ROUND*  *new*  *moon*
  “a (new) moon”

C  *yi*  *lun*  *(măn)*  *yuè*
  *one*  *CL:ROUND*  *full*  *moon*
  “a (full) moon”
A system may also adopt other functions, which fall into the domain of discourse. These additional functions include referent identification by anaphoric or deictic means, ‘re-presentation’, i.e. the introduction of new perspectives on a referent, and the contribution to the discourse-internal reference management through a classifier-sided manipulation of the discourse status, persistence, or discourse prominence (cf. Contini-Morava and Kilarski 2013: 278–290). The most salient function is referent identification by anaphoric reference of the classifier. A referent is introduced to the discourse by a classifier-noun-pair and in the further discourse, the classifier may occur and denote the referent in question on its own (cf. 2.1 – 2.37).

(2.37) Hmong Daw (cf. Mottin 1978: 29)

<table>
<thead>
<tr>
<th>Lab</th>
<th>tsev</th>
<th>tug?</th>
<th>Lab</th>
<th>no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL:GENERAL</td>
<td>house</td>
<td>which</td>
<td>CL:GENERAL</td>
<td>DEM</td>
</tr>
</tbody>
</table>

“Which house? That one.”

Classificatory nature As indicated earlier, it is not originally the noun as a linguistic item that is categorised by a classifier, but rather the extralinguistic referent that a noun refers to. Only by conventionalisation, such a categorisation may give rise to reference patterns that classify linguistic form rather than experience patterns (Lucy 2000: 326ff.; cf. also Sefdt 2000: 27). An example for a partly conventionalised system that does not categorise the referent semantically, but employs a ‘correct’ classifier for a given noun, is Mandarin. Cf. 2.38 for an invariable classification of linguistic form: The noun ku(zi) “pants” always takes the classifier tiao CL:LONGISH; this classifier-noun pair has been conventionalised. Compounds that take ku(zi) “pants” as their compositional head take the same classifier independent of their actual shape and physical properties.

(2.38) Mandarin (cf. Zhang 2007: 47)

A yi tiao neku | one CL:LONGISH underwear 
“one pair of underwear”

B yi tiao sanjiaoku | one CL:LONGISH brief 
“one pair of briefs”

C yi tiao youyongku | one CL:LONGISH swimming.trunk 
“a swimming trunk”

D yi tiao duanku | one CL:LONGISH shorts 
“a pair of shorts”

6Note that these functions are not primary and, depending on their saliency within a given system, may indicate a system’s way out of the realm of nominal classification. For this reason, they are not treated in detail here, but will only be discussed in the specific cases from the sample where secondary functions occur.
2.2. Classifier Systems

Classifier types

The literature offers many terms for specific kinds and types of classifiers, which are briefly introduced and discussed in the following. Classifiers can be grouped in two types, those which constitute individual classifier systems (e.g. numeral, noun, possessive classifiers), and those which render the classifier’s classificatory quality (e.g. sortal, mensural, and proper classifiers). The ‘quality’ classifiers are introduced first and the ‘systemic’ classifiers afterwards. The quality distinction will only play a minor role in this study; the main distinction made here is the one of systemic classifiers.

The ‘quality’ classifiers consist of three pairs that are fairly alike in terms of function respectively: Classifiers proper vs. quantifiers, sortal vs. mensural classifiers, and class vs. measure nouns. Senft (2000: 21ff.) distinguishes classifiers proper, which designate and specify features that are semantically inherent to the nominal referent and thus are rather invariable (cf. wood(en), long), and quantifiers (or quantificational classifiers), which temporarily quantify nominal referents and thus are rather variable (cf. group, flock, bunch). Furthermore, he distinguishes sortal classifiers, which serve the function of individuation of a noun with respect to the kind of entity that it is (cf. animal, tree, rock), and mensural classifiers, which indicate quantities (e.g. handful, basket, dozen). Finally, measure and class nouns are frequently brought in relation with classifiers, though these are not classifiers, since they are also found in languages with gender/noun class systems and even in languages without nominal classification devices. They represent the main lexical source for classifiers in that measure and class nouns may grammaticalise to mensural and sortal classifiers respectively (cf. Grinevald 2000: 58ff. and figure 3.5 on page 69). Measure nouns serve the expression of quantities; they divide mass terms in non-quantitative sortal units (e.g. a dash of strychnine) and count nouns in quantitative measures or arrangements (a cloth of cotton; cf. Grinevald 2000: 58). Class terms are morphemes that take part in a language’s lexicogenesis and put the items in certain categories (e.g. police-man, fire-man, chair-man; straw-berry, black-berry, blue-berry etc.; cf. ib.: 59).

The ‘systemic’ classifiers are mainly distinguished by their morphosyntactic host, as indicated earlier. This study distinguishes numeral classifiers, which occur in quantificational constructions (cf. 2.26 = 2.39), noun classifiers, which occur adjacent to the noun they categorise (cf. 2.25 = 2.40), and possessive classifiers, which occur in possessive constructions (cf. 2.27 = 2.41).


\[
?an \ ?ui \ khwan \ that \ phoon \ lny
\]

1 have child four CL:PERSON

“I have four children.”


\[
ma-\ \ pemoa- \ si- \ na \ yena
\]

DEM- CL:ROTTEN- PL- DEM fish

“These rotten fish”
(2.41) Panare (Mattei-Muller 1974, cited in Grinevald 2000: 66)

```
yu -kon ka’ka
GEN/1 CL:WEAPON bow
```

“my bow”

Figure 2.1 displays the types of classifier systems that are commonly found in the literature. Here, they are grouped into numeral and noun classifiers, possessive classifiers (possessed and relational classifiers), ‘minor’ classifiers (locative and deictic classifiers), and verb classifiers (affixed and incorporated classifiers) according to the descriptions in Aikhenvald (2000) and Grinevald (2000). As mentioned earlier, systems that allow their classifiers to attach to more than one morphosyntactic host are referred to as “multiple (classifier) systems”. For instance, the same Mandarin classifier occurs with different morphosyntactic hosts in 147 = 2.42 which renders Mandarin a multiple classifier system.

(2.42) Mandarin (cf. Sun 2006: 159)

```
A nà yì-pí hēi mā
DEM:REMOTE one-CL:_HORSE black horse
“That black horse.”

B na pí hēi mā
DEM:REMOTE CL:_HORSE black horse
“That black horse.”

C māi pí hēi mā
buy CL:_HORSE black horse
“Buy one black horse.”
```

Note that this study only considers noun, numeral, and possessive classifiers to be classifiers ‘proper’; the motivation for the exclusion of the other classifier types is provided in the following section.

Numeral classifiers are the most common and well-researched classifier type and thus are included in this study. They are well-attested as an areal phenomenon in (South) East Asia and Oceania and South America and frequently found in North American indigenous languages and the Indian subcontinent.
They are not commonly found in Australia, Central and Western Asia, Europe, and Africa (cf. Aikhenvald 2000: 121ff.).

Noun classifier systems are also object of this study; they are commonly found in Mesoamerica, South East Asia, and Australia, but occur seldom or not at all in Europe, Africa, North and South America, and Mainland Asia (cf. Aikhenvald 2000: 96ff.). Noun and numeral classifiers are the textbook examples of classifiers: they occur adjacent to the noun and a quantifier respectively and in the form of morphosyntactically independent or clitic elements and classify the nominal referent with respect to its inherent properties (cf. 2.43 and 2.44).

(2.43) Jakaltek (cf. Craig 1977: 140)

\[
\text{xcam \ no'} \text{\ hìn \ cheh} \\
\text{died \ CL:ANIMAL \ my \ horse}
\]

“My horse died.”

(2.44) Nùng (cf. Saul and Wilson 1980: 14)

\[
\text{lẹọ \ hään \ tű \ má' \ lụhc} \\
\text{then \ kill \ two \ CL:ANIMATE \ dog \ child}
\]

“Then kill two puppies.”

Possessive classifiers are common in Oceanic languages and in North and South America; they are not common in Australia, Africa, Europe, and Mainland Asia (cf. Aikhenvald 2000: 147f.). They are said to distinguish two functional variants, on the one hand possessed classifiers, which categorise the possessed with respect to its inherent properties (cf. 2.45 A, where the inherent property of the shape (of the referent of “head”) serves as classificational basis, and Aikhenvald 2000: 126–132), and ‘relational classifiers’, which categorise the relation between possessor and possessed (cf. 2.45 B and C, where the classifier renders the relationship between possessor and possessee, and ib.: 133–138).

(2.45) Baniwa (Aikhenvald 2000: 143)

\[
\text{A (nu-hwida) \ nu-dza-da} \\
\text{1SG-POSS-CL:ROUND}
\]

“The head (lit. ‘my head’) is mine”

\[
\text{B \ nu-tfimn-\\n} \\
\text{1SG-dog-POSS.1}
\]

“my dog (the one I brought up; close relation of possessor and possessed)”

\[
\text{C \ nu-tfimn-te} \\
\text{1SG-dog-POSS.2}
\]

“my dog (the one I found; less close relation of possessor and possessed)’

This latter categorisation of the relation between possessor and possessed is unique among classifiers (cf. Aikhenvald 2000: 133), which in all other cases categorise the nominal referent with respect to its inherent characteristics. The notion of relationality includes also a not further specified classification of function, such as the different ways in which food is processed (cf. 2.46).

A  ya  -án  Pen  wuur  
   CL:GENERAL  POSS.3SG  Ben  banana
   “bananas raised, owned, or given by Ben”

B  wor  -áy  Pen  wuur  
   CL:FOODraw  POSS.3SG  Ben  banana
   “raw bananas eaten by Ben”

C  yán  -án  Pen  wuur  
   CL:FOODcooked  POSS.3SG  Ben  banana
   “cooked bananas eaten by Ben”

However, the data does not support this assumption, and supposed ‘relational’ classifiers prove to be actual possessed classifiers, as section 2.3 demonstrates. Next to ‘relational’ classifiers, ‘verb’ classifiers and minor classifier types, are discussed and excluded from this study in section 2.3.

Sources of classifier systems  Classifiers are quite uniform with respect to their etymological origin: They virtually all grammaticalise from lexical nouns. They are either freely distributed or from a subset of the nominal lexicon, such as body part terms in Mundurukú — (Aikhenvald 2000: 353 and 366) and frequently gain a more abstract quality than their lexical source possesses. Classifiers are in general assumed to be transparent with respect to their lexical source (cf. table 2.3).


<table>
<thead>
<tr>
<th>classifier</th>
<th>lexical source</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ta</td>
<td>tā “single, unit”</td>
</tr>
<tr>
<td>-ten</td>
<td>tan “body”</td>
</tr>
<tr>
<td>-jeld</td>
<td>jald “volume”</td>
</tr>
<tr>
<td>-dune, -dane</td>
<td>dānah “grain”</td>
</tr>
<tr>
<td>-rā’s, -sar</td>
<td>sar “head”</td>
</tr>
</tbody>
</table>

According to Aikhenvald (2000: 362–365), certain kinds of verbs (e.g. verbs of handling, posture, or motion) can also give rise to classifiers, namely verb, deictic, and some possessive and numeral classifiers. Deverbal nominalisation also may give rise to possessed, relational, and sometimes numeral classifiers (ib.: 365). The interested reader may refer to Aikhenvald (2000: 352–412) for a more in-depth treatment on the sources and pathways of dissolution of noun classification devices.
2.2. Classifier Systems

It is not only the case that certain classificational systemic types prefer certain lexical areas as a source for their markers, but there are also area-specific, language-, and family-specific preferences (cf. also Aikhenvald 2000: 353). The most common areas are body parts, denotations of highly animated entities (kinship terms, terms for humans and higher animates), generics (cf. also Sands 1995: 252), measure nouns (or ‘unit counters’), and culturally salient objects. Repeaters also play an important role in coining classifiers (cf. Sands 1995: 253 and Aikhenvald 2000: 361). By their productive use as “ad hoc classifiers” (Aikhenvald 2000: 361), they can conventionalise into ‘common’ classifiers (numeral, deictic, possessed, relational, incorporated verb classifiers, and concordial classes, but no closed grammaticalised concordial class systems; cf. Aikhenvald 2000: 362), which show expected phenomena that are related to grammaticalisation (such as the semantic expansion, phonological erosion, cliticisation, etc.).

Definition After the examination of earlier descriptions and studies on classifiers and their properties, we are able to define the notion of “classifier” for this study. The choice of criteria was based on the earlier introduced definitions and descriptions of classifier systems, which were tested against my own language sample with respect to their validity and reliability. According to this preliminary study, prototypical classifier systems have the following properties:

- Their markers, labeled “classifiers”, occur as independent or clitic markers.
- Classifiers occur once and overtly within the borders of the NP, the overt head noun of which they categorise.
- Classifiers serve the categorisation or individuation of the nominal referent, as long as the system has not conventionalised or lexicalised given classifier-noun-pairings.
- Classifiers are assigned by the speaker according to his or her discoursal demands, as long as the system has not conventionalised or lexicalised given classifier-noun-pairings.
- The same noun may take a range of different classifiers according to the speaker’s discoursal demands, as long as the system has not conventionalised or lexicalised given classifier-noun-pairings.
- Classifiers contribute to their host NP’s semantics, as long as the system has not conventionalised or lexicalised given classifier-noun-pairings.
- Classifiers occur in predictable morphosyntactic constructions. The predictability is achieved by a given morphosyntactic host of the classifier (i.e., a numeral, possessive, or a bare noun).
- Classifier systems generally employ a largish set of categories (i.e. more than two dozen classifiers).
• A classifier system’s scope may not cover the whole nominal lexicon.

• Due to their employment only in given morphosyntactic constructions within a language, nouns remain uncategorised outside those constructions.

Furthermore, but not obligatorily, classifiers tend to be etymologically transparent; they may also be involved in the activation or satisfaction of given language-internal linguistic registers. In the following section, the supposed classifier types of ‘relational’, ‘verb’, and ‘minor’ classifiers are presented, discussed, and excluded from this study.

2.3 Exclusion of ‘Verb’ and ‘Relational Classifiers’ and ‘Minor Types’

A number of postulated classifier types are not considered in this study. This section discusses these types (‘relational’, ‘verb’, and ‘minor’ classifiers) in separate subsections and provides the respective arguments for their exclusion.

2.3.1 ‘Relational Classifiers’

As I argue in Passer (2016) and elaborate in the following, the category of relational classifiers does not exist and thus is not adopted for this study. Supposed relational classifiers in some cases turn out to be ‘mere’ possessive markers, which distinguish inalienable possession from alienable possession (cf. 2.47 and 2.45) but evidently do not classify a noun or its referent. In other cases, the supposed ‘relational’ classifiers turn out to be possessed classifiers, which categorise a noun’s referent with respect to inherent characteristics (cf. 2.46, 2.49, and 2.50, where the semantic categories of the classifiers denote inherent properties and not a possessive relation).

(2.47) Hawaiian (cf. Lichtenberk 1983a: 162)

\[
\begin{align*}
\text{A} & \quad k & \quad -o & \quad \text{-na} & \quad \text{lima} \\
& \quad \text{ART} & \quad \text{INAL} & \quad \text{POSS.3SG} & \quad \text{hand} \\
& & & & \quad \text{“his hand”}
\end{align*}
\]

\[
\begin{align*}
\text{B} & \quad k & \quad -\tilde{a} & \quad \text{-na} & \quad \text{keski} \\
& \quad \text{ART} & \quad \text{\ddot{A}LI} & \quad \text{POSS.3SG} & \quad \text{child} \\
& & & & \quad \text{“his child”}
\end{align*}
\]


\[
\begin{align*}
\text{A} & \quad \text{nu} & \quad \text{-limu} & \quad \text{-ni} \\
& \quad 1SG & \quad \text{dog} & \quad \text{INAL} \\
& & & \quad \text{“my dog (the one I brought up)”}
\end{align*}
\]
2.3. Exclusion of ‘Verb’ and ‘Relational Classifiers’ and ‘Minor Types’

B nu -fimu -te
1SG dog ALI
“my dog (the one I found)”


A tapu'o 'e- i Bogi
tobacco CL:EDIBLE POSS Bogi
“Bogi’s tobacco”

B na me- na moli
ART CL:DRINKABLE POSS.3SG citrus
“her citrus”

C a -i sele we- i Sepo
ART POSS knife CL:GENERAL POSS Sepo
“Sepo’s knife”

(2.50) Palikūr (cf. Aikhenvald and Green 1998: 460)

A gi -pig pewru/mutom
3M CL:ANIMAL domesticated dog/sheep
“his dog/sheep”

B nu -win arudiki
1SG CL:ANIMAL catch tapir
“my tapir (I caught)”

C pi -mana uwas
2SG CL:PLANT food orange
“your orange (to eat)”

D nu -amutra pilatno
1SG CL:PLANT planted banana
“my banana plant”

E nu -kankayh awayg/tino
1SG CL:CHILD boy/girl
“my son/daughter”

In Passer (2016), I demonstrate that the assumed underlying relationality is expressed by means of construction, whereas classificatory means in these environments are classifying on the basis of inherent properties of the nominal referent. This makes them possessed classifier systems. This analysis is supported by experimental studies on possessive classification in Vanuatu languages (cf. Franjieh 2012, 2014, 2015, 2016) and is exemplified with the case of Kilivila, which marks four different possessive categories (cf. 2.51), which are labeled “possessive pronouns” (Senft 1986: 47) or “relational classifiers” (Aikhenvald 2000: 136f.).


A kapyu -te tu
my CL:SERIES1 yam
“my yams (to eat)”
Nominal Classification

B agu  tetu
my.CL:SERIES2  yam
“my yams (planted in the garden)”

C ula  tobaki
my.CL:SERIES3  tobacco
“my tobacco”

D kaba  -gu
bed  my
“my bed”

Note that the categories distinguish three independent forms of markers (2.51 A–C) and an affixed form (2.51 D). The analysis as a quadripartite set of pronouns and the analysis as a quadripartite set of classifiers does not capture the nature of possessive marking in Kilivila properly, since the markers can be segmented into pronominal and classificatory elements. Table 2.4 lists the markers’ paradigms for all four categories.

Table 2.4: ‘Possessive pronoun’ series in Kilivila (cf. Senft 1986: 48–52).

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>ka-gu</td>
<td>a-gu</td>
<td>ula</td>
<td>-gu(-)</td>
</tr>
<tr>
<td>2SG</td>
<td>ka-m</td>
<td>a-m</td>
<td>m</td>
<td>-m(-)</td>
</tr>
<tr>
<td>3SG</td>
<td>ka-la</td>
<td>a-la</td>
<td>la</td>
<td>-la(-)</td>
</tr>
<tr>
<td>DLincl</td>
<td>ka-da</td>
<td>a-da</td>
<td>da</td>
<td>-da(-)</td>
</tr>
<tr>
<td>DLexcl</td>
<td>ka-ma</td>
<td>a-ma</td>
<td>ma</td>
<td>-ma(-)</td>
</tr>
<tr>
<td>1Plincl</td>
<td>ka-dasi</td>
<td>ka-desi</td>
<td>dasi</td>
<td>-dasi(-)</td>
</tr>
<tr>
<td>1Plexcl</td>
<td>ka-masi</td>
<td>ka-mesi</td>
<td>masi</td>
<td>-masi(-)</td>
</tr>
<tr>
<td>2PL</td>
<td>ka-mi</td>
<td>a-mi</td>
<td>mi</td>
<td>-mi(-)</td>
</tr>
<tr>
<td>3PL</td>
<td>ka-si</td>
<td>a-si</td>
<td>si</td>
<td>-si(-)</td>
</tr>
</tbody>
</table>

A comparison with the pronominal paradigm of Kilivila shows that the forms of category III and IV are pronominals, just like the suffixed elements in category I and II (cf. table 2.5).

Table 2.5: Bound pronominals in Kilivila: the possessive category IV and emphatic pronouns (cf. Senft 1986: 52 and 54f.).

<table>
<thead>
<tr>
<th></th>
<th>possessive category IV</th>
<th>emphatic pronouns I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>-gu(-) “my”</td>
<td>titole-gu “myself”</td>
</tr>
<tr>
<td>2SG</td>
<td>-m(-) “you”</td>
<td>titole-m “yourself”</td>
</tr>
<tr>
<td>3SG</td>
<td>-la(-) “his, her”</td>
<td>titole-la “himself, herself”</td>
</tr>
<tr>
<td>DLincl</td>
<td>-da(-) “our”</td>
<td>titole-da “ourselves”</td>
</tr>
<tr>
<td>DLexcl</td>
<td>-ma(-) “our”</td>
<td>titole-ma “ourselves”</td>
</tr>
<tr>
<td>1Plincl</td>
<td>-dasi(-) “our”</td>
<td>titole-dasi “ourselves”</td>
</tr>
<tr>
<td>1Plexcl</td>
<td>-masi(-) “our”</td>
<td>titole-masi “ourselves”</td>
</tr>
<tr>
<td>2PL</td>
<td>-mi(-) “our”</td>
<td>titole-mi “yourselves”</td>
</tr>
<tr>
<td>3PL</td>
<td>-si(-) “our”</td>
<td>titole-si “themselves”</td>
</tr>
</tbody>
</table>
2.3. Exclusion of ‘Verb’ and ‘Relational Classifiers’ and ‘Minor Types’

This allows a reanalysis of Kilivila’s possessive marking, in which the construction determines the relation between possessor and possessed, and the elements ka- and a- represent classificatory means in that ka- classifies alimentary properties and a- a generic classifier (cf. 2.2). This classificatory pattern fits a widespread Oceanic pattern of possessive classification. Figure 2.2 illustrates the reanalysed model of possessive marking of Kilivila.

**Figure 2.2: Reanalysis of the possessive construction in Kilivila.**

Category IV marks inalienable possession by the most bound form of possessive marking, in which the pronominal attaches to the noun directly (cf. 2.51 D). Category three marks alienable marking by a free pronominal form (cf. 2.51 C). Both these observations match with Haspelmath’s (2006) observation that inalienable possession is always at least equally, but in most cases more grammaticalised than alienable marking. The remaining categories I and II encode alimentary and “a kind of intermediate degree of possession, intermediate between intimate [i.e. inalienable; MP] and more distant [i.e. alienable; MP] possession” (Senft 1986: 49) and thus can be analysed as classifying an inherent property (edibility) in the case of ka- and a residue category in the case of a-. All other supposed ‘relational classifier systems’ can be analysed in the same way as Kilivila or as ‘mere’ possessive markers (cf. again 2.47), which suggests that there is no evidence for the existence of a classifier type of ‘relational classifiers’. The only remaining category of possessive classifiers is thus the one labelled “possessed classifiers” by Aikhenvald (2000); henceforth, these are referred to by the term “possessive classifiers” in analogy to noun and numeral classifiers, which are labelled by their morphosyntactic host. The following subsection discusses the phenomenon of ‘verb classifiers’, which are also excluded from this study.
2.3.2 ‘Verb Classifiers’

‘Verb classifiers’ constitute a similar case as ‘relational classifiers’; they are a commonly assumed classifier type, yet the data suggests a more considerate treatment of supposed verb classifier systems. ‘Verb classifiers’ constitute an areal phenomenon in South East Papua New Guinea; beyond that, they are found as isolated systems in Australia and North and South America. They are not common in Africa, Europe, and Mainland Asia (cf. Aikhenvald 2000: 169ff.). Note that there are two terms which are used differently by different authors in order to refer to ‘verb classifiers’: the more commonly used term is verbal classifiers, the less commonly used is verb classifiers. Here, I use the term ‘verb classifiers’ in order to refer to the assumed classifier type that attaches to the verb and supposedly categorises an argument of the verb; this use is analogical to the other common classifier types of noun, numeral, and possessive classifiers. I use the term ‘verb classifiers’ in order to refer to classificatory markers which classify or manipulate the concept expressed by the verb; this use is in analogy to the concept of nominal classification. In Passer (2015), I demonstrate that a reanalysis of supposed verb classifier systems results in three possible outcomes: Multiple classifier systems, verbal classifier systems, and argument marking systems. According to this analysis, only the multiple classifier systems in question correspond to the definition of nominal classification devices without a doubt. Verbal classifier systems serve the classification and/or modification of the concept that is expressed by the predicate. Argument marking systems serve a pronoun-like function and the lexical restriction of the concept expressed by the verb. Verb classifiers as a distinct phenomenon and individual classifier type, i.e. a classifier system that attaches its classifiers to verbs and verbs only, could not be observed. Each type is addressed in the following.

Multiple classifier systems are classifier systems that allow the same set of classifiers to attach to a variety of morphosyntactic hosts, including the verb. Cf. 2.52 for an example of classifiers from Motuna, which “are combined with the bound forms of the article, the demonstrative, possessive pronouns, numerals, the [universal; MP] ‘all’ quantifier, etc.” and “with free nouns, adjectives, and verbal and participle clauses” (Onishi 1994: 162); the classifier combines with the verb in 2.52 A and B, with the article in 2.52 C, and with the noun in 2.52 D.


A ong topo inak -i -heeto -no uru
DEM.M well look.after 3O:2A FUT LINKER CL:HUMAN
“This is the one (human male) you will look after well.”

B . . . hoo kato honna rii -kui -no -wori
ART:M up big be:3SG IMAG LINK CL:ANIMATE
“...the elder (animate) one who would be big”
2.3. Exclusion of ‘Verb’ and ‘Relational Classifiers’ and ‘Minor Types’

C kai, ho -ko nommai ho -ko nupi ti
boy ART.M EMPH person ART.M EMPH my.grandchild ART
-woiri -ngo arap -ah.
CLE:ANIMATE EMPH be.bad PART

“Boy, that person, that (particular) grandchild of mine (over there) is bad.”

D ...ong -ngori Nagowisi -no -sta mi -i -na.
DEM.M LOC Nagowisi LINKER CL:SIDE go:1SG RMPST F

“...I went to the Nagowisi side.”

2.52 demonstrates the ‘regular’ behaviour of the classifiers in question (except for their NP-external occurrence): They occur once and overtly, and together with an overt noun, the referent of which is categorised by the classifier. Logic suggests that in the case of multiple classifier systems, the verbal classifier host is a byproduct of the process of categorial extension of hosts. Example 2.53 from Mundurukú, another multiple classifier language, demonstrates the usage of these classifiers as carriers of anaphora, which can be observed in other types of classifier systems as well (cf. 153).

(2.53) Mundurukú (Mithun 1986a: 381)

Ti dojot puye, o'-ti-mog ip bascy'a be.
water bring when they-CL:WATER-place in

“When they brought water, they placed it in a basin.”

(2.54) Mandarin (cf. Zhang 2007: 47)

Zhe jian wuzi you liang zheng shuzhuo, na zheng shi wode.
this CL room have two CL desk that CL be mine

“This room has two desks, that (one) is mine.”

Note that the classifiers in 2.53 and 153 are not used kataphorically, i.e. it is always the classifier-noun pair which occurs first, before the classifier may be reused for means of reference tracking. This is not the case in the other two types of outcomes, verbal classifier systems and argument marking systems, which will now be discussed. Verbal classifier systems categorise and/or manipulate the concept expressed by the verb; since these classificatory elements do not affect the semantics of a noun phrase, but the verb phrase, we are not dealing with a case of nominal classification. Cf. table 2.6 which displays the classificatory inventory of the verbal classifiers of Manam and Diegueño; the semantics of the classificatory inventory clearly does not serve the categorisation of the referent of a noun which is an argument of the verb where the classifier occurs and therefore, these systems cannot be analysed as nominal classification devices.
Table 2.6: The inventory the verbal classifier systems of Manam and Diegueño (cf. Lichtenberk 1983b: 214ff. and Langdon 1970: 80ff.).

<table>
<thead>
<tr>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manam</td>
<td></td>
</tr>
<tr>
<td>?tn-</td>
<td>“action performed with fingers”</td>
</tr>
<tr>
<td>?ara-</td>
<td>“action performed with teeth”</td>
</tr>
<tr>
<td>?oro-</td>
<td>“cutting action”</td>
</tr>
<tr>
<td>dua-</td>
<td>“action performed with sole of the foot”</td>
</tr>
<tr>
<td>tara-</td>
<td>“chopping action”</td>
</tr>
<tr>
<td>zap-</td>
<td>“hitting and breaking action”</td>
</tr>
<tr>
<td>rau-</td>
<td>“hitting action (but not necessarily breaking)”</td>
</tr>
<tr>
<td>tata-</td>
<td>“throwing and breaking action”</td>
</tr>
<tr>
<td>rova-</td>
<td>“throwing action”</td>
</tr>
</tbody>
</table>

| Diegueño   |           |
| a-         | “one long/large/hard/sharp object” |
| c-         | “indefinitely large number of small/soft/pliable/liquid objects” |
| cu-        | “(repetitive) action performed by the mouth” |
| i-         | unclear |
| k-         | “action/state broadcast in an indeterminate direction” |
| k(u/a)-    | “action performed by the foot or leg” |
| i-         | “action distorting an object out of its normal state/shape” |
| p-         | “abnormal behaviour of humans” |
| m-         | “inherent qualities” |
| n-         | “completion of an action” |
| n9-        | “(mainly) physiological states” |
| n-         | unclear (hapax legomenon) |
| p-         | “action involving the application of pressure” |
| r-         | unclear |
| s-         | “change in condition” |
| su-        | “playful activities” |
| s(u/a)-    | “action performed manually” |
| t(a)-      | “causative context” |
| tu-        | “small/solid/round object” |
| u-         | “causative action” |
| u-         | “causative action directed to a human” |
| x-         | “unclear semantic value” |
| ?(i)-      | “inherent quality” |

Those cases which are analysed as argument marking systems are the biggest part of supposed verb classifier systems. In these argument marking systems, the supposed classifiers actually serve a pronoun-like function: They occur without an overt nominal the referent of which they categorise, and there is no earlier occurrence of an overt classifier-noun pair, which introduces the classifier for a later usage as carrier of anaphora. This means that the classifiers themselves refer to the referent (cf. 2.55 and 2.56). This is not the case in ‘common’ classifier systems, where a nounless classifier construction is only possible if the classifier-noun pair was earlier introduced (cf. again 153).
2.3. Exclusion of ‘Verb’ and ‘Relational Classifiers’ and ‘Minor Types’

(2.55) Terêna (cf. Ekdahl and Butler 1979: 185, cited, glossed and translated to English in Aikhenvald 2000: 152)

\[\text{oye} \quad -\text{pu'i} \quad -\text{co} \quad -\text{ti}\]

\(\)CL:ROUND THEME PROGR

“He is cooking (round things).”

(2.56) Terêna (cf. Ekdahl and Butler 1979, cited in Derbyshire and Payne 1990: 252)

A
\[\text{movó} \quad -\text{cava} \quad \text{CL:BRANCH}\]

“It (branch) is dry.”

B
\[\text{neve} \quad -\text{nó'e} \quad -\text{co} \quad -\text{ti} \quad \text{nica}\]

\(\)CL:LIQUID PL DUR food

“They are picking food (from the water).”

Both the analysis of supposed verb classifier systems as verbal classifier systems and as argument marking systems are backed by observations of studies on noun incorporation, which closely resembles classifier insertion by allowing only certain generic terms to incorporate. This employment of a smallish set of generic terms results in an underlying cognitive classification, but serves a different purpose than nominal classification (which naturally also is based on a cognitive classification): On the one hand, incorporated nouns serve the lexical restriction of the concept expressed by the verb (cf. Barrie 2010: 292, Mithun 1986b: 34, and Postal 1979: 285), which fits the analysis of supposed verb classifiers that actually are verbal classifiers. On the other hand, incorporated nouns can substitute a free noun (cf. 2.57).

(2.57) Mohawk (cf. Bonvillain 1973: 21)

A
\[\text{wa?-k-ytho} \quad \text{ojíja}\]

AOR-1SG.A-plant flower

“I planted a flower.”

B
\[\text{wa?-jiya-ytho}\]

AOR-1SG.A-flower-plant

“I ‘flower-planted’.”

These incorporating constructions replacing a free noun “are restricted in nearly identical ways to occurrence with third person nominals and thus third person verb prefixes” (Postal 1979: 292) and thus fit the argument marking analysis. Just like these verb prefixes, “the general rule of noun deletion[…] acts as a kind of pronominalization” (Postal 1979: 396). Due to this pronoun-like nature of the incorporated nouns, it is possible for them to introduce new referents to the discourse (cf. Baker et al. 2005: 146 and Mithun 1984: 871), which is not the case for common nominal classifiers. Baker et al. (2005: 146f.; cf. also Baker 1996) further argues that incorporated nouns can be referential themselves. Barrie (2010) offers a slightly different syntactic analysis than Baker,
which still allows for the same conclusions. Finally, the incorporated supposed
'classifiers' can be specified by verb-external modifiers (cf. 2.58 and 2.59), which
is not observed in any other classifier system that I am aware of.

(2.58) Mohawk (cf. Mithun 1984: 870)

\[
\begin{array}{cccc}
\text{Kanekwar} & \text{u} & \text{k-} & \text{akya'tawishe} \\
\text{3SG.dotted.DIST} & \text{PST} & \text{1SG} & \text{make}
\end{array}
\]

"I [dress]-made a polka-dotted one."


\[
\begin{array}{cccc}
\text{wayah} & \text{hak} & \text{k'uh} & \text{isa} \\
\text{a.lot} & \text{PROGR} & \text{grass} & \text{be PROGR}
\end{array}
\]

"There is a lot of grass."

2.3.3 ‘Minor Classifier Types’

Besides ‘verb classifiers’ and ‘relational classifiers’, there are three other types
of nominal classifiers which are often mentioned together with the already
introduced classifier types, but are not commonly regarded nominal classification
devices in their own right (cf. Seifart 2010: 721 and Grinevald 2000: 68f.). These
types are locative classifiers (cf. 2.60), deictic or demonstrative classifiers (cf.
2.61) and repeaters (i.e. nouns that are used as their own classifiers, also la-
belled "self-" or "auto-classifiers"; cf. 2.32 = 2.62) (cf. Aikhenvald 2000: 172ff.,


\[
\begin{array}{cccc}
\text{pis} & \text{keh} & \text{puha-t} & \text{arab} \\
\text{2SG} & \text{make} & \text{one-CL:VERT} & \text{shield}
\end{array}
\]

"You make a shield on your arm."


\[
\begin{array}{cccc}
\text{an-toñi-igi} & \text{dit7-m7e} & \text{dole} \\
\text{2SG-warm-MOD} & \text{CL:ORIZ-DEM} & \text{fire}
\end{array}
\]

"Warm yourself up by the fire."


\[
\begin{array}{cccc}
\text{prathêet} & \text{saaam} & \text{prathêet} \\
\text{land} & \text{three} & \text{CL:LAND}
\end{array}
\]

“three countries”
Locative classifiers are a rare phenomenon that is only attested in South American Indian languages (Palikúr, Lokono, Dâw and Carib languages; cf. Aikhenvald 2000: 172). Due to this isolated status and the fact that data on these languages became available only recently, they did not play a major role in research so far. They occur NP-internally and categorise the locative’s argument with respect to its inherent semantic characteristics (cf. 2.63 from Palikúr, “the only clear-cut example of locative classifiers”; Aikhenvald 2000: 172).


\[
\begin{align*}
\text{ig-kis} & \quad \text{ute-c-qi ig} & \quad \text{mote} & \quad \text{ay-h-te} & \quad \text{a-peru} \\
3M-PL & \quad \text{find-COMPL-3M} & \quad \text{wasp:3M} & \quad \text{there-INT-DIST} & \quad \text{3N-on:BRANCH.LIKE} \\
\text{ah} & \quad \text{tree} & \quad \text{“They found the wasps on the tree.”}
\end{align*}
\]

They always appear in fusion with a locative modifier in pre- or postposition to the argument, which earned them the epithet “classificatory adposition” (cf. Aikhenvald 2000: 172). Since they categorise the locative modifier’s argument in terms of inherent physical properties, they seem to serve the function of classifiers. Locative classifiers form a morphological unit with their locative pre- or postposition and therefore cannot be isolated from them. Locative classifiers are ignored in this study, since they are an extremely rare phenomenon and occur only in a few indigenous languages of South America. They are fused with locational adpositions, and this fusion violates the commonly assumed criterion of categorical independence of classifiers (cf. Dixon 1986). Locative classifiers are not considered in this study due to their marginality and areal boundedness.

Deictic classifiers are found in deictic constructions, i.e. constructions containing an article or a demonstrative. The referents of the deixis are categorised in terms of their shape, animacy and position in space (cf. Aikhenvald 2000: 176). Suchlike deictic classifier constructions are found in the North American language Yuchi and the Siouan family, the South American language Guai-curuan and in Eskimo. The North American deictic classifiers show a close relation to stance verbs or are even grammaticalised forms of stance verbs (cf. ib. and 2.64).


A \text{de-mäk}
\begin{align*}
\text{this-CL:DEICT.LYING} \\
\text{“this one (lying)”}
\end{align*}

B \text{de-näk}
\begin{align*}
\text{this-CL:DEICT.SITTING} \\
\text{“this one (sitting)”}
\end{align*}
The Siouan languages Ponca and Mandan on the one hand and Eskimo on the other hand have two sets of articles which are fused with the classifier’s categorisational function (cf. 2.65 and 2.66): the so-called classifier forms an undividable morphological unit with the article.


A nî $\La$‘c ART:STANDING.COLLECTIVE
“the water”
B nî čá water ART:ROUNDED
“the handful of water (cupped)”
C nî k‘e water ART:HORIZONTAL
“the (line of) water, the stream”


unaaq pagna aiguk
harpoon EXTENDED:one.up.there get
“Get that [visible, extended] harpoon up there.”

Deictic classifiers are not part of this study for the same reasons as locative classifiers: They occur only in American indigenous languages and are extremely rare. Like locative classifiers, they often display a fusion the supposed classificatory element and an article or deictic (cf. again 2.65), which again violates the commonly assumed categorial independence of classifiers. Deictic classifiers were thus excluded from this study due to their marginality and limited areal distribution.

Finally, nominal classification frequently involves so-called repeaters, i.e. nouns which can be used as their own classifiers (cf. again 2.60 C). These may be an alternative for the employment of a general classifier for otherwise ‘uncategorisable’ nouns, for instance. repeaters are not considered an individual type of classifier in this study, since they may occur in a variety of classifier systems, but only next to ‘common’ nominal classifiers, while there is no classifier system that employs only repeaters (Aikhenvald 2000: 104). repeaters of an individual system may thus occur in this study, but only as a part of that system and not for their own sake.

After having discussed the exclusion of a number of supposed types of classifier systems, this study’s taxonomy of classifier systems can be reorganised;
its reduced version is shown in figure 2.3 below. The remaining classifier types (noun, numeral, and possessive classifiers) behave in a fairly uniform way and have the morphosyntactic classifier host as their main distinctive criterion. Section 2.4 provides a theoretical and typological background on the second basic type of nominal classification system next to classifiers, namely gender/noun class systems.

2.4 Gender/Noun Class Systems

This section is dedicated to the introduction of a general and theoretical background on gender/noun class systems, which have been subject to linguistic study ever since early the earliest grammarians such as Panini. This section will introduce various aspects of these systems in the following order: Frequency and location, their various definitions in the literature, their formal properties, their functional properties, the nature of their classification, their different systemic types, their common sources, and a concluding definition that serves as the working definition of this study.

**Distribution** Gender and noun class systems are frequently found in the world’s languages. They occur on every continent (cf. Aikhenvald 2000: 78), but depend on the language type. Gender/noun class systems are defined as classificatory systems which display grammatical agreement; naturally, fusional or agglutinating languages are best suited for the employment for agreement by means of affixes, while isolating languages such as Mandarin Chinese are unlikely to display agreement by affixes (ib.: 20). Europe (and most other Indo-European languages outside of Europe) and Africa are almost fully covered with gender/noun class languages, while the Americas, Australia, and Asia each have larger linguistic areas of gender/noun class systems next to isolated instances. Despite their distinction by name, both gender and noun class systems fulfil the same functions and show the same formal properties (with noun class systems tending to employ a slightly bigger class inventory). Since the labels “noun class system” and “gender system” are rather depending on the terminological tradition of a certain language’s research community than on specific distinctive criteria, the usage and definition of the terms ‘gender’ and ‘noun class’ is

\[\text{CLSs} \quad \text{numeral CLSs} \quad \text{possessed CLSs} \]

Figure 2.3: Taxonomy of classifier systems.
Nominal Classification

varying from author to author. Both may be seen as areas on a functional scale that differ by their degree of formalism. In order to avoid confusion, researchers have gone as far as dropping one of the terms completely: For instance, Corbett (1991) only uses “gender” when referring to any kind of noun class, while the current trend tends to “noun class” as the hypernym for both noun classes and genders (e.g. Aikhenvald 2000; cf. ib.: 19). For reasons of unambiguity, I henceforth use the term “concordial class systems” (CCSs) to refer to gender/noun class systems; I reserve both “gender (system)” and “noun class (system)” for an eventual ad hoc disambiguation.

Earlier definitions There is a number of earlier definitions of concordial class systems: Allan (1977) mentions ‘concordial classifiers’, which employ classifying formatives that occur in the form of affixes (mostly prefixes) on nouns and their satellites. Dixon (1986: 105–107) renders them as obligatory and closed grammatical systems (in the sense that each noun of a given language inherently belongs to a concordial class) with a small class inventory, which is marked by a prefix or portmanteau suffix on the noun and other NP-internal elements and clausal constituents (~agreement). Dixon (ib.) also notes that there is only little variation in a noun’s concordial class among the speakers of a language and that register or style are typically not involved in these systems. Grinevald (2000: 55–60) also defines gender as a grammatical category with a small inventory, which is assigned to all nouns and marked obligatorily. However, the markers occur not always on the noun itself, but on the following elements (decreasing in likelihood): NP-internal constituents (adjectives, articles, determiners, numerals, possessives, etc.), independent pronominal forms (personal pronouns, relative pronouns, wh-words, etc.), within the predicate as an agreement marker of the core argument, and eventually on other clausal constituents such as adverbs, adpositions, and complementisers. A noun’s concordial class is semantically hardly motivated besides for a frequent sex distinction, but it can be predictable by morphological means. Based on Corbett (1991), Corbett and Fraser (2008: 293f.) define concordial class systems as agreement systems. Those can be subdivided into systems the assignment of which is semantics-based and systems the assignment of which is partially based on semantics and supplemented by morphological and/or phonological information. According to Aikhenvald (2000: 19), concordial systems are fairly grammaticalised systems, which correlate with certain semantic characteristics and are realised on modifiers within and outside the noun phrase. Concordial class systems are a device for linguistic classification and not for a biological classes' the larger systems which are considered more 'exotic' from a Eurocentric point of view” (Grinevald 2000: 57). The first of these two criteria is always avoided within this study: If the term “gender system” is used here, it never is used in a manner to distinguish between sex-based concordial class systems and non-sex-based concordial class systems. This tradition can be found e.g. in linguistic works on Sumerian, where we often find the statement that Sumerian does not employs a ‘gender system’ (cf. Hayes 1990: 31, 1997 and 1999: 15, Thomsen 1984: 55), while it actually does employ a (non-sex-based) concordial class system.
classification in terms of sex. Seifart (2010: 720) defines concordial class systems as always characterised by agreement marking, which occurs in the form of affixes or clitics outside the noun on a (language-specific) range of targets. The markers are often portmanteau morphemes, which also encode e.g. person, case, or number. Contini-Morava and Kilarski (2013: 266) define concordial class systems as smallish and closed nominal classification devices, which are overtly marked for agreement (in contrast to classifier systems with the noun itself being potentially marked). A noun’s concordial class membership is “mutually exclusive” (ib.) and the assignment may be carried out upon semantic, morphological, or phonological criteria.

**Formal properties** The formal properties of prototypical concordial class systems are as follows: Concordial class systems affect all nouns of a language in that they pattern the whole nominal lexicon in formally differentiated agreement classes. The information about the obligatory class membership of a noun is located in the lexicon and realised outside the noun itself by agreement on a range of nominal satellites. This makes concordial class systems a ‘lexico-syntactic’ phenomenon. The number of concordial classes is limited (with an average between two and some two dozen classes) and the systems are closed.

Each noun is assigned to only one class and the speaker has no influence on the assignment of a certain class to a certain noun; therefore, a variation in register or semantic individuation as we find it in classifier systems is not possible, but rather expressed by syntactic or lexical means. The class assignment always involves at least partially semantic principles and may be further complemented by morphological and/or phonological assignment; depending on a system’s degree of grammaticalisation, these purely formal assignment principles may prevail if assignment principles are in conflict, semantic assignment typically rules out other assignment principles. Each concordial class provides distinct agreement markers or paradigms (cf. 2.67).

(2.67) Latin

<table>
<thead>
<tr>
<th></th>
<th>me-o</th>
<th>maxim-o</th>
<th>stupor-e</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>my-ABL.M.SG</td>
<td>biggest-ABL.M.SG</td>
<td>stupidity-ABL.M.SG</td>
</tr>
<tr>
<td></td>
<td>“through my enormous stupidity”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>my-ABL.F.SG</td>
<td>biggest-ABL.F.SG</td>
<td>fault-ABL.F.SG</td>
</tr>
<tr>
<td></td>
<td>“through my most grievous fault”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>my-ABL.N.SG</td>
<td>biggest-ABL.N.SG</td>
<td>pursuī-ABL.N.SG</td>
</tr>
<tr>
<td></td>
<td>“through my most intense pursuit”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

8I was made aware of the Bukiyip concordial class system, which is assumed to be solely based on phonological assignment criteria by some scholars (Fortune 1942, Conrad 1978:92). Aikhenvald (2000: 26) disagrees with that assumption and refers to Nekitel (1985 and 1986), who “has convincingly argued against a ‘purely phonological’ assignment of noun classes in Abu’ Bukiyip in favour of a semantic assignment to two human concordial classes, which is accepted ‘by most scholars’ (Aikhenvald 2000: 26).
Nominal Classification

Marking of class membership on the noun itself is possible, but optional (in the sense that it is determined in the system if class marking appears on the noun, not in the sense that the speaker may decide whether to mark the noun or not) and never restricted to the noun itself. The markers occur on a range of language-specific nominal satellites. This agreement exceeds the borders of the noun phrase and the clause (cf. 2.68 and 2.69).

(2.68) German

Sylvia Maria Kristel was eine niederländische Schauspielerin. Sie spielte in mehr als fünfzig internationalen Spielfilmen.

“Sylvia Maria Kristel was a Dutch actress. She played in more than fifty international movies.”

(2.69) Dutch

Sylvia Maria Kristel was een Nederlandse actrice. Ze speelde in meer dan vijftig internationale speelfilms.

“Sylvia Maria Kristel was a Dutch actress. She played in more than fifty international movies.”

Markers may occur in the form of affixes ablaut or vowel changes (cf. Aikhenvald 2000: 58f.) and constitute terms in a grammatical system which are fixed in their meaning (Dixon 1986: 108). If the markers are affixed to the noun stem, the affix may also contain information on other grammatical categories such as number, case or person. These portmanteau morphemes typically provide no information about their etymological origin (cf. again 2.67 and 2.68).

---

9 Dixon (1986: 108) also lists articles and separate grammatical words as concordial class markers, while most other sources restrict the range of markers to affixations and clitics. As concordial class systems are supposed to imply a much more formalised structure than classifier systems do and articles may be interpreted as separate grammatical words (that include a (possibly fused) inflectional affix), separate grammatical words are not taken into account as concordial class markers. This view is also shared by Aikhenvald (2000: 58), who clearly states that concordial classes are never marked by free morphemes. It also serves the limitation of the fuzzy borders within the assumed continuum of nominal classification towards the definitional ‘grey area’ that connects concordial class systems and classifier systems. So-called concordial class systems that are marking class membership by separate constituents have to prove that their function is clearly that of a concordial class systems and not rather that of a classifier system or a system of nominal classification that is at an intermediate stage between classifier systems and concordial class systems.
2.4. Gender/Noun Class Systems

**Functional properties** In contrast to claims that concordial class systems represent ‘meaningless baggage’ (e.g. Adelung, Ockham; cf. Kilarski 2013: 327f.), they constitute a functional linguistic phenomenon. By the provision of several formally differentiated concordial classes and their marking on a range of language-specific agreement targets, they support what Corbett (1991: 322) calls “reference tracking” throughout discourse and attributes to concordial class systems as their primary function (cf. 2.17 = 58, where the pronoun’s concordial class identifies the intended discourse referent).

(2.70) German

| Ein ART:NOM.N.SG | Auto car:NOM.N.SG | stieß push:IPF3SG | mit with | einem ART:DAT.M.SG |
| Zug train:DAT.M.SG | zusammen together | Er 3SG.M | trug carry:IPF3SG | erheblichen considerable:ACC.M.SG |
| Schaden damage:ACC.M.SG | davon away |

“A car collided with a train. It [the train] took considerable damage.”

Besides this primary function of reference tracking, which all concordial class systems and a broad range of classifier systems serve, there is a range of secondary functions that a concordial class system may serve. These are also shared with classifier systems. Again, we find the semantic function of expansion of the lexicon, mostly carried out by the creation of new lexical items (e.g. a root than combines with a concordial class marker in order to derive a new stem or by separating homonyms through assignment to different classes; cf. German *der Schild* m. “the (battle-)shield” vs. *das Schild* n. “the sign”). Beyond that, concordial class marking can serve as a means of individuation of a referent (cf. 2.71 where the switch in concordial class marking indicates semantic agreement with the individuated object in question; cf. also Audring 2013: 39, cited in Contini-Morava and Kilarski 2013: 276).

(2.71) Dutch

| Heb have:2SG | je you | mijn my book:N.SG | gezien? seen | Ik I held:IPF3SG | hem 3SG.M | net just nog yet |
| in handen in hand:PL |

“Have you seen my book? I held it in my hand just a moment ago.”

Besides their primary discourse function of referent identification and reference management, concordial class marking can serve as a mean for a referent’s re-presentation, which changes the perspective on the given referent and adds information to it (Contini-Morava and Kilarski 2013: 287). Cf. 2.14 = 2.72

where a joking reference to a hermaphrodite is made by the noun for “man”, which is categorised by the (feminine) class II marker.
Classificatory nature  The classificatory nature of concordial class systems always contains a semantic component, the most prominent of which is the distinction of the biological sex of a nominal referent (cf. e.g. Latin domin-us m. “master” vs. domin-a f. “mistress”). Nevertheless, the semanticity of concordial class systems is only of a secondary (and, in most cases, of a partial) character. The primary function of concordial class systems is the classification of a noun’s linguistic form; cf. e.g. Latin arbor f. “tree”, navis f. “ship”, nix f. “snow”, or mons m. “mountain”, where no synchronic ‘real world’ semantic motivation for the respective noun’s concordial class is available. Concordial class systems provide formal means to distinguish formally different classes of nouns — to what degree these classes are determined semantically or morphologically/phonologically, is a language-specific question.

Concordial class ‘types’ As indicated at the beginning of this section, the research tradition distinguishes gender systems, which are typically concordial class systems with an inventory of two or three and seldom four or five concordial classes, and noun class systems, traditionally those concordial class systems that employ a bigger class inventory of up to some two dozen concordial classes (often with singular and plural classes counted individually; cf. the Chichewa system in Mchombo 2004 and Watkins 1937) and/or that employ more ‘exotic’ assignment criteria to the western interpretation, such as concordial classes for vegetable food or dangerous/harm-related entities. Since both systemic types display the same functional behaviour and only minor tendencies with respect to their formal constitution, these subtypes are only distinguished in situations where this is not avoidable. In this study, both systems are treated equally and are referred to by the term “concordial class systems”.

Sources of concordial class systems One of the findings of Corbett (1991: 312) is that nouns are “the ultimate source” for concordial class systems. ‘Ultimate’ means that concordial class markers may develop from lexical nouns, but also more or less closed grammatical systems such as classifiers, determiners, articles, and similar sources, which themselves virtually always are rooted in lexical nouns with ‘classificatory potential’ (i.e. generics; cf. Sands 1995: 250 and Corbett 1991: 311f.). Closed grammatical classes result in concordial classes and rather not in classifier systems (Aikhenvald 2000: 367). Classes that were proposed in the literature as a source for concordial class marking are demonstratives (Greenberg 1978a; cf. also Aikhenvald 2000: 367 and Sands 1995: 252), locatives (cf. the case of Maasai, where a third concordial class
is about to arise with currently only one nominal member, *wwéjì* “place”; cf. e.g. Payne 1998: 160 and Corbett 1991: 313f.), and third person pronouns (cf. Heine 1997, cited in Aikhenvald 2000: 368). There are even documented instances where a closed class gave rise to a concordial class system through a reanalysis of the speakers (namely the Kiwa-Tanoan nouns, which have an inherent number conditioning, the concordial class assignment; cf. Aikhenvald 2000: 369 and 377).

Luraghi (2011: 451ff. and 454ff., based on Fodor 1959: 32ff.) proposes a twofold genesis of concordial class systems, the creation “from above”, which corresponds to the already introduced grammaticalisation of e.g. classifier systems, and the creation “from below”, which corresponds to a conventionalisation of “special patterns of case marking” (Luraghi 2011: 454). These special case marking patterns prefer to combine with a given part of the nominal lexicon; for instance, subject markers may prefer to combine with animate nouns, and instrumental markers may prefer to combine with inanimate nouns. This pattern may be reanalysed: The subject marker becomes an animate marker, and the instrumental marker becomes an inanimate marker. The primary function of concordial class systems “from above” is nominal classification, whereas the primary function of those systems “from below” is reference tracking (Luraghi 2011: 459). However, Luraghi’s proposal leaves crucial questions open (Kilarski 2013: 272): On the one hand, the origin of the systems of Bora and the Niger-Congo languages are not taken into account, which is distinctly stated; on the other hand, the functionality of classifier systems is reduced to semantic categorisation, which does not account for e.g. aged classifier systems.

**Definition** After considering earlier works and findings on concordial class systems, a working definition of concordial class systems for this study can be provided. The choice of criteria here is based on earlier propositions in the literature, and is refined by findings from my own preliminary sample study. Properties that were optional or did not prove to be reliable or valid have been ignored. A prototypical concordial class system has the following properties:

- Concordial class markers occur in the form of affixes or clitics.
- The markers occur on a language-specific range of nominal modifiers (and possibly on the noun itself), which are not necessarily restricted to the borders of the noun phrase.
- Concordial class systems primarily serve the formal patterning of the nominal lexicon.
- A noun’s concordial class membership is fixed in its lexical entry and not determined by the individual speaker according to his or her discoursal demands.
- A given noun may generally not change its concordial class; its class membership is inflexible.
• A concordial class system does not primarily contribute to its lexicon’s semantics. However, it may employ partial semantics in grouping conceptually related nouns in the same concordial class.

• Concordial class marking occurs on a language-specific range of nominal satellites.

• Concordial classes employ a smallish class set (up to 24 classes).

• A concordial class system affects all nouns of a given language’s nominal lexicon.

• Since its concordial class membership is inherent, a noun always belongs to a concordial class, no matter in which constructional type it occurs.

Beyond that, it may be noted that concordial class markers are generally not etymologically transparent. Unlike classifiers, they are typically not involved in the activation or satisfaction of linguistic registers. After having provided an overview over nominal classification and definitions of its systemic manifestations, chapter 5 provides an overview over grammaticalisation theory and its relevance for this study.
This section presents grammaticalisation and its implementation in this study. As indicated beforehand, the main hypothesis investigated in this study is that systems of nominal classification evolve and develop in a certain manner: Classifier systems evolve from lexical items and mainly serve the discoursal and semantic purpose of specifying nouns and their referents. The potential further development of classifiers is the subject of this study: Concordial class systems may constitute largely automatised or mechanised classifier systems (not only in terms of their function, but also in the sense that concordial class systems may stem from classifier systems) and serve the formal patterning of the nominal lexicon. In this study, grammaticalisation is assumed to be the vehicle for the emergence and development of nominal classification systems.

After a general introduction of the application of grammaticalisation to nominal classification, section 3.1 introduces the general concept and theory of grammaticalisation, which is further elaborated in section 3.2 where the specific form of the grammaticalisation hypothesis and involved properties of nominal classification devices are presented. Section 3.3 puts the specific expectations from the grammaticalisation hypothesis in a more general theoretical context and assembles a package of properties of nominal classification systems. This package serves as a comparative measure of the grammaticalisation of classification systems in the sample study.
3.1 Grammaticalisation: Its Nature and Constitutive Forces

This subsection introduces general theory on grammaticalisation. After a brief elaboration on grammaticalisation as the most probable vehicle for the assumed shift CLS>CCS, the principle behind grammaticalisation and the processes possibly involved are presented. Furthermore, the more specific topics of unidirectionality, decategorisation, and grammaticalisation cycles are discussed. Re-analysis, analogy, and lexicalisation are briefly addressed and contrasted with the grammaticalisation processes in order to clarify the exclusive role of grammaticalisation in the assumed shift CLS>CCS. This section concludes with the introduction of general shift-related phenomena.

Grammaticalisation is “a process of conventionalization that crucially involves the expansion of the semantic-pragmatic usage contexts, host-class expansion, and the broadening of the syntactic context in which a given item occurs” (cf. Himmelmann 2004: 33, cited in Barðdal 2013: 3). It thus comprehends a shift in grammaticality of linguistic constituents (morphemes, syntactic constructions, discourse units; cf. DeLancey 2004: 1590). The shifting phenomena involve a shift from an item with lexical or grammatical function to an item with a more grammatical and less lexical function. Grammaticalisation hence involves changes in the phonological, morphological, semantic, pragmatic and syntactic properties of a constituent by phonological reduction, loss of morphosyntactic independence, semantic bleaching, categorial extension, paradigmatisation, etc. A common example for grammaticalisation is the emergence of the future tense markers in the transition from Latin to the Romance languages (cf. 3.1, which shows the grammaticalisation of the lexical verb habe-re “to have” into the future tense paradigm of the Romance daughter languages; taken from Roberts 1993, cited in Roberts and Roussou 2003: 49f.) and the development of French negation; both cases are described more in detail in the further course of this section.

(3.1) Latin

A future tense: affixed marker:

<table>
<thead>
<tr>
<th>Latin</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>ama-re: ama-b-o, ama-b-is, ama-b-it</td>
<td>“to love: I will love, you will love, he/she/it will love”</td>
</tr>
<tr>
<td>love:PRS.INF love:FUT1SG love:FUT2SG love:FUT3SG</td>
<td></td>
</tr>
</tbody>
</table>

B full verb habe-re "to have" > future auxiliary (cf. English *will*):

<table>
<thead>
<tr>
<th>Latin</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>amare habe-o, amare habe-s, amare habe-t</td>
<td>“I will love, you will love, he/she/it will love”</td>
</tr>
<tr>
<td>love:PRS.INF have:PRS1SG love:PRS.INF have:PRS2SG love:PRS.INF have:PRS3SG</td>
<td></td>
</tr>
</tbody>
</table>
C Reanalysis as syntactic affix:

\[
\begin{align*}
\text{amare} & \quad \text{habeo} & \quad \text{amar-aio} \\
\text{love:PRS.INF} & \quad \text{have:PRS1SG} & \quad \text{love:FUT1SG} \\
\end{align*}
\]

“I will love”

D Reanalysis as lexical affix in the individual Romance languages:

\[
\begin{align*}
\text{amar-aio} & \quad \text{Italian} & \quad \text{amer-ò, French aimer-ai etc.} \\
\end{align*}
\]

“I will love”

As 3.1 illustrates, grammaticalisation involves a construction change, which “selectively seize a conventionalized form-meaning pair of a language, altering it in terms of its form, its function, its frequency, its distribution in the linguistic community, or any combination of these” (cf. Hilpert 2013: 460; cited in Barðdal 2013: 7). It is necessary for the purpose of this study to disambiguate the grammaticalisation of form, which is directly visible by the loss of formal independence of a classificatory element, and the grammaticalisation of function, which expresses itself e.g. in the quality of its semantic contribution to the noun phrase where it occurs. This functional grammaticalisation is supposed to occur prior to formal grammaticalisation and (at least in theory) independent from it. The split into formal and functional properties, which are separately evaluated in the sample study, is discussed at the end of this chapter (cf. page 79ff.).

Beyond the simple shift of linguistic items from less to more grammatical functionality, grammaticalisation is able to give rise to functional categories that are language-internal innovations, which segregates it from other processes of grammatical change. The rise of articles in the Romance and Germanic languages by the grammaticalisation of demonstrative pronouns is just one example for the creation of a novel functional category. Due to this creative potency, it is assumed to be the vehicle for changes in systems of nominal classification (cf. also Aikhenvald 2000: 91 and 372ff.): Only grammaticalisation is able to give rise to functional innovations that are not earlier existing in a language. The following description is mainly based on DeLancey 2004, Hopper and Traugott 1993, Lehmann (2004, 2005, 2008a and 2008b), Roberts and Roussou 2003, and Traugott 2011.

Why Grammaticalisation? Grammaticalisation is subject to discussion and severe criticism since its rediscovery in the 1970s\footnote{The term “grammaticalisation” first appeared in Meillet 1912 and soon established as a label for change phenomena that contained a rise in grammaticality of the affected linguistic material. Due to a lack of research interest on diachronic language change and philological studies, where grammaticalisation attracted most attention, and the upcoming generative grammar around the middle of the twentieth century, grammaticalisation was neglected for several decades. With an increasing interest in the research of linguistic discourse and universals from the 1970s on, grammaticalisation was “rediscovered” and integrated in modern theoretical linguistics.} So why should research be based on such a ‘problematic’ theoretical concept? There are several other
mechanisms that can give rise to new functional/grammatical material (e.g., reanalysis and analogy). These are outlined and distinguished from grammaticalisation later in this section; the reference to them at this point serves the reasoning why they cannot be a part of the development of systems of nominal classification, while grammaticalisation can.

A crucial misconceptualisation of grammaticalisation that has often been made is that grammaticalisation constitutes a routine or process of mechanisms of change. This rather traditional view on grammaticalisation cannot be maintained nowadays. Modern approaches towards grammaticalisation understand it as the result of various processes such as reanalysis, univerbation, semantic bleaching, metonymy, (metaphorical) extension or phonological reduction (Barðdal 2013: 9). Not all potentially involved processes are necessarily involved in individual cases of grammaticalisation, which causes its rich variety of appearance. In this view, grammaticalisation may give rise to new grammatical constructions by the reorganisation of existing constructions, which happens in a “specific local context characterized by the coming together of several different factors, which jointly facilitate a shift in meaning and a new conventionalization” (cf. Barðdal 2013: 19).

The crucial difference between other mechanisms that can give rise to the degree of a linguistic item’s grammaticality and grammaticalisation is that the latter one is the only mechanism of grammatical change that is unidirectional and cannot be reversed (cf. Lehmann 2004: 22) — only some phonological changes are also unidirectional. As any other aspect of grammaticalisation, unidirectionality is doubted and questioned by critics of grammaticalisation. Against all scepticism, there are strong arguments in favour of the unidirectionality of grammaticalisation (cf. Lehmann 2004: 15 and 19ff.), thus this principle is adopted for this study. According to the working hypothesis, systems of nominal classification know only one natural direction of development. If any other kind of grammatical change than grammaticalisation would be involved in the development of nominal classification systems, a development from a concordial class system to a classifier system or from classifiers to lexical items would be expected. I am not aware of any suggestion of suchlike development in the literature (cf. also Aikhenvald 2000: 372). In addition, only grammaticalisation is able to create language-internal innovations in terms of functional categories (cf. Lehmann 2004: 10), while other categories may give rise to new material, but not to language-internal new functional categories. Therefore, only grammaticalisation qualifies as a changing mechanism for nominal classification systems.

**Working Principle** Meillet defined grammaticalisation as a process by which lexical items enter the grammatical system (le passage d’un mot autonome au rôle d’élément grammatical “the transition of an independent word to the role of a grammatical element”; Meillet 1912, cited in DeLancey 2004: 1590). More in detail, “[g]rammaticalization consists in the increase of the range of a mor-
Grammaticalisation

Pheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one" (Kuryłowicz 1965: 69, cited in DeLancey 2004: 1590; 3.2 for the grammaticalisation of lexical material and 3.3 for the increase of grammaticality of already grammatical material).

(3.2) Minangkabau (Marnita 1996: 107 and 131; cited in Aikhenvald 2000: 375)

A jari
"finger" > jari
"CL:CLOTH.MEASURE"

B ikua
"tail" > ikua
"CL:ANIMAL"

(3.3) German

der, diu, das
"this (m., f., n.; demonstrative pronoun)" > der, die, das
"the (m., f., n.; article)"

The concept of a grammaticalisation process includes the pathway (or shifting operative domains of the grammaticalising item) displayed in figure 3.1; the cline in this sequence equals a gain in grammaticality (Givón 1979: 209, cited in Traugott 2011: 97).

Constitutive Processes

Grammaticalisation involves several processes that are outlined in the following. While grammaticalisation is sometimes perceived as a kind of script or routine that mechanically executes one process after

2This transition can be approached from two different directions: As a formal and structural reduction accompanied by an increase in dependency of the grammaticalised constituent (“traditional” approach; cf. Traugott 2011: 97) or as an expansion (in terms of the host-class, syntactic and semantic-pragmatic relation; “extended” approach; cf. Traugott 2011: 98). The difference between the two is mainly the perception of the process itself: The traditional approach sees the grammaticalised form as a mere reduction with respect to its “substantive, contentful, referential, or lexical meaning” (Traugott 2011: 99) and phonological body. A conceptualisation of this solely reductive process is provided by example 3.20 C: Compared to the status as a demonstrative pronoun, the German definite article der, die, das lost its deictic character and (partly) its autonomy, as it depends on its nominal host. The extended approach by contrast expects the grammaticalised form to gain abstract semantic property (ib.), which can be illustrated in example 3.20 A and B: The lexical items jari “finger” and ikua “tail” may have ‘lost’ their specific meaning during their transition to a classifier status, but in return they gained the abstract semantics of a numeral classifier.

Another, it is important to remember that grammaticalisation may stop at any point of its development and does not necessarily involve all processes.\footnote{For instance, the processes of semantic bleaching and decategorisation are virtually always involved in grammaticalisation, whereas phonological erosion and cliticisation/affixation are less frequently found (cf. the description of the grammaticalisation of the French negator *pas* in this section).} Therefore, it is better conceptualised as a number of related processes than a strict one-way road that has to be passed by making one distinct step after another. Figure 3.2 lists the single processes in their typical order (based on DeLancey 2004: 1593f.).

![Diagram of grammaticalisation processes](image)

Figure 3.2: Grammaticalisation: (possibly) involved processes.

Grammaticalisation relies on a productive syntactic structure (DeLancey 2004: 1593), e.g. the negation in Old French and its successive stages. In Old French, negation consisted of the negator *ne*, a proclitic form of Latin/Old French *non* that precedes the verb (cf. 3.4 A). Optionally, the negation could be reinforced by an adverbial noun such as *pas* “step”, *point* “point”, *mie* “crumb”, *gote* “drop”, *amonde* “almond”, *arest* “fish-bone”, *beloce* “sloe” or *eschalope* “pea-pod” (cf. 3.4 B and Hopper and Traugott 1993: 117).

\[(3.4)\]
\[
\begin{align*}
A &:\ ne + V \\
B &: ne + V (+ pas, point, mie, gote, amonde, areste, beloce or eschalope)
\end{align*}
\]

These nouns serve the indication of a least quantity. One or a few instances of the productive structure, usually the ones with the most general semantic content, occur more often than others. Because of their frequent use, they are functionally specialised, which means they gradually supersede the other possible instances of this structure, which hence becomes less and less productive. This specialisation results in the semantic bleaching of the more frequent instances, which can be used in contexts where the original instance would not have been appropriate. In the case of French negation, these are *pas* “step”,...
point “point”, mie “crumb” and goutte “drop” (by the sixteenth century; cf. 3.5 A); later, only pas “step” and less frequently point “point” (cf. 3.5 B and Hopper and Traugott 1993: 117)) were kept in use. While e.g. pas “step” originally served as means for emphasis of negated verbs of motion (e.g. je ne vais (pas) “I am not going (a step)”), it lost this restriction by generalisation and the loss of the other means of reinforcement. With its increasing generalisation, the original meaning “step” fades away; hence a construction as je ne chante pas must be interpreted as “I am not singing (at all)” and not “I am not singing (a step)”.

(3.5)  
A ne + V (+ pas, point, mie, gote)  
B ne + V (+ pas, point)

The contextual extension is accompanied by the petrification of the instance’s appearance: The original instance’s formal manifestation cannot be varied according to its original category’s characteristic paradigmatic properties anymore. This loss of ‘character’ is labelled “decategorisation”. The loss of the nominal adverbial’s characteristic properties in French negation manifests itself in the mandatoriness of its occurrence in negative constructions (cf. 3.6 A), where it gains the qualities of a negator, which finally leads to omission of the original negator ne (cf. 3.6 B).

(3.6)  
A ne + V + pas  
B V + pas

After the loss of its original category’s significant properties, the form may be subject to phonological erosion, which means it is reduced in phonological weight. This process may be accompanied by further recategorisation, i.e. the element’s reassignment to another morphosyntactic category that seems more appropriate (in terms of the altered function, use and appearance of the form) than the original one. The form may further cliticise or fully fuse with a lexical head. After the fusion, the grammaticalised form may further erode to an unrecognisable state or its complete loss. This triggers another rise of new functional material in a new round of the grammaticalisation cycle.

Unidirectionality  Unidirectionality is a controversial topic within the discussion on grammaticalisation. The previous description of the single processes

4Cf. the English construction on top of all the houses (where top is a relational noun), which can be replaced by the ordinary noun as in on the tops of all the houses, while it is not possible to pluralise the relational noun: *on tops of all the houses (cf. DeLancey 2004: 1593). The relational noun is clearly a noun, while it lost the inflectional properties of its original category. Another visible difference is the semantics of the construction: While on top of denotes the uppermost surface of an item, its top does not have to be the uppermost surface. If the item is e.g. lying on its side, the top lies on the side as well, while on top of the stem denotes its momentary upper side (ib.: 1593f.).

5This is not the case in the example of the development of the French negation, where the grammaticalisation process ended at the recategorisation of pas as negator.
of grammaticalisation communicates a tendency towards a natural sequence of these processes (lexeme in specific linguistic context > syntactic construction > morphological element; cf. Hopper and Traugott 1993: 100). Recall that grammaticalisation (as the result of a process of linguistic change that involves an increase in grammaticality) does not predict a firm chain of individual processes that are involved in a specific order. Therefore the whole discussion on unidirectionality (as a core argument for or against grammaticalisation) is misleading and based on a misconception of grammaticalisation. Speakers tend towards a shift from functionally determined structures to formally determined ones instead of establishing new semantic or functional bases of the same structure. This tendency has been observed and described, and may be related to a preference for structural simplification (cf. Fischer 2007: 470 and Roberts and Roussou 2003: 234) in terms of an “upward reanalysis that gives rise to new functional material” (Fischer 2007: 472):

“To summarize, structural simplification can be defined as a way of avoiding feature syncretism. [...] Once the cue (that is morphology mainly) becomes obscure or ambiguous the conservative nature of the language acquirer will opt for a simplified structure: maximize the correspondence between structure and lexical items. This, of course, yields new exponents for functional features, as is indeed what we get in grammaticalization cases.” (Roberts and Roussou 2003: 205)

This natural preference points to unidirectionality as an important property of grammaticalisation. The whole progress of grammaticalisation includes three stages: The use of lexical elements in specific linguistic contexts, the reanalysis of these patterns as syntactic constructions and a further reduction of these constructions to morphological elements (cf. Hopper and Traugott 1993: 100). Since grammaticalisation may stop at any point of its development, we often find ‘unfinished’ grammaticalisation phenomena. This is explicable by the motivation for grammaticalisation: A certain linguistic device that serves a certain function lacks its functional transparency. In order to regain transparency, the formal (and functional) shift takes place and the grammaticalisation process stops as soon as a satisfactory degree of transparency has been achieved. Due to the cyclic nature of language change, an earlier grammaticalised structure may be subject to a later, secondary grammaticalisation mechanism. As Lehmann (2004: 15) sums it up: “Grammaticalization is unidirectional in the sense that changes that go in the opposite direction of grammaticalization are observed very rarely” (cf. also Aikhenvald 2000: 374).

Another important factor that is related to both grammaticalisation and unidirectionality is divergent layering (cf. Hopper and Traugott 1993: 118-122 and 124-126): A linguistic item is situated on a certain layer and begins to grammaticalise. At the begin of this grammaticalisation process, a second layer is created for the grammaticalising item. In Palikûr, words for body parts (located on the first layer) are taking the function of numeral classifiers (e.g. uku/wok “hand”, biyu/biy “mouth”), ‘verb classifiers’ (e.g. kig “nose”) or both (e.g. kat
“vertical”, which originates from the denotation “trunk, stem of a plan”; all examples from Aikhenvald and Green 1998, cited in Aikhenvald 2000: 355 and 375). Hereby, a second layer arises, where the grammaticalised elements are located. This second layer separates the elements that are about to be grammaticalised from their original source. The fate of the lexical source of the first layer is independent from this new layer, hence it may continue its existence (as it is the case for our Palikûr examples) or be eliminated (in favour of new lexical material that takes over the source’s function).

**Decategorisation**  Decategorisation in general comprehends the transition from a (longer and) more specific unit to a (shorter and) less specific unit, which typically includes the loss of paradigmatic characteristics, phonological erosion and triggers recategorisation in a unidirectional manner. The creation of elements of a minor category (functional material such as prepositions, conjunctions, auxiliary verbs or pronouns) always involves decategorisation of elements of a major category (lexical material such as nouns or verbs; cf. Hopper and Traugott 1993: 106f. and [3.7] and [3.8]):

(3.7) German

\[
\begin{align*}
zi \text{ rucke} & \quad \rightarrow \quad \text{zurück} \\
\text{“to [one’s] back”} & \quad \rightarrow \quad \text{“back (ADV)“}
\end{align*}
\]

(3.8) Dutch (cf. van Wijk 1912: 695)

\[
\begin{align*}
te \text{ rugge} & \quad \rightarrow \quad \text{terugghe} \quad \rightarrow \quad \text{terug} \\
\text{“to [one’s] back”} & \quad \rightarrow \quad \text{“back, off oneself”} \quad \rightarrow \quad \text{“back (ADV)”}
\end{align*}
\]

While recategorisation (the assignment of an element to a new category) will not take place before an element’s decategorisation (the loss of characteristic paradigmatic properties), phonological reduction does not necessarily have to be involved (as shown by the English adverb *back*, which still takes the full form of its lexical root *back* “(human) rear”, and again the French negation particle *pas*).

**Grammaticalisation Cycles**  Once a grammaticalisation process is finished and the grammaticalised material has adopted its new function, this new category is not immune to further development. Since language is flexible and constantly changing, the grammaticalised material can either be further grammaticalised at a later step or further erode to zero. Its (imminent) loss can give rise to the creation of new successive functional material by new grammaticalisation processes. Similarly, newly created functional material can trigger the loss of older one (cf. again the French negation, which omits the original

---

6Cf. e.g. the noun classifier system of Chalcatongo Mixtec, which is extinct and only partially reconstructable on the basis of regular stem-initial phoneme constellations.

negator *ne* after the reanalysis of *pas* “step” as a negator). It is possible that this new material arises at a stage when the older functional material is still used by the speaker community (cf. DeLancey 2004: 1594f.).

The development of language can be of cyclic nature, as speakers show a need to vary the construction of a concept they wish to express in order to e.g. put emphasis on it. There are two main possibilities for means by which a certain grammatical function is expressed, and both tend to alternate at synchronic stages and to relate to each other in their diachronic development: Periphrastic and affixal constructions. Given an unmarked affixal construction and a more marked periphrastic construction (serving the same function), the affixal construction may fall into disuse, while the periphrastic construction becomes the unmarked one and may grammaticalise (cf. the future tense in Latin and the Romance languages). A grammaticalisation process (including the possible total erosion of the involved material) would then be one possible ‘round’ in this cycle, displacing the older construction at its beginning and making room for a succeeding one. After (or while) the formerly periphrastic construction gained its status as the unmarked construction, a new (more marked) construction may arise in order to express emphasis. This of course may start a new grammaticalisation cycle.

Reanalysis and analogy Two mechanisms of grammatical change often interact with grammaticalisation: Reanalysis and analogy (cf. Hopper and Traugott 1993: 39-70). Both may occur in a context independent from grammaticalisation (cf. ib.: 69), and grammaticalisation processes may operate independently from both.7 For reasons of clarity, I first present the commonly accepted view on the relation between reanalysis, analogy and grammaticalisation, before their independence from one another is discussed.

“Reanalysis” denotes a “change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface structure” (Langacker 1977: 58, cited in Hopper and Traugott 1993: 51. Cf. the reanalysis of German constructions that involve *um... zu...* “for... to...” in (3.9).

(3.9) German (Lehmann 2004: 9)

\[
\begin{align*}
A \quad & \text{ging} \quad [\text{um Wasser} \quad \text{zu holen}] \\
& \text{went} \quad \text{for water} \quad \text{to fetch} \\
& \text{“went in order to fetch water”}
\end{align*}
\]
\[
\begin{align*}
B \quad & \text{ging} \quad [\text{um} \quad \text{Wasser zu holen}] \\
& \text{went} \quad \text{for water to fetch} \\
& \text{“went in order to fetch water”}
\end{align*}
\]

7Reanalysis is based on an analogical model. Since there are cases of grammaticalisation that are not related to analogy, a one-sided dependency of grammaticalisation on analogy and reanalysis can be excluded (cf. Lehmann 2004: 7ff.).
Referring to nominal classification, this would mean the functional shift of a class of expressions (classifiers/concordial class markers) from semantics towards morphology. As soon as this shift has established, analogy makes the (invisible) structural change a visible one (cf. Hopper and Traugott 1993: 68). analogy as an organisational unit of a language’s paradigmatic apparatus unites the products of reanalysis with existing constructions and hence establishes systematic changes in the surface structure of a language. According to this view, new grammatical structures are created by reanalysis only, while only analogy makes them overt (cf. the analogical change of German constructions that involve um... zu... “for... to...” in [3.10] and Hopper and Traugott 1993: 63f. and 68).

(3.10) German

A Um (eine Zigarette) zu rauchen müssen Sie das Gebäude verlassen.
for a cigarette to smoke must you the building leave
“In order to smoke (a cigarette), you have to leave the building.”

B Um (Indogermanistik) zu studieren empfehlen sich
for Indo-European linguistics to study recommend REFLEX
Kenntnisse klassischer Sprachen wie des Altgriechischen
knowledge classical languages:GEN.PL like the Ancient Greek:GEN.SG
oder Védischen.
or Vedic Sanskrit:GEN.SG
“In order to study (Indo-European linguistics), knowledge of classical languages such as Ancient Greek or Vedic Sanskrit is recommended.”

However, Lehmann (2004) strongly argues against this view and emphasises the importance of a distinction of the processes of grammaticalisation, analogical change, lateral conversion, reanalysis and lexicalisation. Due to “the enormous diversity of opinion, numerous contradictions and weak points in the literature on grammaticalization” (ib.: 1), this issue is explored more in detail in the following. Grammaticalisation is not the only mechanism that creates grammatical items (ib.: 3), and the abovementioned processes involve a functional overlap with grammaticalisation in their own respective way. analogy may be a part of grammaticalisation ("analogically-oriented grammaticalisation"; ib.: 8), as it is the case in the development of the Latin full verb habere “to have” into the Romance future auxiliary (cf. [3.1] = [3.11] and Lehmann 2004: 8): [3.11]

A is the pre-grammaticalisation stadium, [3.11] B the initial stage of grammaticalisation in form of an analogical compensation of a reanalysis of the full verb habere “to have”. The periphrastic construction is reduced to and reanalysed as an affix due to reasons of linguistic economy, which is represented in [3.11] C. In [3.11] D, this affix is finally established as the canonical future marker of the respective Romance daughter language.

(3.11) Latin

A future tense: affixed marker:

\[ \text{ama-rc: ama-b-o, ama-b-is, ama-b-it} \]
\[ \text{love:PRS.INF love:FUT1SG love:FUT2SG love:FUT3SG} \]

“to love: I will love, you will love, he/she/it will love”

B full verb habe-re "to have" > future auxiliary (cf. English will):

\[ \text{amare habe-o, amare habe-s, amare} \]
\[ \text{love:PRS.INF have:PRS1SG love:PRS.INF have:PRS2SG love:PRS.INF} \]
\[ \text{have:PRS3SG} \]

“I will love, you will love, he/she/it will love”

C Reanalysis as syntactic affix:

\[ \text{amare habe-o} \]
\[ \text{love:PRS.INF have:PRS1SG} \]

“I will love”

D Reanalysis as lexical affix in the individual Romance languages:

\[ \text{amar-aio} \]
\[ \text{Italian amer-ò, French aimer-ai etc.} \]

“I will love”

In contrast, there are grammaticalisation processes that cannot rely analogically on an existing structure within the individual language, since the target function is not yet existing in the respective language ("pure grammaticalisation"; ib.: 7f.). This innovative potency is involved in the genesis of articles in the Germanic and Romance languages (cf. 3.3 = 3.12, 3.13, and 3.14):

(3.12) German

\[ \text{der, diu, das} \]
\[ \text{“this (m., f., n.; demonstrative pronoun)”} \]
\[ \text{der, die, das} \]
\[ \text{“the (m., f., n.; article)”} \]

(3.13) French

\[ \text{ille, illa} \]
\[ \text{“this (m., f.; demonstrative pronoun)” (Latin)} \]
\[ \text{le, la} \]
\[ \text{“the (m., f.; article)”} \]

(3.14) Italian

\[ \text{ille, illa} \]
\[ \text{“this (m., f.; demonstrative pronoun)” (Latin)} \]
\[ \text{il, la} \]
\[ \text{“the (m., f.; article)”} \]

Reanalysis is defined as “the assignment of a different grammatical structure” to a construction, while “[e]xpressions instantiating the construction are thereby not changed” (ib.: 8). The main difference between grammaticalisation and reanalysis is that grammaticalisation can be creative, whereas it has been
argued that reanalysis is closely related to analogy and thus cannot create a structure that is not yet existing in a language. The literature is somewhat contradictory at this point: While reanalysis is characterised as the preceding stage of analogy by some (e.g. Hopper and Traugott 1993: 68), others point out that “it has always been assumed that the output of a reanalysis presupposes an analogical model” (Lehmann 2004: 11). Reanalysis as a covert process is to a certain degree more likely to prepare the overt marking of its change by analogy. Still, it cannot be claimed that either view is right or wrong, since systematic changes that involve reanalysis and analogy often are of more complex nature and tend to interact, as the development of the English auxiliary be going to illustrates (cf. [3.15] and Hopper and Traugott 1993: 69).

(3.15) English

I. be progressive going V(directive) to visit Bill purpose clause
II. [be going to] visit Bill tense V(action) (by reanalysis)
III. [be going to] like Bill tense V (by analogy)
IV. gonna like/visit Bill tense V (by reanalysis)

While reanalysis thus generally precedes analogy, analogy and reanalysis may occur at the same time and interact once analogy has entered the process. Besides its inability to create ‘brand-new’ material, reanalysis does not involve any particular direction of development and constitutes a categorical process, whereas grammaticalisation involves unidirectionality and is a gradual process (ib.: 10).[8]

Lexicalisation Finally, another process that fuses juxtaposed elements and includes the loss of autonomy has to be distinguished from grammaticalisation: Lexicalisation (cf. Lehmann 2004: 13f.). The difference is that the target module of grammaticalisation is the grammar, while lexicalisation results in transferring an item into the lexicon (cf. [3.16]).

(3.16) Italian

all’ arme > allarme
to the arms alarm
“To arms!” > “alarm”

[8] Reanalysis is reversible. As an example, consider the English plural suffix. On the one hand, the OE [Old English; MP] plural form trew-es is reanalyzed as the singular truce. On the other hand, the Old French cerise(s) was borrowed into English in the form /ʃriːz/, and this was reanalyzed as being the plural form cherri-es.” (Lehmann 2004: 22).
The properties that distinguish grammaticalisation from all other kinds of grammatical change are its innovative potency and unidirectionality.

**General Shift-Related Phenomena** After several specific considerations on grammaticalisation, this section concludes by their integration in a more general, theoretical level: What general characteristics do the development of nominal classification (lexical item > classifier > concordial class marker) and the processes of grammaticalisation have in common? A brief answer to this question was given by Lehmann (2004): “Where degrees of freedom shrink, automatization comes in.” (Lehmann 2004: 25) The development of systems of both grammaticalisation and nominal classification usually involves a certain loss of specificity. The hypothesis of a natural tendency of systems of nominal classification developing from a semantic-pragmatic tool towards a more or less mechanical and automated ordering principle of linguistic form is consistently rendered by the following remark of Lehmann (2004: 26) on grammaticalisation: “Decrease of information means reduction of choices; reduction of choices means automatization.” This includes the loss of speaker-sided choice (which is present in classifier systems) by conventionalising pairs of nouns and classifiers, while there is little purely speaker-sided choice in concordial class assignment.  

The automatisation of a system of nominal classification has an important side effect: Classifier systems mainly serve a semantic-pragmatic function, while concordial class systems do not. Instead, the latter pattern the nominal lexicon with respect to its formal behaviour (cf. also Irslinger 2009: 3f.). In other words, there is a crucial point in the development of systems of nominal classification where it becomes more important that each noun is assigned to a or class rather than categorising only certain nouns or nouns in certain morphosyntactic environments. What kind economic advantage arises from such-like functional shift? Experimental research on speech production in languages that employ grammatical gender (cf. Plemmenou/Bard/Branigan 2002, Monpiou/Metz-Lutz/Wioland/Brock 1995, Carstairs-McCarthy 1994) show that the procession of information is accelerated if the subject is a native speaker. The procession of information was retarded if the subject was a second language speaker (cf. Irslinger 2009: 3f.). In other words, unconscious access to concordial class categorisation promotes speech procession, while conscious access hampers it.

In his research on cross-linguistic categoriality (also “category determinacy” or “categorial specificity”) of roots, stems, and word classes, Lehmann found different levels of categoriality in the individual languages: Languages with a high categoriality of roots and stems did not allow much speaker-sided influence, while languages with low categoriality of roots and stems did. Lehmann explains the existence of varying levels of categoriality between languages as follows: “[…] the raison d’être of (virtual) langue is to render (actual) parole possible. Consequently, the categorization of units of langue, like roots and stems, is, so to speak, a preliminary categorization or precategorization whose purpose is to unburden categorization in parole” (Lehmann 2008b: 17). The mentioned (pre-)categorisation of roots and stems happens at a lower (or more automatised) level of language procession and is opposed to the categorisation of word forms, which happens on a higher (or less automatised) level and may be influenced by the speaker.

Note that Lehmann’s view does not represent a commonly accepted view on language. While his view on language focuses on language as a medium for mainly communicative purposes,
The catchphrase “from discourse to syntax” renders the natural development that is assumed in the working hypothesis. The relation of grammaticalisation and information structure, a branch of discourse pragmatics, is explored in Lehmann (2008a). Since classifier systems are assumed to be the initial stage of nominal classification and to mainly serve discoursal function, Lehmann (2008a) is relevant for this study’s theoretical frame. The basic proposition of Lehmann (2008a: 2f.) is that information structure has two sides: While the structural means by which information structure is expressed are of a lexical and grammatical nature (which serve cognitive functions), the function of information structure itself is of a discoursal nature. This is reflected in classifier systems: A grammatical element (the classifier) serves the discoursal individualisation of the referent. The grammatical constructions that encode information structure are the item that undergoes grammaticalisation. The same accounts for a classifier in case of grammaticalisation: It is the classifier as a linguistic unit that is grammaticalised. Since “any construction that may serve as the input to some grammaticalization channel is shaped by information structure” (ib.: 15), grammaticalisation and information structure influence each other: “grammaticalization constrains and formalizes information structure, and information structure conditions and directs grammaticalization” (ib.: 16). If this is true, a specific circumstance under which systems of nominal classification start to develop cannot be isolated. Neither would the loss of knowledge about the discoursal function of a classifier trigger its grammaticalisation, nor would an incipient grammaticalisation process of a classifier trigger the loss of knowledge about its discoursal function on its own. Instead, both processes would emerge parallel and reinforce each other.

If a linguistic unit undergoes grammaticalisation, it solidifies by conventional use, loses structural autonomy, and becomes subject to the systematic constraints of a specific language (Lehmann 2005: 2). Imagine a lexical word becoming a classifier: While the former lexical word would take a position as the lexical head of a noun phrase, the new context which may host the classifier is restricted to a sublevel of the noun phrase. The same applies for the transition of a classifier system to a concordial class system: The classifier, originally an independent NP-constituent, becomes obligatory and enters a formal liaison with its lexical head and/or its modifiers and therefore further loses structural autonomy.

As a last point, the gradualness of the assumed change is addressed. Grammaticalisation is per se a gradual process (cf. Lehmann 2004: 10). One argument for the gradualness of change within the speakers’ community as a whole is the parametric change in the context of child language acquisition, which is assumed to take part in the development of nominal classification. It is closely connected to the parallel occurrence of two layers of grammaticalisation levels at one synchronic stage (cf. Aikhenvald 2000: 375). The initial stage of a generative notion of language such as the Chomskyan one would strongly disagree, ignore the communicative function of language and “focus attention upon the purely structural properties of languages” (Lyons 1981: 71f. on Chomsky 1957).

Grammaticalisation provides another argument for the assumption of gradualness in the development of nominal classification: As outlined before, grammaticalisation relies on a productive syntactic structure. One or a few instances of this structure occur more often than the other possible instances and adopt their new function bit by bit; this also happens not in an abrupt, but a gradual manner. The expected development of a grammaticalising structure occurs in the form of a sigmoid curve or S-curve (cf. figure 3.3), where the number of grammaticalising constructions increases between the stages of the productive original structure and the established grammaticalised structure (cf. also Weerman 2005: 1).

![Figure 3.3: Sigmoid curve.](image)

Gradualness pervades the concept of language change and thus may be specified for the context of nominal classification. Intrasystemic changes are of a gradual nature: A classifier for example may develop slowly and individually or change its specific classifying function within the classifier system; this is expected to be mirrored by a step-by-step shift in its properties. This may even result in local or speaker-individual manners of classifier use (cf. Lakoff 1986: 27f.) or code-specific classifiers, which are only used if a certain linguistic code (or register) is chosen for by a speaker (e.g. to address a socially superior or inferior opponent).

Section 3.2 elaborates the presented general theory on grammaticalisation with respect to the grammaticalisation hypothesis: It addresses the processes that are involved in grammaticalisation processes and relates them to the development of properties of nominal classification devices, before section 3.3 puts these specific theoretical conjectures into a more general context and filters out a package of properties that allow for a comparative study of nominal classification devices.
3.2 The Grammaticalisation Hypothesis: What Changes When and Why?

The by far most common conceptualisation of nominal classification involves two systemic subtypes, classifier and concordial class systems, constituting a functional continuum. This continuum is characterised by two functional poles, a semantics-based one on the one hand and a morphosyntax-based one on the other hand. The semantics-based pole is typically associated with the largely interactive and operationally speaker-sided classifier systems, whereas the morphosyntax-based pole is largely associated with the more automaticised and operationally system-sided concordial class systems. The glue that holds this continuum together is grammaticalisation. Take 3.17–3.19: 2.16 = 3.17 features a typical classifier construction, which displays a high systemic semanticity and a contribution to the noun phrase’s semantics. 3.18 features a highly intermediate system of nominal classification: It allows the speaker to classify with a rather big inventory of 72 classes, which is considered a property typical of classifier systems, but at the same time seems to employ agreement, which is considered an exclusive property of concordial class systems (cf. e.g. Dixon 1982 and Grinevald 2000: 62). Note that the speaker may mark the nominal modifiers in 3.18 either as masculine or 2-dimensional and round; this classificatory choice is generally assumed to be an exclusive property of classifier systems, not of concordial class systems. In 3.19 a typical example of a concordial class system, the nouns’ concordial class membership does not contribute to the semantics of the noun phrase, but is overtly marked on a language-specific range of nominal modifiers.

(3.17) Jakaltek (cf. Craig 1986b: 263)

\[ \textit{naj} \text{CL:MAN} \textit{witz} \]

“the spirit of the mountain”

(3.18) Bora (cf. Seifart 2005: 169)

\[ \textit{e:-hi mähwe:bc kähnu:-hi} \]


“that big turtle”

(3.19) German

\[ \textit{Wär} \textit{niht das menschliche Auge \textit{sonnenhaft,}} \]

\[ \text{was not \textit{the: NOM.N.SG human: NOM.N.SG eye: NOM.N.SG sunnish}} \]

\[ \textit{die Sonne könnt’ es nie erblicken.} \]

\[ \text{the: ACC.F.SG sun: ACC.F.SG could it: NOM.N.SG never spot} \]

“Would the human eye not be sunlike, it never could see the sun.”
3.2. The Grammaticalisation Hypothesis

The formal properties of the classificatory markers in 3.17–3.19 reflect the assumed formal development of a grammaticalising element: An independent element marks class membership in 3.17, a cliticised marker in 3.18, and an affix in 3.19. This suggests that the semantics-based pole constitutes a natural entry point for systems of nominal classification. Once part of the nominal classification continuum, a system may further grammaticalise, which pushes it into the direction of the morphosyntax-based pole and away from the semantics-based pole (cf. figure 3.4). A more detailed description of this shift, its individual stages, and the properties that we expect to find there is provided in section 5.3.

Grinevald (2000: 61) defines classifier systems as overt systems of nominal classification, which are of a clear lexical origin and used in specific morphosyntactic constructions. They are different from gender/noun class systems in that classifier systems are incompletely grammaticalised, interact with only specific morphosyntactic constructions, and maintain a lexical nature to a certain degree. They are also different from the nominal lexicon and lexical systems (such as measure and class terms; cf. e.g. German drei Glas Bier “three glasses [of] beer”) in that they denote nominal categories beyond the noun itself (ib.; cf. e.g. 3.2 = 3.20, where the classifier developed a broader semantics than its source noun).

(3.20) Minangkabau (Marnita 1996: 107 and 131; cited in Aikhenvald 2000: 375)

A  jari  >  jari
   “finger”    “CL:CLOTH.MEASURE”

B  ikua  >  ikua
   “tail”     “CL:ANIMAL”

Figure 3.4: The functional continuum of nominal classification and its relation to grammaticalisation.
Classifier systems thus display a state of intermediacy between the lexicon and a purely grammatical means such as gender/noun class systems (cf. figure 3.5).

Figure 3.5: Position of classifier systems opposed to lexical items and concordial class systems on a lexico-grammatical axis (adapted from Grinevald 2000: 61).

Note that the scenario CLS>CCS is not the only one possible — I hypothesise neither that all classifier systems develop into concordial class systems, nor that all concordial class systems stem from classifier systems. The continuum may have different entry points, which is also suggested by studies that identify e.g. determiners or articles (and thus not classifiers) as the sources of concordial class markers (cf. Corbett 1991: 312) or by the approach that considers gender “from below” (cf. Luraghi 2011).

In the following, the relation of the specific properties of nominal classification to specific subprocesses of grammaticalisation will be elaborated. An order of the individual systems of nominal classification along the lines of grammaticalisation is expected to mirror a prevailing particular order of the occurring grammaticalisation processes. Figure 3.2 = 3.6 displays the processes that are involved in grammaticalisation in their preferred order, which are expected to be reflected in the properties of the individual systems from the language sample.

1 It is important to remember that grammaticalisation may stop at any point of its development and thus does not necessarily involve all processes listed in figure 3.2 = 3.6. Therefore, it is better conceptualised as a number of related processes than a strict one-way road. The processes are expected to show the preferred order as it is given in figure 3.6 but this order is not carved in stone, since individual processes may be omitted.
3.2. The Grammaticalisation Hypothesis

The processes are discussed individually below and related to the earlier introduced properties of nominal classification. It is not possible to relate all properties to distinct subprocesses of grammaticalisation, since the subprocesses do not allow compulsory predictions on phenomena such as the locus operandi (i.e. the morphosyntactic context in which the classificational markers occur). However, a number of properties displays a logical relation to processes of grammaticalisation; these expected relations are discussed below.

**Productive syntactic structure**  This stage represents the subject of grammaticalisation before this process starts. As an example, take the ‘unitisation’ of German mass nouns. Example (3.21) demonstrates the use of fully lexical nouns as a means of measure or a unitising instrument. This ‘unitisation’ is a fully productive process, which includes the insertion of a noun whose properties determine an entity’s size, amount, volume, etc. The full lexical status of the nouns in question is determinable by the occurring plural inflection.

(3.21) German

\[
\begin{align*}
A \quad & \text{drei } \underline{\text{Flaschen/Gläser/Fässer}} \quad \text{Bier} \\
& \text{three } \underline{\text{bottles/glasses/barrels}} \quad \text{beer} \\
& \text{“three bottles/glasses/barrels [of] beer”} \\
B \quad & \text{drei } \underline{\text{Scheiben/Krumen/Läber}} \quad \text{Brot} \\
& \text{three } \underline{\text{slices/crumbs/loafs}} \quad \text{bread} \\
& \text{“three slices/crumbs/loafs [of] bread”} \\
C \quad & \text{drei } \underline{\text{Blätter/Bögen}} \quad \text{Papier} \\
& \text{three } \underline{\text{sheets\textsubscript{1}/sheets\textsubscript{2}}} \quad \text{paper} \\
& \text{“three sheets [of] paper”} \\
D \quad & \text{drei } \underline{\text{Meilen}} \quad \text{Umfang} \\
& \text{three } \underline{\text{miles}} \quad \text{perimeter} \\
& \text{“a perimeter of three miles”}
\end{align*}
\]
Semantic Bleaching  The first assumed indication of a beginning process of grammaticalisation is the loss of semantic content of the structure in question. It generalises its meaning and is used in contexts that would not qualify for this structure in its original meaning (cf. \(3.22\) where German *Stangen* “rods” captures the shape of the dynamite and the celery in question, but does not quite capture the shape of a carton of cigarettes). The process of semantic bleaching may continuously apply to classifier systems, where the semantic scope of individual classifiers changes (cf. e.g. Aikhenvald 2000: 409f.).

(3.22) German

<table>
<thead>
<tr>
<th></th>
<th>drei</th>
<th>Stangen</th>
<th>Dynamit/Sellerie</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>three rods</td>
<td>“three rods [of] dynamite/celery”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>drei</td>
<td>Stangen</td>
<td>Zigaretten</td>
</tr>
<tr>
<td>B</td>
<td>three rods</td>
<td>“three cartons [of] cigarettes”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>eine</td>
<td>Stange</td>
<td>Geld</td>
</tr>
<tr>
<td>C</td>
<td>a</td>
<td>rod</td>
<td>money</td>
</tr>
</tbody>
</table>
|   | “a packet [of] money” (only singular)

A grammaticalising structure is expected to reduce its productivity: Where earlier a relatively free integration of items is possible, a grammaticalised structure is expected to display delimited possibilities (cf. again the grammaticalisation of the French negation as it is described on p. 56 and Hopper and Traugott 1993: 117). In the case of the grammaticalisation of a classifier system, the system’s inventory size is expected to reduce due to the expansion of the semantic scope of a few classifiers and the reduction or loss of the use of other classifiers. Note that the semantic potency of a classifier system is not necessarily affected by the semantic bleaching effects, as these only expand the original semantic scope of a linguistic item. Cf. e.g. the development of the Mandarin...
classifier "small branch" > "sticks" > "snakes, lengths of cloth, strings of gold ingots" > "string, clothing" > "long things" (Aikhenvald 2000: 410), which clearly carries semantics despite the far progressed grammaticalisation of the Mandarin classifier system.

Even if the system as a whole loses most of its semantic content (or the existing semantics are obsolete due to changes in the socio-cultural speaker environment; cf. Aikhenvald 2000: 347f.), the semantic remotivation of the system by the speakers is possible (cf. the case of Dyirbal, which has lost the transparency of its system due to the obsolescence of its speakers’ genuine culture, and consequently was reanalysed; cf. section 8.2). A system’s semantic transparency is one of the crucial criteria for the speaker’s ability to assign nouns freely at his or her own will. If a system semantically bleaches, this ability may decrease. Since a stark semantic bleaching is not obligatorily involved in grammaticalisation, a productive speaker-sided assignment does not necessarily vanish at all and thus does not have to be expected.

**Decategorisation** The decategorisation of the structure in question comprises the loss of morphosyntactic characteristics of its category. In 3.23, this happens by the decrease of the structure’s ability to form a plural in the specific context of quantification.

(3.23) German

<table>
<thead>
<tr>
<th></th>
<th>drei</th>
<th>Pfund</th>
<th>Fleisch</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>three</td>
<td>pound</td>
<td>meat</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“three pounds [of] meat”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3.23</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>dre</td>
<td>drei</td>
</tr>
<tr>
<td>drei</td>
<td>drei</td>
</tr>
<tr>
<td>Loaf/Loaves</td>
<td>Paar</td>
</tr>
<tr>
<td>three</td>
<td>three</td>
</tr>
<tr>
<td>loaf/loaves</td>
<td>pair</td>
</tr>
<tr>
<td>“three loaves [of] bread”</td>
<td>“three pairs [of] shoes”</td>
</tr>
</tbody>
</table>

Pfund “pound” and Paar “pair” has a distinct plural form Pfunde and Paare and thus is identifiable as a singular form in 3.23. Note that the process of decategorisation is a gradual one. In German, measure nouns often can occur in both the singular (i.e. grammaticalised) or the plural (i.e. in its original lexical) form, but there are only a few examples of measure nouns that occur only in the singular (e.g. those in 3.23). All examples in 3.24 are acceptable in both the singular and the plural, but the original lexical plural forms are all less acceptable than the grammaticalised singular forms.

(3.24) German

<table>
<thead>
<tr>
<th></th>
<th>drei</th>
<th>Laib/Laibe</th>
<th>Brot</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>three</td>
<td>loaf/loaves</td>
<td>bread</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“three loaves [of] bread”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>drei</th>
<th>Stück/Stücke</th>
<th>Fleisch</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>three</td>
<td>piece/pieces</td>
<td>meat</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“three pieces [of] meat”</td>
</tr>
</tbody>
</table>
This phenomenon may also be sensitive to the constructional context (cf. where only singular forms are possible, not the plural forms Stücke and Männer respectively).

A potential trigger for these forms may be Scharnierformen (“hinge forms”), which employ the same surface form in the singular and the plural (cf. 3.26).

The flexibility of a noun’s class membership is related to both the grammaticalising item’s semantic bleaching and its decategorisation: The more the classifiers expand their semantic scope and certain classifiers are conventionised in their combinability with certain nouns, the less flexible the members of the nominal lexicon are in their class membership.
3.2. The Grammaticalisation Hypothesis

Phonological erosion, cliticisation, and morphologisation  The grammaticalising element has already partly lost content and function that relate it to its linguistic source item; in the above examples, this would be the status and full meaning of a lexical noun and its ability to form a plural. This may lead to a reanalysis of the grammaticalising element and may result in the assignment of this element to a language-internally already existing or even new category. For instance, measure nouns e.g. might be reanalysed as numeral classifiers. Since the bonds with its source material are weak or even vanished, the grammaticalising element may lose formal weight and erode phonologically (cf. 3.27).

(3.27) German

A drei Mand voll Suppe > drei Mand voll Suppe
three mouth full soup three mouthful soup
“three ‘mouthfuls’ [of] soup”

B drei Hand voll Erde > drei Hand voll Erde
three hand full earth three handful earth
“three ‘handfuls’ [of] earth”

The item may also enter a liaison with its constructional neighbours (e.g., a quantifier in numeral classifier languages; cf. 147 B = 3.28 A). It is however not neccessary for the item to erode phonologically before it becomes decategorised, a clitic or an affix (cf. 3.28 B for and 3.20 = 3.29 for examples from Mandarin and Minangkabau).


A nà yí-pì hēi mā
DEM:REMOTE one-CL:HORSE horse
“That black horse.”

B lun > (=)lun
wheel > CL:ROUND
“wheel, round”


A jari > jari
“finger” “CL:CLOTH.MEASURE”

B ikua > ikua
“tail” “CL:ANIMAL”

The stage of phonological erosion is a gradual one (cf. table 2.3 = 3.1 for an example from Persian, where the individual classifiers display varying degrees of erosion).
A second indicator for the gradualness of the cliticisation and morphologisation of a grammaticalising category (e.g. of classifiers) are systems that employ both free and bound forms of this category’s members. This is due to the set-internal fluctuation of a classifier system, where the individual classifiers may compete with each other: Specific classifiers grow semantically more and more general, general classifiers may drop out of the system, new lexical items enter the set and turn into (specific) classifiers; cf. Aikhenvald (2000: 409f.) for an overview over the rise, expansion, and decay of individual Mandarin classifiers. Suchlike classifier sets thus synchronically employ classifiers of a different age. Cf. [3.30] A and [147] – [3.30] B, where the same Mandarin classifier may occur in both a bound and a free form according to their constructional environment and function, and table 3.2 from Korean, where the genuine Korean classifiers are consequently free and Sino-Korean loan classifiers are consequently bound.


A  yì-pī  bēi  mǎ  one-CL:HORSE  black  horse  “a black horse”
B  mǎi  pī  mǎ  buy  CL:HORSE  black  horse  “Buy one black horse.”

Table 3.2: The inventory of Korean numeral classifiers (selection; cf. Martin 1992: 179ff.).

<table>
<thead>
<tr>
<th>Korean form</th>
</tr>
</thead>
<tbody>
<tr>
<td>phil</td>
</tr>
<tr>
<td>chay</td>
</tr>
<tr>
<td>tay1</td>
</tr>
<tr>
<td>kuleu</td>
</tr>
<tr>
<td>pen</td>
</tr>
<tr>
<td>pāng</td>
</tr>
<tr>
<td>chung</td>
</tr>
<tr>
<td>phil</td>
</tr>
<tr>
<td>-ho, -tong1</td>
</tr>
<tr>
<td>-tay</td>
</tr>
<tr>
<td>-cwu</td>
</tr>
<tr>
<td>-hoy</td>
</tr>
<tr>
<td>-pal</td>
</tr>
<tr>
<td>-chung</td>
</tr>
<tr>
<td>category</td>
</tr>
<tr>
<td>horse</td>
</tr>
<tr>
<td>buildings</td>
</tr>
<tr>
<td>vehicles, machines</td>
</tr>
<tr>
<td>trees, shrubs</td>
</tr>
<tr>
<td>times</td>
</tr>
<tr>
<td>(gun)shots</td>
</tr>
<tr>
<td>floors, stories</td>
</tr>
</tbody>
</table>

Marking is expected to occur multiply only after the grammaticalising element has passed the stage of cliticisation and/or morphologisation. A pre-
dictable obligatory classification requires an obligatory means to mark the obligatory category - this obligatory mean is only provided by inflection. Therefore, an obligatory classification can be expected only after the marker’s morphologisation. Accordingly, we find systems of nominal classification that employ clitics as their referential class markers, which are optional at least in certain contexts (cf. 3.31).

(3.31) Great Andamanese (cf. Abbi 2011: 762)

A  tokbi  t-ot-omu
    turtle  INANIMATE-CL:4=flesh
    “turtle meat (cut up for consumption)”
B  tokbi  at=omu
    turtle  CL:4=flesh
    “turtle meat (still on its body)”
C  tokbi  omu
    turtle  flesh
    “turtle meat (general, non-specific)”

The following subsection merges the presented theory on nominal classification, grammaticalisation, and the grammaticalisation hypothesis and discusses it in a more general theoretical frame. Based on this discussion, section 3.3 establishes a package of properties that serves as a measure of the grammaticalisation of systems of nominal classification.

3.3 Conclusion: A Paradigmatic Shift of Properties

This subsection puts the presented theory on grammaticalisation and nominal classification in a more general context. It assembles a package of properties that together constitute a measure for nominal classification systems; this measure marks characteristics of nominal classification devices as either classifier-like or concordial class-like.

Each of the three stages of the assumed shift [lexical item > classifier system > concordial class system], whose second and third stage are in the main focus of this study, is expected to show a higher degree of grammaticalisation than the previous one. A higher degree of grammaticalisation equals an increase in morphosyntactic dependency and a decrease in semantic content of the respective system’s marker. It is also assumed that this difference in the degree of grammaticalisation is visible in the diachrony of a system as well as in

\footnote{An obligatory classification can occur before the establishment of agreement, if an individual system’s semantics allow a categorisation of all members of the nominal lexicon and the speakers’ conventions demand it. However, obligatoriness that is established in this manner cannot be predicted by argumentation, it can only be expected by the linguistic tendency of conventionalisation and generalisation.}
the comparison of synchronic properties of various existing systems at different developmental stages.

What happens if systems of nominal classification develop? In the beginning, we find lexical nouns that change their category and become classifiers (Senft 2000: 39f.; cf. 3.20 – 3.32).


A jari “finger” > jari “CL: measure for cloth”

B ikua “tail” > ikua “CL: animal”

Culture-specific practices and specific types of classifiers determine, which nouns grammaticalise into classifiers (cf. Seifart 2010: 728 and Bisang 1999: 22f. for an evaluation of the cognitive foundations of classification). According to the literature, classifier systems again may develop into concordial class systems (cf. ib.: 722). This case was reconstructed for e.g. Proto-Australian (cf. 3.33; Aikhenvald 2000: 372f.).

(3.33) Proto-Australian

I. generic-specific pairings as a discourse device (generic: mayi)

II. generic > obligatory NP-component as noun classifier (omissible under specific conditions)

III. noun classifier > noun class affix, subsequent phonological reduction:

1. Dyirbal: ‘agreement’ marker -m “vegetable food” (attached to deictics and determiners)

2. North Australian languages: nominal prefix and agreement marker ma- (mi-, m-):

   - Ngandi: “vegetable food”
   - Yanyuwa “non-flesh food, firesticks, articles made of fibrous materials, etc.”
   - Wardaman: “flora”

This means that the source of nominal classification typically is a lexical item; this item loses its specific lexical denotation in order to gain the function of a classifier (cf. 3.34–3.37).


14 The three-stage transition [lexical item > classifier system > concordial class system] is not meant to constitute one continuous process of grammaticalisation. Lexical items grammaticalise into classifiers, and may further grammaticalise to concordial class markers by a secondary process of grammaticalisation. This may also be projected on the distinction of an “extended” grammaticalisation approach (where an item’s host-class relation as well as its syntactic and semantic-pragmatic relation are expanded) and an “traditional” grammaticalisation approach (where an item’s grammaticalisation is interpreted as a formal and structural reduction, which is accompanied by an increasing degree of dependency; cf. Traugott 2011: 97f. and p. respectively.)
3.3. Conclusion: A Paradigmatic Shift of Properties

A lexeme urang “person, human” > generic classifier urang “person, human”
B lexeme ikan “fish” > generic classifier ikan “fish”
C lexeme buruang “bird” > generic classifier buruang “bird”

(3.35) Kana (Ikoro 1996: 90 and 94f., cited in Aikhenvald 2000: 360)
A lexeme ákpó “length” > numeral classifier ákpó “inanimate objects with a trunk”
B lexeme kpó “head” > numeral classifier kpó “objects with irregular or head-like shape”

A lexeme mwar “title” > possessed classifier mware “title/garland”
B lexeme ahd “name” > possessed classifier ede “name”

A lexeme jil “wild animal” > noun classifier jil “non-human”
B lexeme ko “heart” > incorporated verb classifier ko “round object”

This semantic property may be further developed into a lexico-syntactic category and thus gain further grammatical function while losing its lexical qualities. This phenomenon constitutes the focus of this study. Based on Dixon (1982, 1986) and a number of later typological works on nominal classification (e.g. Grinevald 2000, Seifart 2010), I collected a number of properties that are characteristic for both systemic types of nominal classification; each property has two values, corresponding to either a prototypical classifier system or a prototypical concordial class system. Due to their binary nature, they serve as a basic means to capture a given system’s position in the nominal classification continuum and especially the grey area.

In a nutshell, the prototypical properties of classifiers and concordial classes are as follows: Ideal classifiers are optional (in that they categorise nouns only in specific morphosyntactic contexts; in these contexts, they are obligatory), semantically potent operators that shape their referent noun’s meaning in terms of the discoursal needs of the speaker. They occur as variable, independent or clitic morphemes within the noun phrase and adjacent to a predictable modifier of the head noun and are selected by the speaker. Ideal concordial classes constitute an obligatory (in that each noun inherently belongs to a concordial class under any circumstances) system whose markers are morphological operators that structure the nominal lexicon with respect to the nouns’ form. They occur NP-internally and NP-externally as predetermined (and thus invariable and non-selectable) morphemes that may be fused with other grammatical categories and are bound to predictable modifiers of a noun and optionally the noun itself. The markers agree with the concordial class of the referent noun. The properties that were just listed are summarised in table 3.3.
Table 3.3: Properties of nouns and prototypical devices of nominal classification.

<table>
<thead>
<tr>
<th>assignment</th>
<th>nouns</th>
<th>classifiers</th>
<th>concordial classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>function</td>
<td>free choice</td>
<td>free choice</td>
<td>inherent to N</td>
</tr>
<tr>
<td>appearance</td>
<td>reference to entity</td>
<td>semantic specification</td>
<td>formal patterning</td>
</tr>
<tr>
<td>semantics</td>
<td>contentful</td>
<td>contribution to NP</td>
<td>bound</td>
</tr>
<tr>
<td>agreement</td>
<td>none</td>
<td>none</td>
<td>vain</td>
</tr>
<tr>
<td>locus</td>
<td>NP-head</td>
<td>predictable (NP)</td>
<td>obligatory</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>obligatory (clause)</td>
<td>optional (NP)</td>
<td>predictable</td>
</tr>
</tbody>
</table>

These are the components of nouns and nominal classification systems that are relevant for this study and common to all systemic manifestations investigated here. Since each property’s shift to the right corresponds to its expected behaviour when undergoing grammaticalisation, this represents theoretical evidence for the grammaticalisation hypothesis. They make a comparison of formal shifts between the single stages of development of nominal classification with respect to the grammaticalisation hypothesis possible. The functionality of the systems is only captured by the single criterion of semantic contribution to the noun phrase in this Dixon-based typology. Each property is individually treated below with respect to its application to lexical nouns, classifiers, and concordial classes.

**Interactive freedom:** A broad choice is present in the case of lexical items and classifier systems. The speaker is assumed to possess interactive influence on the assignment or choice of a class or category for a given noun in classifier systems. This choice ends as soon as a system enters the state of a concordial class system, since concordial class systems are responsible for the assignment of nouns to classes and not the speaker. 3.38 illustrates the choice of lexical items in order to refer to the same entity in a concordial class language. The choice of a certain lexical item highlights different facets of the entity.

(3.38) English: The dog/the collie/Lassie ran back home.

The same process in a classifier language is carried out by the choice of a certain classifier, as 2.10 = 3.39 illustrates. In classifier languages, it is not the choice of a lexical item that highlights different facets of the same entity, but the choice of a classifier.


A myi? ta ya?
river one CL:PLACE

“one river (e.g. destination for a picnic)”
3.3. Conclusion: A Paradigmatic Shift of Properties

Note that in concordial class languages, it is virtually always the choice of a lexical item that may be chosen by the speaker, whereas this lexical item’s concordial class is prototypically and invariably determined by the system itself (cf. 3.40 where all nouns belong to one concordial class only and cannot alter it).

(3.40) Polish

M: kołek, młotek, królik, grzech, bój, zmysł, róg,...
“stake, hammer, rabbit, sin, tallow, sense, horn,...”

F: książka, mgęka, nietrwałość, elektrownia, pokuta,...
“book, torture, perishability, power plant, repentance,...”

N: słońce, dziecko, auto, akwarium, czoło,...
“sun, child, car, fish tank, forehead,...”

**Function:** The raison d’être of nouns is their ability to refer to the mental representations of entities. These referential labels partake in discourse.

---

3.40 Nouns themselves are labels of mental concepts. These mental concepts again refer to (individual) real life counterparts. Nouns do not directly refer to individual real life referents. Taking an example like linguist, this becomes obvious: There is no direct relation to a specific individual, whereas all potential real life referents of course are individuals such as Kees Hengeveld, Heide Zeijlstra and Noam Chomsky. Nouns such as linguist, German or billiard cue therefore refer to groups of individuals (cf. von Stechow 2007: 16). Which individual actually is referred to is determined by the context of the discourse: The heads of the humanities’ departments had a meeting today. The linguist was chairing the meeting. Proper
There is no restriction on what a noun might refer to, hence they can denote any concept either physically existing in the reality (as a *Wirklichkeitsausschnitt* “extract of reality”; e.g., a stone, street, tree, house, etc.) or as an abstract concept in the human mind (e.g., alien, time travel, purgatory, etc.). This referential freedom is not given for classifiers, whose function consists in the restriction of a noun’s semantics. A noun thus denotes something within the scope of the universe or the human mind, while a classifier denotes something within the scope of the noun it classifies (either a taxonomic role of the referent or one of its inherent properties; cf. again (3.39). A concordial class marker does at best marginally relate to a semantic concept or aspect, while its characteristic relation is the one to a distinct morphological pattern; i.e., it copies and displays properties of the head noun.

**Formal appearance:** Nouns are independent constituents and head their own phrase (cf. (3.41)–(3.43); the respective NP-heads are marked by underlining in the examples (3.41) to (3.45)).

(3.41) French

<table>
<thead>
<tr>
<th><em>La lune</em></th>
<th><em>cette nuit</em></th>
<th><em>veillé</em></th>
<th><em>mon amie</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>the moon</td>
<td>that night</td>
<td>have:PRS3SG</td>
<td>watch:PRT.PASSÉ</td>
</tr>
</tbody>
</table>

“It the moon watched over my friend that night.”

(3.42) Latin

| *Meum* | *propositum* | *in taberna* | *mori.* |
| my      | intention    | in tavern    | die:INF.PRS |

“It is my intention to die in the tavern.”

(3.43) German

Nach *einem kurzen, aber heftigen Gewitterregen* besuchen wir after a short but severe thundershower visit:PRS1PL we zusammen *die Wiese* und *das Venusbild.* together the meadow and the Venus figurine

“After a short, but severe thundershower, we visited together the meadow and the Venus figurine.”

names (e.g., *Noam Chomsky*) and names of unique entities (e.g., *homomonument*) also do not refer to an individual or a specific entity, but to the mental representation of that individual or entity on the speaker’s or hearer’s mind. For simplicity’s sake, the indirect reference of nouns to real life referents via the mental representations of these referents is not always explicitly mentioned here. A phrase like *the noun refers to X* must be understood as “the noun refers to the mental representation of X on the speaker’s or hearer’s mind, which again refers to the actual entity X.”
Classifiers are typically morphosyntactically independent, but depend on the occurrence of a nominal modifier that triggers their insertion. They are situated lower in the NP-internal structure as compared to free nouns. Note that noun classifiers occur as direct nominal modifiers (cf. 3.44 where the classifier categorises the noun without a coocurring modifier).


\[
xil [\text{naj} \text{n}u]w\text{n} \text{xuwan}\text{NP} \ [\text{no7} \text{lab'}a]\text{NP}.
\]

“Juan saw the snake.”

An aged system that develops into a concordial class system may allow the classifiers to cliticise with its adjacent nominal modifier (cf. 3.45). Concordial class markers are bound morphemes and occur on a range of nominal modifiers and (optionally) on the noun itself (cf. 2.67 B = 3.46).


\[
[pajkuhua-na yahui]NP
\]

“one dog”

(3.46) Latin

\[
[me-a maxim-a culp-a]\text{NP}
\]

“Through my most grievous fault.”

**Semantics:** The transition from a noun to a classifier to a concordial class marker is accompanied by a decline of semantic content. Nouns as lexical items have a full semantic content (cf. 3.47). Classifiers are semantically potent; this means that they contribute to the distinct semantics of the noun phrase by narrowing down its semantic scope (or by highlighting a salient property of the referent that is of importance in the specific discourse context) in a manner that enables the hearer to identify this NP’s intended referent (cf. 3.48 and 3.49).

(3.47) snake,

\[
[\text{+animate, -human, +reptile, +carnivore, +squamate, etc.}]
\]

brick

\[
[-\text{animate, +artefact, +solid, +cuboid, etc.}]
\]


\[
[\text{no7 lab’a}]
\]

“snake”
(3.49) Minangkabau (Marnita 1996, cited in Aikhenvald 2000: 84)

A batang limau
CL:TREE lemon
"lemon tree"
B buah limau
CL:FRUIT lemon
"lemon fruit"

Concordial class markers do not primarily contribute to the semantics of the NP. A noun’s mere membership of a concordial class does not necessarily provide information on the referent (cf. (3.50) A - G, where no further information on the referent is provided by its membership of a concordial class).

(3.50) Indo-European

A Schlange (f.) “snake” (German)
B serpent (m.) “snake” (French)
C wąż (m.) “snake” (Polish)
D φίδος (n.) “snake” (Greek)
E змее (f.) “snake” (Russian)
F slang (f.) “snake” (Dutch)
G serpente (m.) “snake” (Italian)

It can however not be excluded that a certain semantic content has fossilised and was inherited to the concordial class system, since concordial class markers are assumed to stem from classifiers. A paragon of this principle is the distinction of sex by grammatical gender; cf. (3.51) where in A the mother language Latin allows a distinction of the referent’s sex, whereas Latin’s daughter languages have conventionalised either the masculine or the feminine gender as a default for both male and female snakes. In this case, the semantic heritage of the Latin concordial class system has been given up.

(3.51) Romance

A serpens (m./f.) “snake” (Latin)
B serpent (m.) “snake” (French)
C serpente (m.) “snake” (Italian)
D serpiente (f.) “snake” (Spanish)
E serpente (f.) “snake” (Portuguese)
F şarpe (m.) “snake” (Romanian)
G sèrp (f.) “snake” (Occitan)

Another example for an inherited semantic fossil with respect to nominal classification are default genders for certain classes of entities: For instance, trees always take feminine gender in Latin (even if the declensional class advises masculine gender; cf. (3.52) and nominalised infinitives take the neuter class in German (cf. (3.53)).
3.3. Conclusion: A Paradigmatic Shift of Properties

(3.52) Latin

*arbor, quercus, fagus, platanus, abies,* etc. (all F.)
“tree, oak, beech, plane tree, fir, etc.”

(3.53) German

*Hauen und Stechen, Beten und Arbeiten, Geben und Nehmen*
“the punching and stabbing, the praying and working, the giving and taking”

The default gender of such conceptual fields are semantically less transparent than e.g. the sex of animates. These semantic assignment principles have to be learned and actively used by the speaker. Therefore, the default gender is more likely to undergo changes or disappear at all, as it was the case for the originally feminine tree names in Latin, that are resumed with an individual gender for the individual tree name for each daughter language (cf. 3.54).

(3.54) Romance

A *quercus* (f.) “oak”, *fagus* (f.) “beech”, *abies* (f.) “fir” (Latin)
B *chêne* (m.) “oak”, *hêtre* (m.) “beech”, *sapin* (m.) “fir” (French)
C *quercia* (f.) “oak”, *faggio* (m.) “beech”, *abete* (m.) “fir” (Italian)
D *roble* (m.) “oak”, *haya* (f.) “beech”, *abeto* (m.) “fir” (Spanish)
E *carvalho* (m.) “oak”, *faia* (f.) “beech”, *abeto/pinheiro* (m.) “fir” (Portuguese)
F *ghindă* (f.) “oak”, *fag* (m.) “beech”, *brad* (m.) “fir” (Romanian)
G *casse/garric* (m.) “oak”, *fau/fag* (m.) “beech”, *avet* (m.) “fir” (Occitan)

If a system grows semantically opaque, speakers may undertake one or multiple processes of resemanticisation or reclassification, “which over time may lead to a considerable degree of overall semantic opacity” (Seifart 2010: 729).
In general, a coexisting semantically transparent and opaque form is in accordance with the assumed natural pathway of a nominal classification marker (cf. figure 3.5 = 3.7 below).
Note that this is the only criterion in the Dixon-based typology that accounts for the functionality of nominal classification. Contini-Morava and Kilarski (2013) show that the semanticity of nominal classification devices manifests itself in manifold ways. The sample study furthermore confirms that the not further specified formulation of a mere semantic contribution to the noun phrase is not sufficient to account for the richness of semanticity that nominal classification devices possess (cf. section 6.7). Therefore, a more fine-grained account for a system’s semanticity and transparency is presented in subsection 7.2.2.

Multiple marking: There are no obligations in terms of multiple marking by means of agreement for nouns (in the function of a lexical item, cf. (3.55)) and classifiers (cf. (3.56)).

\[(3.55)\]

A Die Katze ist tot. “The cat is dead.” (German)
B Kot jest zmarły. “The cat is dead.” (Polish)
C De kat is dood. “The cat is dead.” (Dutch)
D Katten är död. “The cat is dead.” (Swedish)

\[(3.56)\] Mandarin (Hashimoto 1971: 24f., cited in Rijkhoff 2002a: 333)

A nei ge dai yanjing de xiaohair hen
that CL:GENERAL wear glasses NOMINALISER child very
pang
fat
“The child, who wears glasses, is very fat.”

B dai yanjing de nei ge xiaohair hen
wear glasses NOMINALISER that CL:GENERAL child very
pang
fat
“The child who wears glasses is very fat.”

One might argue that all the nouns in (3.55) actually do agree in number and person with the finite verb. While this is true, it has nothing to do with
the noun in its role as a lexical item, but with it being the respective clause’s subject. As one can see in (3.57), no other clausal constituent agrees with die Katze/Katzen “the cat/cats” if it serves the function of the direct object.

(3.57) German

\[
\begin{align*}
\text{Der Hund} & \quad \text{biß} & \quad \text{die Katze/Katzen} & \quad \text{tot}, \\
& & \quad \text{the dog:NOM.M.SG} & \quad \text{bite:IPF3SG} & \quad \text{the cat:ACC.F.SG/PL} & \quad \text{dead}
\end{align*}
\]

“The dog bit the cat/cats to death.”

In concordial class systems, a range of language-specific modifiers of a noun (as a member of a concordial class) must mirror this noun’s concordial class (cf. 2.67 = 3.58 and 3.59[3.62]).

(3.58) Latin

A me-\text{o} maxim-\text{o} stupor-e
my-ABL.M.SG biggest-ABL.M.SG stupidity-ABL.M.SG

“through my enormous stupidity”

B me-\text{a} maxim-\text{a} culp-\text{a}
my-ABL.F.SG biggest-ABL.F.SG fault-ABL.F.SG

“through my most grievous fault”

C me-\text{o} maxim-\text{o} studi-\text{o}
my-ABL.N.SG biggest-ABL.N.SG pursuit-ABL.N.SG

“through my most intense pursuit”

(3.59) French

\[
\begin{align*}
\text{Mon} & \quad \text{amie} & \quad \text{la} & \quad \text{rose} & \quad \text{est} \\
& & \quad \text{my:NOM.F.SG} & \quad \text{friend:NOM.F.SG} & \quad \text{the:NOM.F.SG} & \quad \text{rose:NOM.F.SG} & \quad \text{be:PRS3SG} \\
& & & & \quad \text{morte} & \quad \text{ce matin}.
\end{align*}
\]

“My friend the rose has died this morning.”

(3.60) German

\[
\begin{align*}
\text{Jede} & \quad \text{Form} & \quad \text{von} & \quad \text{Amputation} & \quad \text{an} & \quad \text{einer} \\
& & \quad \text{every:NOM.F.SG} & \quad \text{form:NOM.F.SG} & \quad \text{of} & \quad \text{amputation} & \quad \text{at} & \quad \text{a:DAT.F.SG} \\
\text{Extremität} & \quad \text{beleuchtet} & \quad \text{für} & \quad \text{denn} & \quad \text{betroffenen} \\
& & \quad \text{extremity:DAT.F.SG} & \quad \text{mean:PRS3SG} & \quad \text{for} & \quad \text{the:ACC.M.SG} & \quad \text{affected:ACC.M.SG} \\
\text{Menschen} & \quad \text{den} & \quad \text{irreversiblen} & \quad \text{Verlust} \\
& & \quad \text{person:ACC.M.SG} & \quad \text{the:ACC.M.SG} & \quad \text{irreversible:ACC.M.SG} & \quad \text{loss:ACC.M.SG} & \quad \text{loss:ACC.M.SG} \\
\text{seiner} & \quad \text{körperlichen} & \quad \text{Integrität} \\
& & \quad \text{his:GEN.F.SG} & \quad \text{physical:GEN.F.SG} & \quad \text{integrity:GEN.F.SG} & \quad \text{his:GEN.F.SG}
\end{align*}
\]

“Every form of amputation at an extremity means to a person the irreversible loss of his physical integrity.”

\[
\text{kun-ekke kun-kku kun-bo-gimuk}
\]

4-That water:4 4-CL:LIQUID-big

“That river is big.”


\[
\text{uku-dla kw-ethu ko-nke ku-phelile}
\]

15-eat 15-our 15-all 15-is.finished

“All our food is finished.”

Since it is a range of nominal modifiers that agrees with the controller noun, the multiple occurrence of the same classificatory marker in the same classificatory context is a strong indication of the presence of agreement. During the collection of the sample data, multiple class marking thus serves as a criterion for the identification of markers as agreement markers. Note however that this preliminary notion of agreement is revised in section 6.6.

**Locus operandi:** The position of nouns, classifiers and concordial class markers are all predictable. Nouns occur in the argument position of a verb (cf. 3.63). Classifiers occur adjacent to the nominal modifier that is determined by the classifier type (cf. 3.48 = 3.64 (noun classifier), 2.26 = 3.65 (numeral classifier), and 2.27 = 3.66 (possessive classifier)).

(3.63) English

A X gave Y Z.
Santa gave the children their presents.

B X sleeps.
The dog sleeps.

C X is eating (Y).
The dog is eating (the cat).

(3.64) Jakaltek (Craig 1986b: 264, cited in Grinevald 2000: 65)

\[
\text{no7 lab’u}
\]

CL:ANIMAL snake

“snake”


\[
\text{?an ?ai khwan that phoon lwg}
\]

I have child four CL:PERSON

“I have four children.”

3.3. Conclusion: A Paradigmatic Shift of Properties

Concordial class markers occur on a range of nominal modifiers which is specified by the respective language’s grammar (cf. 3.67 for gender systems and again 3.59 C and D for noun class systems). Beyond that, they may occur on the noun they modify.

(3.67) German

Der, the:NOM.M.SG dicke, plump:NOM.M.SG Mann, man:NOM.M.SG welcher, who:NOM.M.SG

durch den Schornstein gekommen war, schien kein
through the chimney come:PRT.PST be:IPF3SG seem:IPF3SG no

Einbrecher zu sein. Er, burglar to be he:NOM.M.SG verteile:IPF3SG distribute:IPF3SG presents

“The plump man who had come through the chimney did not seem to be a burglar.
He distributed presents.”

These examples show that the respective item’s environment indicates a loss of autonomy and saliency from a clausal view: From the argument position in a VP as a lexical noun to an NP-internal position as a classifier to the position of a pre- or suffix of nominal satellites as a concordial class marker.

Obligatoriness: The status of obligatoryness of nouns, classifiers and concordial classes depends on the perspective on the respective item (cf. 3.63 – 3.68 – 3.70). Nouns are obligatory in most kinds of clauses, because they and only they can refer to the universe or concepts of the human mind. Since a discourse is inevitably about either the universe or concepts of the human mind, nouns are a fundamental linguistic category, which is often perceived as a universal one.

(3.68) English

A X gave Y Z.
B Santa gave the children their presents.
C *gave the children their presents.
D *Santa gave the children.
E *Santa gave their presents.

(3.69) English

A X sleeps.
B The dog sleeps.
C *sleeps.
**Grammaticalisation**

(3.70) English

A  X eats (Y).
B  The dog eats (the cat).
C  *eats (the cat).

Classifiers are optional in a sense that they do not automatically occur in a noun phrase. However, they do occur obligatorily next to the nominal modifier that determines the classifier type: If a noun is quantified by a numeral in a language with a numeral classifier system, a classifier will accompany it (cf. 3.71A). If it is an adjective that modifies the noun, it will not be accompanied by a classifier (cf. 3.71B).


A  Wo zhe-ge zueqi shangle san men ke.
   "I took three courses this semester."
B  jiā hé wàn shi xìng
   family harmony myriad thing rise
   "A harmonious family will prosper in everything it does."

In a concordial class system, each noun must belong to a concordial class without exception (cf. 3.40 = 3.72).

(3.72) Polish

M.:  kółek, młotek, królik, grzech, łój, zmysł, róg,…
    “stake, hammer, rabbit, sin, tallow, sense, horn,…”
F.:  książka, męka, nietrwałość, elektrownia, pokuta,…
    “book, torture, perishability, power plant, repentance,…”
N.:  słońce, dziecko, auto, akwaria, czoło,…
    “sun, child, car, fish tank, forehead,…”

In conclusion, the above criteria apply to both classifier systems and concordial class systems. Therefore, they can function as a comparative measure of the grammaticalisation of nominal classification devices. Table 3.8 lists the individual criteria. Their specific application in this study is embedded in the following chapter 4, which presents the focus and method of this study.

**systems of nominal categorisation**

<table>
<thead>
<tr>
<th>CLSs</th>
<th>CCSSs</th>
</tr>
</thead>
<tbody>
<tr>
<td>optional assignment</td>
<td>obligatory assignment</td>
</tr>
<tr>
<td>flexible assignment</td>
<td>inflexible assignment</td>
</tr>
<tr>
<td>assignment by speaker</td>
<td>assignment by system</td>
</tr>
<tr>
<td>large class set</td>
<td>delimited class set</td>
</tr>
<tr>
<td>marked once</td>
<td>multiple marking</td>
</tr>
<tr>
<td>NP-internal occurrence</td>
<td>NP-internal and -external occurrence</td>
</tr>
<tr>
<td>semantically potent</td>
<td>semantically impotent</td>
</tr>
</tbody>
</table>

Figure 3.8: The basic properties of systems of nominal classification.
After having introduced a theoretical background on nominal classification and grammaticalisation, this chapter turns to the specific focus and method of this study. Section 4.1 introduces the research questions, hypotheses, and predictions that are addressed. Section 4.2 tackles the original problem of the grammaticalisation hypothesis: How can a diachronic hypothesis be studied in the absence of diachronic data? Section 4.3 presents the sample method and the sample languages. Section 4.4 elaborates on the data this study builds upon and presents the structure of the individual language dossiers, which are attached to this study as an appendix (cf. page 355).

4.1 Research Questions, Hypotheses, and Predictions

After having introduced a general theoretical background and a preliminary definition of nominal classification systems and grammaticalisation, it is in order to address the specific orientation of this study. Classifier systems and concordial class systems are assumed to stand in relation with each other; yet, they are neither the same phenomenon nor are they fully different from each other. The central question of this study thus is:

Q₁ To what degree are classifier systems and concordial class systems inherently similar or different?

Since it is commonly accepted that some inherent similarity is the case, we
may preliminarily assume this question to be answered in favour of the similarity and continue asking in what manner and to what extent classifier and concordial class systems are related. Since both systemic types are interpreted as devices of nominal classification, it is rather obvious that a mutual similarity should be found in their classificatory function. Classification can be exerted in two ways: Productively, i.e. in a spontaneous and by the speaker, or the classification is conventionalised, i.e. the speaker has no influence on class assignment, since a given noun’s class membership has become one of its inherent linguistic properties. Ad hoc classification is typically associated with classifier systems, whereas conventionalised classification is typically associated with concordial class systems. Since classifier and concordial class systems are assumed to be inherently similar, we can further ask:

Q2 Can we observe spontaneous classification and conventionalised classification in both systemic types?

Classifier systems are claimed to operate by spontaneous assignment, concordial class systems are claimed to employ conventionalised assignment. If we observe both classificational manners in both systemic types, we can further conclude that the assumption of a close inherent similarity is supported with respect to the system’s function. Note that productive classification includes not only a spontaneous class assignment, but also a certain degree of speaker-sided choice which class a noun is assigned to.

Another aspect that is closely related to the similarity of systems of nominal classification is their diachrony. It is commonly assumed that the source of classifier systems typically are lexical nouns and that classifier systems are less grammaticalised than concordial class systems (cf. again figure 3.3 on page 69 and Grinevald 2000: 55–62). It is thus in order to further ask:

Q3 Where do concordial class systems come from?

The research community has not yet come to a conclusive answer about the function of concordial class systems. It is commonly accepted that virtually all concordial class systems seem to be involved in reference tracking (Corbett 1991: 322). Yet there is a broad range of functions that can be observed to play a role in concordial class systems (attribution of properties, expansion of the lexicon’s referential power, reference management in discourse, etc.; cf. Corbett 1991: 322 and Contini-Morava and Kilarski 2013: 268-290) and the discussion on the functionality of concordial class systems is far from reaching an end. Throughout its research history, nominal classification has been accused timed and again to represent “a useless fossile” (Bonfante 1946: 849, cited in Kilarski 2013: 329) and “not relevant to a rational grammar” (Leibniz, cited in Kilarski 2013: 107). This legitimates the following question:

Q4 Why do concordial class systems exist, if they are functionwise seemingly vain?
Focus and Method of this Study

Two specific hypotheses address this question. These are elaborated in the following together with other hypotheses that derive from the above questions. Their relation to one or more of the beforementioned research questions is indicated at the respective hypothesis.

H1.1 Classifier and concordial class systems constitute a functional continuum. (ad Q1, Q3)

Classifier and concordial class systems constitute a functional continuum. Since they display formal differences (e.g. their respective set size and their morphologically independent, clitic, or affixal markers), it is most likely that they constitute different types of systems of nominal classification despite their connection through the continuum. Furthermore, the ongoing discussion on their rich and diverse functionality suggest that we are dealing with at least two different systemic types. These types of nominal classification devices are connected only by a certain overlap:

H1.2 Classifier and concordial class systems are different systemic types, which display formal and/or functional differences. (ad Q1, Q2, Q3, Q4)

H1.2 emphasises that we investigate two distinct, yet linked phenomena. This clarification is necessary, since a more recent approach to nominal classification, canonical typology (cf. Corbett and Fedden 2015), takes a slightly different view on the continuum of nominal classification. The difference lies in the conceptualisation of the composition of the continuum, which is determined by only one canonical systemic type instead of two prototypical poles of the continuum. In the canonical view, every system of nominal classification thus represents a more or less close instance of this canonical type, which is determined by a high degree of formalisation (marking by agreement, large number of classificatory hosts, system-sided assignment) and semantically transparent. In this case, we are not dealing with a continuum that is defined by two (functionally and formally determined) poles, but by one (formally determined) center. This has a direct effect on the concept of the earlier stated “inherent similarity” of the systems: In the “canonical approach”, deviations from the canonical instance lie in the properties themselves (e.g., if a concordial class system has a restricted range of only one or two classificatory hosts). In the “prototype approach”, deviations from the prototypes are based on the composition of properties, i.e. intermediacy is determined by properties that are themselves prototypical for either classifier or concordial class systems, but in combination display a mix of both prototypes (e.g. by employing a classifier-like large inventory of a few dozen class markers and concordial class-like agreement).

This study investigates nominal classification in general and draws its conclusions on basis of observations from the language sample. Therefore, it preliminarily follows the “prototype approach”. According to the “prototype approach”, the continuum consists of two ideal systems, classifier and concordial class systems. These systems have a number of characteristic properties. If a
nominal classifier system develops, it gradually adopts characteristic properties of concordial class systems; systems that undergo this process, but did not (yet) finish it, form an intermediate ‘grey area’. This grey area constitutes the link between the ideal systems. The grammaticalisation hypothesis is the third hypothesis that is considered for this study and refers to this process:

\[ H_2 \] Classifier systems may grammaticalise into concordial class systems.

(“grammaticalisation hypothesis”, “CLS$\rightarrow$CCS”; ad Q1, Q3, Q4)

\[ H_2 \] poses a clear answer to Q3 and Q4: Concordial class systems may represent the output of a grammaticalisation process of other (formally more independent and semantically more motivated) linguistic systems, in this specific case classifier systems (cf. Aikhenvald 2000: 91 and 372ff., Greenberg 1978a: 78ff., Sands 1995: 249ff., Seifart 2010: 722 and Craig 1986b: 281). The theoretical reflections underlying the grammaticalisation hypothesis make this scenario fairly likely; chapter 3 presents the argumentation and theory that advocate this course of systemic shift in detail. In a nutshell, the prototypical classifier system, which is highly semantically motivated and interactive and operates by means of independent constituents, loses its semantic motivation, interactivity, and independence of its markers by the mechanisms that constitute grammaticalisation (semantic bleaching, phonological erosion, cliticisation and morphologisation, decategorisation) during its shift towards a concordial class system. The prototypical concordial class system, the target for this shift, is characterised by a high degree of system-sided mechanism and automatism and does not involve much interactive opportunity for the speaker. During such a shift, the system’s operative domain is shifting from semantics to morphosyntax.

The main problem of this diachronic claim is that there is no documented case available which could prove the assumed shift from a classifier system to a concordial class system. For this reason, this study aims at eliciting evidence for or against it from a broad typological investigation. Except for the underlying theory of the grammaticalisation hypothesis, which is sound in itself, a number of supposed intermediate systems (e.g. the ones of Ngan’gityemerri or Bora) argue for the grammaticalisation hypothesis. These systems seemingly represent the ultimate intermediate case between classifier and concordial classes in displaying core properties from both systemic types. Since we lack both data from older stages of these languages and proof that they reached the final stage of a concordial class system, these cases suggest the probability of the grammaticalisation hypothesis, but do not prove it. The same applies to reconstructional attempts, which are available for e.g. Proto-Indo-European and Australian languages: While the likeness of the reconstructed classifier systems in these languages will not be questioned here, it may be noted that these reconstructional attempts, too, are not a proof for the grammaticalisation hypothesis, but possess only an illustrative force due to their reconstructional character.
These questions and hypotheses, the preliminary definitions of both classifier and concordial class systems, and several theoretical arguments suggest a number of predictions. These are sought to be verified by the sample data and consequently be used to answer the abovementioned hypotheses and questions. Classifier systems systemically differ from concordial class systems. Given that classifier and concordial class systems constitute a functional continuum, we can predict that:

$P_1$ The different systemic (sub-)types of nominal classification cover different areas on the functional continuum.

Earlier studies have shown that the nominal lexicon is by far the most common source for classifiers; both classifier and concordial class systems typically recruit their markers from there (cf. Corbett 1991: 312 and Aikhenvald 2000: 353). Considering $H_2$, this predicts in a potential development of nominal classification devices in three steps:

$P_2$ The development of nominal classification devices consists in (at least) three steps: Lexical nouns, which grammaticalise into classifiers, which may further grammaticalise into concordial class markers.

Note that this study focuses on the latter grammaticalisation process, where classifiers turn into concordial class markers. If accurate, $P_2$ should be reflected in the typological sample study. As for the next prediction, consider that grammaticalisation is a gradual process. We thus can predict that:

$P_3$ There are intermediate systemic types which display properties of both classifier and concordial class systems.

These intermediate systemic types constitute a loose group that is located somewhere in the overlapping area between classifier and concordial class systems. Since there are so far no closer definitions concerning this overlapping area in the functional continuum of nominal classification to my knowledge, we can predict that:

$P_4$ These intermediate types form a coherent and so far not further specified grey area in the functional continuum.

The following predictions focus on the process assumed by the grammaticalisation hypothesis and specify individual phenomena, which are expected to occur, if the grammaticalisation hypothesis is accurate:

$P_{5.1}$ By the course of $CLS > CCS$, the classificatory elements (partially or fully) lose their potential of semantic contribution to the noun phrase.

---

1 It is e.g. not yet defined up to what point in the continuum a classifier system can be called a classifier system and from which point it may (or must) be called a concordial class system.
A core property of classifier systems is their semantic contribution to their noun phrase. This is not a required property in concordial class systems. Since grammaticalisation promotes semantic generalisation and erosion, we can predict a decline in the semanticity of classifier systems in the case of their grammaticalisation.

P₅.₂ By the course of CLS→CCS, the set of classificatory elements decreases in size.

Grammaticalisation reduces the productivity of a linguistic construction. This is displayed in the reduction of (potentially) employed tokens and corresponds to the largish sets of classificatory elements which are employed in classifier systems and the relatively small class sets of concordial class systems.

P₅.₃ By the course of CLS→CCS, the classificatory elements cliticise and later morphologise to agreement markers.

Grammaticalisation promotes formal dependency of the grammaticalising linguistic unit. Classificatory elements are thus expected to occur in the form of independent or clitic elements in the case of classifier systems and in the form of clitics or affixes in the case of concordial class systems.

P₅.₄ By the course of CLS→CCS, a classificatory element’s loss in semantic contributive force precedes its functional reanalysis as an agreement marker.

A given system’s semanticity is expected to decrease by the course of grammaticalisation and we expect the classificatory elements to turn into agreement affixes that pattern the nominal lexicon. If a system of nominal classification thus employs agreement markers, these markers are expected to lack semantic specificity and instead copy information of their head noun. The following section introduces the methodological background, which serves the addressation of the abovementioned questions, hypotheses, and predictions.

### 4.2 Diachronic Problems, Synchronic Solutions

If a system of nominal classification develops, it changes in its appearance and its function. According to the hypothesis about grammaticalisation being the vehicle for the development of nominal classification systems, we are able to observe systematic changes in their morphosyntactic and pragmatic behaviour, since grammaticalisation captures and describes changes within these domains. The sparse record situation does not allow an extensive and interlinguistic diachronic investigation of multiple independent systems of nominal classification: For instance, the Indo-European languages and Chinese are documented for a period of more than 3000 years and therefore allow a diachronic investigation;
other languages (e.g. the Semitic and Dravidian ones) are also well-documented over a long time span. Yet most other languages or language families know written records only over a period of at best several hundred years. However, the records of the Indo-European, Semitic, and Dravidian daughters and Chinese do not display an intersystemic change (i.e. from the stage of a classifier system to the one of a concordial class system). In order to be able to draw valid conclusions about the relation of grammaticalisation and the development of systems of nominal classification, a typological method is employed, which does not rely on diachronic data. Subsection 4.2.1 discusses the crucial problem of the grammaticalisation hypothesis, namely the lack of an affirmative documentation, and presents a few reconstructional attempts that involve the grammaticalisation hypothesis. Subsection 4.2.2 discusses the typological method that is employed in order to compensate for the lack of diachronic documentation.

4.2.1 Lack of Diachronic Data & Reconstructional Attempts

The grammaticalisation hypothesis makes a strong claim about the diachrony of systems of nominal classification; a diachronic study thus seems to be the most obvious strategy to verify the grammaticalisation hypothesis. The available records do not document the full course of events that is described by the grammaticalisation hypothesis; also, the most important stage of the assumed process, which is the switch from a classifier system to a concordial class system, is not documented as such. To my knowledge, there are two descriptions of systems of nominal classification that are not included in the sample, but might come into question for a current suchlike shift: Claudi 1985 and Dixon 1970.

In his short description of Olgolo (Pama-Nyungan; ISO 639-3: olk; cf. Lewis et al. 2013: s.v.) and its syllable structure, Dixon (1970) notes a phenomenon that he identifies as a potentially incipiating concordial class system. In contrast to most Australian languages, where a word must begin with a consonant and contain at least two syllables, Olgolo words always begin and mostly end in a vowel (ib.: 273). In order to stabilise the phonological instability (which is mirrored by the massive dropping of consonants), Olgolo allows its speakers to prefix an n-, w-, or y-. These consonants may combine with any noun on semantic grounds: n- goes together with animals and insects, w- with “certain inanimate things such as trees, grasses, sun, fire, language, etc.” (ib.: 275), and y- with fish, oyster, and eel. This phenomenon thus displays a nature that is close to a classifier system and it is conceivable that by grammaticalisation of these prefixes, a classifier or even concordial class system might arise. A further potential marker for a vegetable food class is y-, but Dixon’s data is not sufficient to support this idea (ib.: 276).

Nevertheless, Dixon (ib.: 276) emphasises “that there is not yet a system of

\[\text{I would like to thank Jenny Audring and Greville Corbett for making me aware of these studies.}\]
noun classification in Wolgolo \[sic\], since it is a matter of the speaker’s will to attach the prefixes or omit them. Since “the language is at present [1970; MP] spoken by a handful of old people, and is bound to become extinct in the next decade or so” (ib.), Olgolo does not represent a case in support of the grammaticalisation hypothesis.

Zande, a Niger-Congo language of the Ubangi phylum (ISO 639-3: zne; cf. Lewis et al. 2013: s.v.), employs a concordial class system with four categories (masculine, feminine, animal, and neuter; cf. Claudi 1985: 87), which are marked in the third person of the personal pronoun (where the masculine and feminine class conflate in the plural; ib.: 89). Further concordial elements are rather combinations of linguistic items with the pronominal paradigm (e.g. the demonstratives, which are rather demonstrative adjectives plus the personal pronoun; attributive adjectives in general do not agree in Zande; ib.: 98 and 106).

The masculine and the feminine class are semantically bijective (\textit{eineindeutig}): All terms for adult male human beings are assigned to the masculine class, and all terms for female adult human beings are assigned to the feminine class. The animal and neuter class are semantically injective (\textit{eindeutig}): All terms for animals are members of the animal class (there are some 80 members denoting inanimates), but not all members of the animal class denote animals. The same holds for the neuter class (ib.: 115ff. and 120). In general, the system shows typical semantic extensions of the original class semantics (e.g. celestial objects that are assigned to the animal class due to mythological relations (cf. Claudi 1985: 121), which suggests this class to be better labeled “non-human animate”). Most of the markers can be etymologised and have a common noun as their source (e.g. the pronominal neuter class marker \textit{he}/-hà relates to the noun \textit{he} “item, object” (Claudi 1985: 128; for all etymologies, see also ib.: 127-137). Since the system is most probably an internal development and not a loan (Claudi 1985: 139), the basis of the Zande system of nominal classification is not a classifier system, “but a classification of the environment that is inherent to the language as well as to the mind” (Claudi 1985: 140; own translation).

There is a small number of ‘hybrid’ systems (Nasioi, Bora, Mundurukú) which were included in this study and seem to employ a system at the most intermediate stage possible, i.e. that their large number of classificatory markers employ classifier-like semantics, while they formally behave like agreement markers that are found in concordial class systems. However, these languages were not documented at their earlier stages, where they might have employed a classifier system proper, and thus cannot verify the grammaticalisation hypothesis. Nevertheless, they constitute the most promising potential evidence for the grammaticalisation hypothesis in being the most likely candidate for the ‘missing link’ and thus are investigated more in detail in section 5.4.

Another way of going into the diachrony of languages is the reconstruction of this language’s proto-stages, which is typically carried out by a comparison of a language and its closest relatives and a careful evaluation of the most likely change phenomena that the language and its siblings have undergone. As
this description suggests, a reconstruction also cannot count as ‘hard proof’ for
the likeliness of the otherwise undocumented grammaticalisation hypothesis. Instead, reconstructions have an illustrative force, which tracks, describes, and depicts the assumed processes.

There are a few reconstructional attempts for the history of nominal classification systems, of which a selection is discussed in the following in order to illustrate the potential course of the grammaticalisation hypothesis with the help of concrete examples. Sands (1995) provides a comprehensive treatment of nominal classification on the Australian continent, including a broad comparative reconstructational study of how the Proto-Australian system of nominal classification might have looked like (ib.: 285–317). This tentative reconstruction is proposed as “a good starting point for further research, despite it is likely proof from the strict application of the comparative method will never be possible” (Sands 1995: 288). Sands’ (1995: 317) reconstruction delivers a contrastation of masculine and feminine animates in the pronominal paradigm and “a small set of generics that could be used in classifierlike constructions”. Only two of these generics could be reconstructed with some certainty (the words for “vegetable food” and “animal”); a third potential generic in this function is the word for “tree”. Depending on the respective system of nominal classification, these generics grammaticalised into either a classifier or a concordial class marker denoting the same semantic value for its referential class (page 77), as table 4.1 illustrates.

Table 4.1: The development of the vegetable food class in Australian (cf. Sands 1995: 310).

<table>
<thead>
<tr>
<th>stage</th>
<th>marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Proto-Australian generic-specific noun pairings *mαγι “vegetable food”</td>
</tr>
<tr>
<td></td>
<td>grammaticalise into classifiers:</td>
</tr>
<tr>
<td>II</td>
<td>Ngan’gityemerri mγι “vegetable food”</td>
</tr>
<tr>
<td></td>
<td>Yankunytjavara mγι “vegetable food”</td>
</tr>
<tr>
<td>or concordial class markers:</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Mayali mαn- “vegetable food”</td>
</tr>
<tr>
<td></td>
<td>Dyirbal -m “vegetable food”</td>
</tr>
</tbody>
</table>

3A number of authors of reconstructions for earlier stages point to the hypothetical character of their reconstructs themselves (cf. e.g. Balles 2004: 55, Sands 1995: 267f.).
While the value as a starting point for the study of Australian nominal classification is out of question, the full reconstruct as such is not quite convincing for several reasons. The study considers about 250 Australian languages, which includes genetically unrelated phyla (cf. Dixon 2002: 659) and phyla that are diachronically distanced towards each other (cf. Sands 1995: 288). This not only increases the likelihood of e.g. borrowed forms or forms of a different age being considered equal bases for reconstruction, but also makes the reconstruction invalid a priori, since two languages which are provedly not genetically related to each other cannot be assumed to share the same earlier stage. Beyond that, the actual proto-stage employs a mere two or three generic nouns, which is in stark contrast to the expectations of the grammaticalisation hypothesis. Recall that the initial inventory of classificatory elements is expected to be rather large and, by the course of grammaticalisation, to decrease in size, semantic specificity, and formal independence. Despite the more frequent loss of formal independence in a number of languages, we cannot observe an overall loss in the specificity of meaning e.g. in *mayi “vegetable food”. The reconstructed Proto-Australian class inventory as a whole also is unexpectedly small with only four or five categories that hardly constitute a comprehensive means to classify the real world.

Most reconstructional attempts for the predecessor of the Proto-Indo-European concordial class system aim at a semantically more transparent and potent form of a concordial class system, which is typically compared to the noun class systems found in the Bantu, Niger-Congo or other African languages (e.g. Balles 2004: 49, Erhart 1993: 39, Froschauer 2003: 39). For instance, Froschauer (2003, based on Leiss 1997 and Lehmann 1958) proposes a Proto-Indo-European threefold derivational system that allows each nominal root a ‘nominal aspectualisation’ by taking one out of three stem-forming affixes (that later grammaticalised into the commonly accepted Indo-European gender markers for the masculine, feminine, and neuter gender). The three affixes used in the Proto-Indo-European stem formation are *-s “singulativity/countability”, *-H “collectivity”, and *-m “continuativity”, whereas an affixless stem expressed the general semantics of the noun root (Froschauer 2003: 32f.). A supposed reflection of this system can be found in a threefold Vedic Sanskrit noun (cf. Lehmann 1958, cited in Leiss 1997: 35): hima-h > hima “winter” contrasts with himá-s “cold, frost” and himam “snow”. Despite not aiming at a typologically more diverse classifier system, but at a typologically relatively close noun class system, reconstructions like the presented one are of a highly hypothetical character (cf. Balles 2004: 55) and lack at times sound theoretical and methodological requirements for a valid comparative reconstruction (cf. e.g. Passer 2014: 319ff. on Erhart 1993; Irslinger 2009: 16f. and 19f., Passer 2014: 323ff., and Schwink 2004: 66ff. on Leiss 1997 and Froschauer 2003).\footnote{For instance, the Vedic examples given are falsified by Irslinger (2009: 16f.), who also convincingly argues that virtually no other ‘unproblematic’ threefold examples from an Indo-European language can be provided. Among a variety of other problematic prerequisites of Froschauer’s (2003) and Leiss’ (1997) approach, they severely violate established the-}
There are also diachronic conjectures on earlier stages of African systems of nominal classification: For instance, Claudi (1985: 139f.) rejects the idea that the Zande (Ubangi; Niger-Congo) gender system has grammaticalised from an earlier and richer classification system and suggests its language-internal genesis on the cognitive grounds of “a classification of the environment which is inherent to both language and thought” (ib.: 140). Pasch (1986: 372) on the other hand studies the development of the systems of nominal classification in the Ubangi languages (Niger-Congo) and comes to the rather paradoxical conclusion that the Proto-Mba (Ubangi) system originated from a two-class distinction, expanded up to “at least ten singular and five plural classes” employing agreement, which again were given up partly or as a whole in all languages besides Dongo-Ko.

The outcomes of the presented reconstructional attempts conform with the grammaticalisation hypothesis, but cannot be accounted for as evidence for it because of their tentative character. The following subsection introduces a typological model that is used to elicit evidence for the grammaticalisation hypothesis.

4.2.2 Typology Supplementing Diachrony

Established theory argues for the likelihood of the grammaticalisation hypothesis, and both tentative reconstructions and synchronic systems that seem to mirror different stages along the lines of grammaticalisation can illustrate its course of process. There is no evidence for the grammaticalisation hypothesis within a documented system, nor are there broader (diachronic or typological) comparative studies of the suggested developmental shifts, which could be related to change in systems of nominal classification. Given the lack of diachronic documentation, a broad typological study thus is an obvious opportunity to gain insights into these shifts: “[T]he typological description of languages has two sides, namely, synchronic and diachronic” (Skalička 1968: 444, cited in Greenberg 1978b: 66) Both diachronic and synchronic linguistic research methods consist of “a constant interaction between the explanation of specific cases and deductive hypotheses” (Greenberg 1978b: 63). The investigation is in accordance with the premise that not ‘language types’, but types of linguistic change (“processes”) should be subject to comparison (Greenberg 1978b: 65). The grammaticalisation hypothesis renders the process of a classifier system developing into a concordial class system and does not render a language type. Therefore, it can be tested by synchronic instances of systems of nominal classification, which should be historically unrelated (cf. Greenberg
1978b: 65). Since each system of nominal classification supposedly represents a certain spot on the continuum of grammaticalisation of nominal classification, a language sample can be compiled. This enables an examination of whether these systems’ properties, to be exact those properties that are expected to change with grammaticalisation (e.g. the rise in obligatoriness and decline in flexibility of the class assignment, the decreasing semantic contribution to the noun phrase, and the shrinking size of the class inventory), correspond to the expectations derived from the grammaticalisation hypothesis. These properties are listed and explained in the previous chapter in detail (cf. page 77 and figure 3.8 = 4.1); their expected correspondence to the grammaticalisation hypothesis is elaborated in the following.

If the hypothesis is correct, systems of the same type and developmental stage feature a comparable state, i.e. the same behaviour and properties. For this reason, a typological taxonomy of nominal classification devices was created (cf. figure 4.2): Each terminal node of the taxonomy is expected to cover a limited range of the continuum of grammaticalisation of nominal classification. In order to find out of which size this range is, a number of systems of nominal classification that corresponds to the type of the node (e.g. all numeral classifier 5

At this point, a more general question that is related to the methodology of this study rather than the study itself may be addressed briefly. The scenario suggested by the grammaticalisation hypothesis appears to be a fairly rare phenomenon, since it is not (fully) documented in a single language (including the documentation of numerous languages that abandoned their classifier systems). If the grammaticalisation hypothesis indeed is accurate, though very rare, to what degree can we draw reliable and valid conclusions about an infrequent phenomenon from a language sample that mainly consists of presumably average systems? Logic suggests that the data from average systems will not elicit much supportive data in favour of the grammaticalisation hypothesis. Greenberg (1987b: 75f. and 85) comments on the probability of the occurrence of certain phenomena and suggests that certain linguistic phenomena are common and unmarked, since these are coined by frequent processes and are of an inherent stability, whereas other phenomena are uncommon, marked, or even non-existent, since the processes that lead to their genesis are infrequent and/or the phenomena themselves are inherently instable. Beyond that, marked phenomena are always expected to occur parallel to their unmarked variant. Speaking of a whole linguistic category, these unmarked parallels have to be expected intralinguistically rather than interlinguistically, of course.

Let us project this conjecture onto the grammaticalisation hypothesis: Both classifier systems and concordial class systems appear to be common and unmarked statuses of nominal classification systems, given their frequency and stability even over millennia (cf. e.g. Mandarin). The process assumed by the grammaticalisation hypothesis on the other hand is obviously not frequent, since it was not yet observed or even documented. A potential explanation for the low frequency may be found in the precondition for the process to be carried out: In order for a classifier system to ‘switch’ into a concordial class system, we expect the (earlier fairly semantic) system to semantically bleach to a degree where most semantics are lost. However, many classifier systems (e.g. Mandarin; cf. Aikhenvald 2000: 410) display a grammaticalisation of individual classifiers rather than the whole classifier set. This means that individual classifiers bleach semantically and eventually until a degree where they drop out of the system. Similarly, new classifiers may enter the productive system and keep it transparent. It is thus conceivable that most classifier systems are continuously remotivated rather than pushed into the functional direction of a concordial class system. If this is indeed the case, the lack of evidence is explained by the scarcity of the source material for the assumed shifting process.
Focus and Method of this Study

systems of nominal categorisation

CLSs | CCSs
---|---
optional assignment | obligatory assignment
flexible assignment | inflexible assignment
assignment by speaker | assignment by system
large class set | delimited class set
marked once | multiple marking
NP-internal occurrence | NP-internal and -external occurrence
semantically potent | semantically impotent

Figure 4.1: The basic properties of systems of nominal classification.

systems of the language sample) is examined in chapter 5 in terms of their behaviour and properties. Concordial class systems are expected to exhibit a higher degree of grammaticalisation than classifier systems.

The origins of classifier systems are in most cases quite transparent and well-researched and are not subject to the specific claim of the grammaticalisation hypothesis. This study thus focuses on eliciting system-internal evidence for or against the assumed shift. By comparing the differently grammaticalised states of systems of nominal classification, we can draw conclusions on the processes that occur between those states (“state-process model of language”; cf. Greenberg 1978b: 67). Especially the ‘hybrid’ systems are promising objects for study: They cause difficulties in their assignment to an established typology and thus can be assumed to be either instances which are currently undergoing the assumed process or they may serve as a basis for a further subtypology (Greenberg 1978b: 73f.). Though there is a number of descriptive works on the functional variety within systems of nominal classification, this perspective has in general been understudied. A broader and typological study of function (and not form) as their “most obvious connection” (Senft 2007: 686) yet remains to be carried out.

Figure 4.2: Taxonomy of the systems of nominal classification.
In the following, I present the composition of the fused bundle of functional, formal, and mixed or hybrid properties by Dixon (1982) and Contini-Morava and Kilarski (2013), which are considered in the creation of the dossiers on the nominal classification systems of the sample languages. Table 4.2 contains the properties that serve to distinguish between nominal classification devices and non-nominal classification devices, as is elaborated on the following pages. Table 4.3 contains the properties that serve to distinguish whether nominal classification device is more classifier-like or more concordial class-like; these will also be elaborated in the following.

Table 4.2: Functional and hybrid properties of nominal classification devices (based on Contini-Morava and Kilarski 2013).

<table>
<thead>
<tr>
<th>functional properties</th>
<th>referent differentiation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>individuation</td>
</tr>
<tr>
<td></td>
<td>property attribution</td>
</tr>
<tr>
<td>hybrid properties</td>
<td>reference identification</td>
</tr>
</tbody>
</table>

Table 4.3: Formal and hybrid properties of nominal classification devices (based on Dixon 1982, 1986).

<table>
<thead>
<tr>
<th>formal properties</th>
<th>obligatoriness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>class membership</td>
</tr>
<tr>
<td></td>
<td>class number</td>
</tr>
<tr>
<td></td>
<td>locus operandi</td>
</tr>
<tr>
<td></td>
<td>agreement</td>
</tr>
<tr>
<td>hybrid properties</td>
<td>assignment</td>
</tr>
</tbody>
</table>

All sample languages are analysed with respect to their functional, formal, and hybrid properties. The individual setup of systemic properties is analysed and compared on the scale of the whole sample in the chapters 5 and 6 with respect to commonalities and differences within and beyond given systemic types. Since the established typologies are mostly based on the systems’ morphosyntactic properties, we expect the formal behaviour of individual systemic types (e.g. numeral classifier systems, noun classifier systems, etc.) to be rather homogeneous and to correspond to the predictions of the grammaticalisation cline that was rendered in chapter 3. The comparison of the functional side of the sample languages is expected to be more varied even within one and the same systemic type, since the functional aspects of nominal classification have so far not enjoyed much research; also, Contini-Morava and Kilarski (2013) show that all sorts of nominal classification systems may serve all sorts and combinations of functions. Considering this general functional variability of nominal classification systems, it is most likely that a system’s form and function are at
least not directly depending on each other; a system’s formal properties may e.g., be mainly determined by a given language’s type and characteristics of its genetic phyllum. The same system’s functional properties on the other hand may be rather depending on cultural and cognitive characteristics that are not strictly genetically determined: Recall e.g., the Australian languages, which – independent of their genetic relation or systemic classification type – quite consistently employ a referential or concordial class for non-flesh food. I presuppose the mutual independence of form and function, which is also argued for (not in relation to nominal classification, but to grammaticalisation in general) by van Rijn (2016) and Hengeveld (2012).

**Functional properties:** Note that the Dixon-based property semantic potency is not included in this list. Semantic potency is a rather vague property which only states if the classificatory marker adds meaning to its noun phrase, not the kind of semantic contribution. This criterion is subcategorised in Contini-Morava and Kilarski (2013) into a number of subcriteria rendering the type of semantic contribution. This includes the class markers’ use as expanders of the nominal lexicon (by their productive participation in derivational noun formation) and variable classification for reference to different entities (“referent differentiation”), which constitute the joint criterion “referent differentiation” in this study. Classificatory markers can also serve as ‘unitisers’ (cf. Lucy 2000: 334); their related function is individuation. Class markers can also be used to express attitudes of the speaker by assigning the noun denoting the respective referent to a given referential class, e.g., to express one’s own disgust towards the referent; this is accounted for by the criterion property attribution.

In this study, the functional properties proposed by Contini-Morava and Kilarski (2013) serve as a filter for the determination of whether a given system is a system of nominal classification: If a candidate system does not employ a single semantic or discourse function, it is not regarded a system of nominal classification (cf. figure 4.3).

**Formal properties:** The formal properties listed in table 4.3 render a system of nominal categorisation with respect to its application to the nominal lexicon as a whole (obligatoriness), the persistence of its assignment (class membership), and the size of its class inventory (class number). Furthermore, the formal properties account for the respective system’s way of class marking, i.e., whether the markers are bound to the borders of the noun phrase or may exceed them and whether they occur multiply or once. Note that the origin of nominal classification markers is not included in this list as a property. The consulted sources in many cases did not provide sufficient data for a justified conclusion on the etymological origin of the classificatory markers, which is why this property cannot be considered reliable for the sample as a whole.

After having determined whether a system can be considered a system of

---

6Consider e.g., the Indo-European languages, which have a strong tendency to employ a gender system, or South East Asian languages, which are more likely to employ a classifier system.
nominal classification (with the help of the functional criteria proposed by Contini-Morava and Kilarski 2013), the formal properties serve as a tool to determine whether a system is formally rather a classifier or a concordial class system (cf. figure 4.3).

Figure 4.3: Formal and functional properties of nominal classification and their use in the sample study.

Hybrid properties: Finally, there are two properties that do not clearly serve either a formal or a functional purpose. These are the manner of the assignment between noun and referential class (speaker-sided vs. system-sided and mechanic vs. interactive) and reference tracking in discourse (also “referent identification”; e.g. by agreement marking on a number of nominal satellites or by the anaphoric usage of classifiers). Properties like a system’s assignment are crucial for its functionality in that it reveals the manipulability of the system with respect to the speaker, which is why they are included in the study. Their behaviour is observed and compared, which promises not only a better understanding and more in-depth description of these systems.

The outcomes of the synchronic sample study are tested against the diachronic background of selected and well-documented sample languages and language families in order to confirm their reliability and validity. Section 4.3 describes the sampling method and composition of this study.

4.3 Sampling Method and Composition

This section is dedicated to an overview and a genetic and typological description of the sample and the sampling method. Beyond this, the functions that are employed by the sample systems are evaluated. Careful reflection concerning the kind and size of the language sample used for one’s research is imperative for reliable and valid research findings. If one brings to mind that a mere third
of the approximately 7000 extant languages is described and that the document-
ation situation is even worse if one takes extinct languages into account, it becomes obvious that any language sample consisting of a part of these few recorded ones can hardly ever be called representative (cf. Bakker 2010: 1ff.): “The maximum we might get out of a sample is a fair idea about what is possible in the languages of the world, though not a very reliable idea about what is not possible” (Bakker 2010: 3). In order to make this ‘fair idea’ as sound as possible, several factors have to be considered and the most suitable sample method has to be chosen. These factors and sample types are described in the following.

The language sample is designed by the criteria for composing variety samples described in Rijkhoff et al. (1993) and its succeeding elaborations Rijkhoff and Bakker (1998) and Bakker (2010). The core strategy of this approach to language sampling is finding the greatest possible diversity, which includes the rarest typological instances of the investigated phenomenon and maximum genetic diversity (Rijkhoff et al. 1993: 171ff.). The sampling is primarily based on genetic diversity, since most of the other possible kinds of bias are related to a language’s genome (ib.: 172). It can be achieved by including one language from each known phylum of a classification of language, all linguistic isolates and a pidgin or creole language. According to the classification of Ruhlen (1987), this results in 27 sample candidates (17 families, nine isolates and one pidgin or creole). If a larger sample is desired (such as in the case of this study), the family-internal linguistic diversity must be warranted by a consideration of the phenomenon under investigation. A calculation of the family-internal diversity is established by an analysis of the tree structure of the respective language family. While the mere number of levels (generations), nodes or daughters within the family tree does not tell us a lot about the genetic diversity within a language family, a more sophisticated account for the internal structure of a language tree can give us information about the internal diversity of this language family. The computation of the “diversity value” (cf. Rijkhoff et al. 1993: 179ff. and 187ff.) enables the researcher to compare the internal diversity of different language families and therefore serves the maximum diversity of the language sample as a whole. The diversity value accounts for the structural depth and the width of a language tree. After having chosen a language from each phylum, the diversity value determines how many languages from each phylum have to be added. If a phylum contributes with more than one language to the sample, a consideration of both the availability of reliable data on the phenomenon

7In this study, a language phylum is understood as “the largest spectrum of languages for which a genetic relationship can be demonstrated. e.g. the Indo-European languages” (Bußmann 2006: 643). Language isolates are interpreted as individual phyla, or singleton phyla, as it is suggested in Rijkhoff et al. (1993: 173 and 197; cf. also Bakker 2006: 268), whereas all pidgin and creole languages are interpreted as members of one phylum.

8Note that “[s]ome of Ruhlen’s genetic (sub)groupings, however, are not uncontroversial and in some cases there is even serious disagreement among the experts” (Rijkhoff et al. 1993: 173). Nevertheless, Ruhlen’s classification allows the composition of smaller variety samples, which this study strives for.
under investigation in the respective languages and of the family-internal diversity value decide which languages are added to the sample (cf. Rijkhoff et al. 1993: 197).

Some languages drop out of the sample (cf. Rijkhoff et al. 1993: 191): There may be no adequate description available (as for Etruscan and Meroitic, two isolates that otherwise had to be included), or the phenomenon under investigation is not present (for instance in Basque, a number of supposed ‘relational classifier’ languages, and supposed ‘verb classifier’ languages). If this is the case, the resulting slots in the sample must not be filled by members of other phyla in order to preserve the proportions of diversity. The ideal sample size is therefore in most cases reduced to an actual sample size. The account for the fourty language sample of this study results in an actual sample size of 32 languages.

In the following, I explain and itemise the composition of the language sample. According to Rijkhoff et al. (1993: 186), a language sample of fourty languages must contain as many languages from each phylum as shown in table 4.4, where grey colour marks those phyla that drop out of the sample due to a lack of data or the lack of a nominal classification system.

Table 4.4: Number of languages from the basic phyla in a fourty-language sample.

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Language Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afro-Asiatic</td>
<td>2</td>
</tr>
<tr>
<td>Altaic</td>
<td>1</td>
</tr>
<tr>
<td>Amerind</td>
<td>5</td>
</tr>
<tr>
<td>Australian</td>
<td>2</td>
</tr>
<tr>
<td>Austric</td>
<td>4</td>
</tr>
<tr>
<td>Caucasian</td>
<td>1</td>
</tr>
<tr>
<td>Chukchi-Kamchatkan</td>
<td>1</td>
</tr>
<tr>
<td>Elamo-Dravidian</td>
<td>1</td>
</tr>
<tr>
<td>Eskimo-Aleut</td>
<td>1</td>
</tr>
<tr>
<td>Indo-European</td>
<td>1</td>
</tr>
<tr>
<td>Indo-Pacific</td>
<td>3</td>
</tr>
<tr>
<td>Khoisan</td>
<td>1</td>
</tr>
<tr>
<td>Na-Dene</td>
<td>1</td>
</tr>
<tr>
<td>Ket</td>
<td>1</td>
</tr>
<tr>
<td>Nahali/Nihali</td>
<td>1</td>
</tr>
<tr>
<td>Hurrian</td>
<td>1</td>
</tr>
<tr>
<td>Pidgins &amp; Creoles</td>
<td>1</td>
</tr>
<tr>
<td>Burushaski</td>
<td>1</td>
</tr>
<tr>
<td>Meroitic</td>
<td>1</td>
</tr>
<tr>
<td>Sino-Tibetan</td>
<td>1</td>
</tr>
<tr>
<td>Uralic-Yukaghir</td>
<td>1</td>
</tr>
</tbody>
</table>

In order to ensure the internal maximum variety of the sample, each listed phylum must be concerned and integrated into the sample. There are several languages and language families that drop out of the sample, either because they are not documented well enough or because they do not employ a system of nominal classification. Two isolates drop out due to their insufficient record situation: Etruscan (which only distinguishes the sex of human adults, while other nouns are not categorised; cf. Bonfante and Bonfante 2002: 82) and Meroitic (cf. Bechhaus-Gerst 1985: 310f.). The sparseness of the data on these languages is confirmed by Rijkhoff et al. (1993: 191).

Several of the isolates do not employ a system of nominal categorisation: These are Hurrian (cf. Diakonoff 1971: 87), Basque (cf. de Rijk 2008: 18), and Gilyak (or Nivkh; cf. Gruzdeva 1998: 16ff.). For the Eskimo-Aleut family, the
Focus and Method of this Study

Case is not as clear as for the other languages: It is generally accepted that it does not employ a concordial class system (cf. Bergsland 1997: 47, Greenberg 2000: 185, Mithun 1999: 95). There is no further indication for employed classifiers in the relevant linguistic literature (cf. e.g. de Reuse 1988: 34ff. and Mithun 1999: 104-117), yet nouns are marked for their kind of possession. I briefly discuss the reasons for the exclusion of the Eskimo-Aleut phylum: In Aleut languages, inalienability of body parts and kinship is marked by a suffix optionally: “In a normal context”, the suffix occurs, in others it may be omitted (Bergsland 1997: 55 and 350f.). In Eskimo, nouns are marked for case and number. If a noun’s referent is possessed, it is further specified for number (singular, dual and plural) and person (first, second, third (cf. A) and third reflexive (cf. B)) of its possessor (cf. de Reuse 1985: 35):


- aghnagh -m naa -(ng)an atkug -(ng)a
  woman REL.CASE mother REL.CASE.3SG parka ABS.CASE.3SG
  “the woman’s mother’s parka”

- meaning- (ng)ite fte- ul- qun-
  meaning lack.NOUN.BASE apparently.VERB.BASE IND.3PL see
  igagh -yagmegtenqnga
  tattoo TRANSITIVE.PARTICIPLE.OBLIQUE.MOOD.3PL.REFL.1SG
  “At the time that they tattooed me, they [tattoos] did not have meaning.”

These elements do not behave like classifiers in the definition of this study due to various reasons: First, the elements in question are of a suffixal nature. Second, the relation between possessor and possessee determines the occurrence of the suffix, while the suffix itself does not express any information about this relation or the kind of possession (as we would expect from a ‘relational’ or possessive classifier). Third, the suffix can be combined with other nominal suffixes (cf. 4.1A) as well as with suffixes of verbal nature (cf. 4.1B), which contradicts the definitional property of classifiers as elements occurring “in certain syntactic environments” (Dixon 1986: 105).

As for the language isolate Nahali (ISO 639-3: nll; cf. Lewis et al. 2013: s.v.), two factors are responsible for it dropping out of our sample: The only specific source on Nahali available to the author is Kuiper 1962, where no indication of any system of nominal classification could be found. This lack of nominal classification is confirmed by Dryer and Haspelmath (2013: s.v., with reference to Kuiper 1962 and Nichols 1992: 297), where the existence of possessive and numeral classification is distinctly excluded for Nahali. Nahali is thus excluded for two reasons, its sparse documentation and the lack of any indication of a system of nominal classification.

The case of pidgins and creoles is a special one if it comes to nominal classification. As Aikhenvald (2000: 389) notes, “[t]he only example of a creole language with numeral classifiers appears to be [Chavacano] Zamboangueño, a Spanish-based creole spoken in the Philippines (Carl Rubino, p.c.)”. According
to Anthony Grant (p.c.), Chavacano Zamboangueño does not employ classifiers, “though proper nouns are marked with different articles from other nouns”. Since proper nouns are not in the focus of this study, this confirms the common assumption that “nouns in creoles are not specified for gender” (Bobyleva 2013: 109); one can even put it stronger in stating that “the reduction and loss of class/gender distinctions is a universal feature of the pidginization and creolization of languages” (Heine and Reh 1984: 42, cited in Aikhenvald 2000: 388). Beyond that, no further indication of a pidgin or creole employing nominal classification could be found in the studied data and literature.

After elaborating the exclusion of a number of systemic types, isolates, and language families, it is in order to present this study’s sample. Table 4.5 displays the genetic distribution of the sample and is further commented on in the following.

Table 4.5: The language sample.

<table>
<thead>
<tr>
<th>Language Family</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afro-Asiatic</td>
<td>Masri (Semitic)</td>
</tr>
<tr>
<td></td>
<td>Hausa (Chadic)</td>
</tr>
<tr>
<td>Altaic</td>
<td>Korean</td>
</tr>
<tr>
<td>Amerind</td>
<td>Kiowa (Tanoan)</td>
</tr>
<tr>
<td></td>
<td>Bora (Ge-Pano-Carib)*</td>
</tr>
<tr>
<td></td>
<td>Jakaltek (Mayan)</td>
</tr>
<tr>
<td></td>
<td>Mundurukú</td>
</tr>
<tr>
<td></td>
<td>Tohono O’odham (Pimic)</td>
</tr>
<tr>
<td>Australian</td>
<td>Dyirbal (Pama-Nyungan)</td>
</tr>
<tr>
<td></td>
<td>Ngan’gityemerri (Daly)</td>
</tr>
<tr>
<td>Austric</td>
<td>Hmong (Hmong-Mien)</td>
</tr>
<tr>
<td></td>
<td>Boumaa Fijian (Oceanic)</td>
</tr>
<tr>
<td></td>
<td>Nùng (Tai)</td>
</tr>
<tr>
<td></td>
<td>North Ambrym (Oceanic)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>Ingush (Nakh)</td>
</tr>
<tr>
<td>Chukchi-Kamchatkan</td>
<td>Chukchi</td>
</tr>
<tr>
<td>Elamo-Dravidian</td>
<td>Tamil (Dravidian)</td>
</tr>
<tr>
<td>Indo-European</td>
<td>German (Germanic)</td>
</tr>
<tr>
<td>Indo-Pacific</td>
<td>Tainae (Angan)</td>
</tr>
<tr>
<td></td>
<td>Great Andamanese (Andamanese)</td>
</tr>
<tr>
<td></td>
<td>Nasiol (Bougainville)</td>
</tr>
<tr>
<td>Khoisan</td>
<td>Hadza*</td>
</tr>
<tr>
<td>Sumerian</td>
<td>(isolate)*</td>
</tr>
<tr>
<td>Ket</td>
<td>(isolate)*</td>
</tr>
<tr>
<td>Burushaski</td>
<td>(isolate)*</td>
</tr>
<tr>
<td>Na-Dene</td>
<td>Tlingit</td>
</tr>
</tbody>
</table>
Focus and Method of this Study

After having chosen for the required number of languages from each phylum with due regard to phylum-internal genetic variation, the preliminary composition is further adjusted, since the sample must include all types of nominal classification in their respective full spectrum: Smallish and largish classifier systems (e.g. Tlingit (two numeral classifiers) and Korean (~197 numeral classifiers)) and concordial class systems (e.g. Sumerian (two concordial classes), Burushaski (four concordial classes) and Chichewa (10 concordial classes)), all terminal subtypes of the taxonomy of the systems of nominal classification, systems that employ multiple categorisation devices (e.g. Hmong and Nung (both multiple classifier systems) and Bora (hybrid system)), systems of the same categorisational type that are more based on semantic principles (e.g. Ket and Tamil, cf. Aikhenvald 2000: 22f.) and ones that are more based on formal principles (e.g. German). Furthermore concordial class systems that employ markers in form of different forms of affixes (e.g. Maasai (prefixes) and Ngan’gityemerri (both independent and bound forms)) were considered as well as classifier systems that employ classifiers in form of free morphemes (e.g. Korean; cf. Aikhenvald 2000: 99), in form of clitics (e.g. Boumaa Fijian) and both bound and independent classifiers (e.g. Korean and Mandarin Chinese; cf. 4.2).


A sān ˈge
three CL:GENERAL people
“three people”

B nèi-ˈge ˈkǐ
that-CL:GENERAL course of food
“that course of food”

Finally, systems were included that may represent the most extreme instances of nominal classification and thus are generally not perceived as such. These include for instance the system of Chukchi, which divides its nominal lexicon into declensional classes that are depending on the nominal referent’s animacy, but is not assumed to be a nominal classification device by most researchers (cf. Aikhenvald 2000: 438f.). Hungarian also is generally assumed not to employ a system of nominal classification. However, there are optional classifier-like elements that occur in quantificational and demonstrative contexts with inanimate objects. The objects are categorised with respect to their shape and form (Aikhenvald 2000: 102). Beyond that, Hungarian employs set
nouns (cf. de Groot 2005), which are related to nominal aspectualisation and therefore occur in classifier languages (cf. Rijkhoff 2002a: 347). Due to this, Hungarian was adopted for this study.

If a language proved to employ phenomena that potentially fell into the realm of nominal classification (even though they might not commonly perceived as such in the specific language) during the data collection, these phenomena were documented and considered as well (e.g. the Jakaltek numeral modifiers, which are referred to as classifiers in the literature, but do not employ a classifier function). Table 1 displays the typological distribution of the sample languages; languages marked by an asterisk are multiple classifier systems and occur under the respective system’s commonly assumed primary morphosyntactic host. Note that Dongo-Ko employs both a possessive classifier system and a concordial class system.

![Figure 4.4: Sample coverage of the terminal subtypes of nominal classification systems.](image)

The sample systems’ functional properties are used to determine whether or not the investigated phenomenon is an actual system of nominal classification or not. The systems investigated in this study naturally do employ at least one of the functions presented in chapter 2 (cf. page 11f.); their functional distribution is displayed in table 4.6. The functional account for the individual systems is provided in the appendix (cf. page 353).

All systems except for nine serve both the discourse function of reference identification (C₄) and the semantic function of referent differentiation (C₁). Out of the seven that do not employ both, it is the classifier systems of Great Andamanese (9), Boumaa Fijian (7), and Dongo-Ko (13) which serve only the
Focus and Method of this Study

Table 4.6: The functional properties of the language sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- property is employed, blank: property is not employed;

function of referent differentiation and the classifier systems of Hungarian (8), Tlingit (12), Tohono O’odham (14), and Hmong Daw (15) as well as the concordial class systems of Moro (26) and Sumerian (22) which serve reference identification. The data in Table 4.6 already verifies H1.1 (classifier and concordial class systems form a functional continuum). Beyond the function of reference differentiation and identification, a number of systems (the classifier systems of Jakaltek (10) and Hmong Daw (15), the intermediate systems of Mundurukú (5) and Ngan’gityemerri (6), and the concordial class systems of Maasai (21), Dyirbal (18), Tamil (24), and Chichewa (29)) allows their speakers to attribute properties to nominal referents by classificatory means; four systems (the classifier systems of Hmong Daw (15) and Nùng (16), the intermediate system of Bora (4), and the concordial class system of Kiowa (23)) serve reference identification and differentiation and additionally individuation. Finally, there are the classifier systems of Korean (1) and Mandarin (2) and the concordial class system of Burushaski (20), which employ all four functions accounted for in this study.

As this overview demonstrates, the functions are fairly well-balanced among the systemic types of nominal classification. Yet there seems to dominate a preference for concordial class systems to employ the functional combination of reference differentiation and identification without additional functions; this combination of C4 and C1 is found in eleven concordial class systems, but only
4.4 The Data: Nature and Dossiers

The individual dossiers on the sample languages all follow the same structure, which allows an easy comparison of the presented data. Their structure and the data sources are discussed in this section.

The main data source consists in secondary literature on the individual sample languages and their classification systems. The individual systems are assessed typologically and, if necessary, reclassified: For instance, the Mandarin and Korean classifier systems are commonly described as numeral classifier systems. However, both systems allow their classifiers to attach e.g. to the numeral and the demonstrative, which renders them multiple classifier languages in the definition of this study. Their respective dossier consequently lists the Mandarin and Korean classifiers as multiple classifiers.

Similarly, some languages feature additional grammatical phenomena next to their nominal classification system, which bear parallels to nominal classification devices. These phenomena are integrated into the dossiers and assessed with respect to their potential as nominal classification devices. For instance, Jakaltek is commonly described as a noun classifier language (cf. e.g. Craig 1986b). Day (1973: 59) describes additional ‘numeral classifiers’, which are described and assessed in the dossier; these ‘numeral classifiers’ e.g. are not considered a nominal classification device in the definition of this study.

Each dossier consists of four content sections:

1 Background
2 Classification Typology
3 Analysis
4 Grammaticalisation

The first section provides a general background: It features the genetic affiliation according to Lewis et al. (2013) and Ruhlen (1987), describes the area where the language in question is spoken and potential alternative names,
Focus and Method of this Study

names the main sources that were used for the compilation of the dossier, and comments briefly on the kind of classification system that a language employs. It concludes with a figure that displays the type of classification system according to this study’s analysis. Figure 4.5 summarises the first section of the dossiers.

Figure 4.5: Overview: Structure of the first section of the individual language dossiers.

1 Background
[genetic affiliation (Lewis et al. 2013, Ruhlen 1987)]
[language area, main sources, type of classification system]

2 Classification Typology
2.1 Noun classifiers: present/absent?
2.2 Numeral classifiers: present/absent?
2.3 Possessive classifiers: present/absent?
2.4 Multiple classifiers: present/absent?
2.5 Concordial class system: present/absent?

The second section addresses classification typology of the language in question, more specifically the individual types of classification systems. Each systemic type is either included or excluded from the language in question on basis of the consulted sources and/or examples. Figure 4.6 summarises the structure of the dossiers’ second section.

Figure 4.6: Overview: Structure of the second section of the individual language dossiers.

The third section comprises the main part of the dossiers: Here, the three approaches to nominal classification are applied to the sample language in individual subsections. All subsections of section three includes a table or figure summarising the respective account’s findings, and an elaboration and/or exemplification of the individual criteria. The first subsection contains this study’s adaption of the functional typology of nominal classification (cf. figure 4.7 and Contini-Morava and Kilarski 2013). The account for a system’s application to
the functional typology serves the determination of whether or not a system is a nominal classification system: If it employs one or more functions from Contini-Morava/Kilarski (2013), it is considered a potential system of nominal classification. If it does not employ at least a single function from Contini-Morava/Kilarski (2013), it is not considered a nominal classification system.

Figure 4.7: Overview: Structure of the first subsection of section three of the individual language dossiers.

3 Analysis 3.1 Functional setup

| differentiating referents (C₁) | productive derivation, variable classification |  
| individuation (C₂) | unitisation for countability |  
| attribution of properties (C₃) | speaker-sided values |  
| reference identification (C₄) | reference tracking via AGR, anaphora |  

ad C₁: discussion/exemplification
ad C₂: discussion/exemplification
ad C₃: discussion/exemplification
ad C₄: discussion/exemplification

The second subsection contains the application and exemplification of the prototype approach of the form-based typology (cf. figure 4.7 and Dixon 1986). The properties of prototypical classifier systems and concordial class systems are presented in the sections 2.2 and 2.4 and their expected correspondence to the subprocesses of grammaticalisation are addressed in section 3.2. The structure of this section is depicted in table 4.7; each property is addressed, discussed, and exemplified individually.

Table 4.7: Overview: Structure of the second subsection of section three of the individual language dossiers.

| 3.2 Formal setup |  
| obligatoriness (F₁) |  
| CLS-like: nouns may remain uncategorised |  
| CCS-like: obligatory assignment of each noun |  
| flexibility (F₂) |  
| CLS-like: flexible class membership |  
| CCS-like: single class membership |  
| inventory (F₃) |  
| CLS-like: largish class number |  
| CCS-like: limited class number |  
| marking (F₄) |  
| CLS-like: marked once |  
| CCS-like: multiply marked |  

|
Focus and Method of this Study

The third subsection presents the mapping account that is presented in detail in the chapters 7 and 8 of this study. This account includes the calculation for two dimensions of nominal classification, the formal one and the dimension of transparency. For this purpose, a number of properties are evaluated and translated into an approximation value, which allows the graphic representation of each system in a two-dimensional graph (cf. figure 4.8).

The fourth section briefly summarises the observations that were made with respect to the individual systems’ grammaticalisation. It furthermore provides an overall characterisation of the respective system.

This chapter outlined the focus of this study in form of its research questions and the related hypotheses and predictions. Beyond this, it argued for the feasibility of a typological study of the grammaticalisation hypothesis by means of a sample of synchronic systems. The sampling method and composition was introduced, and the structure of the individual language dossiers, which are provided in the appendix (cf. page 353) presented. The following chapter 5 comprises the first part of the sample study.
Figure 4.8: Overview: Structure of the third subsection of section three of the individual language dossiers.

3.3 Mapping

<table>
<thead>
<tr>
<th>$x$-crit.</th>
<th>Value</th>
<th>$y$-crit.</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size $X_1$</td>
<td>semantic structure $Y_1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>host number $X_2$</td>
<td>rule types $Y_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>locus operandi $X_3$</td>
<td>rule number $Y_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>obligatoriness $X_4$</td>
<td>categorial independence $Y_4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>boundedness $X_5$</td>
<td>discreteness of markers $Y_5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>marking $X_6$</td>
<td>redundancy $Y_6$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>exhaustivity $X_7$</td>
<td>flexibility $Y_7$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\text{ad } X_1$: discussion/exemplification  \hspace{1cm}  \text{ad } Y_1$: discussion/exemplification

$\text{ad } X_2$: discussion/exemplification  \hspace{1cm}  \text{ad } Y_2$: discussion/exemplification

$\text{ad } X_3$: discussion/exemplification  \hspace{1cm}  \text{ad } Y_3$: discussion/exemplification

$\text{ad } X_4$: discussion/exemplification  \hspace{1cm}  \text{ad } Y_4$: discussion/exemplification

$\text{ad } X_5$: discussion/exemplification  \hspace{1cm}  \text{ad } Y_5$: discussion/exemplification

$\text{ad } X_6$: discussion/exemplification  \hspace{1cm}  \text{ad } Y_6$: discussion/exemplification

$\text{ad } X_7$: discussion/exemplification  \hspace{1cm}  \text{ad } Y_7$: discussion/exemplification
The synchronic analysis evaluates the validity of this study’s initial hypotheses and the taxonomy that was created on the basis of them. According to these hypotheses, a nominal classification system may start as a classifier system and develop towards a concordial class system, and this should be mirrored in the properties of the individual types of nominal classification systems. The actual properties of the subtypes of nominal classification are best studied in a synchronic analysis of various known systems of the same type. A possible outcome of the synchronic analysis is a clearer conceptualisation of the ‘grey area’ between classifier systems and concordial class systems. Due to the prototypical and canonical definition of nominal classification systems (cf. e.g. Corbett 2012, Corbett and Fedden 2015, Dixon 1982 and 1986), they are not expected to obligatorily share all features of the prototype. This allows them to take a position that is somewhat different from the prototype. Along with the gradient nature of the development of nominal classification systems, this creates a fuzziness towards the respective ‘anti-pole’ of the continuum, giving rise to the question: What can (still) be considered a classifier system and what can (yet) be called a concordial class system? Or to state this in a different manner: How grammaticalised can a system of nominal classification be in order to be considered a classifier system and from which point of its development must a system of nominal classification be called a concordial class system?

In order to get a grip on the definitory fuzziness and to study a variety of mixed systems, which most likely deliver insights into the nature of nominal classification, this thesis’ definition of the subtypes of nominal classification is a narrow one. There are various systems that can be assigned neither the category of classifier nor the category of concordial class systems due to the relatively
strict definitions of both categorisation types. These highly intermediate ‘hy-
brids’ are of special interest for this investigation, since their inspection may
result in a clearer image of the functions of nominal classification and possi-
ibly even in a reduction of this intermediate grey area inbetween classifier and
concordial class systems.

Section 5.1 provides an introductory overview on the full sample. The clas-
sifier systems of Korean, Mandarin, Boumaa Fijian, Hungarian, Great An-
damanese, Jakaltek, North Ambrym, Tlingit, Dongo-Ko, Tohono O’odham,
Hmong Dow, and Nùng (1, 2, 7–16) are addressed in section 5.2 the con-
cordial class systems of Chukchi, Dyirbal, Tainae, Burushaski, Maasai, Sume-
rian, Kiowa, Tamil, Dongo-Ko, Moro, Ket, Ingush, Chichewa, German, Hadza,
Hausa, and Masri (17–33) are addressed in section 5.3. The hybrid systems
of Nasioi, Bora, Mundurukú, and Ngan’gityemerri (3–6) are discussed in sec-
tion 5.4 and the grey area (which includes all systems except the classifier
system of Korean and the concordial class systems of Dongo-Ko, Moro, Ket,
Ingush, Chichewa, German, Hadza, Hausa, and Masri; 2–24) is focused on in
section 5.5.

5.1 Overview

The following paragraphs constitute a basis for the subsequent detailed study
in providing a schematic overview over the properties of the systems in the
language sample. The raw data from the sample is presented and put in a
first order of concrete typological patterns, which then are analysed in a more
detailed manner. After the discussion of the big picture of the sample, classifier
systems, highly intermediate ‘hybrid’ systems, concordial class systems, and
the grey area in general are addressed in separate subsections. The systems’
properties are presented in tabular form and discussed with respect to their
 correspondences to the grammaticalisation hypothesis. This overview concludes
with a discussion of the preliminary findings, which also addresses the research
questions, predictions, and hypotheses from section 4.1. Table 5.1 gives an
overview over the form-based properties of the whole sample.

Table 5.1: The form-based properties of the language sample.

<table>
<thead>
<tr>
<th>Korean</th>
<th>Mandarin</th>
<th>Nasioi</th>
<th>Bora</th>
<th>Boumaa Fijian</th>
<th>Hungarian</th>
<th>Great Andamanese</th>
<th>Jakaltek</th>
<th>North Ambrym</th>
<th>Tlingit</th>
<th>Dongo-Ko (CLS)</th>
<th>Tohono O’odham</th>
<th>Hmong Dow</th>
<th>Nùng</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The systems in table 5.1 display a gradual and layered increase in concordial class-like properties from the left to the right, which only is blurred by two bigger islands in the left half of the table (Mandarin (2)–Ngan’gityemerri (6) and Great Andamanese (9)–Dyirbal (18); these disturbances of a clear grammaticalisation cline are the more interesting cases for this study and are discussed below. In the following, the ‘vertical’ distribution of properties along the columns is addressed, i.e. the distribution of properties from the perspective of the individual systems.

The observable rise in grammatical formality fits the assumed tendency of nominal classification systems to generally develop along the lines of grammaticalisation and thus supports the grammaticalisation hypothesis. All the systems to the left of Chukchi (17) are those systems that are referred to as classifier systems in the literature, all the systems at its right are commonly referred to as concordial class systems. This confirms H1.2 (classifier and concordial class systems are different systemic types, which display formal and/or functional differences) and P2 (the development of nominal classification devices consists in classifiers grammaticalising into concordial class markers); it also confirms the underlying assumption of the grammaticalisation hypothesis that concordial class systems are more grammaticalised than classifier systems.

There are four cases of hybrid systems. These are systems that cannot be assigned the status of either a classifier or a concordial class system without any doubt, since they display a fairly classifier-like inventory and semanticity, but a formwise fairly concordial class-like behaviour. These are the systems of Nasioi (3), Bora (4), Mundurukú (5), and Ngan’gityemerri (6), and occur as a group among the classifier-like languages (cf. table 5.1). They are addressed in detail in section 5.4.

The expected rise in grammaticality (cf. again pages 70ff.) according to table 5.1 consists in a reduced classificatory set, which subsequently becomes obligatory and develops agreement markers. The markers’ ability to transcend the borders of the noun phrase is closely related to the morphological means provided by affixed multiple markers. The assigning agent seems not to depend on a certain degree of grammaticalisation, since a number of systems with a low degree of grammaticalisation displays a system-sided assignment.
Similarly, the degree of grammaticalisation does not necessarily entail a given system’s semanticity (semantic contribution to the noun phrase and/or a noun’s flexibility in class membership), since there are fairly grammaticalised systems which display productive semantics. For instance, the concordial class system of Burushaski (20) possesses a semantically productive mass-count distinction, and thus the class for masses and the class for collectives share the same noun stem; cf. čhumár c. “(metal) pan”/r. “iron”, bayú c. “leavened bread”/r. “sourdough”, bağýndo c. “rock salt”/r. “grained salt” (Berger 1998: 35). Beyond that, fairly ungrammaticalised systems partly display a low degree of semantic contribution (e.g. the Hungarian (8) numeral classifier system, where classifiers occur optionally even in quantificational contexts; cf. 5.1).

(5.1) Hungarian (cf. Csirmaz and Dékány 2010a: 12)

A  hét (féj) saláta
    seven CL:HEAD lettuce
    “seven lettuces”

B  hét (szem) cukor
    seven CL:EYE candy
    “seven pieces of candy”

C  hét (szál) gyertya
    seven CL:THREAD candle
    “seven candles”

The distribution of the diverging systems is quite stunning: The obvious deviations comprise almost exclusively classifier systems, whereas it is only the concordial class system of Chukchi (17) and Dyirbal (18) that form the offset of the larger fields of diverging systems in table 5.1. A number of concordial class systems (Maasai (21), Sumerian (22), and Kiowa (23)) break the grammaticalisation pattern in table 5.1 by displaying a slightly different order of grammaticalised properties, which results in a line of single classifier-like properties that penetrates an otherwise stable concordial class-like property ‘block’. However, since grammaticalisation is not a strict script, a light variation in the order of grammaticalising properties may be expected (cf. again page 56 and 57) and does not contradict the grammaticalisation hypothesis as such or a general tendency of nominal classification devices to grammaticalise.

Furthermore, it is quite unexpected that not only the bulk of systems is of an intermediate nature due to the employment of both classifier-like and concordial class-like properties, but also that only a single classifier system from the sample, Korean (1), shows the full range of prototypical classifier properties. This surpasses the predictions of P3 (there are intermediate systemic types which display properties of both classifier and concordial systems) and P4 (these intermediate systemic types form a coherent and so far not further specified grey area in the functional continuum). Most of the systems deviate from the grammaticalisation cline and there is only one prototypical classifier system; both these facts emphasise the need for a close inspection of form and
function and a redefinition of classifier systems, since our understanding of them seems to lag far behind the one of concordial class systems. Concordial class systems in contrast are represented in the sample by a comparatively high number of eight fully prototypical systems (Moro (26), Két (27), Ingush (28), Chichewa (29), German (30), Hadza (31), Hausa (32), and Masri (33)). They also generally fit the grammaticalisation pattern without deviating from it a lot.

The deviating islands in table 5.1 are briefly addressed in the following. Due to the limits of this overview, a detailed investigation of these is carried out in chapter 6. The first deviating island comprises all deviating languages that do not employ a small class inventory, i.e. the systems at the left of the table. This ‘hybrid island’ is constituted by the four hybrid systems Nasioi (3), Bora (4), Mundurukú (5), and Ngan’gityemerri (6), and the classifier system of Mandarin (2; cf. table 5.2).

Table 5.2: The ‘hybrid island’.

<table>
<thead>
<tr>
<th></th>
<th>Mandarin</th>
<th>Nasioi</th>
<th>Bora</th>
<th>Mundurukú</th>
<th>Ngan’gityemerri</th>
</tr>
</thead>
<tbody>
<tr>
<td>set size</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>obligatoriness marking</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>locus operandi flexibility</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>assignment</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>sem. contribution</td>
<td>blank: classifier-like, ●: concordial class-like, n: both</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The hybrid systems obviously are exceptional in employing a large class inventory (except for Ngan’gityemerri (6), which possesses a set of ‘only’ 15 marked classes) and a non-obligatory classification, which are both classifier-like properties, but (partly) occur in form of multiple markers. This is distinctly defined as a core property of concordial class systems and by definition not a property of classifier systems. Since these four cases are the most exceptional ones out of the sample, their modus operandi and nature are discussed in section 5.4. Mandarin (2) represents an instance of class assignment which is partly carried out by the system (depending on the register), not the speaker; beyond that, Mandarin nouns are inflexible with respect to their class membership in the higher registers, which means that the ‘correct’ classifier has to be learned for each noun (cf. Li and Thompson 1981: 105, Lin 2001: 106, and Zhang 2007: 48 and 52f.) in order to satisfy those registers.

The second deviating island consists of a number systems (the classifier sys-
tems of Great Andamanese (9), Jakaltek (10), North Ambrym (11), Dongo-Ko (13), Tohono O’odham (14), Hmong Daw (15), and Nùng (16), and the concordial class systems of Chukchi (17) and Dyirbal (18); cf. table 5.3 which each display one concordial class-like property that seems to have formalised ‘too early’ according to the sequence of processes that is predicted by the grammaticalisation hypothesis. The Tlingit system differs even in three properties. Due to their ‘premature’ mechanisation, the sum of these deviations constitutes the ‘mechanised island’.

Table 5.3: The ‘mechanised island’.

<table>
<thead>
<tr>
<th>System</th>
<th>Great Andamanese</th>
<th>Jakaltek</th>
<th>North Ambrym</th>
<th>Tlingit</th>
<th>Dongo-Ko (CLS)</th>
<th>Tohono O’odham</th>
<th>Hmong Daw</th>
<th>Nùng</th>
<th>Chukchi</th>
<th>Dyirbal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set size</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Obligatoriness</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Marking</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Locus operandi</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Flexibility</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Assignment</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Semantic contribution</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

blank: classifier-like, •: concordial class-like, b: both

Most of the systems of the mechanised island (Dongo-Ko (13), Hmong Daw (15), Nùng (16), and Chukchi (17)) differ in that they show no semantic contribution to the noun phrase, while the only other already grammaticalised properties are their small class inventory (and the obligatory class membership in Chukchi (17)). Great Andamanese (9), Jakaltek (10), and North Ambrym (11) also have a very restricted class inventory, which combines with a (partly) system-sided assignment, while all other properties are classifier-like. Dyirbal (18), one of the textbook examples of Australian concordial class systems, surprises in its richness in classifier-like properties: Next to the small set size and obligatoriness of assignment, which occur in the ‘right order’ according to the grammaticalisation cline, only the rigid class membership of its nouns correspond to the general understanding of a concordial class system. For this reason, Dyirbal is another promising case for a hinge system between the ‘hybrid’ systems and concordial class systems. Finally, Tlingit (12) and Tohono O’odham (14) stand out among the other classifier systems, since both employ more concordial class-like properties than classifier-like ones. Since these concordial-class like properties occur not in the order which is predicted by the grammaticalisation hypothesis, Tlingit deviates from the grammaticalisation pattern. It is characterised by a small classifier inventory, a rigid and system-
sided assignment, and a lack of semantic contribution to the noun phrase by classificatory means, which all are concordial class-like properties. At the same time, its markers occur once and within the noun phrase, which together with the non-obligatoriness across morphosyntactic contexts is typical for classifier systems.

To sum up, the sample displays a clear grammaticalisation cline, distributed over less grammaticalised classifier and hybrid systems and more grammaticalised concordial class systems. Considering the sample as a whole, two issues attract attention: Concordial class systems correspond a lot better to their prototype than classifier systems, and many sample systems deviate from their prototype with respect to their inventory size and semantic contribution to the noun phrase. A lack of understanding of the underlying nature of classifier systems may explain shortcomings in our definition of classifier systems. Consequently, careful and data-directed study may gain us a refined and sound definition for a prototypical classifier system.

This chapter investigates the sample ‘vertically’, i.e. from the perspective of the individual systemic types column by column, before the systems’ individual properties are studied in chapter 6. It further explores what the sample tells us about the grey area inbetween classifier and concordial class systems and whether or not the systemic types display a preferred setup of properties. One of the most obvious observations from table 5.1 is the relative homogeneity of concordial class systems with respect to their properties’ degree and order of grammaticalisation, whereas classifier systems display a more heterogeneous picture in that the two deviating islands consist almost exclusively of classifier systems. In the following, the basic systemic types are inspected more closely for reasons for this divergence, and also those systems constituting the deviating islands and other individual systems that display distinct deflections from their expected behaviour. It is shown that the bulk of deviations is due to the conventionalisation of a classifier system rather than its actual grammaticalisation: A given system loses its semanticity and conventionalises noun-class pairs. If these formalised systems can be shown to overlap with the deviating islands from the expected grammaticalisation pattern, the grammaticalisation hypothesis can be verified as an alternative process of change to the conventionalisation of a system of nominal classification. Table 5.4 gives an overview over the distribution of systemic types within the sample, which are treated in separate subsections below.

Table 5.4: The systemic distribution within the language sample.
5.2 Classifier Systems

This subsection provides an overview over the properties of the sample systems which are treated as classifier systems in the literature. Table 5.5 lists those sample systems that are commonly referred to as classifier systems in the literature, except for the ‘truly’ intermediate systems of Nasiol (3), Bora (4), Mundurukú (5), and Ngan’gityemerri (6), which are discussed in section 5.4. The systems are ordered with respect to their morphosyntactic host (noun, numeral, possessive, or multiple).

Table 5.5: The classifier systems from the sample.

| noun CLSs | Jakaltek | Tlingit |
| possessive CLSs | Dongo-Ko | Boumaa Fijian | Tohono O’odham | North Ambrym |
| multiple CLSs | Mandarin (num, deic, n) | Ning (num, deic, poss) | Great Andamanese (poss, n, adj, v, adv) | Hmong Daw (num, deic, poss, n) | Hungarian (num, dem) | Korean (num, dem, adj) |

Their distribution of properties, which is provided in table 5.6, makes it evident that classifier and concordial class systems display formal differences despite their commonalities that result from them being part of the nominal classification continuum (cf. H1,2).
Table 5.6: The form-based properties of the classifier systems from the sample.

<table>
<thead>
<tr>
<th>Set size</th>
<th>Obligatoriness marking</th>
<th>Locus operandi flexibility</th>
<th>Assignment</th>
<th>Semantic contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>2</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>7</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>8</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>9</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>10</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>11</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>12</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>13</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>14</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>15</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>16</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

blank: classifier-like, •: concordial class-like, b: both

Table 5.6 shows that all sample systems except for the one of Korean are intermediate in employing both classifier-like properties and at least one concordial class-like property; yet most of the systems display prevailing classifier-like properties in their individual setup with only one or two deviations. Only Tlingit (12) displays four concordial class-like and three classifier-like properties. When comparing table 5.5 and 5.6, it becomes evident that the systemic classifier types are quite evenly distributed. Of course, twelve systems are not sufficient to reliably generalise from, but there seems to be no strict correlation between a classifier system’s type and its preference in grammaticalised properties.

The overall optionality of the classification and its NP-internal marking by a single marker are consistent properties throughout the sample. These three properties thus appear to be useful for the definition of classifier systems. The other properties display a more mixed picture: While most systems allow a flexible classification, Tlingit (12) and partly Mandarin (2) have conventionalised classifier-noun pairs. Note that spoken Mandarin allows a greater flexibility, while higher registers of Mandarin demand a more rigid classification. Both these systems and the ones of Jakaltek (10) and North Ambrym (11) employ a (partly) system-sided assignment, i.e. the speaker has no free choice on the assignment. The assignment of Great Andamanese (9) classifiers is partly performed by the system, but also allows the speaker to choose a classifier; this may be due to an ongoing grammaticalisation process. Five systems (Dongo-Ko (13), Tohono O’odham (14), Hmong Daw (15), Nùng (16), and Tlingit (12)), which is almost half of the investigated classifier systems, do not semantically contribute by classificatory means\(^\text{1}\). All of these systems employ in addition a

\(^{1}\)Note that the numeral classifier systems of Nùng and Hmong partake in individuation, which is a semantic function. However, these systems do not distinctly add new or more specific semantics to their host NP, whereas in other numeral classifier languages, semantics
delimited class set, which points to an advanced degree of grammaticalisation. The inventory size appears to be the most susceptible property for grammaticalisation: Only two out of twelve classifier systems from the sample (Korean (1) and Mandarin (2)) employ a ‘largish’ inventory, while all other systems do not. In conclusion, the following properties appear the most instable ones in the case of a grammaticalisation process in a classifier system: The set size, the semantic contribution, and the freedom of speaker-sided assignment. This conjecture is investigated more in detail in chapter 6.

There are two systems that deviate in only one property from the prototype (Boumaa Fijian and Hungarian, which employ both a small classifier set). These properties again fit the expectations according to the grammaticalisation hypothesis, since semantic bleaching and inventory reduction are supposed to be the first visible symptoms of a grammaticalisation process. The system that differs most from the defined classifier prototype, namely in four properties, is the numeral classifier system of Tlingit, which consists of only two categories and is semantically vain. Table 5.6 suggests an interaction of the development of a classifier system’s set size and the markers’ semantic contribution to the noun phrase. The assigning agent (system or speaker) and the flexibility of the assignment also interact when it comes to grammaticalisation.

There is a number of deviant classifier systems which stick out of the sample due to their nonconform appearance or systemic composition. These are the systems of Mandarin (2), Great Andamanese (9), and Tlingit (12), and those possessed classifier systems that are elsewhere analysed as ‘relational’ classifier systems (Boumaa Fijian (7), North Ambrym (11), and Tohono O’odham (14)). Note that with the exception of Mandarin, which employs at least 187 classifiers, these are all the sample systems with an inventory smaller than ten classes. The systems are studied individually in the following.

The classifier system of Mandarin behaves like a typical classifier systems in lower spoken registers, which means that the speaker variably assigns classifiers to nouns at his or her will (cf. Bisang 1993: 23). However, in higher registers, there are invariant classifier-noun pairs: The ‘correct’ classifier for a given noun has to be acquired with the learning of the noun (Zhang 2007: 54–56). The rigid conventionalisation of classifier-noun pairs in higher registers is of a strong concordial class-like character: Recall that the German concordial class system requires a noun’s class to be learned together with the noun; the situation in Mandarin is the same, except that the inventory is with at least 187 classes much larger than the German concordial class system. A full concordial class-like character of Mandarin nominal classification in higher registers is only restricted by the large inventory and Mandarin’s isolating language type, which

\[\text{Note that Jakaltek (10) possesses an inventory of 24 classifiers and thus is borderlining the ‘largish’ systems. The reason for its inventory being classified as delimited is the narrow definition of a delimited set as employing up to 24 categories. The preliminary working definition of the individual properties is re-evaluated in the course of chapter 6.}\]
Grammaticalising Systems

prevents it from marking class membership by means of agreement.

The Mandarin case represents a challenge to the grammaticalisation hypothesis in that there is no observable tendency towards a reduction of the inventory; a semantic category expansion takes place only in individual classes (such as méi “trunk of a bamboo tree”, which semantically eroded to a general classifier until about 900 AD and subsequently “virtually dies out”; Aikhenvald 2000: 410), whereas the bulk of classes displays a fairly specific meaning. Though the classifiers may be bound clitics, the markers in general maintain a morphosyntactically independent status (cf. again 6.4 on 188, where the same marker occurs in clitic and independent form). This seems rather natural for an isolating language. A grammaticalisation process is less likely to be involved of the creation of the higher registers with their invariably classifier-noun pairs than a process of lexicalisation.

The Great Andamanese (9) system fits the expected grammaticalisation cline well. The only divergent property is the assignment, which is partly system-sided and partly speaker-sided. A salient trait of the morphosyntactic behaviour of the system is its potential to combine with a broader range of hosts, for instance nouns, possessives, adjectives, adverbs, and verbs. The marking is classifier-like in that a classifier occurs once per classificatory context; even if e.g. an adjective is marked for a given class, the class assignment is carried out with respect to the adjectival semantics and does not depend on the noun (cf. 71; note that the noun in B remains uncategorised, while the adjectives themselves are categorised by their respective classifier).

(5.2) Great Andamanese (cf. Abbi 2013: 244)

<table>
<thead>
<tr>
<th>A</th>
<th>aka=p^up</th>
<th>ut=tgy</th>
<th>e=caγ</th>
<th>his=sputum CL:4=odour CL:5=bad</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“His bad smelly sputum.”</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>k^ŋ</th>
<th>e=mele</th>
<th>e=me</th>
<th>bhi=caγ-o</th>
<th>soil CL:5=fine CL:5=good earth.OBJ=find-PST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“He found good and fine soil on the ground.”</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Despite the fairly unobtrusive distribution of their formal properties within the sample, the classifiers of Great Andamanese divert from a prototypical classifier system in that they display semantic variation of the individual classes in different morphosyntactic environments. This is not common for nominal classification devices, where a class typically displays stable semantics across contexts and semantic variation is achieved by class variation. Semantic extension can also be found in individual classes in nominal classification systems; yet in Great Andamanese, different morphosyntactic hosts may trigger different semantic paradigms of the same set of classifiers, as is illustrated in table 5.7.
5.2. Classifier Systems

Table 5.7: The semantics of the Great Andamanese classifiers in different morphosyntactic contexts (cf. Abbi 2011: 773 and 775f.).

<table>
<thead>
<tr>
<th>class</th>
<th>with adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a-</td>
<td>“mouth-related attribute”</td>
</tr>
<tr>
<td>2 er-</td>
<td>“external attribute”</td>
</tr>
<tr>
<td>3 or-</td>
<td>“limb-related attribute”</td>
</tr>
<tr>
<td>4 ut-</td>
<td>“negative attitude”</td>
</tr>
<tr>
<td>5 e-</td>
<td>“inherent attribute”</td>
</tr>
<tr>
<td>6 aru-</td>
<td>“belly-related attribute”</td>
</tr>
<tr>
<td>7 o-~o-</td>
<td>“attribute of shape/texture”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>class</th>
<th>with adverbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a-</td>
<td>“deixis front/back, anteriority of action”</td>
</tr>
<tr>
<td>2 or-</td>
<td>“deixis of adjacency”</td>
</tr>
<tr>
<td>3 or-</td>
<td>“haste, hurriedly done action”</td>
</tr>
<tr>
<td>4 ut-</td>
<td>“directional deixis”</td>
</tr>
<tr>
<td>5 e-</td>
<td>“deixis of internal space”</td>
</tr>
<tr>
<td>6 aru-</td>
<td>“deixis of immediate vertical/horizontal space”</td>
</tr>
<tr>
<td>7 o-~o-</td>
<td>“temporal deixis”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>class</th>
<th>with verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a-</td>
<td>“mouth-related activity, origin”</td>
</tr>
<tr>
<td>2 or-</td>
<td>“action involving front part of body”</td>
</tr>
<tr>
<td>3 or-</td>
<td>“hand-related activity”</td>
</tr>
<tr>
<td>4 ut-</td>
<td>“directional, experiential”</td>
</tr>
<tr>
<td>5 e-</td>
<td>“action involving interior of an object”</td>
</tr>
<tr>
<td>6 aru-</td>
<td>“action involving middle portion of body”</td>
</tr>
<tr>
<td>7 o-~o-</td>
<td>“resultative state”</td>
</tr>
</tbody>
</table>

Especially the pairing of classifiers with adverbs and verbs results in a function that evidently differs from nominal classification, since it is the verbal and not the nominal semantics which is manipulated (cf. 5.3).

(5.3) Great Andamanese (cf. Abbi 2011: 770)

A
ut=file, e=file, ek=file
CL:4=aim CL:5=aim OBJ=aim
“aim from above, aim to pierce, aim at”

B
aru=p³o, or=p³o, ut=p³o
CL:6=cut CL:2=cut CL:4=cut
“cut down/fell, hit with a stick (in the front), separate from the source”

C
er=bate, ek=bate, ut=bate
CL:2=slap OBJ=slap CL:4=slap
“slap on the face, slap suddenly, slap hard”

Even in the same morphosyntactic environment, the classifiers may entertain a variety of semantic domains depending on the context and intended contribution to the noun phrase. Cf. [5.4] where a combination of the classifiers with the same kind of morphosyntactic host, the noun, entails a different underlying semantic basis, namely bodily reference (cf. [5.4] A, where p³“back”
is categorised as a major external body part), spatial reference (cf. 5.4 B, where cala “scar” is locally specified by the classifier), and inherent relation (5.4 C).

(5.4) Great Andamanese (cf. Abbi 2011: 748 and 764f.)

A buruj p^b uf "the back of the mountain"

B er= cala Cl:2 scar "scar on the head"

C t^b= er= tpe 1SG Cl:2 bone "My bone."

The tables 44 and 45 list the possible semantic domains, which derive from bodily reference and extended to spatial and possessive (or inherently related) reference.

Table 5.8: The semantic domains of bodily and possessive reference of the Great Andamanese classifiers (cf. Abbi 2011: 745 and 763).

<table>
<thead>
<tr>
<th>class</th>
<th>bodily reference</th>
<th>possessive axioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a-</td>
<td>“mouth and semantic extensions”</td>
<td>“(extension of) mouth, source”</td>
</tr>
<tr>
<td>2 er-</td>
<td>“major external body parts”</td>
<td>“external organs, deictic”</td>
</tr>
<tr>
<td>3 og-</td>
<td>“extreme ends of the body”</td>
<td>“(extension of) hand, extremities”</td>
</tr>
<tr>
<td>4 uf-</td>
<td>“bodily products, part-whole”</td>
<td>“(extension/products of) self”</td>
</tr>
<tr>
<td>5 e-</td>
<td>“inner organs”</td>
<td>“internal organs”</td>
</tr>
<tr>
<td>6 ara-</td>
<td>“round shape, sexual organs”</td>
<td>“relational, circular, curved”</td>
</tr>
<tr>
<td>7 o-~.</td>
<td>“parts for legs and related terms”</td>
<td>“lower body part”</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>class</th>
<th>spatial reference</th>
<th>reference point</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a-</td>
<td>“surface”</td>
<td>“front”</td>
</tr>
<tr>
<td>2 er-</td>
<td>“anterior, exterior”</td>
<td>“front, out”</td>
</tr>
<tr>
<td>3 og-</td>
<td>“posterior, superior”</td>
<td>“up”</td>
</tr>
<tr>
<td>4 uf-</td>
<td>“interior, centre”</td>
<td>“in”</td>
</tr>
<tr>
<td>5 e-</td>
<td>“periphery”</td>
<td>“edge”</td>
</tr>
<tr>
<td>6 ara-</td>
<td>“inferior”</td>
<td>“down”</td>
</tr>
</tbody>
</table>

The pathway of semantic extension and consequently the original state of the system can be reconstructed: According to Abbi (2011: 745f.), the primary
Classifier Systems

context of classification is the possessive construction, where inalienable relations are encoded by body part terms. This categorisation is then extended to “a variety of other terms including kin terms, spatial relation terms, closely related object terms, human attribute/propensity terms, and terms concerning actions, manner and states” (Abbi 2011: 747). The semantic extension of the class set from bodily reference to spatial reference is standing to reason, as body parts may convey locational information within themselves (cf. again 5.4 B). As 5.3 shows, it is the spatial information rather than the bodily reference, which contributes to the semantics of the concept expressed by the verb. One can thus conclude that once the spatial reference has established as a secondary semantic basis of the classifier set, the classifiers’ spread towards elements such as adverbs and verbs is comprehensible and the classifiers’ spatial reference can unfold its full productive contribution to the verbal domain. By undergoing this functional extension, the system transcends the borders of nominal classification and adopts the additional function of verbal classification by divergent layering; yet nominal classification remains an operative domain for the system. Those parts of the Great Andamanese system that correspond to nominal classification are conform with the classifier systems from the sample: As mentioned above, only the assignment does not align with the expected grammaticalisation pattern in that it is partly conventionalised (Abbi 2011: 761).

There is at least one parallel case of a sample-external system of nominal classification which underwent a comparable extension while maintaining its nominal classification character: The Mawng (Yiwaidjan) system of nominal classification consists of five classes; these are marked by affixed agreement markers and occur on a range of NP-internal and -external nominal satellites (cf. Singer 2010: 392ff. and 2012: 966). This renders the system a concordial class system. It serves common functions of nominal classification devices such as referent differentiation (cf. kapaja, which means “wild cotton tree” if assigned to the vegetable class and “wild cotton” if assigned to the masculine class; cf. Singer 2006: 168), nominalisation by means of the classificatory markers (cf. 5.5 A, where the class marker nominalises the adjective), and reference tracking by means of the classificatory markers (cf. 5.5 B, where the class marker serves as a carrier of anaphora).

(5.5) Mawng (cf. Singer 2006: 171 and 178)

A  

\[ \begin{align*} 
\text{k-i-mi-n} & \quad \text{nuyu} \quad \text{nung-murrunti} \\
\text{PRS-3M-do-NON PST} & \quad 3M.OBL \quad \text{LIQUID-bad} \\
\end{align*} \]

“He is doing bad things to him”

B  

\[ \begin{align*} 
(tuka) \quad \text{ta} \quad \text{kurrula.} & \quad \text{inji} \quad \text{kurrum-ta-∅} \\
\text{DEM.PL} \quad \text{LIQUID} & \quad \text{saltwater:LIQUID} \quad \text{NEG} \quad \text{2PL/3L-drink-NON.PST} \\
\end{align*} \]

“(This is seawater.) Don’t drink it.”

In addition to these common functions of nominal classification systems, Mawng allows its class markers to attach to the verb in order to mediate se-
lectional restrictions in the verbal semantics (cf. [5.6] where the class marker specifies the verb semantics).

(5.6) Mawng (cf. Singer 2012: 968)

A  inyi-  wa  -ny  
3M/3F  consume  PST.PUNCTUAL
“He ate it (e.g. a crab).”

B  ani-  wa  -ny  
3M/3LIQUID  consume  PST.PUNCTUAL
“He drank it (liquid).”

C  ati-  wa  -ny  
3M/3VEG  consume  PST.PUNCTUAL
“He ate it (plant food).”

D  ini-  wa  -ny  
3M/3M  consume  PST.PUNCTUAL
“He ate it (animal).”

The combination of a given class marker and verb are prone to lexicalisation (Singer 2010: 403ff.): A (former) class marker occurring on the verb does not represent an instance of grammatical agreement, but an individual lexical unit, which overrides class agreement. For instance, if a speaker wants to express that someone drank maningul “blood” (vegetable gender), he or she phrased it in the form of 5.6 B, where not the expected class marker k-ami-la-∅ “PRS-3M/3VEG-consume-NON.PST” occurs. Instead, the lexicalised variant k-an-la-∅ “PRS-3M/3L-consume-NON.PST” is being used, which carries the class marker for liquids, ani-. In these cases, a function of nominal classification is not anymore employed.

The Great Andamanese system as a whole shows various traits of grammaticalisation. The classifiers most probably derive from lexical items and individually grammaticalised to a greater or lesser degree (Abbi 2011: 778). This involves semantic bleaching, a loss of morphosyntactic independence and phonological weight due to cliticisation, and decategorisation due to the spread of the classifiers across morphosyntactic contexts. The degree of grammaticalisation varies system-internally not only between the individual classifiers (for instance, the classes 6 and 7 are more grammaticalised than class 5), but also between the individual morphosyntactic hosts of one and the same classifier (Abbi 2011: 777f.). On average, the classification occurring on adverbs shows a notably higher degree of grammaticalisation than the classification on adjectives.

---

3Given that the cases of Mawng and Great Andamanese show origins in nominal classification and maintain their nominal classification functions, the synchronically parallel occurring manipulation of verb semantics may be caused by a later process of grammaticalisation, which expanded the systems’ operative nature beyond the limits of nominal classification. Since the source material of a grammaticalising linguistic element may maintain its original function and coexist with the grammaticalising unit (“divergent layering”; cf. page 58 and Hopper and Traugott 1993: 118-122 and 124-126), a suchlike functional split may well have occurred in both Mawng and Great Andamanese.
or verbs. The degree of grammaticalisation in this case is mainly determined by the semantic abstractness of the classes (ib.), since formal properties such as phonological shape and boundedness do not differ across morphosyntactic contexts.

Tlingit is typically labeled a numeral classifier system (cf. Aikhenvald 2000: 121, Dryer and Haspelmath 2013: s.v., Sherzer 1976: 74)\footnote{This assumption seems to be based on Sherzer (1976: 74), where Tlingit is listed among others as a language in which numerals are “classified according to the form or shape of objects”, albeit no further reference or data is provided to support this statement. The consulted sources could not confirm a form or shape classification of Tlingit numerals.} however, this system presents itself in an unusual visual appearance for a classifier system. Tlingit numerals differentiate a plain form tléix’, déix, nás’k, etc., and an infixed form tléi-ná-x, dás-ná-x, nás’g-ná-x, etc.. The infixed form occurs with human referent nouns (Edwards 2009: 17). The first nonconform property is the boundedness of the markers: Classifiers are generally defined as morphosyntactically independent or clitic elements (cf. Dixon 1986, Grinevald 2000: 62, and Rijkhoff 2002a: 74), whereas Tlingit infixes the human class marker -ná-.

The second deviating property is the inventory itself, which consists of a conceivably minimal distinction of a marked class for human referents and an unmarked residue class. This is not typical for classifier systems, and there is only one other supposed classifier system in the sample: The possessive classifiers of Tohono O’odham (cf. again table \[6.4\] on page \[181\]), which feature a binary class distinction and are discussed further below in this section\footnote{The semantic basis of this distinction also largely coincides with the Tlingit system in that animates are marked by \textit{soí} and the residue category by \textit{ını} (Saxton 1983: 186); the difference to Tlingit is that it is both classes that are marked.}. Lichtenenberk (2009: 269f.) argues for the recognition of (possessive) classifier systems that consist of a single marked class; he states that “systems with single classifiers are historical reductions of systems with multiple classifiers” and observes the ‘single classifier’ construction contrasting with a non-classifier construction, which leads to the conjecture that the ‘single classifier’ bears a classifying function. While I do not consider an alternation of construction instead of class marking a true means of nominal classification, there is little formal or logical reason to exclude systems that have only one marked category available; the Tlingit system so far thus represents an exceptionality, but can be argued to represent a case of nominal classification.

The class markers do not add information to the noun phrase, but copy semantic information about the referent’s humanness. This information is marked by means of infixes on only one word class. This closely resembles the Dyirbal system, which is generally referred to as a concordial class system with four referential classes that are assigned on semantic grounds and marked on only one element, the so-called ‘noun marker. Considering the twofold class inventory, the infixal nature of the classificatory marker, and that its semantic information is copied from the head noun and not added to the noun phrase, the Tlingit system is better analysed as a concordial class system. Like Dyirbal,
the most grave concern against treating the Tlingit system as an instance of a concordial class system ‘proper’ is that a noun’s class membership is only expressed on one nominal satellite, which was defined as a characteristic property of classifier systems, not concordial class systems. A more detailed study of the Tlingit and Dyirbal class marking is carried out in section 6.3.

The possessive classifier systems of Tohono O’odham, Boumaa Fijian, and North Ambrym all feature a small inventory of two, four, and five classes respectively. All three inventories are characterised by semantically specific classes, which are complemented by one or two residue or general classes: Tohono O’odham distinguishes animacy from inanimacy, Boumaa Fijian an edible and a drinkable category from general and intimate possession, and North Ambrym a category for edibles, drinkables, fire, and baskets, which it distinguishes from a residue category. Table 5.10 lists the inventories and their semantic specification.


<table>
<thead>
<tr>
<th>Tohono O’odham</th>
<th>Boumaa Fijian</th>
<th>North Ambrym</th>
</tr>
</thead>
<tbody>
<tr>
<td>class semantics</td>
<td>class semantics</td>
<td>class semantics</td>
</tr>
<tr>
<td>soi- [+animate]</td>
<td>‘e- [+edible]</td>
<td>a(n), ye [+edible]</td>
</tr>
<tr>
<td>me- [+drinkable]</td>
<td>‘e- [+related]</td>
<td>ma(n) [+drinkable]</td>
</tr>
<tr>
<td>më- residue</td>
<td>we-, o- residue</td>
<td>to(n) residue</td>
</tr>
<tr>
<td>bo(n) [+basket]</td>
<td>mwen(n) [+flammable]</td>
<td></td>
</tr>
</tbody>
</table>

Most languages of the sample feature a small class inventory of less than 25 classes, which involves all systemic types. Together with Tlingit, the mentioned possessive classifier systems constitute the smallest classifier systems within the sample. Even if the inventory size itself proves to not directly relate to the

Note that the marker ‘e- marks two distinct categories according to Dixon (1988: 137): “There appears to be no connection between them; the two ‘e- may conceivably have different historical origins”. Even if they are etymologically no distinct forms and we are dealing with an actual tripartite system the “edible” category of which has semantically expanded in a way that we cannot immediately comprehend, this does not pose a direct problem for our analysis. Typological parallels can be found in Mota (Vanuatu; ISO 639-3: mtt), where the possessed classifier “ga denoted close belongings and edible possessions” (Franjieh 2016: 2), in Manam, where articles of clothing are “possessed inalienably when being actually worn, but are otherwise possessed generally” Lichtenberk (1983a: 159), and in Kilivila, where the same principle holds (Senft 1986: 54). Lichtenberk (1983a: 174) argues that a researcher’s analysis is necessarily an outsider’s interpretation that “need[s] not fully coincide with that of the speakers of the language, and that some items that to him are exceptional would no longer appear as such if he were able to view the world through the eyes of the members of that culture”. It is thus reasonable to assume that Boumaa Fijian distinguishes only three possessive categories.
typology of nominal classification devices or their degree of grammaticalisation, this allows for a closer inspection of the semantic structure underlying these systems.

The inventories of Tohono O’odham and Tlingit are fairly similar in that they distinguish animacy and humanness respectively by their classificatory marking. The difference between both systems is that Tlingit marks only the human category overtly, whereas the residue category remains unmarked; Tohono O’odham has distinct markers for each class (cf. 5.7).

(5.7) Tohono O’odham (cf. Langacker 1977: 91f.)

A huan gogs sai-ga
Juan dog CL:ANIMATE-POSS.ALI
“Juan’s dog”
B huan uus ñi-ga
Juan tree CL:GENERAL-POSS.ALI
“Juan’s tree”

Like in the Tlingit case, there are a number of arguments that suggest the Tohono O’odham system to be typologically closer to a concordial class system than to a classifier system: The minimal inventory of only two classes, the bound status of its affixed markers, and the classificatory markers mirroring semantics of the head noun instead of adding semantic content to the noun phrase. The most obvious deviation from a typical concordial class system consists in the marking of a nominal referent’s class membership on only one constituent, the possessive marker for alienability -ga (cf. again 5.7).

The inventories of Boumaa Fijian and North Ambrym relate to an Oceanic ancestral proto-system, which distinguished alimentary objects from other posses-sibles. Boumaa Fijian mirrors this original state well, while North Ambrym has expanded the system by two highly specific classes, one for flammables and one for the single lexical item arrbol “basket” (Franjieh 2012: 242). Additionally, the edible and drinkable classes have been expanded: The edible category includes food, animals, tools, and units of time, and the drinkable category hosts items for liquids and their containers, buildings, holes, mats, and domestic property (cf. Franjieh 2012: 241, 244 and 2014: 5). What makes these systems special within the sample is the lexicalisation of class membership; like a number of other systems (cf. e.g. Aikhenvald 2000: 133ff.), they are typically analysed as ‘relational classifier systems’. These systems are used in possessive constructions that distinguish alienability from inalienability, and this distinction is

---

7The exact composition of the Proto-Oceanic inventory is debated: Lichtenberk (1983a: 172) suggests that a single alimentary category existed at the proto-stage, was inherited to the daughters of Proto-Oceanic, and occasionally subspecified into an edible and a drinkable category. Franjieh (2012: 206f., drawing on Lynch 1996) discusses the reconstruction of an originally threefold model (*ka- “edible and subordinate”, *ma- “drinkable”, and *na- “general”), which was inherited to the daughters of Proto-Oceanic and occasionally fused the edible and drinkable category into one alimentary category.

8Cf. again the discussion on page 28f., Franjieh (2016), and Passer (2016) for a discussion in favour of an exclusion of supposed ‘relational classifiers’ from the typology of nominal classification devices.
“a grammatical rather than a purely semantic property” (Bickel and Nichols 2013). Franjieh (2014, 2015, and 2016) convincingly shows that the class membership of North Ambrym nouns is lexicalised and inflexible: “possessed nouns generally occur with just one classifier regardless of context” (Franjieh 2016: 1). This renders the supposed ‘relational classifiers’ a phenomenon that is closer to the status of a concordial class system than to classifiers, and like in the Tlingit and Tohono O’odham case, the most noticeable aspect of these systems is the occurrence of the classificatory markers with one morphosyntactic host only, whereas ‘proper’ concordial class systems are renowned for expressing class membership on a range of nominal satellites.

In summary, the vast majority of deviating classifier systems is characterised by an inventory of up to a handful of classes, which do not semantically contribute to the noun phrase and tend to be invariable for a given noun. These systems resemble concordial class systems that mark class membership on one NP-internal element only. Only the system of Great Andamanese represents a genuinely different case of deviation in that it extended its functional repertoire beyond the limitations of nominal classification and additionally serves e.g. the manipulation of verbal semantics.

So far, it has been shown that concordial class systems, classifier systems, and hybrid systems have a distinct combination of formal properties, which does not come as a surprise, since their respective definition is based on their way of formal expression. It is argued that the way classifiers are expressed is too varied to even account for them as a single phenomenon (Corbett p.c.; cf. also Corbett and Fedden 2015: 2); this subsection is consequently dedicated to the question whether individual systemic subtypes display a distinct setup of properties. Recall that gender and noun class systems are not analysed as distinct systemic subtypes, since their major distinctive criterion is their set size, with noun class systems featuring slightly larger inventories than gender systems. This is in accordance with the state of research of the past decades (cf. e.g. Aikhenvald 2000, Corbett 1991, Grinevald 2000, etc.). We are thus left with the individual types of classifier systems, which are addressed in the following. Table 5.11 displays the properties of the classifier systems from the sample, which are ordered according to their systemic type.
### 5.2. Classifier Systems

Table 5.11: The property composition of the systemic classifier types.

<table>
<thead>
<tr>
<th>Language</th>
<th>Korean</th>
<th>Mandarin</th>
<th>Hungarian</th>
<th>Tlingit</th>
<th>Hmong Daw</th>
<th>Boumaa Fijian</th>
<th>North Ambrym</th>
<th>Dongo-Ko</th>
<th>Tohono O’odham</th>
<th>Great Andamanese</th>
<th>Jakarta</th>
</tr>
</thead>
<tbody>
<tr>
<td>set size</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>12</td>
<td>15</td>
<td>16</td>
<td>7</td>
<td>11</td>
<td>13</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>marking</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>locus operandi</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>flexibility</td>
<td>b</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>assignment</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>sem. contribution</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

numeral CLs | possessive CLs | noun CLs

Table 5.11 shows that there is not much variation occurring between the different types of classifiers. The variation between the types is comparable to the type-internal variation. The individual systemic types thus seem not to prefer a specific and distinct constellation of properties; the main difference between them seems to be the classificatory host. This allows for the question whether there is a noticeable difference in the constellation of properties between classifier systems ‘proper’, i.e. classifier systems that employ only one kind of morphosyntactic host for their markers, and multiple classifier systems. Table 5.12 contrasts the properties of ‘single host’ systems and multiple classifier systems.

Table 5.12: The property composition of ‘single host’ and multiple classifier systems.

<table>
<thead>
<tr>
<th>Language</th>
<th>Boumaa Fijian</th>
<th>North Ambrym</th>
<th>Tlingit</th>
<th>Dongo-Ko</th>
<th>Tohono O’odham</th>
<th>Korean</th>
<th>Mandarin</th>
<th>Hungarian</th>
<th>Great Andamanese</th>
<th>Indonesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>set size</td>
<td>7</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>marking</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>locus operandi</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>flexibility</td>
<td>b</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>assignment</td>
<td>B</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>B</td>
</tr>
<tr>
<td>sem. contribution</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

‘single host’ CLSs | multiple CLSs

blank: classifier-like, •: concordial class-like, n: both
Table 5.12 shows a similar picture as table 5.11 in that the internal variation in multiple and ‘single host’ classifiers is not that different from the variation between the two types of systems. This further supports the observation that the main formal difference between the individual types of classifiers consists in the classifier host, whereas the formal composition of a classifier system does not tell much about its systemic type. After the discussion of the classifier systems from the sample, section 5.3 addresses the concordial class systems from the sample.

5.3 Concordial Class Systems

In this section, the properties of concordial class systems are discussed. Their distribution in the sample is displayed in table 5.13; note again that the hybrid systems of Nasioi (3), Mundurukú (5), Bora (4), and Ngan’gityemerri (6) are discussed in the following subsection.

Table 5.13: The form-based properties of the concordial class systems from the sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chukchi</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>Dyirbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tainae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maasai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sumerian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiowa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingush</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chichewa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hadza</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hausa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hausa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Blank: classifier-like, •: concordial class-like, s: both.

There is no noticeable difference between gender and noun class systems in the sample: Both display either a full correspondence with the definition of a prototypical concordial class system or show a fairly low correspondence. In addition, it is striking that there are five out of eighteen systems which fully correspond to the definition of a prototypical concordial class system, while there is only one classifier system which corresponds to the definition of a prototypical classifier system. This calls for a critical revision of the present notion of a prototypical classifier system.

The grammaticalisation cline is well-mirrored in the distribution of the properties of the concordial class systems from the sample. The top four properties appear to be relatively stable and to correspond to a grammaticalisation cline without exception. There are two properties which are consistently employed by all concordial class systems; these are the obligatoriness of classification.
5.3. Concordial Class Systems

and the limitation of their class inventory. Furthermore, all systems except for Dyirbal (18) and Chukchi (17) mark a noun’s concordial class on more than one nominal satellite at the same time. The markers occur NP-internal and NP-external in all concordial class systems except for the ones of Chukchi (17) and Dyirbal (18). The last three properties are less stable and contain deviations from the grammaticalisation cline. Five out of eighteen systems (Dyirbal (18), Tainae (19), Burushaski (20), Maasai (21), and Tamil (24)) are able to semantically contribute to their host NP. Among the fairly little grammaticalised concordial class systems, Chukchi (17) is the only one that does not allow its markers to semantically contribute to the NP and thus blurs the cline. Seven systems (Chukchi (17), Dyirbal (18), Burushaski (20), Ingush (28), Chichewa (29), Hadza (31), and Hausa (32)) allow at least a part of their nominal lexicon to be flexibly assigned to categories, which is not expected from a prototypical concordial class system. Among the other less grammaticalised systems, only the system of Dyirbal (18) and Maasai (21) display an inflexible assignment and break the line of classifier-like behaviour to the left side of the table. Finally, the speaker is the assigning agent in the systems of Chukchi (17), Dyirbal (18), Tainae (19), Burushaski (20), Sumerian (22), Kiowa (23), and Tamil (24). This pattern is interrupted by the system of Maasai (21), which has a mixed assignment by both the system and the speaker. These three properties are all connected to the semanticity of the systems not only in a productive contribution of the classification to the NP semantics, but also by providing a transparent semantic assignment pattern for the speaker. This suggests that a semantically motivated system may stay semantically motivated, eventually by a remotivation.

There is a number of systems (Chukchi (17), Dyirbal (18), Maasai (21), Sumerian (22), and Kiowa (23)) that slightly deviates from the sample’s grammaticalisation pattern in one property. These deviations concern the semantic contribution, the assignment, and its flexibility, which are partly grammaticalised, while the above criteria are not. Grammaticalisation consists of a number of related processes that occur in most instances of grammaticalisation and that tend to occur in a given order. As this order is not set in stone, a minor variation in the order of grammaticalising properties is to be expected.

Table 5.13 suggests that there are two correlating aspects of concordial class systems that (presumably independent from each other) tend towards change: The flexibility and manipulability of the assignment and the system’s semanticity on the one hand, and the multiple marking and locus operandi on the other hand. These correlations are studied in section 6.8. The individual non-prototypical concordial class systems deviate in one to four form-based properties from the prototype. Sumerian (22) and Kiowa (23) allow for a speaker-sided assignment; Tamil allows for a partly speaker-sided assignment and contributes semantically to the NP. Maasai (21) also semantically contributes to the NP and displays a great deal of classificatory flexibility. It furthermore displays a mostly speaker-sided assignment. Burushaski (20) combines a speaker-sided assignment with the semantic contribution to the NP.
and furthermore shows a flexible classification for a larger part of the nominal lexicon. Tainae (19) employs a fully flexible and speaker-sided assignment with semantic contribution to the NP. Dyirbal (18) and Chukchi (17) both mark a noun’s class membership only once and NP-internally; Chukchi furthermore allows a flexible and speaker-sided assignment, whereas the Dyirbal system has the potential to contribute to the NP semantics and allows a speaker-sided assignment.

A number of concordial class systems show operative traits that deviate from the other systems in the sample, namely Chukchi (17) and Dyirbal (18), Maasai (21), Sumerian (22), and Kiowa (23). These are discussed in the following. Chukchi and Dyirbal display commonalities in employing a small inventory of two and four classes respectively, a semantics-based and speaker-sided assignment, a lack of ‘proper’ multiple marking, and an obligatory classification of nouns. The differences between the system lie in their marking and semanticity: The Chukchi classes are distinguished by different case paradigms (cf. Comrie 1981: 246f. and Weinstein 2010: 40f.) and thus do not possess independent markers for class membership. The markers attach only to nouns and pronouns. Also, a noun’s class membership does not add semantic content to the noun phrase: Human referent nouns constitute one class, the remaining nominal lexicon the other class. The lack of semantic contribution can be recognised by common nouns, which may be assigned to either class (cf. Comrie 1981: 246 and Dunn 1999: 65). Dyirbal nouns have a distinct marker for each of its four classes, and these markers occur on the so-called ‘noun marker’ in its plain, demonstrative, or interrogative form (Dixon 1968: 107 and 120). The ‘noun marker’ constitutes an article-like element that occurs adjacent to the noun. Class membership expresses semanticity, which can be exemplified by the class assignment of spears: Fighting spears are assigned to class II, which hosts nouns referring to dangerous items (fire, fighting), water, and females; fishing spears are assigned to the natural class for fishes, which is class I, and big short spears (supposedly for hunting purposes) are assigned to the residue class IV (Dixon 1968: 122). Since both systems’ assignment is semantics-based and since both mark class membership on only one morphosyntactic host, they resemble classifier systems rather than concordial class systems with respect to their formal expression. This is reflected in their sample-internal position, where they are the systems that are the closest to classifiers (cf. again table 5.1 on page [120]).

The Maasai concordial class system fits the formal pattern of a concordial class system at large, yet there are at least two aspects that set it apart from a typical concordial class system. The system possesses three concordial classes: The masculine class, the feminine class, and a class which is reserved for a single lexical item, *wwéji* “place” (cf. Payne 1998: 160 and Tucker and Tompo Ole Mpaayer 1955: 15). It is highly unusual even for classifier systems and their tendentially bigger class inventories (that consequently host fewer items per class) to employ classes with only one member, but examples such as the North Ambrym classifier *ba(n)*, which combines only with the noun *arrbol*
“basket” (Franjieh 2012: 242) can be found. The reason for the existence of the ‘place class’ in Maasai lies in its diachrony: It is usually analysed as an incipient class of a very young age (cf. Corbett 1991: 313f.) and thus does not yet display the size of a ‘true’ concordial class. However, if the Maasai system introduces a new class to its binary gender system – which is paralleled e.g. in Proto-Indo-European, where a concordial class system with a common and a neuter class diversified into a concordial class system with a masculine, a feminine, and a neuter class –, this means a semantic specification and motivation takes place, which stands in contrast to the expected reduction of a system’s inventory. It may be noted that the Maasai concordial system is formally at the possible end of its grammaticalisation and thus cannot be expected to further grammaticalise in a way that the grammaticalisation hypothesis would predict. Due to its high semanticity, a subspecification of the class inventory is even the most natural development that one would expect.

The other deviation from the grammaticalisation pattern is the degree of flexibility in class membership of the Maasai nominal lexicon. Most of the Maasai nouns “can choose either the masculine or the feminine gender prefix” (Payne 1998: 159f.) due to the underlying connotation of size or strength with these classes, which can also result in a pejorisation of the reassigned noun’s semantics. The size/strength connotation can be applied to virtually any entity, which makes the high degree of flexibility in Maasai standing to reason.

The Sumerian concordial class system almost fully correlates with the defined prototypical concordial class system, except for one property: Its assignment is fully based on transparent semantics, which rigidly classify each noun according to its referent’s humanness. One needs only two rules in order to master the Sumerian gender assignment: First, all human referent nouns are assigned to the common class; second, all remaining nouns are assigned to the neuter gender. Even though concordial class systems always have a semantic core (Corbett 1991: 307), such an absolute instance of transparent assignment is unparalleled among the concordial class systems in the sample.

Finally, the Kiowa system stands out among the concordial class systems of the sample for its unusual way of marking a noun’s concordial class. Kiowa possesses only one marker, which attaches to the noun itself, adjectives, and demonstratives (Wonderly et al. 1954). This marker divides the nominal lexicon into concordial classes (four or nine depending on the analysis; cf. Wonderly et al. 1954: 2ff. and Harbour 2007: 105 and 2011: 578) in that it only occurs together with a noun, adjective, or demonstrative in predetermined numbers. Kiowa nouns are assumed to possess an inherent number value, i.e. a noun remains unmarked in the singular and/or dual and/or plural (cf. table 81 on

---

9Even though not a straight typological parallel, the Dutch diminutive suffix -tje may serve as an example for a comparable productivity. Mentioned suffix is productively attached to a vast number of lexical items in daily speech, which at times results in elsewhere highly uncommon expressions such as zonne-tje ‘sun-ling’ ~ “sun”. Teachers of Dutch even recommend to second language acquirers of Dutch to use the suffix, which triggers agreement with the neuter class, in cases where a noun’s original class is not known for sure.
If a noun appears in a number that is not inherent to it, the marker -go attaches to it. The individual concordial classes of Kiowa are composed by those nouns that have the same inherent number value.

This way of marking class membership is unparalleled in the sample, where different classes are typically marked by distinct markers. Yet Kiowa fits the formal definition of concordial class marking: It is obligatory and a closed grammatical system with a small inventory, expresses itself by affixation on the noun and other NP elements, and is assigned by the speaker on semantic grounds. An interaction or even fusion of nominal classification devices with the category of number is also highly common and found in all types of systems. The main function of the Kiowa system lies in the individuation of nominal referents. Other common functions of concordial class systems such as the expansion of the lexicon’s referential power cannot be observed, though reference tracking may be assumed to play a role in the system due to the marking of demonstratives. Kiowa certainly represents a highly uncommon form of nominal classification, and while it fulfils the formal requirements of a concordial class system, it can be discussed whether or not it actually represents one. In conclusion, its synchronic state does not suggest the Kiowa system to stem from a less grammaticalised classifier stage, but rather the reinterpretation of a number marker for given lexical items, which resulted in the conversion of the marker into a ‘non-inherent number marker’. Note that the sample’s concordial class systems do not deviate as much as its classifier systems, both with respect to quality and quantity.

After having shown that concordial class systems in general conform to an expected grammaticalisation cline and are fairly homogeneous with respect to their inventory size, obligatoriness of classification, potential for multiple marking, and their locus operandi, section 5.4 discusses the sample’s ‘hybrid’ systems.

### 5.4 Hybrid Systems

This subsection addresses the overview over the ‘truly’ intermediate or hybrid systems of Bora (4), Mundurukú (5), Ngan’gityemerri (6), and Nasioi (3). Their properties are listed in table 5.14.
Table 5.14: The form-based properties of the ‘truly’ intermediate systems from the sample.

<table>
<thead>
<tr>
<th></th>
<th>Naso</th>
<th>Bora</th>
<th>Mundurukú</th>
<th>Ngam’gityemerri</th>
</tr>
</thead>
<tbody>
<tr>
<td>set size</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>b</td>
</tr>
<tr>
<td>marking</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>locus operandi</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>b</td>
</tr>
<tr>
<td>flexibility</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>assignment</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>semantic contr.</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

blank: classifier-like, •: concordial class-like, b: both

Table 5.14 displays fairly classifier-like systems with large inventories and productive semantics. They have in common with concordial class systems that they allow their markers to occur NP-externally and employ multiple marking, which is a defining criterion for the exclusion from classifier systems. Therefore, they do not correspond well to either prototype and are treated as ‘hybrid’ systems; not however that the seeming occurrence of multiple marking in these systems is reanalysed in section 6.3. In the case of Ngam’gityemerri (6), the multiple marking is optional (either all nominal modifiers are marked for concordial class or none is), which may indicate a state of emerging multiple marking. The classification of nouns is not obligatory in any of these languages. Furthermore, the intermediate systems are all characterised by a high degree of semanticity: All of them contribute semantically to the NP, and they allow a flexible and speaker-sided assignment. In summary, these systems formally resemble concordial class systems due to the seeming presence of multiple marking, while their high semantic productivity resembles to be classifier-like.

A surprising finding of table 5.14 is the position of the ‘truly’ intermediate systems: As intermediate ‘hinge systems’, they are expected to be located inbetween classifier and concordial class systems. However, they are grouped among the fairly lowly grammaticalised classifier systems towards the left of table 5.1. These supposed ‘hinge systems’ thus do not represent a ‘hinge’ in the sense that they are located somewhere inbetween classifier and concordial class systems, but in combining a high degree of semanticity, speaker-sided involvement, and typically a large inventory with the formal preconditions of a concordial class system, i.e. multiple class marking. This combination of properties locates them at the left outset of the sample, close to the prototypical classifier systems (cf. again table 5.1 on page 120).

In the following, the hybrids are individually discussed, before they are re-evaluated at the end of this section. Note that all hybrid systems share the same
constellation of classifier-like and concordial class-like properties, with only Ngan’gityemerri additionally employing a small class inventory (cf. table 5.14 = 5.15). Therefore, the individual properties are not presented point by point, but the individual nature and characteristics of the classifier- and concordial class-like traits are discussed with reference to the corresponding properties.

Table 5.15: The form-based properties of the ‘truly’ intermediate systems from the sample.

<table>
<thead>
<tr>
<th></th>
<th>Nasioi</th>
<th>Bora</th>
<th>Mundurukú</th>
<th>Ngan’gityemerri</th>
</tr>
</thead>
<tbody>
<tr>
<td>set size</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>●</td>
</tr>
<tr>
<td>marking</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>●</td>
</tr>
<tr>
<td>locus operandi</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>●</td>
</tr>
<tr>
<td>flexibility</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>●</td>
</tr>
<tr>
<td>assignment semantic contribution</td>
<td>blank: classifier-like, ●: concordial class-like, ●: both</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nasioi presents itself in the guise of a concordial class system like the other hybrid systems, but at the same time operates much like a classifier system. Therefore, its classifier-like character are first presented, before its formal concordial class-like traits are discussed. Nasioi employs between 40 and 50 referential classes “according to the size, shape or nature of the object being counted” (Capell 1969: 115, cited in Serzisko 1982: 113f.; cf. also Hurd and Hurd 1966: 20; “several dozen agreement classes” according to Aikhenvald 2000: 79). A selection of these classes is provided in table 109.

Table 5.16: A selection of Nasioi class markers (cf. Hurd 1977: 115-127).

<table>
<thead>
<tr>
<th>category</th>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>social</td>
<td>-varang</td>
<td>“generation, peer group”</td>
</tr>
<tr>
<td></td>
<td>-vuntu’</td>
<td>“village, community, kingdom”</td>
</tr>
<tr>
<td></td>
<td>-ni</td>
<td>“feminine, woman-one”</td>
</tr>
<tr>
<td></td>
<td>-rung</td>
<td>“masculine, man, neutral or generic-one”</td>
</tr>
</tbody>
</table>

10The inventory comprises more than 100 ‘classifiers’ (Aikhenvald 2000: 219, Foley 1986: 83, Terrill 2002: 292). Hurd (1977: 162ff.) lists 118 of these elements (which are labeled “case-gender-number markers” or “CGNs”; Hurd 1977: 114 and Terrill 2002: 292) and groups them in 22 semantic categories (a default one and 21 specific ones; cf. table 109 for a selection), which results in “more concord classes than genders”. Some of these ‘classifiers’ include measures such as -vu’ “piece”, -marang “part”, or -mare’ “year” (cf. Hurd 1977: 122 and Serzisko 1982: 114). The number of 40-50 referential classes is received when one subtracts the “about fifty classes of numbers or count nouns” (Hurd and Hurd 1966: 20).
5.4. Hybrid Systems

There is no information available on the flexibility of classification in Nasioi (Serzisko 1982: 114), except for the general understandability of specific classifier replacement by a general classifier (Hurd and Hurd 1966: 20). However, markers like those listed under the category “bananas” in table 109 typically combine with the same lexical unit and specify it semantically, as (5.8) and (5.9) from Yucatec Maya and Bora illustrate. This allows for the conjecture that a variable classification beyond the alternation of general and specific class markers is also possible in Nasioi.

(5.8) Yucatec Maya (Lucy 1992: 74, cited in Lucy 2000: 329)

A ‘un-ts’íit há’as
   one-CL:1D banana
   "one banana fruit"

B ‘un-wáal há’as
   one-2D banana
   "one banana leaf"

C ‘un-kúal há’as
   one-CL:PLANTED banana
   "one banana tree"

D ‘un-kúuch há’as
   one-CL:LOAD banana
   "one bunch of bananas"

E ‘um-p’íit há’as
   one-CL:BIT banana
   "one bit of a banana"

(5.9) Bora (cf. Seifart 2007: 418)

A áhí
   banana
   "banana(s) (fruits, plants, etc.)"

B áhí -¿ɑ/-ko
   banana CL:LONG.OBJ/CL:POINTED
   "a banana fruit, a banana plant"
The noun itself carries a class marker in the context of classification (cf. 5.10 A), while not all nouns carry a class marker in all contexts (cf. 5.10 B, where toire’ “children” lacks class marking). Nouns thus seem to be categorised non-obligatorily in Nasioi.


A A-u 
murasini-n-u 
DEM-CL:TABLET medicine-DERIV-CL:TABLET head
siipa-kee-n-u. sick-part-DERIV-CL:TABLET
“This pill is for headaches.”

B Teni toire’ 
tareuri-ma∅-i 
ART.F children care.for-PRS.HAB-DERIV-CL:F
bau‘uri-ma∅-i. feed-PRS.PROGR-DERIV-CL:F
“The lady who cares for the children (is) the one who is feeding them.”

In addition, the Nasioi classificatory markers may amongst others modify nouns (Hurd 1977: 111, cited in Terrill 2002: 293) and thus clearly contribute semantically to their noun phrase. There is no information available on the assigning agent of the markers; yet the size and semantics suggest that the speaker chooses the class marker that he or she deems the most appropriate for the respective context. These properties give Nasioi its classifier-like character.

At the same time, Nasioi seemingly employs multiple class marking and allows its classificatory markers to attach to NP-external elements such as verbs (Terrill 2002: 292). Recall that multiple and NP-external marking go hand in hand within the sample. It is thus not surprising that we find the combination of those two specific properties to be concordial class-like. It is however unexpected to find a system that is classifier-like to such a strong degree to employ multiple marking, which is by definition a property of concordial class systems and by definition not a criterion of classifier systems. 5.10 and 5.11 illustrate multiple class marking in Nasioi.


A Te-varo 
bau-i-na-varo 
ART-CL:PART he.wove-NEAR.PST-DERIV-CL:PART
tamp-a-varo. good-DERIV-CL:PART
“The part that he wove was (a) good (one).”
Bora employs a system with similar properties. The bulk of them shows traits of classifiers; these are presented first, before the concordial class-like properties of multiple marking and the locus operandi are discussed. The Bora class inventory consists of 72 class markers and 53 repeaters (Seifart 2005: 86ff. and 96ff.), which clearly renders it a large one. Table 5.17 provides a selection of the Bora class markers.

### Table 5.17: Selected polysyllabic classifiers of Bora (cf. Seifart 2005: 91ff.)

<table>
<thead>
<tr>
<th>classifier</th>
<th>glossing</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-bohí</td>
<td>POND</td>
<td>&quot;stationary pond of liquid&quot;</td>
</tr>
<tr>
<td>-gwajhko</td>
<td>HOOK</td>
<td>&quot;hook&quot;</td>
</tr>
<tr>
<td>-hpajko</td>
<td>LIQUID</td>
<td>&quot;liquid&quot;</td>
</tr>
<tr>
<td>-htoi</td>
<td>CURVE</td>
<td>&quot;curve&quot;</td>
</tr>
<tr>
<td>-hái:to</td>
<td>PALMLEAF</td>
<td>&quot;palm leaf&quot;</td>
</tr>
<tr>
<td>-kaba</td>
<td>CREEK</td>
<td>&quot;little creek&quot;</td>
</tr>
<tr>
<td>-pahtsí</td>
<td>RING</td>
<td>&quot;roll, ring&quot;</td>
</tr>
<tr>
<td>-ra:ra</td>
<td>BROKEN</td>
<td>&quot;broken, rotten&quot;</td>
</tr>
<tr>
<td>-ro?dá:</td>
<td>VERY.TWISTED</td>
<td>&quot;very uneven and twisted slender objects&quot;</td>
</tr>
<tr>
<td>-tohko</td>
<td>CORNER</td>
<td>&quot;corner, turning point&quot;</td>
</tr>
<tr>
<td>-htsua:to</td>
<td>BUNDLE</td>
<td>&quot;objects (folded up and tied together)&quot;</td>
</tr>
<tr>
<td>-ai:toj</td>
<td>GRAINS</td>
<td>&quot;grains&quot;</td>
</tr>
<tr>
<td>-ai:to</td>
<td>CLUB</td>
<td>&quot;object used for beating&quot;</td>
</tr>
<tr>
<td>-bi:au</td>
<td>SLICE</td>
<td>&quot;slice of big, long, and round objects&quot;</td>
</tr>
<tr>
<td>-fii:ða</td>
<td>PUNCTUATED</td>
<td>&quot;object with one hole in it&quot;</td>
</tr>
<tr>
<td>-ʔahtsí</td>
<td>CLEARING</td>
<td>&quot;clearing&quot;</td>
</tr>
<tr>
<td>-ʔa:mi</td>
<td>LEAF</td>
<td>&quot;thin and flexible&quot;</td>
</tr>
<tr>
<td>-ʔda:pi</td>
<td>PIECE</td>
<td>&quot;small piece&quot;</td>
</tr>
<tr>
<td>-ʔo:ba</td>
<td>GABLE</td>
<td>&quot;gable of a roof&quot;</td>
</tr>
<tr>
<td>-ʔo:gwa</td>
<td>DOORWAY</td>
<td>&quot;opening&quot;</td>
</tr>
<tr>
<td>-ʔaran:u</td>
<td>SHALLOW.HOLE</td>
<td>&quot;shallow and narrow whole&quot;</td>
</tr>
<tr>
<td>-makkei</td>
<td>THIN.PART</td>
<td>&quot;dented part in slender object&quot;</td>
</tr>
<tr>
<td>-tahigwa</td>
<td>VERY.FLAT</td>
<td>&quot;flat objects&quot;</td>
</tr>
<tr>
<td>-ʔo:kue</td>
<td>PROTUBERATION</td>
<td>&quot;round protuberation&quot;</td>
</tr>
</tbody>
</table>

The same noun can be assigned to a variety of classes, even in the same context of classification. Cf. [5.12] (3.18 = 5.12 B), where the same noun and referent are assigned to and marked by of two different classes; as the example shows, the markers also can vary with respect to their distribution within the noun phrase.
(5.12) Bora (cf. Seifart 2005: 169)

A: \( v \cdot hí \) mähau-hí
DISTANT-CL:2D.ROUND be.big.SUB-CL:2D.ROUND
kú:muu-hí
turtle-CL:2D.ROUND
“that big turtle”

B: \( v \cdot hí \) mähau-he
DISTANT-CL:2D.ROUND be.big.SUB-CL:M.SG
turtle-CL:2D.ROUND
“that big turtle”

C: \( aj \cdot di \) mähau-hí
DISTANT-CL:M.SG be.big.SUB-CL:2D.ROUND
turtle-CL:2D.ROUND
“that big turtle”

D: \( aj \cdot di \) mähau-he
DISTANT-CL:M.SG be.big.SUB-CL:M.SG
turtle-CL:2D.ROUND
“that big turtle”

The markers attach to a variety of nominal modifiers (cf. again 5.12) and also to the noun itself, where they function among others as a means of derivation (cf. 2.24 = 5.13).

(5.13) Bora (cf. Weber 2002: 3)

A: muts\#ts\#t\#  -pá
pear.apple  CL:FRUIT
“pear apple fruit”

B: muts\#ts\#t\#  -q\#
pear.apple  CL:TREE
“pear apple tree”

C: muts\#ts\#t\#  -páh\#uí
pear.apple  CL:GROVE
“pear apple grove”

Since the noun is marked for class membership, it is easily observable that nominal classification in Bora is not obligatory. On the one hand, there are ‘non-countable’ nouns such as ūtfí “basket” or pí:ka “manioc”, which generally remain unclassified except for the event of their unitisation by classificatory means (Seifart 2005: 113f.). On the other hand, there are instances of unclassified nouns in discourse (cf. 5.14 where \( aki \) “palm” remains unclassified).

(5.14) Bora (cf. Seifart 2005: 424)

\( (ná:ni) \) g\#w\#á:ðá:ñú:rí  aki
my.uncle cut-CL:M.SG palm
“(My uncle,) he cuts (the) palm tree(s).”
The class markers carry semantics and clearly contribute to their host stem and noun phrase, as 5.9 and 5.13 demonstrate. The assignment is almost fully semantically motivated and allows the speaker to vary class membership and class marking (cf. again 5.12); only “a minority of inanimate” nouns and animal names deviate from this pattern (Seifart 2005: 220).

The Bora system marks a nominal referent’s class membership on a range of nominal satellites, and also on the verb (cf. again 5.14, where the class marker on the verb matches the masculine class of *nda:n* “uncle”). Therefore, Bora marks class membership inside and outside of the classified noun’s phrase. The class markers typically occur more than once within the same construction (cf. again 5.12), which renders the marking as a multiple one on first sight. These two concordial class-like properties are further complemented by the degree of boundedness of the markers, which are suffixes: In the case of the markers being used as a means of nominal derivation, they enter a liaison with a noun, which results in a new noun stem (Seifart 2007: 417f.). In summary, the Bora system acts a lot like a classifier system, which makes use of the formal properties of a concordial class system.

Mundurukú shares the exact same constellation of properties with Nasioi and Bora. Its classifier-like character is presented first, followed by the demonstration of its concordial class-like traits. Mundurukú employs about one hundred or more classificatory markers. Most of them convey body part semantics, which at times are semantically extended; for instance, (*’*a) categorises the front part of the head, but also round things, and *’*a categorises the arm, long objects, and round and rigid items (Gonçalves 1987: 24f.). The exact inventory size cannot be determined, not only since the individual sources vary with respect to the indicated size but also due to the compositional character of the bulk of the markers. Table 5.18 lists a number of obviously composed classifiers, which all carry the classificatory *’*a “mouth, opening” as their initial component and are semantically related.

Table 5.18: Selected classificatory markers of Mundurukú (cf. Gonçalves 1987: 25).

<table>
<thead>
<tr>
<th>marker</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>*’*a</td>
<td>“mouth, opening”</td>
</tr>
<tr>
<td><em>’<em>a</em>di</em></td>
<td>“saliva”</td>
</tr>
<tr>
<td><em>’<em>a</em>do</em></td>
<td>“breath”</td>
</tr>
<tr>
<td><em>’<em>a</em>xe</em></td>
<td>“lips”</td>
</tr>
<tr>
<td><em>’<em>a</em>og</em></td>
<td>“soul, portrait”</td>
</tr>
</tbody>
</table>

Aikhenvald (2000: 355) assumes about 120 or about 100 forms (ib.: 160). Derbyshire and Payne (1990: 261; also based on Gonçalves 1987) count 125 classifiers. The bulk of them (“at least 96”; cf. Gonçalves 1987: 24-29 and Aikhenvald 2000: 355; 98 according to Derbyshire and Payne 1990: 261) is derived from body part terms, 13 forms refer to plants, six forms refer to natural elements, and nine forms refer to “culturally significant items” (Derbyshire and Payne 1990: ib.).
Despite the clear compositional nature of these markers, a general judgement about their status as independent classes or subclasses cannot be made, since a greater number of the involved components is not present elsewhere in the inventory: While for instance $bi^3 di^2$ “saliva” can be traced back to the individual classificatory elements $bi^2$ “mouth” and $di^2$ “water, liquid”, the element -$do^3$ in $bi^2 do^3$ “breath” does not occur independently. For this reason, the common estimations of about one hundred categories are adopted for this study.

There is no information available on the assigning agent of Mundurukú class markers; however, the semantics underlying the class inventory strongly suggest a speaker-sided assignment. The same noun can take different class markers, and these markers contribute semantically to the noun phrase by specifying the referent (cf. 5.15). Note that the semantics of the Mundurukú class markers in general is highly specific, which does not suggest a greater productive classificatory variability, if not for the purpose of semantic alternation.

(5.15) Mundurukú (cf. Derbyshire and Payne 1990: 261)

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>kape</em>-di</td>
<td><em>kape</em>-da</td>
</tr>
<tr>
<td>coffee-CL:LIQUID</td>
<td>coffee-CL:SEED</td>
</tr>
<tr>
<td>“coffee (drink)”</td>
<td>“coffee (seed)”</td>
</tr>
</tbody>
</table>

The markers attach to a range of nominal modifiers and the noun itself; this allows an identification of unclassified nouns (cf. 170, where *bekitkit* “children” remains uncategorised) and the conclusion that nominal classification in Mundurukú is non-obligatory.

(5.16) Mundurukú (cf. Derbyshire and Payne 1990: 261)

<table>
<thead>
<tr>
<th><em>bekitkit</em></th>
<th>ako-ba</th>
<th><em>3'-su-ba-dobuzik</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>child</td>
<td>banana-CL:LONG</td>
<td>3-REF-CL:LONG-find</td>
</tr>
<tr>
<td></td>
<td>“The child found the banana.”</td>
<td></td>
</tr>
</tbody>
</table>

Despite these strong similarities to classifier systems, Mundurukú employs class marking across a range of nominal satellites, namely the numeral, the adjective, and the demonstrative (cf. 167).

(5.17) Mundurukú (cf. Derbyshire and Payne 1990: 260f.)

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>*xepxep-*ã</td>
<td>*wezik-*ã</td>
</tr>
<tr>
<td>two-CL:ROUND</td>
<td>potato-CL:ROUND</td>
</tr>
<tr>
<td>“two potatoes”</td>
<td>“red (stick, pencil, etc.)”</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>wu-su-men-na</em></td>
<td>FOR#ATIVE-CL:STICK-red-ADJ</td>
</tr>
<tr>
<td>“red (stick, pencil, etc.)”</td>
<td></td>
</tr>
</tbody>
</table>
Class marking on nominal satellites seems to occur in the form of multiple marking in that modifying elements carry a marker that corresponds to the head noun’s class (cf. 5.18); the consulted sources however did not feature examples with more than one nominal modifier at a time.


As demonstrated in [5.18] A, the markers also attach to the verb and thus occur NP-externally, which in combination with the multiple class marking gives the system the formal appearance of a concordial class system, while its other properties resemble a classifier system.

Ngan’gityemerri differs in one aspect from the other hybrid systems: It is the only systemic hybrid that employs a small class inventory of fifteen marked classes and one unmarked class, which results in a total of three concordial class-like and four classifier-like properties. The sixteen classes of Ngan’gityemerri comprise a number of general and specific classes, the semantic basis of which is comparable to those of other Australian languages: It contains a class for vegetable food, animals, spears, etc. Cf. table [113] for the class inventory and its markers. Note that the unmarked sixteenth class is not listed in table [113] it comprises “mostly terms for natural objects such as sun, rocks, ground, stars, clouds etc.” (cf. Reid 1997: 172).
Table 5.19: Class marking in Ngan’gityemerri (cf. Reid 1997: 173).

<table>
<thead>
<tr>
<th>Class</th>
<th>Head Marking</th>
<th>AGR Marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Free</td>
<td>Dependent</td>
</tr>
<tr>
<td>1 male</td>
<td>(mipurr)</td>
<td>wa= ~ ∅-</td>
</tr>
<tr>
<td>2 female</td>
<td>(falms)</td>
<td>wur= ~ ∅-</td>
</tr>
<tr>
<td>3 group</td>
<td>awa=</td>
<td></td>
</tr>
<tr>
<td>4 animal</td>
<td>gagu</td>
<td>a- ~ ∅-</td>
</tr>
<tr>
<td>5 vegetable</td>
<td>mugi</td>
<td>mi- ~ ∅-</td>
</tr>
<tr>
<td>6 body parts</td>
<td>da- ~ ∅-</td>
<td>a- ~ a-</td>
</tr>
<tr>
<td>7 canines</td>
<td>wu-</td>
<td></td>
</tr>
<tr>
<td>8 trees/things</td>
<td>yawurr</td>
<td>yerr= ~ ∅-</td>
</tr>
<tr>
<td>9 bamboo spears</td>
<td>yawul</td>
<td>∅-</td>
</tr>
<tr>
<td>10 fire</td>
<td>yenggi</td>
<td>yenggi ~ ∅</td>
</tr>
<tr>
<td>11 strikers</td>
<td>syiri</td>
<td>syiri ~ ∅</td>
</tr>
<tr>
<td>12 canegrass spears</td>
<td>kurum</td>
<td>kurum ~ ∅</td>
</tr>
<tr>
<td>13 drinks</td>
<td>kuru</td>
<td>kuru ~ ∅</td>
</tr>
<tr>
<td>14 womerases</td>
<td>tyin</td>
<td>tyin ~ ∅</td>
</tr>
<tr>
<td>15 digging sticks</td>
<td>kini</td>
<td>kini ~ ∅</td>
</tr>
</tbody>
</table>

The system’s nature and semanticity is comparable to the other hybrid systems. In the following, its classifier-like characteristics are discussed first and the concordial class-like characteristics second. The same noun can be assigned to variable classes, which has an influence on the noun semantics, as illustrates.

(5.19) Ngan’gitymerri (cf. Reid 1997: 178)

A syiri yawurr stick “stick for fighting”
B kini yawurr DIGGING stick “stick for digging”
C yenggi yawurr FIRE stick “stick for burning”

Nouns may also take different markers in order to distinguish different entities (cf. 5.20). These distinctions are lexicalised, yet speakers typically have an association of the respective class assignment. However, these explanations vary between speakers. For instance, the class pairing in 5.20 is motivated by the similarity of the fruits of the red plum tree and the fresh water crocodile by one speaker, while another speaker emphasises the similarity of the tree’s bark and the crocodile’s skin (Reid 1997: 200). The most common variations are between the animal, vegetable, and tree category (Reid 1997: 201); this phenomenon is fairly frequent in Ngan’gitymerri.
5.4. Hybrid Systems

(5.20) Ngan’gityemerri (cf. Reid 1997: 198)

A  a-
  △ANIMAL  weemisye
  “fresh water crocodile”
B  mi-
  △VEGETABLE  werrmisye
  “red plum”

The classification is not obligatory in that there are nouns that are not overtly marked for a given class and at the same time do not trigger class marking on other clausal elements.\[12\] In addition, the Ngan’gityemerri system employs both free and bound class markers, as table 113 shows: the morphosyntactically independent markers (syiri “strikers”, kurum “canegrass spears”, yenggi “fire”, kuru “drinks”, tyin “woomeras”, and kini “digging sticks”) occur optionally (cf. 5.21 and Reid 1997: 177).

(5.21) Ngan’gityemerri (cf. Reid 1997: 167)

(yawurr)  ganbi  kide
  TREE  didgeridoo  where
  “Where is the didgeridoo?”

Despite its high semanticity, the freedom of class assignment appears to be rather limited; the frequent ‘crossclassed’ nouns seem to be lexicalised in many cases (which is why different speakers came up with different folk etymologies for certain terms; cf. Reid 1997: 200), and the class semantics are fairly specific. The most ‘liberal’ and speaker-sided assignment possibilities appear to be situated in the realm of metaphor (cf. 184), while the other class variations are lexicalised and/or related to other circumstances such as the discourse-active linguistic register (cf. Reid 1997: 200f.).

(5.22) Ngan’gityemerri (cf. Reid 1997: 201)

ngatya  ngkin  fi-npuwu-pe  ya  mi=  dityunggurr-wurr
  father  shit  puff-do:1SG.SUBJ-FUT  hey  VEG  short-UNSATISFACTORY
  “Hey daddy, let me have a puff of that shit... Oh, it’s too short now.”

However, it is common for all types of nominal classification systems to have conventionalised the class membership of a part of the nominal lexicon, as is shown in the discussion of class assignment in the sample (cf. pages 216f.). As the assignment principles in the remaining cases are exclusively semantic, the primary assigning agent in the Ngan’gityemerri system is the speaker. The

\[12\] Only nouns that fulfil both these conditions constitute the unmarked residue class, since each class contains a few nouns which are not marked themselves, while they do trigger class marking on other clausal elements (Reid 1997: 173).
semantic assignment is bound to the systemic inventory’s semantic structure, like e.g. in the case of the Jakaltek physical classifiers.

This brings us to the concordial class-like character of the Ngan’gityemerri system, which is represented by its supposed multiple marking and the locus operandi. Ngan’gityemerri allows its class markers to attach to verbs and even whole clauses. Mostly, this serves the creation of a functional equivalent to a relative clause (cf. Reid 1997: 203), but there are also instances where a class proclitic is attached to a verb while referring to one of its arguments (Reid 1997: 210). In this usage, they serve as carriers of anaphora for the purpose of reference tracking. Since markers in this function do not necessarily categorise one argument of the verb as a whole, but only a part of it (cf. 5.23), these instances can be analysed as an instance of NP-external categorisation of a referent.

(5.23) Ngan’gityemerri (cf. Reid 1997: 210)

```
ticket kinya ngninmem pay’m-meny watypela-ninggi wa=
ticket this already pay-do:3SG.SUBJ whitefella-AGENT M
[wunu-pefi Melbourne ngadde-nine-tye]
there-THITHER Melbourne go:1DL.EX.SUBJ-trial-PST
```

“This plane ticket has already been paid for by that whitefella, that one [(of) the three of us] who went there to Melbourne.”

The Ngan’gityemerri marking is exceptional in two respects: First, it employs not only bound markers, but also morphosyntactically independent ones, which is a property that is found in classifier systems such as Mandarin (cf. 147 = 5.24).

(5.24) Mandarin; cf. Sun 2006: 159f.)

```
A nà yì-pí hêí mā
DEM:REMOTE one-CL:HORSE black horse
“That black horse.”

B na pí hêí mā
DEM:REMOTE CL:HORSE black horse
“That black horse.”

C māi pí mā
buy CL:HORSE black horse
“Buy one black horse.”
```

Ngan’gityemerri goes one step further than these systems in that it expresses both the bound and independent markers in the form of multiple marking, and not only once and overtly. 5.25 illustrates multiple marking by means of bound markers, and 182 multiple marking by means of independent markers.
(5.25) Ngan’gityemerri (cf. Reid 1997: 169)

A wa= yedi wa= munu wa= lenger
  M man M that M bad
  “that bad man”

B wa- yedi wunu wa- lenger
  M man that M bad
  “That man is bad.”

(5.26) Ngan’gityemerri (cf. Reid 1997: 177)

(syiri) magulfu (syiri) marrgu
strike cylindrical.fighting.stick strike new
dem-wurity-dim
AUX:3SG.SUBJ-make:2SG.SUBJ
“He is making a new cylindrical fighting stick.”

The second remarkable trait of Ngan’gityemerri’s class marking is its optionality: As (5.25) illustrates, class marking by independent markers is optional in all cases (Reid 1997: 177). In addition, class marking on the targets in general “is not strictly obligatory. The tendency is that all targets show agreement, or none of them do” (Reid 1997: 168), which includes the bound forms (cf. 5.27). Note however that the head noun is typically marked for its class, and that the optionality affects only the occurrence of the markers on non-nouns.

(5.27) Ngan’gityemerri (cf. Reid 1997: 174)

A wa= tyerrmusye (wa=) marrisyarra perrety- meny
  M old.man M blind die do:3SG.SUBJ
  “The old blind man has died.”

B wur= wedimuy (wur=) ngay ngusyak- mem
  F child F mine sick do:3SG.SUBJ
  “My little girl is sick.”

C awa= purrupurrk (awa=) ngurra ngunu
  COLL little’uns COLL your how.about
  “How about your mob of kids...?”

This indicates a low degree of grammaticalisation of the Ngan’gityemerri system among the hybrid systems, since concordial class marking ‘proper’ demands regular marking of non-nominal elements without exceptions, not an optional one. The Ngan’gityemerri system might represent an instance of incipient multiple marking; however, section 6.3 reanalyses the marking of all ‘hybrids’ as not multiple.
**Discussion** Hybrid systems such as those discussed above do not fit the established dichotomy of classifier systems and concordial class systems, since they combine core properties of both: Supposed multiple marking, which is a defining property of concordial class systems and a defining non-property of classifier systems, and a high degree of semanticity and classificational variability, which is a common trait in classifier systems. Note that a high degree of semanticity does not exclude a system from being a concordial class system, yet concordial class systems are reputed for their rigid class membership. It is thus the combination of a semantically productive and variable assignment and contribution to the noun phrase, complemented by a large class inventory, which attributes these systems their strong classifier-like character, while their formal expression by means of multiply occurring markers rules out an analysis of these systems as classifier systems on first sight. None of these systems can be proven to actually undergo a shifting process from a classifier to a concordial class systems, and they might very well represent a stable state. This typological conflict calls for a revision of the commonly assumed basic systemic types; an attempt for a more elaborated account for nominal classification devices is provided in the next chapter. The next section discusses the grey area of intermediate systems and its composition.

### 5.5 The Grey Area

After having treated the properties of the sample’s classifier systems, concordial class systems, and hybrid systems, it is in order to briefly turn towards the part of the continuum of nominal classification which comprises the grey area. The grey area is different from the divergent islands, which comprise the deviations from the observable grammaticalisation cline in the sample: Recall that the nominal classification continuum is determined by two poles, one corresponding to the fairly ungrammaticalised prototypical classifier system, the other corresponding to the fairly grammaticalised prototypical concordial class system. Theoretically, all systems that do not feature a fully prototypical setup of properties fall under the scope of the grey area, which thus is larger than the two divergent islands. What can the sample tell us about the not further defined connection between the prototypical classifier and concordial class systems? The respective sample systems are marked in table 5.20.

<table>
<thead>
<tr>
<th>Korean</th>
<th>Mandan</th>
<th>Ngaat</th>
<th>Mwe &amp; Hajasal</th>
<th>Nunivak</th>
<th>Inupiak</th>
<th>Magdallan</th>
<th>Tlingit</th>
<th>Dongo-Ko (CLS)</th>
<th>Tohono O’odham</th>
<th>Hmong</th>
<th>Nùng</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 5.20: The grey area in its widest definition.

---

The previous sections have shown that there is not much absolute correspondence to the prototypical systems in the sample, but that most systems deviate from either prototype in at least one property (cf. again table 5.1 on page 120). Nevertheless, the vast majority of the deviating sample systems can be attributed either the state of a classifier system or a concordial class system without problems; it seems thus not helpful to define the grey area as all systems that deviate in one or more properties from one of the two systemic prototypes, but to find a more fine-grained approach. Considering the concept of the nominal classification continuum, which can be projected onto an axis of grammaticalisation, two central questions are most promising for a useful delimitation of the grey area:

1. Up to which point in the continuum can a classifier system be considered as such?

2. From which point on can a concordial class system be considered as such?

Both questions consider the potential pathway of a classifier system down the grammaticalisation scale into the direction of a concordial class system; since the point of origin of this system is a classifier system and the logical endpoint is a concordial class system, there must be a point in the changing process at which the system ‘switches’ from classifiers to concordial classes. In order to identify this point, it is useful to recapitulate the properties and systemic traits that are characteristic for the respective prototype. As discussed in section 6.1, the inventory size is not a useful indicator for a system’s type, since both classifier and concordial class systems tend to employ small class inventories of less than two dozen classes (cf. again table 6.4 on page 181; note however that none of the sample’s concordial class systems employs a large inventory). The same is true for the semantic contribution of the markers to the noun phrase, since both classifier and concordial class systems may or may not add specific semantic information on the nominal referent (cf. again table 6.19 on page 223). The remaining properties represent more reliable diagnostics; the above questions are thus addressed from their perspective in the following.

What characterises the classifier systems in the sample? First and foremost, it is the absence of obligatory classification, multiple marking, and NP-external marking. This means that nouns may remain uncategorised in a classifier language, and that the classifiers occur once and overtly in case of categorisation.

14Recall that classifiers that are used as carriers of anaphora in discourse occur outside of
Despite the sample languages not featuring their classifiers outside of the NP borders, the single and overt occurrence appears more important than the restriction to the noun phrase itself — it is easy to think of e.g. a multiple classifier system that allows its classifiers to attach to an element outside of the noun phrase, which is less a violation of classifier principles than e.g. a system that features the same classifier more than once within a context of classification. In general, the most salient difference between the subtypes of classifier systems is their respective morphosyntactic host, while their remaining properties are fairly similar; an NP-external classifier host is thus easily conceivable, as long as the classifier occurs once and overtly and categorises a noun or its referent.

As their name indicates, concordial class systems are characterised by the occurrence of concordial markers on a range of nominal satellites, which can include NP-external elements. Furthermore, concordial class systems classify the nominal lexicon obligatorily and exhaustively. Multiple marking is found in the bulk of concordial class systems in the sample, yet there are two exceptions (Chukchi and Dyirbal) which break this pattern. Beyond this, the multiple marking in the hybrid systems displays several peculiar properties, which is why multiple marking is more closely inspected in section 6.3.

The two other properties, flexibility and assignment, are less unambiguous for the determination of nominal classification systems, but nevertheless can play an active role in it. Not all classifier sets allow for classificatory variation, for instance because their categories’ underlying semantics are mutually exclusive (e.g. in the Jakaltek physical classifiers or the Tlingit numeral classifiers). If however the semantics are not mutually exclusive, classifier variation is virtually always met, for instance in the variation between a general and a specific classifier, but also between individual specific categories. The general availability of classificatory variation within a system, or more specifically the possibility for each noun (and not just a fraction of the nominal lexicon) to change between two or more classes without changing the nominal referent, is thus a strong indication for the presence of a classifier system. A system’s assignment does not necessarily tell us much about this system’s type, but it can tell us when a system can be excluded from the status as a classifier system: Virtually all nominal classification systems have a semantic core or niche in their assignment rules, and nearly all of them have lexicalised noun-class pairs. If a system employs semantic assignment and lexicalised class membership, it can be either a classifier or a concordial class system; it is however only concordial class systems that can employ a morphological or phonological class assignment in addition to the semantic one. If a system thus possesses other assignment mechanisms than semantic and conventionalised ones, it is to be analysed as a concordial class system.

Like with classifiers, a concordial class system’s classificatory and assignment rule types are not as expressive as identification criteria as its obligatoriness their original head noun’s phrase, but nevertheless occur within the borders of a noun phrase that would otherwise host the categorised noun.
and agreement. While the prototypical concordial class system assigns a noun to one class and one class only, many concordial class systems reserve a certain flexibility for at least a part of the nominal lexicon, e.g. for the distinction of plants and their fruits (e.g. Burushaski; Berger 1998: 36), concrete and mass nouns (e.g. Burushaski; Berger 1998: 36ff. and Munshi 2006: 163), or the biological sex of the nominal referent (e.g. Ingush; Nichols 2011: 145). If a system thus is largely rigid in its nouns’ class membership and only allows a smaller fraction of the nominal lexicon to productively shift in class, this may indicate the presence of a concordial class system. This is however not a hard criterion for the identification of a system’s systemic type; recall e.g. the Maasai system, where practically every noun can shift in class due to a size and strength connotation with the respective concordial classes. A system’s assignment rules are a more reliable indicator for a given system’s type: While all concordial class systems know semantic assignment rules and most of them lexicalised class-noun pairs, like classifier systems do, there is no classifier system in the sample and beyond that employed morphological and/or phonological assignment rules (e.g. Moro, where a noun’s phonological shape determines its class membership; Gibbard et al. 2009: 110f.). If a system thus displays morphological or phonological assignment, it is to be analysed as a concordial class system.

There are systems that do not correspond well with both the definition of a classifier system and a concordial class system; especially the ‘hybrid’ systems (cf. section 5.4) elude a clear identification as either type. The sample data suggests a further refinement of the typology of nominal classification, since classifier and concordial class systems are not sufficiently capturing the ‘real life situation’ in their current understanding. The hybrid systems in their synchronic ‘real life’ state also do not conform to their expected outcomes: They are argued to constitute transitional systems, i.e. systems at an intermediate stage of a shift from classifier to concordial class systems, yet they are among the semantically most productive systems of the whole sample (cf. again table 5.1 on page 120). They are thus not necessarily diachronic ‘hinges’, but rather typological ones in that they combine abovementioned ‘core properties’ of both prototypical classifier and concordial class systems. These core properties are largely mutually exclusive in the prototypes’ definition. Therefore, a more elaborated typological analysis of these systems and their properties is carried out in chapter 6.

Next to the ‘hybrid’ systems, there is one other phenomenon that does not fit the established notion and definition of nominal classification devices: Fully optional systems, i.e. systems that have the option to classify their nouns, but do not have to classify them at all. There is only one such case in the sample, the numeral classifiers of Hungarian. Recall that classifier systems in general are non-obligatory in that not every noun is categorised in every context. Certain contexts however, like quantificational ones in numeral classifier languages, demand an obligatory classification. Hungarian however leaves the use of nominal classification up to the speaker (cf. 5.1 = 5.28).
(5.28) Hungarian (cf. Csirmaz and Dékány 2010a: 12)

A hét (fej) saláta
seven CL:HEAD lettuce
“seven lettuces”

B hét (szem) cukor
seven CL:EYE candy
“seven pieces of candy”

C hét (szál) gyertya
seven CL:THREAD candle
“seven candles”

This closely resembles the numeral classifiers of Persian (ISO 639-3: pes; Indo-Iranian), which occur only in quantificational contexts, where they are optional (cf. 5.29).

(5.29) Persian (cf. Mahootian 1997: 195)

A do (-ta) pesér -a
two CL:ITEM boy PL
“[the] two boys”

B cel (-ra’s) gusfænd
twenty CL:HEAD sheep
“twenty sheep”

Due to their general optionality and a few other observations, both the Hungarian and the Persian system seem to be not a part of the grey area inbetween classifier and concordial class systems, but rather to represent incipient systems. For instance, the Hungarian numeral classifier system displays a very low degree of grammaticalisation: Its classifiers take the same form as their source nouns and they occur optionally. There are semantically more or less fixed combinations of specific classifiers and nouns that they categorise. Only several classificatory elements such as szál “thread” and szem “eye” have a more general semantic scope and are complemented by a general classifier darab “piece”. The Persian classifiers evolved from count nouns, which underwent phonological weakening and cliticisation. They appear to have been developing to clitics only recently, since several grammars describe them as unbound count nouns (cf. Rastorgueva 1964: 32 and Mace 2003: 170f.), whereas Mahootian (1997: 195), one of the more recent reference grammars, describes them as clitics and clitics only. A comparison with older Persian dictionaries (Tucker 1850, Stein- gass 1892) shows that these count nouns are used as independent lexical nouns. Both systems possess a smallish inventory of 18 classes (Hungarian) and five classes (Persian).

There is one more sample-external so-called concordial class system, which also classifies its nouns optionally. In Awtuw (ISO 639-3: kmn; Sepik), object suffixes and third person singular pronouns (demonstratives, personal pronouns, and obliviative pronouns) can be marked for the feminine class, if the
referent is female; in all other cases, the noun is marked for the non-feminine class (henceforth “masculine”; cf. table 5.21 and Feldman 1986: 41, 44f., and 107f.).

Table 5.21: Class markers in Awtuw (cf. Feldman 1986: 41, 44f. and 108).

<table>
<thead>
<tr>
<th>target</th>
<th>[+female]</th>
<th>[-female]</th>
</tr>
</thead>
<tbody>
<tr>
<td>personal pronoun</td>
<td>tey</td>
<td>rey</td>
</tr>
<tr>
<td>demonstrative pronoun I (close)</td>
<td>(t)ader</td>
<td>(t)ader</td>
</tr>
<tr>
<td>demonstrative pronoun II (remote)</td>
<td>(t)opor</td>
<td>(t)opor</td>
</tr>
<tr>
<td>obliviative pronoun object</td>
<td>menetey</td>
<td>menerey/meneney</td>
</tr>
<tr>
<td></td>
<td>-te</td>
<td>-re</td>
</tr>
</tbody>
</table>

(5.30) Awtuw (cf. Feldman 1986: 44 and 109)

A rey piyan Kampo-re d-æl-i
3SG.M dog Kampo-OBJ.M FACTIVE-bite-PST
“The dog bit Kampo.”

B Yowmen rey-rey-ke nemet-te du-puy-e
Yowmen 3SG.M-3SG.M-POSS mother-OBJ.F FACTIVE-hit-PST
“Yomen hit his own mother.”

However, feminine class marking for females is optional, just as class marking in general is: Lower empathy objects are always optionally marked (Feldman p.c.). Cf. 5.31 A, where the object marker does not indicate the noun’s class. In B, there is mixed marking on the object suffix (non-feminine marker) and the pronominal form (feminine marker). In C, the non-feminine object marker occurs optionally.

(5.31) Awtuw; cf. Feldman 1986: 44 and 110)

A eywe Eliw menetey -e do -k -o
ancestor Eliw OBLIVIATIVE OBJ FACTIVE get PST
“Ancestor Eliw married what’s-her-name.”

B tey tale-re yaw d-æl-i
3SG.F woman-OBJ.M pig FACTIVE-bite-PST
“The pig bit the woman.”

C tey tale yaw(-re) d-æl-i
3SG.F woman pig-OBJ.M FACTIVE-bite-PST
“The woman bit the pig.”

Like in Hungarian and Persian, the optionality of the Awtuw class marking suggests that the Awtuw system is at an incipient stage and in need of further formalisation, before it can be argued to represent a nominal classification system ‘proper’. Based on these three cases, fully optional systems are not

15The possessive marker -ke is indexed as an object marker in Feldman (1986: 44).
analysed as intermediate systems of nominal classification, but as emerging ones.

Summing up, the grey area does not comprise all instances of deviating systems (i.e. all systems that deviate in at least one property from their prototype), but can be narrowed down on the basis of a number of core characteristics. This better understanding of essential properties of both classifier and concordial class systems provides us with a ‘filter’, which enables the researcher to decide whether a system indeed is a classifier system, a concordial class system, or neither. The latter case occurs in the sample in form of the hybrid systems, which seem to represent a systemic type of their own, precisely because they combine characteristic traits of both other systemic types, which are mutually exclusive within classifiers and concordial classes. Chapter 6 presents a refined analysis of the systemic hybrids, which unravels their underlying classifier-like nature. It has further been argued that fully optional systems are more likely to be developing into nominal classification systems and thus are more likely to be in a state of entering the nominal classification continuum rather than shifting within its boundaries.

In the beginning of this subsection, the coverage of the grey area within the sample was presented in its most exhaustive form, i.e. comprising all systems that deviate in one or more properties from either prototype. After the consideration of the individual properties and their role towards the grey area and influence on it, the grey area can be presented in a revised form. Two kinds of contributing systems can be distinguished: The ‘core criteria’ are a system’s obligatoriness, its way of marking, and its locus operandi. All classifier systems and concordial class systems (besides Chukchi (17) and Dyirbal (18)) show a homogeneous satisfaction of those three criteria (cf. table 5.22). This means that a system with mixed classifier-like and concordial class-like properties amongst those three may be assigned to the grey area. As table 5.22 shows, this affects only the ‘hybrid’ systems (3–6), Chukchi (17), Dyirbal (18), and Tainae (19) within the sample. These systems cannot be clearly assigned the status of either a classifier or concordial class system and thus represent instances of the grey area that I label “first order deviations”.

<table>
<thead>
<tr>
<th>Korean</th>
<th>Mandarin</th>
<th>Nasioi</th>
<th>Bora</th>
<th>Mundurukú</th>
<th>Ngan’gityemerri</th>
<th>Boumaa Fijian</th>
<th>Hungarian</th>
<th>Great Andamanese</th>
<th>Jakaltek</th>
<th>North Ambrym</th>
<th>Tlingit</th>
<th>Dongo-Ko (CLS)</th>
<th>Tohono O’odham</th>
<th>Hmong Daw</th>
<th>Nùng</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>F3</td>
<td>F4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.22: The form-based properties of the language sample.
Furthermore, a system’s inventory and semantic contribution to the noun phrase are no reliable diagnostics for the identification of that system’s type. Its classificatory flexibility and assignment however may provide information about its type, depending on the individual case. A ‘second order deviation’ is thus met if a given system’s ‘core criteria’ (obligatoriness, marking, and locus operandi) are either classifier-like or concordial class-like, but a system’s flexibility and/or assignment are not of the same nature. Finally, ‘third order deviations’ are those systems that display a general either classifier-like or concordial class-like setup of properties, which may be complimented by properties that are both classifier-like and concordial class-like. According to these criteria of first and second order deviations, the revised grey area of the sample shrinks down to the systems marked in table 5.23.

Table 5.23: The grey area.

Furthermore, a system’s inventory and semantic contribution to the noun phrase are no reliable diagnostics for the identification of that system’s type. Its classificatory flexibility and assignment however may provide information about its type, depending on the individual case. A ‘second order deviation’ is thus met if a given system’s ‘core criteria’ (obligatoriness, marking, and locus operandi) are either classifier-like or concordial class-like, but a system’s flexibility and/or assignment are not of the same nature. Finally, ‘third order deviations’ are those systems that display a general either classifier-like or concordial class-like setup of properties, which may be complimented by properties that are both classifier-like and concordial class-like. According to these criteria of first and second order deviations, the revised grey area of the sample shrinks down to the systems marked in table 5.23.

Table 5.23: The grey area.

5.6 Discussion

The sample clearly shows that ‘real life’ concordial class systems correspond well to their definition on the one hand and to the expectations of the grammaticalisation hypothesis on the other hand. All concordial class systems in the sample employ a smallish inventory and an obligatory classification, and all except for the systems of Chukchi (17), Dyirbal (18), and Tainae (19) allow a multiple classificatory marking in- and outside of the NP borders. While the bulk of concordial class systems knows a rigid class membership, a number of systems (Chukchi (17), Dyirbal (18), Burushaski (20), Ingush (28), Chichewa
(29), Hadza (31), and Hausa (32)) provides a certain flexibility in class membership for a delimited part of the nominal lexicon. Only Tainae (19) and Maasai (21) are generally flexible: Tainae nouns can be marked for either a specific or a residue class, depending on the specificity of the nominal referent in discourse; the Maasai classes have a size and strength connotation, which allows the speaker to shift nouns from one class to another. The pattern of classifier-like and concordial class-like properties shows a clear tendency towards an order of grammaticalised properties and thus corresponds to a grammaticalisation cline as it is predicted by the grammaticalisation hypothesis. Yet there are a few systems that call for a conclusive evaluation of their status as a nominal classification device either due to their suspiciously classifier-like property setup (Chukchi (17), Dyirbal, (18), and Tainae (19)) or due to their seeming secondary nature as a system of nominal classification (Kiowa (23)). These are discussed individually in the following.

Recall that Chukchi is commonly not considered to employ a system of nominal classification (cf. Aikhenvald 2000: 438f. and Weinstein 2010: 22 and 38); its system does however meet the definition of a nominal classification system in that it classifies nouns and their referents on the basis of their semantics and by means of overt markers and was therefore incorporated in the sample. The Chukchi classes are distinguished by a number of different case markers and a singular-plural distinction, which is only fully met in class I (Bogoras 1922: 694, Comrie 1981: 246f., Dunn 1999: 65, and Weinstein 2010: 41). This way of class distinction is fairly common in concordial class system. For instance, many Indo-European languages do not know distinct class markers, but specific case marking paradigms for each class: Cf. table 38 for the paradigms of the article in German; note the lack of class distinction in the plural and the indexed numbers for syncretic forms across classes.

Table 5.24: Definite and indefinite articles in German.

<table>
<thead>
<tr>
<th>case</th>
<th>M_def</th>
<th>F_def</th>
<th>N_def</th>
<th>PL_def</th>
<th>M_indef</th>
<th>F_indef</th>
<th>N_indef</th>
<th>PL_indef</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>der²</td>
<td>die⁵</td>
<td>das</td>
<td>die⁵</td>
<td>ein⁶</td>
<td>eine</td>
<td>ein⁶</td>
<td>Φ</td>
</tr>
<tr>
<td>GEN</td>
<td>des²</td>
<td>der¹</td>
<td>des²</td>
<td>der¹</td>
<td>eines⁵</td>
<td>eimer</td>
<td>eines³</td>
<td>Φ</td>
</tr>
<tr>
<td>DAT</td>
<td>dem¹</td>
<td>der¹</td>
<td>dem³</td>
<td>den⁴</td>
<td>einem⁷</td>
<td>eimer</td>
<td>einem⁸</td>
<td>Φ</td>
</tr>
<tr>
<td>ACC</td>
<td>den¹</td>
<td>die⁵</td>
<td>das</td>
<td>die⁵</td>
<td>eimen</td>
<td>eimer</td>
<td>eim⁵</td>
<td>Φ</td>
</tr>
</tbody>
</table>

In Chukchi, class I hosts proper names of humans and animals, certain kin terms, and optionally other nouns denoting humans (Dunn 1999: 64). Its markers for all differing cases (locative, ergative, and dative/allative) take the form -ne and/or -na; their corresponding forms for class II lack a nasal (e.g. locative -k and ergative -e/-a; cf. Bogoras 1922: 719f., Comrie 1981: 246f., and Weinstein 2010: 38 and 40f.). The only parts of speech taking those affixes are nouns and pronouns, and no other nominal satellites show class marking (cf.
5.6. Discussion
the respective sections on the nouns, adjectives, determiners, numerals, etc. in Weinstein 2010). This means that Chukchi does not employ agreement, as the noun cannot agree with itself, and pronouns are no nominal modifiers, but carriers of anaphora and/or deixis, which replace nouns in discourse. Pronouns referring to class I nouns infix the class I marker -na (e.g. *gum-na’n* ISG-ERG:CLASS.I ‘I’; Bogoras 1922: 719), which makes the marker -ne/-na the distinctive criterion between the two classes. Considering the lack of agreement, Chukchi cannot be analysed as a concordial class system. Considering the overt class marking of only one of the two classes and the fact that common nouns denoting humans are only optionally assigned to class I, the Chukchi system is better analysed as a ‘single classifier system’ on the basis of its classificatory flexibility and morphosyntactic expression (cf. Lichtenberk’s argument for the existence of suchlike systems on page 134 as reductions from earlier, more sophisticated classification systems).

One last factor has to be considered: The class I marker -ne/-na shows the same shape in the cases where it occurs, whereas class II distinguishes all cases by means of different affixes (cf. Weinstein 2010: 38). At the same time, the definition of concordial class systems allows class membership to be expressed by means of portmanteau morphemes, while classifier systems are defined as marking class membership by means of markers that express only information about a nominal referent’s class (cf. e.g. Dixon 1986: 105). Is Chukchi thus a systemic hybrid in the sense that it employs portmanteau markers, while the morphosyntactic appearance system is of a fairly classifier-like character? Considering that the obligatorily marked class I members yield a very delimited fraction of the nominal lexicon (proper names and certain kin terms), I propose a different analysis: The Chukchi system appears to be rather aged, since it has no productive semantic function beyond the mere assignment. Its distinctive markers are well-incorporated into the case paradigm, which is marked by means of affixes. Affixes are commonly assumed to be the ultimate consequence of a grammaticalisation process. If the marker -ne/-na is of a classificatory origin, the classification at the synchronic state of this analysis is certainly not productive in terms of semantics. This argument combined with the morphosyntactic appearance of the Chukchi system, which only allows its analysis as a classifier system, results in a former system of nominal classification, the classificational markers of which have been reanalysed and integrated into the category of grammatical case (eventually replacing earlier case markers or fusing with them). This analysis renders Chukchi an extinct classifier system, which does not actively serve a function of nominal classification.17

Dyirbal and Tainae are the only sample-internal supposed concordial class systems which employ markers that are not fused with other grammatical cate-

---

16 It is beyond the scope of this study to explore the likeliness of the existence of ‘single classifier systems’ and the theoretical fruitfulness of the assumption of suchlike systems, which is why they are admitted as systems in their own right.

17 Only if Chukchi developed agreement on nominal modifiers, the system could be analysed as having adopted the status of a concordial class system.
Grammaticalising Systems

Gories, but convey class information. Dyirbal diverts from other concordial class systems in that it marks a noun’s class membership only once per classificatory context, namely on the plain, demonstrative, or interrogative form of the so-called noun marker. The assignment is speaker-sided and flexible for a part of the nominal lexicon, which is another property that attributes the system an interactive character that is expected from a classifier system rather than a concordial class system. By marking only one NP-internal constituent for a given noun’s class, Dyirbal’s visual appearance does not deviate from any ordinary classifier system, while it differs from the visual appearance of a ‘proper’ concordial class system. The strongest argument that can be made for the Dyirbal system to be a concordial class one is the mostly rigid class membership of its nouns. This does however not hold for parts of the nominal lexicon such as nouns denoting humans and exceptionally animals, which are assigned to a class on basis of their biological sex (Dixon 1968: 123). The noun classifier system of Jakaltek features a comparable degree of rigidity: Its ‘social set’ of classifiers, which indicates the relation between the (human) speaker and the (human) classified referent, is flexible, while its ‘physical set’ classifiers, which indicates physical properties of the classified, is not subject to classificatory variation.

There is one last factor to be considered for Dyirbal: The sample’s classifier systems do not employ an obligatory classification, while Dyirbal seems to obligatorily classify its nominal lexicon. It is however difficult to assess the obligatoriness of classification in Dyirbal, since the noun marker does not occur in the ablative, allative, and locative case (Dixon 1982: 163). In the absence of an overt marker, it is impossible to decide whether a classification takes place or not. If Dyirbal obligatorily classifies its nouns, this renders it a systemic hybrid and a first order deviation in the grey area. If one takes the absence of classificatory marking in the ablative, allative, and locative case as evidence for the absence of classification, this renders Dyirbal a fairly formalised instance of a classifier system, which is comparable to Jakaltek (cf. also Sands 1995: 274f.) and a second order deviation.

The Tainae system also exposes itself in a fairly classifier-like body: demonstrates that a noun is not rigidly assigned to a single class, as it is commonly defined for concordial class systems (cf. e.g. Aikhenvald 2000: 21, Corbett and Fedden 2015: 9, Dixon 1986: 106). Instead, the class marker is referential in that it adds semantics to the noun that is not present otherwise.

Suggesting that Dyirbal does not employ a concordial class system may seem controversial, as the system has served as a textbook example of an Australian concordial class system for the last couple of decades. Certainly, a much more in-depth study of the systemic semantics and the class assignment would be necessary to come to a permanent conclusion about the status of the Dyirbal system. Classifier and concordial class systems are defined on basis of their morphosyntactic expression, and Dyirbal’s form-based properties suggest that the system is much closer to a prototypical classifier system than it is to a prototypical concordial class system.
5.6. Discussion

(5.32) Tainae (cf. Carlson 1991: 27f.)

A  

**skypw*na**  
leaf-FLAT  
“leaf”

B  

**skypw*wa**  
leaf-CYLINDRICAL  
“leaf (roll)”

C  

**skypw iliti**  
leaf-THIN  
“leaf (strip)”

D  

**skypw ai**  
leaf-LONG  
“(folded) leaf (rod)”

E  

**skypw ipi**  
leaf-FLUID  
“leaf (ball)”

Beyond instances of free classification like in (5.32) every Tainae noun is generally marked by the residue marker -e, if it does not refer to a specified discourse referent (Carlson 1991: 27). Additional classificatory variation is also found in other lexical fields (e.g. **at-om** m. “grandfather” vs. **at-epi** f. “grandmother”; **iky-iwa** (cylindrical) “(fallen) tree” vs. **ik-ai** “(standing) tree”; cf. Carlson 1991: 25). This renders the Tainae classification even more flexible than e.g. the one of Dyirbal, since each noun is assigned to both a specific category and the residue category. Despite Carlson (1991: 23) noting that class marking occurs in copular constructions (and other, NP-internal material), class marking seems to occur only within the borders of the noun phrase: The “basic building block in nominalized verb forms and copular constructions” (Carlson 1991: 35) is the personal pronoun (cf. 5.33), which does not behave different from nominals with respect to its structure.

(5.33) Tainae (cf. Carlson 1991: 36)

> **t-e-mt-k**  
DEM-RESIDUE-tell.him-PST  
epak  
2SG.F  
*han-epak*  
witch-2SG.F  

“He said this [to his wife]: ‘You [are a] witch’.”

Beyond that, nouns may lack class marking, if they cooccur with a modifier that is marked for class. This means that there is a tendency in Tainae to mark an NP once and not multiply (cf. 5.34).

(5.34) Tainae (Carlson 1991: 29)

> **inky-i**  
1SG.M  
*yaik-a-we*  
bird-EPENTHETIC  
hamte-ek**i**  
big-3SG.M-OBJ  
see-1/2SG.STATIVE

“I see the big bird.”
The assignment of Tainae is productive and carried out by the speaker, who is unaware of the underlying semantics of the system (Carlson 1991: 23 and 27). All these properties favour an analysis of the Tainae system as one that is closer to the status of a classifier system than to the one of a concordial class system. The only phenomenon that theoretically argues for it being a concordial class system is the possible multiple occurrence of classificatory markers in the same classificational context; given the referential nature of these markers, this multiple occurrence can be analysed as an instance of multiple classification instead of multiple marking (cf. again 5.33). If one accepts this analysis, the Tainae system is closer to the status of a classifier system.

The Kiowa nominal lexicon is patterned into four or nine classes, depending on the analysis. The Kiowa system meets the definition of a concordial class system by employing a small class inventory, which obligatorily and rigidly patterns the nominal lexicon and can be distinguished by the formal agreement marking of each class (on the noun, the adjective, and the demonstrative). What makes this system peculiar, is that the distinctive marker is the same for all classes, which takes the form \(-go\), and occurs only in the non-inherent number values of the tripartite category number (cf. 5.35).

(5.35) Kiowa (cf. Wonderly et al. 1954: 4f.)

\[
\begin{array}{c}
\text{é\'-\(\cdot\)go} \\
\text{this-INV.NUMBER}
\end{array}
\quad
\begin{array}{c}
\text{\(\bar{a}\)\(\cdot\)go} \\
\text{apple-INV.NUMBER}
\end{array}
\quad
\begin{array}{c}
\text{é\'-de} \\
\text{this-BASIC.NUMBER}
\end{array}
\quad
\begin{array}{c}
\text{\(\bar{a}\)\(\cdot\)} \\
\text{apple}
\end{array}

\text{“this apple/these apples, these (two) apples”}
\]

This renders the supposed Kiowa system of nominal classification a system that actually marks a noun’s number value. Each noun occurs unmarked, if its referent overlaps with its ‘natural’ inherent number value, which can be singular, dual, plural, or any combination of these; if the entity referred to diverts from the noun’s inherent number value, the marker \(-go\) occurs (cf. table 81).
Table 5.25: The number values of the Kiowa concordial classes (cf. Harbour 2003: 551 and Wonderly et al. 1954: 2ff.).

<table>
<thead>
<tr>
<th>class</th>
<th>singular</th>
<th>dual</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>inherent</td>
<td>inherent</td>
<td>marked</td>
</tr>
<tr>
<td></td>
<td><em>cenbó “a cow”</em></td>
<td><em>cenbó “two cows”</em></td>
<td>*cenbó-ŋ <em>“cows”</em></td>
</tr>
<tr>
<td>II</td>
<td>marked</td>
<td>inherent</td>
<td>inherent</td>
</tr>
<tr>
<td></td>
<td>*łłpʰg <em>“a ribbon”</em></td>
<td><em>łłpʰg “two ribbons”</em></td>
<td><em>łłpʰg “ribbons”</em></td>
</tr>
<tr>
<td>III</td>
<td>marked</td>
<td>inherent</td>
<td>marked</td>
</tr>
<tr>
<td></td>
<td><em>ałł-g</em> “one apple”</td>
<td>*ałł “two apples”</td>
<td>*ałł-g “apples”</td>
</tr>
<tr>
<td>IV</td>
<td>inherent</td>
<td>inherent</td>
<td>inherent</td>
</tr>
<tr>
<td></td>
<td><em>hólda “one dress”</em></td>
<td><em>hólda “two dresses”</em></td>
<td><em>hólda “three dresses”</em></td>
</tr>
</tbody>
</table>

The definition of a concordial class system’s formal requirements lacks a specification that different classes are supposed to be distinguished by formally recognisable class markers. Kiowa thus theoretically meets the formal properties of a concordial class system, yet its supposed noun classes are a byproduct of the overt marking of a noun’s inverted inherent number value. It does not serve the categorisation or classification of nouns and/or their referents in the fashion of a nominal classification system, but in the fashion of a quantificational system. Due to this, the Kiowa system is better analysed as a system that marks number and does not serve nominal classification purposes.

The concordial class systems in the sample can thus be reduced by four systems: The Chukchi system is unproductive, the systems of Dyirbal and Tainae are rather classifier systems than concordial class systems, and the Kiowa system is a byproduct of number marking. Table 5.26 presents the revised concordial class systems of the sample.

Table 5.26: The sample’s revised concordial class systems.

<table>
<thead>
<tr>
<th>Burushaski</th>
<th>Masai</th>
<th>Sumerian</th>
<th>Tamil</th>
<th>Dogon-Ko</th>
<th>Maro</th>
<th>Krt</th>
<th>Jirakh</th>
<th>Girgowa</th>
<th>German</th>
<th>Hadza</th>
<th>Hausa</th>
<th>Masji</th>
</tr>
</thead>
<tbody>
<tr>
<td>set size</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>marking</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>locus operandi</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>flexibility</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>assignment</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
</tbody>
</table>

Blank: classifier-like, •: concordial class-like b: both
The hybrid systems (Nasioi (3), Bora (4), Mundurukú (5), and Ngan’git-yemerri (6)) are characterised by an overall classifier-like character (e.g., large inventories and unobligatory classification), while their way of marking class membership closely resembles class agreement. This way, they combine mutually exclusive core properties of both classifier and concordial class systems. A closer inspection of the supposed multiple marking in those systems suggests that the multiple occurrence of classificatory markers in the same context is not an instance of agreement. In agreement systems, semantic information of the head noun is copied to class markers on a range of nominal satellites and may multiply occur. In contrast, each occurrence of a classificatory marker in the ‘hybrid’ languages represents an own context of classification, and the markers’ morphosyntactic host stand in apposition to each other. Cf. section 6.3 for a detailed analysis.

The sample analysis of the classifier systems has shown that there are no salient differences between the individual subtypes of classifiers except for the kind of host they employ. They all categorise a nominal referent with respect to its inherent characteristics, occur once and overtly within the context of classification, do not obligatorily categorise the nominal lexicon, and most often occur in the form of independent or clitic markers. The assignment is typically carried out by the speaker, though a certain degree of conventionalisation or lexicalisation of classifier-noun pairs is present in virtually all systems (cf. e.g. North Ambrym and the higher registers of Mandarin). The distinction of noun, numeral, and possessive classifiers is thus merely indicating the classificatory context of a given system and does not entail systemic behavioural differences except for the locus of the markers’ occurrence. Figure 6.4 presents the revised form of the taxonomy of nominal classification.

Figure 5.1: Revised taxonomy of the systems of nominal classification.

Two classifier systems from the sample deviate notably from the other classifier systems and thus are assessed more in detail in the following. These are the systems of Great Andamanese (9) and Tlingit (12). The system of Great Andamanese is serving not only nominal classification, but also the manipula-

---

19As indicated in chapter 2, it is not possible to make a clear statement about the specific functional setup of systems of nominal classification, since functionality in nominal classification lacks a deeper understanding and theoretical underpinning. The above statements are thus referring to the formal appearance of systems, their classificatory semantics, and the potential of classificatory manipulability by the speaker.
tion of verbal semantics (cf. 5.3 = 5.36 where the verbal semantics are altered by classifier insertion instead of a categorisation of a nominal referent).

(5.36) Great Andamanese (cf. Abbi 2011: 770)

A  
\[ \text{ut} = \text{file} \quad \text{e} = \text{file} \quad \text{ek} = \text{file} \]
\[ \text{CL:4} \text{= aim} \quad \text{CL:5} \text{= aim} \quad \text{OBJ} \text{= aim} \]
“aim from above, aim to pierce, aim at”

B  
\[ \text{ara} = p^b \text{o} \quad \text{er} = p^b \text{o} \quad \text{ut} = p^b \text{o} \]
\[ \text{CL:6} \text{= cut} \quad \text{CL:2} \text{= cut} \quad \text{CL:4} \text{= cut} \]
“cut down/fell, hit with a stick (in the front), separate from the source”

C  
\[ \text{cr} = \text{bate} \quad \text{ek} = \text{bate} \quad \text{ut} = \text{bate} \]
\[ \text{CL:2} \text{= slap} \quad \text{OBJ} \text{= slap} \quad \text{CL:4} \text{= slap} \]
“slap on the face, slap suddenly, slap hard”

This is the reason why it stands out among its fellow classifier systems. The analysis of the Great Andamanese system focused only on the part of the system that serves nominal classification though, in order to guarantee a valid comparability towards the other systems. This separation of operative domains of the Great Andamanese system is possible without greater problems, since it is the morphosyntactic host of the classifier that determines the operative domain. The attachment of classifiers to adverbs and verbs does not fall under scope of the ‘nominal classification part’ of the system. Note that Great Andamanese is the only sample-internal system that employs a functional distinction of the classifiers with respect to the morphosyntactic host. All other systems employ the same use for classifiers, regardless of the element to which they are attached. The ‘nominal classification part’ of the system displays no eye-catching or exceptional features: Despite a fair degree of conventionalised classifier-noun pairs and a comparatively rather small inventory of seven classes, Great Andamanese classifiers occur once and overtly in the context of classification, are typically assigned by the speaker on the semantic basis of the intended referent’s characteristics, occur NP-internally and may vary for the same noun with respect to the intended referent. Further and more in-depth documentation of the system is necessary to determine the interaction between the nominal and the non-nominal domain of the system, and to allow a judgement about its quality as a nominal classification system.

The Tlingit system stands out among the classifier system due to its formal resemblance to a concordial class system. Its inventory comprises a human and a residue class and its assignment is speaker-sided and semantics-based. Its markers occur in the form of the infix \(-na\) for the human class and by zero marking for the residue class. There is no classificatory flexibility possible due to the rigid class semantics. Beyond that, the class markers do not add semantics to the noun phrase, since the marking depends on the referent’s semantics and not vice versa. Yet the system is considered a numeral classifier system by some (cf. Aikhenvald 2000: 121, Dryer and Haspelmath 2013: s.v., Kinkade 2001, and Sherzer 1976: 74), which is due to the fact class membership
is only marked on the numeral. This elicits two problems for its further analysis: First, the classificatory markers occur only in quantifying contexts, which makes a clear statement about the system’s obligatoriness impossible. Second, this study’s definition of agreement objects the occurrence of the markers on only one element.

This closely resembles the Chukchi case in that the semantically unproductive systems distinguish a human and a non-human category, and only mark the human category overtly by an affixed form -na on a minimalistic range of one host (Chukchi) and two hosts (Tlingit, where the marked elements, the noun and the pronoun, do not occur in the same noun phrase) respectively. Even if these systems are genetically unrelated, their assessment delivers the same result in that the Tlingit system is most likely extinct. The system clearly resembles nominal classification, but misses out on employing agreement on a range of nominal modifiers as well as a semantic ‘fine tuning’ or differentiation potential of the nominal referent. It is likely that the numeral system of Tlingit emerged from a classifier system. At its current state however, there is no greater evidence that it serves a function of nominal classification beyond the possible attribution of ‘reference tracking’ due to the employment of agreement of numeral and noun, which is not very convincing as an argument. Even if it is not possible to provide a list of preconditions that must be met by a system in order to be analysed as a system of nominal classification, it does seem neither fruitful nor theoretically helpful to presume that two paradigmatic sets of numerals in a given language suffice to constitute a nominal classification device. Therefore, I analyse the Tlingit ‘system’ as two paradigms of numerals, which derive from an earlier potential classifier system. This earlier system petrified and is not productive anymore.

Table 5.27 provides an overview of the revised classifier systems of the sample; note that the systems of Dyirbal and Tainae have been added. Finally, the reanalysed system of Tlingit drops out of the sample.
5.6. Discussion

Table 5.27: The sample’s revised classifier systems.

<table>
<thead>
<tr>
<th></th>
<th>Korean</th>
<th>Mandarin</th>
<th>Bouma</th>
<th>Fijian</th>
<th>Hungarian</th>
<th>Great Andamanese</th>
<th>Jakaltek</th>
<th>North Ambrym</th>
<th>Dongo-Ko</th>
<th>Tohono O’odham</th>
<th>Hmong Daw</th>
<th>Nùng</th>
<th>Dyirbal</th>
<th>Tainae</th>
</tr>
</thead>
<tbody>
<tr>
<td>set size</td>
<td>1 2 7 8 9 10 11 13 14 15 16 18 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>obligatoriness</td>
<td>• • • • • • • • • • • • • • • •</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>marking</td>
<td>b b b b b b b b b b b b b b b b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>locus operandi</td>
<td>b b b b b b b b b b b b b b b b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flexibility</td>
<td>b b b b b b b b b b b b b b b b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>assignment</td>
<td>b b b b b b b b b b b b b b b b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

blank: classifier-like, •: concordial class-like, b: both

This chapter has investigated the sample from the perspective of the systemic types, i.e. the dispersion of properties within each type. It has shown that the distribution of classifier systems and concordial class systems is as expected: Classifier systems are less grammaticalised than concordial class systems. Beyond this, they form coherent groups, which display a continuous increase in grammaticalised properties. The assumption of a continuum of nominal classification systems, which is based on the systems’ degree of grammaticalisation, is thus reflected in the data. Similarly, the expectations of the grammaticalisation hypothesis are reflected in the sample in that the coherent group of classifiers is at a lower level of grammaticalisation than the coherent group of concordial class systems.

The specific setup of classifiers, hybrid systems, and concordial class systems is evaluated above, and ‘core properties’ are identified for each systemic type. These serve in the exclusion of the systems of Kiowa and Chukchi from the further study, and in the delimitation of the grey area of intermediate systems. Depending on the kind of divergent properties that a system employs, different types of deviations can be found in the grey area: First order deviations, which employ divergent core properties from their systemic type, second order deviations, which employ divergent non-core properties from their systemic type, and third order deviations, which employ properties that correspond to both systemic prototypes.

Chapter 6 investigates the sample from the perspective of the individual properties, i.e. how the sample languages relate to them and whether or not this distribution corresponds to the expectations of the grammaticalisation hypothesis.
This chapter investigates the sample with respect to the system’s individual properties, which mirror a clear grammaticalisation cline. Each property is discussed in a separate subsection. Table 6.1 displays the distribution of properties in the sample after their revision in the previous chapter.

Table 6.1: The form-based properties of the revised language sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasioi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bora</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mundurukú</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ngan’gityemerri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boumaa Fijian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Great Andamanese</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jakaltek</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Ambrym</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Daw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nùng</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: The form-based properties of the revised language sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyirbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tainae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maasai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sumerian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CCS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingush</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chichewa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hausa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

blank: classifier-like, •: concordial class-like, n: both;  
F_1: inventory size, F_2: obligatory, F_3: multiple vs. single marking, F_4: locus operandi,  
F_5: flexibility, F_6: assignment, F_7: semantic contribution
Only six of the sample systems employ a larger class inventory. This prompts that the inventory size cannot represent a defining criterion of classifier-like systems, since only two classifier-like systems employ a large class inventory, in contrast to ten classifier-like systems which employ a small class inventory. The systems of Nasiio (3), Bora (4), Mundurukú (5), and Ngan’gityemerri (6) are ‘hybrid’ systems and thus are discussed separately, although they group together with classifier-like systems in table 6.1. The classificatory obligatoriness on the other hand appears to be a reliable instrument for the distinction of classifier and concordial class systems: All classifier-like systems and the four ‘hybrid’ systems are not employing an obligatory classification for every noun in every morphosyntactic context, whereas all concordial class systems are. The criterion of obligatoriness aligns well with the criterion of set size in that the grammaticalisation cline is not interrupted by exceptions. Multiple marking and the markers’ locus operandi show a strong correlation. Note that supposed multiple marking on nominal satellites in Ngan’gityemerri (6) is optional. All the systems employing multiple marking are hybrid systems or, unsurprisingly, concordial class-like systems, with only two systems that are commonly referred to as concordial class-like systems (Dyirbal (18) and Tainae (19)) not displaying a ‘proper’ multiple marking. The grammaticalisation cline of marking and the locus operandi is only blurred by the four hybrid systems. A marker’s multiple occurrence in the same classificatory context is the identification criterion of the presence of agreement during the collection of the sample data; the data however suggests that the multiple occurrence of the same marker in the same classificatory context is not sufficient as an analytical tool. The notion of agreement is thus revised more in detail in section 6 (page 197ff.). The hybrids’ type of marking is reanalysed there as an actual multiple instance of classification, not multiple marking. The distribution of the abovementioned four properties across the sample leaves no doubt that grammaticalisation plays an important role in the development of nominal classification devices, and suggests the properties to grammaticalise in the following order: reduction of inventory, obligatory classification, introduction of multiple marking, and introduction of class marking across the NP’s borders.

The remaining three properties (flexibility, assignment, and semantic contribution) are those which blur the grammaticalisation cline (cf. again table 6.1), as a greater number of sample systems employs concordial class-like characteristics, where classifier-like ones would fit the grammaticalisation pattern. The classificatory flexibility does so in individual cases (the classifier system of Mandarin (2), and the concordial class system of Dyirbal (18), and Burushaski (20)), but further fits the grammaticalisation cline. Given the frequent emphasis of the inflexibility of a noun’s class membership in common definitions of concordial class systems, it is worth noting that it is actually fairly common for concordial class systems to allow at least a part of the nominal lexicon to vary with respect to their class membership (cf. the systems marked by b in table 6.1). The assignment of nouns to classes is closely related to the flexibility of class membership: There are two deviating systems (the classifier systems
of Jakaltek (10) and North Ambrym (11)) in table 6.1 which deviate from the
grammaticalisation pattern in displaying a (partly) system-sided assignment,
but at the same time a flexible and thus classifier-like assignment. Finally, the
semantic contribution to the noun phrase by classificatory means can be found
in both classifier and concordial class systems. The distribution is fairly mixed,
with a preference of concordial class systems to not semantically contribute to
the noun phrase. There are no obvious correlations between the semantic con-
tribution and the other properties. Since the criterion of semantic contribution
is a rather vague one, its random distribution is not completely surprising.

This chapter evaluates the validity of the individual properties in detail
and aims at an explanation for the issue of the general heterogeneity of clas-
sifier systems as compared to concordial class systems. This heterogeneity is
expressed among others by the deviating islands in table 6.1, which include
almost exclusively classifier and hybrid systems and only the concordial class
system of Dyirbal (18). Classifier systems constitute a wider and less clear-cut
phenomenon than concordial class systems. This also means that they cover
a bigger range of grammaticalisation, including hardly grammaticalised and
fairly flexible systems as well as almost concordial class-like systems that sim-
ply happen to not behave like concordial class systems with respect to their
formal expression (e.g. the noun classifier system of Jakaltek (10)).

Another issue relates to the overall inconsistency of the properties with the
expectations from the prototype, most of all the classifier systems’ inventory
size and the markers’ semantic contribution to the noun phrase. A semantic
contribution to the noun phrase is expected to be found in classifier systems and
to be absent from concordial class systems; nevertheless, a number of classifier
systems does not semantically contribute to the noun phrase, whereas a number
of concordial class systems displays a semantic contribution. It is thus in order
to elaborate this study’s yet underspecified notion of the not further specified
property of “semantic contribution”. The specification of this property is based
on an investigation of the types of classificatory semanticity that a system
is based on, and on the kinds of observable contribution to the noun phrase
that are found across systemic types. The fairly inconsistent inventory size is
most obviously present in classifier systems: Many classifier languages show a
classificatory inventory that is kept to a manageable extent of some two dozen
classes. It thus seems necessary to rephrase the role of the inventory size in
nominal classification typology.

Finally, correlations between properties are present in the sample. On the
one hand, the interactive semanticity of a system affects the flexibility of the
assignment and the assigning agent as well as the system’s semantic contribu-
tion to the noun phrase. On the other hand, multiple marking and the locus
operandi correlate. These correlations, their dependence on systemic types, and
their role for the grammaticalisation hypothesis are investigated in section 6.8.
Section 6.9 discusses the outcomes of the synchronic study, addresses the pre-
dictions, hypotheses, and research questions, and serves as a basis for chapter 7,
where a more finegrained instrument for mapping nominal classification devices
is presented.

6.1 Inventory Size

The individual sample systems' class inventory corresponds well to the general grammaticalisation cline. Within the language sample, the inventory size is the property that is most grammaticalised: Only five systems employ a clearly large class inventory of more than two dozen classes. These are the classifier-like systems of Korean (1) and Mandarin (2) as well as the hybrid systems of Nasioi (3), Bora (4), and Mundurukú (5; cf. table 6.2).

Table 6.2: The systemic inventory size in the language sample.

<table>
<thead>
<tr>
<th>System</th>
<th>Korean</th>
<th>Mandarin</th>
<th>Nåsioi</th>
<th>Bora</th>
<th>Minangk耐心</th>
<th>Non-pi’imooti</th>
<th>Rammaa Fijian</th>
<th>Hungarian</th>
<th>Akha</th>
<th>North Ambrym</th>
<th>Dongo-Ko (CAS)</th>
<th>Tohono O’odham</th>
<th>Hitung Dwe</th>
<th>Singk</th>
<th>Dyirbal</th>
<th>Turia</th>
<th>Burushaski</th>
<th>Mosasai</th>
<th>Sunucaan</th>
<th>Dingo-Ko (CCS)</th>
<th>Dongo-Ko (CLS)</th>
<th>Maori</th>
<th>Ndo</th>
<th>Ket</th>
<th>Basque</th>
<th>German</th>
<th>Hadza</th>
<th>Hitan</th>
<th>Malay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
</tr>
</tbody>
</table>
| blank: large, •: smallish

All other systems employ between two and twenty-four classes (cf. table 6.3), which thus does not fully meet the expected distribution of larger classifier inventories as opposed to smaller concordial class inventories.
Table 6.3: The sample systems’ form-based properties: inventory size.

<table>
<thead>
<tr>
<th></th>
<th>CLSs</th>
<th>HSs</th>
<th>CCSs</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>NUM</td>
<td>POSS</td>
<td>MULTI</td>
</tr>
<tr>
<td>large</td>
<td>—</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>small</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>mean</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

The following paragraphs address two questions: Which systems prefer larger inventories, and does the composition of inventories across the sample correspond to the expectations of the grammaticalisation hypothesis (i.e. many and fairly specific classes > fewer and fairly generic classes)? Due to the predominant employment of smallish inventories, there is no noticeable trend of individual subtypes to employ a large inventory. It may be mentioned though that the two largish classifier systems are referred to as numeral classifier systems in the literature; however, the Mandarin classifiers may occur not only with quantifiers, but also with demonstratives and the bare noun. This renders it a multiple classifier system according to the definition of this study. Figure 6.1 shows the individual sample systems and their inventory size.
Figure 6.1: The inventory size of the sample languages.
In addition to the prevalence of small class inventories, the inventories of the sample systems display another unexpected feature: There are no 'medium inventories' present in the sample. The large inventories all feature more than a hundred classifiers, but there are no systems that employ a number of between e.g. fifty and eighty classifiers. It is not always possible to determine the exact size of the class inventory of a given system, for instance when there are no formal grounds on which classifier constructions and generic-specific noun pairings can be distinguished (cf. Aikhenvald 2000: 86). Beyond that, a construction that hosts classifiers may host a mixed set of markers, which do not all constitute nominal classifiers (cf. e.g. the Mandarin classifier tiao “long” and the Mandarin measure noun bei “cup”, whose host construction shows the same structure and service; cf. Zhang 2007: 47 and Zhang 2013: 103). In order to receive a better impression of the size of the various classifier systems, table 6.4 provides an overview over the actual size of sample-internal and -external inventories.

Table 6.4: Individual languages’ inventory size.

<table>
<thead>
<tr>
<th>sample-internal languages</th>
<th>sample-external languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Tohono O’odham pCLS</td>
<td>2 Emmi nCLS (A: 84)</td>
</tr>
<tr>
<td>2 Sumerian GS</td>
<td>2 Manam pCLS (g: 277f.)</td>
</tr>
<tr>
<td>2 Hadza GS</td>
<td>2 Kilivila pCLS (l: 12ff.)</td>
</tr>
<tr>
<td>2 Hausa GS</td>
<td>3 Pa’ijamalah nCLS (A: 84)</td>
</tr>
<tr>
<td>2 Masi GS</td>
<td>4 Tape pCLS (F: 5-7)</td>
</tr>
<tr>
<td>3 German GS</td>
<td>4 Paamese pCLS (F: 5-7)</td>
</tr>
<tr>
<td>3 Ket GS</td>
<td>4 Lolovoli pCLS (F: 5-7)</td>
</tr>
<tr>
<td>3 Maasai GS</td>
<td>5 Lenakel pCLS (F: 5-7)</td>
</tr>
<tr>
<td>3 Tamil GS</td>
<td>5 Araki pCLS (F: 5-7)</td>
</tr>
<tr>
<td>4 Boumaa Fijian pCLS</td>
<td>5 Palikûr pCLS (F: 460)</td>
</tr>
<tr>
<td>4 Burushaski GS</td>
<td>5 Bengali qCLS (A: 105)</td>
</tr>
<tr>
<td>4 Dyirbal GS</td>
<td>5 Persian qCLS (H: 195)</td>
</tr>
<tr>
<td>5 North Ambrym pCLS</td>
<td>6 Lounwolwol pCLS (F: 5-7)</td>
</tr>
<tr>
<td>6 Ingush NCS</td>
<td>6 Anejom pCLS (F: 5-7)</td>
</tr>
<tr>
<td>7 Great Andamanese mCLS</td>
<td>6 Sakao pCLS (F: 5-7)</td>
</tr>
<tr>
<td>7 Dongo-Ko NCS</td>
<td>8 Coastal Sarawak qCLS (C: 293)</td>
</tr>
<tr>
<td>10 Chichewa NCS</td>
<td>8 Yucuna qCLS (A: 106)</td>
</tr>
<tr>
<td>11 Nùng mCLS</td>
<td>10 Assamese qCLS (A: 102)</td>
</tr>
<tr>
<td>11 Hmong Daw mCLS</td>
<td>10 Murrinhpata nCLS (A: 84)</td>
</tr>
<tr>
<td>11 Dongo-Ko pCLS</td>
<td>11 Acehnese nCLS (A: 84)</td>
</tr>
</tbody>
</table>

1 For that reason, exact quantities cannot be provided for all languages, and rather vague estimations of individual classifier systems’ inventories occur in the literature. For instance, the possessive classifier system of Puluwat and the numeral classifier system of Mal are claimed to employ “a few dozen classifiers” (cf. Elbert 1974 and Adams 1989, both cited in Aikhenvald 2000: 141 and 101 respectively). Those numbers are not necessarily accurate; this can be exemplified with an assessment of the literature Mandarin on the classifier system, which is also claimed to employ “a few dozen” classifiers (Li and Thompson 1981: 105), while other estimations (e.g. Peiru et al. 2001: preface, Zhang 2007: 44) oscillate between 187 and 902 classifiers, which obviously exceeds a number of a few dozen classificatory elements.

2 Note that the sample-external systems are all randomly selected and classifier systems, since concordial class systems employ by definition small systems and thus do not have the potential to employ ‘medium size’ inventories.
### 6.1. Inventory Size

<table>
<thead>
<tr>
<th>Code</th>
<th>Language</th>
<th>Type</th>
<th>Inventory Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Tainae</td>
<td>NCS</td>
<td>14 Tashkent Uzbek</td>
</tr>
<tr>
<td>16</td>
<td>Ngar’ginyemerri</td>
<td>HS</td>
<td>14 Akatek</td>
</tr>
<tr>
<td>18</td>
<td>Moro</td>
<td>NCS</td>
<td>15 Kanjobal Mayan</td>
</tr>
<tr>
<td>18</td>
<td>Hungarian</td>
<td>qCLS</td>
<td>15 Tibetan</td>
</tr>
<tr>
<td>24</td>
<td>Jakaltek</td>
<td>nCLS</td>
<td>19 Mparntwe Arrerute</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>23 Yidiny</td>
</tr>
<tr>
<td>26</td>
<td>Bahwana</td>
<td>qCLS</td>
<td>26 Bahwana</td>
</tr>
<tr>
<td>38</td>
<td>Malay</td>
<td>qCLS</td>
<td>38 Malay</td>
</tr>
<tr>
<td>118</td>
<td>Nasioi</td>
<td>HS</td>
<td>&gt;40 Yagua</td>
</tr>
<tr>
<td>120</td>
<td>Mundurukú</td>
<td>HS</td>
<td>140 Vietnamese</td>
</tr>
<tr>
<td>125</td>
<td>Bora</td>
<td>HS</td>
<td>150 Japanese</td>
</tr>
<tr>
<td>&gt;187</td>
<td>Mandarin</td>
<td>mCLS</td>
<td>~200 Burmese</td>
</tr>
<tr>
<td>&gt;197</td>
<td>Korean</td>
<td>qCLS</td>
<td>~200 Thai</td>
</tr>
</tbody>
</table>


The sample-external systems also display a strong preference for small class inventories: There are four systems with large inventories, and only three systems that employ ‘medium size’ inventories according to the preliminary definition of this study. Note that the large systems in table 6.4 are all hybrid or numeral classifier systems and are all located in East and South East Asia, the Pacific area and South America. They form a large belt, where obligatory numeral classification occurs as an areal phenomenon (cf. Aikhenvald 2000: 122 and Gil 2013: s.v.). This matches Aikhenvald’s (2000: 103) observation that it is mainly isolating languages that tend to employ greater numbers of specifically numeral classifiers. The East and South East Asian languages are indeed of an isolating type, whereas the South American ones and Nasioi are of an agglutinating language type.

Why is it numeral classifier systems that employ large inventories? Numeral classifiers are assumed to “make count nouns enumerable by individualizing and classifying them” (Bisang 1999: 1), which is the reason why they are not always easy to distinguish from measure nouns, which unitise or group conceptually unbound entities. English for instance is not a classifier language, as it employs measure nouns (e.g. a bottle of beer or a colony of vultures), but no classifiers (e.g. *three fruits of orange* or *three round object of orange*). Numeral classifiers account for a nominal referent’s stable and inherent properties and may constitute a system’s inventory together with non-classificatory elements such as measure nouns, which account for a nominal referent’s temporary constellation. This is the reason for e.g. the Mandarin inventory to be claimed to consist of between a few dozen and 902 classifiers, since the definition of what is a classifier and what is not varies between the authors.

Measure nouns of classifierless languages such as English typically are legion. Many measure noun-noun pairings are conventionalised (except for the unitisers, which can be varied according to the intended unit; e.g. a bottle/sip/glass of beer). It is thus not perceived as ‘proper English’ if one varies in the usage of
measure nouns with a given noun, e.g. *a plump of falcons or *a trip of falcons. Instead, the conventionalised pair a cast of falcons is expected. This system of unitisation is thus highly specified and inflexible. Non-classificatory systems like the one present in English occur also in numeral classifier constructions: Cf. table 6.5 for a selection of Korean measure nouns, which morphosyntactically behave like the Korean numeral classifiers and semantically like English measure nouns.

Table 6.5: A selection of Korean measure nouns (Martin 1992: 181).

<table>
<thead>
<tr>
<th>plants</th>
<th>units</th>
<th>fish</th>
</tr>
</thead>
<tbody>
<tr>
<td>tan “bunches (vegetables, firewood)”</td>
<td>pyeng “bottles”</td>
<td>son “handfuls of fish”</td>
</tr>
<tr>
<td>tapal “bunches (flowers, plants)”</td>
<td>mokum “sips”</td>
<td>mus “10 fish”</td>
</tr>
<tr>
<td>mus1 “sheaves of straw”</td>
<td>can “cups”</td>
<td>mus1 “sheaves of straw”</td>
</tr>
<tr>
<td>cuwd “tied bunches”</td>
<td>mosum “handful”</td>
<td>mokum “sips”</td>
</tr>
</tbody>
</table>

These measure nouns, which display a varying degree of semantic specificity (cf. general a bunch of flowers/people/grapes vs. specific a litter of pups), are complemented by numeral classifiers: There are general classifiers such as kay “objects”, tay1 “vehicles”, and phil “animals” as well as specific classifiers such as pel “suits of clothes” and tay2 “cigarettes, pipefuls” (Martin 1992: 179f.). Considering that the concept of nouns in classifier languages is claimed to be not as specific as in non-classifier languages (cf. Bisang 1996: 519 and Foley 1997: 232, cited in Rijkhoff 2002b: 221), these languages rely on the specification the nominal properties with respect to its spatial dimension (Rijkhoff 2002b: 220), which is achieved by means of nominal classifiers. This also marks the difference from non-classifier languages: Most languages possess collective measure nouns such as flack, bunch, group, or heap, but since these collective terms group singleton entities that are typically well-defined as individual objects (e.g. flower, which denotes a single representative of the real world category of flowers), the non-classifier languages have no need to employ a grammatical means to specify these individual objects, since the specification of the object referred to is part of the lexical entry of the noun in use (cf. 6.1).

(6.1) English

A a bouquet/bunch of flowers  
B a flower  
C a blossom  
D a shoot

In classifier languages, the lexical item that is categorised usually has a broader semantics than in non-classifier languages. It is therefore the classifier specifying which representative from the ‘flower domain’ is referred to, not the lexeme itself (cf. 6.2).
6.1. Inventory Size


A kkoč sëgy tapal
flower three bunch
“three bouquets of flowers”

B kkoč han kalunu
flower one CL:PLANT
“one flower [plant]”

C kkoč han songi
flower one CL:BLOSSOM
“one blossom”

D kkoč han phoki
flower one CL:YOUNG.FLOWER
“one shoot”

The reason for numeral classifier systems to employ larger sets of classifiers is thus the underspecified semantics of their nominals, which are in need of an unambiguous individuation in quantifying contexts. In contrast to non-classifier languages, this individuation task has to be performed also on singleton entities, which invites numeral classifier systems to employ large sets of classifiers that are similar to measure nouns, but semantically account for non-plural items. Since measure nouns in all sorts of languages may be highly sophisticated and comprise a high number of members of differing degrees of semantic specificity, it does not come as a surprise that we find the same phenomenon among numeral classifiers, and in addition a fair number of classifiers complementing the set for the purpose of specification of singleton entities.

Large inventories of many dozen distinct classes are conceivably less economic to manage for the speaker than systems with smaller inventories. Consequently, it is to be expected that we find a larger number of more economic smaller systems. Also, the classificatory range of a classifier system with a large inventory is not always fully exploited: For instance, most Mandarin specific classifiers occur in conventionalised and rigid noun-classifier pairs only in higher registers (Li and Thompson 1981: 105), whereas in the lower registers, a greater flexibility prevails, and the general classifier ge is often used to replace more specific classifiers (Zhang 2007: 55).

The large inventories of the hybrid systems of Nasioi, Bora, and Mundurukú display a similar structure as the large inventories of the numeral and multiple classifier systems: There are a few relatively general markers, which are complemented by a broad range of more specific markers (cf. table 6.6).

3Similarly, Japanese employs a number of about 150 classifiers. However, only 28–36 of them are consistently used by all speakers, and some 47 additional classifiers show a relatively broad use (cf. Downing 1996: 17 and 21). The total of Japanese classifiers which are more or less consistently used across the adult population thus adds up to a ‘mere’ 80.

<table>
<thead>
<tr>
<th>Bora</th>
<th>Mundurukú</th>
</tr>
</thead>
<tbody>
<tr>
<td>general</td>
<td></td>
</tr>
<tr>
<td>-gwa “plank-shaped”</td>
<td>(’)a “round objects, head (front side)”</td>
</tr>
<tr>
<td>-hi “flat and round”</td>
<td>b̄l “round and rigid object, arm”</td>
</tr>
<tr>
<td>-hu “tube”</td>
<td>dip “plant (wild or cultivated)”</td>
</tr>
<tr>
<td>-hpaj “liquid”</td>
<td>di “liquid”</td>
</tr>
<tr>
<td>specific</td>
<td></td>
</tr>
<tr>
<td>-ngua “object with broken edge”</td>
<td>i̲ “nut, chestnut”</td>
</tr>
<tr>
<td>-ŋi:ba “small palm tree”</td>
<td>a³ ô “trunk (upper part)”</td>
</tr>
<tr>
<td>-ŋi:pa “gable of a roof”</td>
<td>dip̂ “blood”</td>
</tr>
<tr>
<td>-fiti: “object with hole in it”</td>
<td>narin “feces”</td>
</tr>
<tr>
<td>-tagi: “object used for beating”</td>
<td>p̱i̲ “interior, pain”</td>
</tr>
<tr>
<td>-hu:i: “palm leaf”</td>
<td>a³ di “placenta”</td>
</tr>
<tr>
<td>Nasioi</td>
<td></td>
</tr>
<tr>
<td>general</td>
<td></td>
</tr>
<tr>
<td>-ni “feminine”</td>
<td></td>
</tr>
<tr>
<td>-voro “animal”</td>
<td></td>
</tr>
<tr>
<td>-ru’ “liquid (in a container; pool, lake)”</td>
<td></td>
</tr>
<tr>
<td>-vari “tree”</td>
<td></td>
</tr>
<tr>
<td>specific</td>
<td></td>
</tr>
<tr>
<td>-veng “broken off end of something hard and flat”</td>
<td></td>
</tr>
<tr>
<td>-veku “coconut frond basket”</td>
<td></td>
</tr>
<tr>
<td>-vou “children of a common grandmother”</td>
<td></td>
</tr>
<tr>
<td>-rang “string of coconuts as found in the tree”</td>
<td></td>
</tr>
<tr>
<td>-ranto “sago frond/shingle”</td>
<td></td>
</tr>
<tr>
<td>-roope “coconut shell bowl or half section with or without its meat”</td>
<td></td>
</tr>
</tbody>
</table>

The ‘hybrid’s’ inventories make use of more than one kind of semantic categorisation. For instance, Nasioi nouns are categorised among others with respect to their animacy, constitution (e.g. -ru “liquid”), and inherent nature (e.g. -vari “tree”), Bora nouns with respect to their shape (e.g. -hi “flat and round”) and constitution (-hpaj “liquid”), and Mundurukú nouns with respect to their shape (e.g. (’)a “round objects, head (front side”)”, inherent nature (e.g. dip “plant (wild or cultivated”)”), and constitution (e.g. di “liquid”). These more general and abstract categories are complemented by a wide range of specific classifiers, which further specify e.g. the arrangement of objects (e.g. Nasioi -rang “string of coconuts as found in the tree”), more specific inherent nature properties (e.g. Bora -hu:i: “palm leaf” and -ŋi:pa “gable of a roof” or Mundurukú a³ o “trunk (upper part)”), or part-whole relations (e.g. Bora -ŋi:ba “gable of a roof” or Mundurukú a³ o “trunk (upper part)”). Instead of a whole lot of semantically highly specific classifiers, smaller inventories may be expected to display a set of more abstract semantic categories,
which together constitute an economically efficient classificatory system. This is indeed mirrored in the sample; table 6.7 displays the semantic domains that are found in the smaller classifier inventories of the sample.

Table 6.7: The semantic domains of the smaller classifier inventories of the sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>General/Residue</th>
<th>Humanness</th>
<th>Animacy</th>
<th>Inanimate</th>
<th>Shape</th>
<th>Function</th>
<th>Material</th>
<th>Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tohono O’odham (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boumaa Fijian (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Ambrym (5)</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong (11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nùng (11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jakaltek (24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Great Andamanese (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian (18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The small systems also categorise on basis of a variety of semantic domains. Almost all classifier sets make a distinction of humans and/or animates from nonhumans and/or inanimates. They also tend to employ a smaller set of classifiers distinguishing general properties such as shape (e.g. Hmong “long” and “flat” and Nùng “stick-like”, “leaf-like”, “bean-like”, and “lump-shaped”), function (e.g. Boumaa Fijian and North Ambrym “edible” and “drinkable”), or material (Jakaltek “stone”), which are complemented by a smaller series of more specific classifiers (e.g. Hmong “song, dream”, North Ambrym “basket”, and Jakaltek “salt”). This means that the bulk of categorisable nouns in one of these languages can be categorised on the basis of two or three principles (typically humanness/animacy and shape, function, or material), and only a smaller fraction of nouns requires a semantically more specific categorisation (e.g. due to their cultural saliency). This classificatory principle also accounts for the large numeral classifier inventories: A smaller number of more general e.g. shape-based classifiers is complemented by a larger number of more specific classifiers, which may be motivated by the classificatory specification and individuation of otherwise underspecified singleton nouns.

The two main questions can thus be answered as follows: Numeral classifier systems or systems that evolved as numeral classifier systems prefer larger inventories. The inventories themselves display properties which indeed fit the grammaticalisation pattern in that both smaller and larger inventories display a number of general classifiers which are complemented by a smaller number of specific classifiers in smaller inventories and a larger number of specific classifiers in larger inventories.
6.2 Obligatoriness

The obligatoriness of classification is one of the more tangible defining properties of classifier and concordial class systems and thus is distributed according to the expectations of this study: All concordial class systems employ an obligatory classification, while all classifier and hybrid systems employ a non-obligatory classification (cf. table 6.8). In the following, the concept of obligatoriness in the respective systemic type is explored and exemplified. The systems of Hungarian, Nganjuyemerri, Dyirbal, and Tainae are looked into more in detail.

Table 6.8: Obligatory classification in the language sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasioi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bora</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mundurukú</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nganjuyemerri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonai Fijiian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Great Andamanese</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sumerian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CCS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Daw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tainai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sumerian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CCS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Daw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

blank: non-obligatory, ●: obligatory

A concordial class system classifies nouns obligatorily in that each and every noun belongs to a concordial class in each and every context, even if this class membership is not overtly marked. Classifiers are not inserted automatically into every context. Their classification is not obligatory in two ways: First, the classifier set as a whole does not necessarily cover a language’s full nominal lexicon, either because their semantic scope cannot cover it (cf. the Jakaltek inventory in table 6.9, which does not cover e.g. abstracts or locative nouns; cf. Craig 1986a: 273), or because certain groups of nouns are excluded from classification despite an available category. For instance, inalienable nouns are not categorised in Hmong Daw (cf. 6.3, where no classifier occurs with kin and household terms despite the availability of e.g. the category “human”).
Table 6.9: The Jakaltek classificatory categories (cf. Craig 1977: 133 and 154, Day 1973: 125f.).

<table>
<thead>
<tr>
<th>Jakaltek</th>
<th>physical set</th>
<th>social interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>animal (product), rubber</td>
<td>male deity</td>
</tr>
<tr>
<td>female deity</td>
<td>corn/wheat (product)</td>
<td>respected nondeity</td>
</tr>
<tr>
<td>infant</td>
<td>stone/metal/glass (product)</td>
<td>infant</td>
</tr>
<tr>
<td>female nondeity</td>
<td>cord/fiber rope (product)</td>
<td>male nonrespected noninfant kin</td>
</tr>
<tr>
<td>infant</td>
<td>cloth (product)</td>
<td>female nonrespected noninfant kin</td>
</tr>
<tr>
<td>female noninfant child kin</td>
<td>plant (product)</td>
<td>male noninfant child kin</td>
</tr>
<tr>
<td>female noninfant nonkin child</td>
<td>water (product)</td>
<td>female noninfant nonkin child</td>
</tr>
<tr>
<td>female noninfant nonkin child</td>
<td>thread</td>
<td>male nonrespected nonchild nonkin</td>
</tr>
<tr>
<td>female noninfant nonkin child</td>
<td>soil</td>
<td>female nonrespected nonchild nonkin</td>
</tr>
<tr>
<td>female noninfant nonkin child</td>
<td>fire</td>
<td>female nonrespected nonchild nonkin</td>
</tr>
<tr>
<td>female noninfant nonkin child</td>
<td>salt</td>
<td>female nonrespected nonchild nonkin</td>
</tr>
</tbody>
</table>

(6.3) Hmong Daw (cf. Bisang 1993: 30)

kuv 1SG.POSS  txiv, father
kuv 1SG.POSS niam, mother
koj 2SG.POSS tsev, house
kuve 1SG.POSS teb, land
kuve 1SG.POSS chaws fields
“my father, my mother, your house, my land, my fields”

Second, classifiers occur only in language-internally selected morphosyntactic environments, and it is only there that they are obligatory. Cf. (6.4) where Mandarin classifiers combine with a quantifier, a demonstrative, or the noun itself, whereas outside of those contexts, they do not occur (cf. (6.5 A and 3.71 B = 6.5 B, where no classifier occurs with the possessive or adjective).

(6.4) Mandarin (cf. Sun 2006: 159f.)

A nà yī-pī hēi mā
DEM:REMOTE one-CL:HORSE black horse
“That black horse.”

B na pī hēi mā
DEM:REMOTE CL:HORSE black horse
“That black horse.”

C māi pī mā
buy CL:HORSE black horse
“Buy one black horse.”


A zhè shǐ nín-de chá
DEM:CLOSE be you:POLITE-POSS tea
“This is your tea.”
Obligatoriness in nominal classification thus comprises an obligatory classification only “in surface structures under specifiable conditions” (Allan 1977: 285), i.e. in at least one specific morphosyntactic type of construction (e.g. numeral or possessive classifier constructions), possibly several morphosyntactic contexts (cf. again 6.4 for the case of Mandarin classifiers), or with all nominal occurrences in the case of concordial class systems. This notion of obligatory classification is met very well in the sample (cf. again table 6.8, where all classifier and hybrid systems display a non-obligatory classification and all concordial class systems display an obligatory classification). Yet, there are two exceptions where the classification is optional even in those contexts where it is predicted to occur. These are the Hungarian (8) numeral classifier system (cf. 5.1 – 105) and the ‘hybrid’ system of Ngan’gityemerri (6; cf. 5.21 – 6.7).

(6.6) Hungarian (cf. Csirmaz and Dékány 2010a: 12)

A hét (fej) saláta
seven CL:HEAD lettuce
“seven lettuces”

B hét (szem) cukor
seven CL:EYE candy
“seven pieces of candy”

C hét (szál) gyertya
seven CL:THREAD candle
“seven candles”

(6.7) Ngan’gityemerri (cf. Reid 1997: 167)

(yawurr) ganbi kide
TREE didgeridoo where
“Where is the didgeridoo?”

Why do these languages not categorise their nominals obligatorily? The numeral classifiers of Hungarian appear to be a rather recent innovation: Hungarian as an Uralic language does not employ other nominal classification devices and has not been described as employing a numeral classifier system until recently (cf. de Groot 2005, Csirmaz and Dékány 2010a and b). The case of the Persian numeral classifiers displays noticeable similarities: Persian as an Indo-Iranian language employed an inherent gender system, which was productively used until the stage of Old Persian (600-300BC; cf. Duke 2009: 77). After that, it decayed and gave way to a numeral classifier system. This numeral classifier system seems to be in the process of establishing grammatically during the 20th
Even in recent studies, the classifier system is described as optional (cf. 5.29 = 6.8).


A  do (-ta) pesør [-a]
   two  CL:ITEM boy  PL
   “[the] two boys”

B  cel (-m’s) gusfænd
   twenty CL:HEAD sheep
   “twenty sheep”

The data from Hungarian and its typological parallel Persian thus suggests that both languages are able to categorise their nouns, but that this system is not (yet) a ‘proper’ system of nominal classification in terms of obligatory classification under “specifiable conditions” (Allan 1977: 285).

Ngan’gityemerri marks its nouns’ class membership by means of optional independent markers (cf. again 5.21) and bound markers. Head nouns tend to be obligatorily marked for their class, while agreement targets either all show class marking or none at all (Reid 1997: 168), except if their are used predicatively (cf. 5.25 = 174, where the predicative adjective is marked for the concordial class of its referent).

(6.9) Ngan’gityemerri (cf. Reid 1997: 169)

A  wa= yedi wa= wunu wa= lenggir
   M  man  M  that  M  bad
   “that bad man”

B  wa= yedi wunu wa= lenggir
   M  man  M  that  M  bad
   “That man is bad.”

The ‘tendency towards obligatory marking’ of the head nouns is puzzling: If they are obligatorily marked, there is no tendency for head marking to be expected; if they are marked in most cases, there is a strong tendency for head marking, but no obligatory marking. I interpret this description of Ngan’gityemerri in a sense that head marking is obligatory, but that exceptions occur (e.g. in cases where an independent marker would be inserted; cf. again 5.21). Though rare, cases like this can be found in other systems: Cf. 6.10 for an example from a Mandarin quantificational phrase, where classification is obligatory (cf. A), but exceptionally no classifier occurs in the synonymous construction in B.

4For instance, the Persian classifiers appear to have been developing to clitics only recently, since several grammarians describe them as independent nominals (cf. Rastorgueva 1964: 32 and Mace 2003: 170f.), whereas Mahootian (1997: 195), one of the more recent reference grammars, describes them as clitics and clitics only. Beyond that, the classificatory markers display a strong similarity with their source nouns, which may indicate their young age (cf. again table 2.3 on page 29).

A  wu  ge  bing  er  tiao  yu
five  CL:GENERAL  loaf  two  CL:LONG  fish
“five loaves and two fish”

B  wu  bing  er  yu
five  loaf  two  fish
“five loaves and two fish”

The optional multiple marking in Ngan’gityemerri on the other hand is important for its positioning towards concordial class systems: “To demonstrate the existence of a gender system evidence is required from agreement, that is, evidence outside the noun itself” (Corbett 1994: 1348). This evidence is not fully present in the case of Ngan’gityemerri, since the targets are marked optionally for a noun’s class. Considering the grammaticalisation hypothesis, this is actually the situation that we expect to find in a system that is about to exert the systemic shift from a classifier to a concordial class system: The classificatory markers are predicted to spread to more than just one morphosyntactic host (e.g. from the numeral to possessives, deictives, adjectives, etc.) and subsequently to become obligatory across contexts. Though it is not a part of the prediction that the variety of chosen targets all take the marking or all drop it simultaneously, this case is not excluded. Ngan’gityemerri might thus represent the only case in the sample where we can observe the introduction of concordial marking in a living and natural environment. However, section 6.3 suggests a different analysis of supposed multiple marking in the ‘hybrid’ systems.

To sum up, obligatory classification of nominals is a well-met defining criterion for nominal classification devices. Classifier systems obligatorily occur in one or more morphosyntactic contexts, and concordial class systems always classify their nominals. The classifier system of Hungarian (and the one of Persian) is likely to be a very young system, which borders the status of a classifier system ‘proper’, but is expected to develop an obligatory categorisation by further grammaticalisation. The ‘hybrid’ system of Ngan’gityemerri might currently develop concordial class marking from an underlying nominal classification that is already present on the noun itself in most cases, but has not yet reached a state where it obligatorily copies the class marker to the targets, i.e. the status of an agreeing system (cf. figure 6.2).
6.2. Obligatoriness

Finally, the obligatory classification of the systems of Dyirbal (18) and Tainae (19) is explored more in depth, since both of them display ‘gaps’ in their marking systems. In addition, they are reanalysed as non-agreeing systems in section 6.3, which questions their status as concordial class systems and obligatorily classifying systems. Why is Dyirbal an obligatory system of nominal classification? Dyirbal nouns are accompanied by the so-called noun marker, which carries a class marker. All consulted sources on Dyirbal (Dixon 1968, 1972, and 1982, and Lakoff 1990) confirm that the Dyirbal system employs class agreement and that each noun is assigned to a typically stable concordial class with respect to the individual class features [± masculine], [± animate], [± feminine], [± fire-related], [± water-related], [± harmful], [± (bearer of) non-flesh food]; if these features do not apply, a noun is assigned to the residual class IV (cf. Dixon 1968: 124). This information strongly suggests a noun’s class membership to be inherent to the noun, which equals an obligatory classification. However, the noun marker, which is the only element that carries class marking, does not occur in the ablative, allative, and locative case (cf. Dixon 1982: 163), which makes it impossible to prove whether or not an actual nominal classification takes place. Since the Dyirbal system displays a prevailing classifier-like character and since there is no overt evidence for the exertion of nominal classification in the context of the ablative, allative, and locative case, an assessment of Dyirbal must result in the recognition that it is impossible to prove that Dyirbal classifies its nouns obligatorily.

Tainae also corresponds a lot better to the definition of a classifier system than it does to a concordial class system. The consulted source on Tainae (Carlson 1991) does not reveal information on the obligatoriness of classification. It is however possible to demonstrate that Tainae employs a non-obligatory classification. First, remember that only nouns that refer to a specific and salient discourse referent are assigned to a specific class, and are otherwise marked by -e, which Carlson (1991: 29) analyses as a plural marker. In that case, classification in general would be optional in Tainae. Since the supposed ‘plural’ marker is also used to refer to singular entities (cf. 206), it is reanalysed as a residue class marker here.

\[
\begin{array}{llll}
Padae-ta & mpa-po & ba-e-yo & f-as-\text{ka} \\
exit-1PL & outside-static & stone-PL-OBJ & one-LONG-two \\
ki-nk-i & f-as-\text{ka} & otiku-\text{e} & mpi-nu \\
3SG-POSS-EPENTHESIS & one-LONG-two & dirt-PL & one-only \\
\end{array}
\]

“We went outside and he took two stones and one [clump of] dirt.”

By reanalysing the marker \(-e\) as a residue class marker, nominal classification in Tainae gains an almost obligatory status, since classificatory marking occurs on almost every noun in discourse. Yet there are various examples of e.g. proper nouns or kinship terms, but also common nouns, which remain unclassified (cf. 6.12).


A to-na-i ankwi mante-yi sakun-e
1SG.M-1PL today Monday-OBJ school-RESIDUE
do-DUR-1PL.DIFFERENT.REFERENT
“We were doing schoolwork today, monday, when[...]”

B as-i
male.cross.cousin-EPENTHETIC 1DL go-1DL.FUT-UNQUOTE
“His cousin replied, ‘Let’s both go.’”

C mi tumpuni stati at-iwa-mi stati[...]
QUOTE ancestor story grandparent-3PL.M-OBJ story
“The ancestral story, (about how people scared the bushknife which customarily did work by itself)[...]”

Even though these cases are not frequent at all, their existence suffices for a reanalysis of the Tainae system as a non-obligatory one. Table 6.10 presents the reanalysed distribution of obligatoriness in the sample.

Table 6.10: Obligatoriness in the sample revised.

<table>
<thead>
<tr>
<th>Language</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesian</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minangkare</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minangkare Banaus</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minangkare Grop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minangkare Sari</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minangkare Lyeusuri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pusimay</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyirbal</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Ambrym</td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CCS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Daw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Nung</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyirbal</td>
<td></td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Ambrym</td>
<td></td>
<td></td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CCS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Daw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Nung</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong Mari</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

blank: non-obligatory, •: obligatory
6.3 Multiple vs. Single Marking

Multiple marking corresponds well to the grammaticalisation cline. This is not surprising, since multiple marking is the most crucial defining criterion for concordial class systems and a defining non-criterion of classifier systems (cf. e.g. Corbett 1994: 1348). However, the sample data calls for a closer inspection of the supposed multiple marking of the ‘hybrid’ systems of Nasioi, Bora, Mundurukú, and Ngan’gityemerri, since it diverts from ‘proper’ multiple marking in a few aspects. This analysis is carried out after a description of single marking in classifier systems and multiple marking in concordial class systems. It shows that the ‘hybrid’ systems allow multiple classification in the same context, which formally closely resembles multiple marking, yet is different from multiple marking by means of agreement. The hybrid systems thus do not employ agreement and are reanalysed at the end of this section.

Agreement in this study has to satisfy the following criteria: The noun’s class information is copied to a class marker that occurs on a range of language-specific nominal satellites, which are different from the noun that controls the agreement. By combining with a variety of those satellites, the markers may occur multiply in the same context. It is supposedly employed in both concordial class and ‘hybrid’ systems, but not in classifier systems (cf. table 6.11).

Table 6.11: Multiple marking in the language sample.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Blank: no agreement, •: agreement, •: optional agreement

The hybrid systems form an outlier of systems that seemingly allow multiple marking; this outlier is located towards the left pole of the continuum. One precondition for a system to be identified as a concordial class system is that it
employs agreement, and this criterion is well-met within the sample. However, the narrow definition of “agreement” in this study elicits one problematic system in the sample: Dyirbal (18). Beyond the classificatory markers on nominal satellites expressing information on the class membership of their head noun, agreement is defined as potentially occurring multiply in the same context. Classifiers in contrast occur only once and overtly. In Dyirbal, each noun phrase’s head noun normally takes a “noun marker”, i.e. a determiner that displays the noun’s case, class, and location or visibility of the noun’s referent (cf. Dixon 1968: 106). The class markers also attach to the demonstrative and interrogative form of the noun marker (again, with the case marker preceding the class marker; cf. Dixon 1968: 107). Class membership thus is marked on an extranominal constituent, but it is the only element that is marked for class.


A  
<table>
<thead>
<tr>
<th>bala-∅</th>
<th>diban-∅</th>
<th>ya-ŋgu-∅</th>
<th>yibi-ŋgu</th>
<th>buran</th>
</tr>
</thead>
<tbody>
<tr>
<td>there-ABS-IV</td>
<td>stone-ABS</td>
<td>here-ERG-II</td>
<td>woman:II-ERG</td>
<td>look.at</td>
</tr>
</tbody>
</table>

“The woman here is looking at the stone there.”

B  
<table>
<thead>
<tr>
<th>yala-∅</th>
<th>diban-∅</th>
<th>ba-ŋgu-∅</th>
<th>yibi-ŋgu</th>
<th>buran</th>
</tr>
</thead>
<tbody>
<tr>
<td>there-ABS-IV</td>
<td>stone-ABS</td>
<td>there-ERG-I</td>
<td>man:II-ERG</td>
<td>look.at</td>
</tr>
</tbody>
</table>

“The man there is looking at the stone here.”

The noun itself is not marked for concordial class. Beyond that, the noun marker as the only marked element does not occur in the ablative, allative, and locative case (Dixon 1982: 163). In terms of formal realisation of the class marking, Dyirbal is thus much closer to the status of a classifier system, which inserts its classifiers by definition once and overtly within the NP-borders (cf. 2.43 = 6.14).

(6.14) Jakaltek (cf. Craig 1977: 140)

<table>
<thead>
<tr>
<th>xcam</th>
<th>[no’] CL:ANIMAL</th>
<th>hin</th>
<th>cheh</th>
</tr>
</thead>
<tbody>
<tr>
<td>died</td>
<td>my horse</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“My horse died.”

A concordial class system in contrast marks its nouns’ concordial class on a range of nominal modifiers simultaneously (cf. 6.15).

(6.15) German

<table>
<thead>
<tr>
<th>Mein-e</th>
<th>Rüstung</th>
<th>ist</th>
<th>ein-e</th>
<th>ganz</th>
<th>aus</th>
</tr>
</thead>
<tbody>
<tr>
<td>my-NOM.F.SG</td>
<td>armour:NOM.F.SG</td>
<td>be:PRS1SG</td>
<td>a-NOM.F.SG</td>
<td>fully</td>
<td>from</td>
</tr>
<tr>
<td>Fehlern</td>
<td>gewebt-e-</td>
<td>mistakes:NOM.M.PL</td>
<td>woven:ǐNOM.F.SG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“My armour is one [that is] woven fully from mistakes.”
6.3. Multiple vs. Single Marking

It may not be a common strategy for classifier systems to attach their markers regularly to the determiner as the main morphosyntactic host. Given that determiners are a very common secondary host for classifiers (e.g. in Mandarin, Nùng, Hmong, and Mundurukú), and that the determiner is the one NP-internal constituent beside the noun that is consistently present in discourse in Dyirbal, it seems natural that the classificatory markers regularly attach to the determiner and not another element. Dyirbal does not meet this study’s definition of a system employing multiple marking and rather resembles a fairly conventionalised classifier system. Since the difference between a fairly grammaticalised noun classifier system and a fairly ungrammaticalised concordial class system mainly lies in the type of morphosyntactic expression, a reclassification of the exceptional case makes more sense than the expansion of the definition of multiple marking, which is otherwise well-met in the sample (and all other instances of concordial class systems that I am aware of). Based on the sample data, the locus operandi does not play a crucial role in the determination of a system being identified as either classifier or concordial class system. Multiple marking may facilitate a class marker’s ability to transcend the borders of the noun phrase, but the case of a system of nominal classification without multiple marking, yet operating outside the noun phrase, is easily conceivable. However, since the correspondence of multiple marking and the locus operandi is strong in the sample, an NP-external occurrence of classificatory marking can be interpreted as a strong indication that the system in question is a concordial class system.

The hybrid systems of Nasioi (3), Bora (4), Mundurukú (5), and Ngan’git-yemirri (6) all seem to employ multiple class marking on a range of nominal modifiers (cf. Nichols 1992: 129, Seifart 2005: 164, Gomes 2006: 185ff., and Reid 1997: 168). This separates them from classifier systems, which feature a classifier once and overtly, whereas the classificatory markers in the hybrid systems occur multiply and outside the noun itself (cf. 5.10 A = 6.16 5.12 A = 6.17 5.18 A = 6.18 and 6.19).

(6.16) Nasioi (cf. Hurd 1977: 140)

\[
\begin{array}{ll}
A-n & \text{marasini-n-u} \\
\text{DEM-CL:TABLET} & \text{medicine-DERIV-CL:TABLET} \\
siipa-koo-n-u. & \text{head} \\
sick-part-DERIV-CL:TABLET
\end{array}
\]

“This pill is for headaches.”

(6.17) Bora (cf. Seifart 2005: 169)

\[
\begin{array}{ll}
c-bi & \text{mûhau-hi} \\
\text{DISTANT-CL:2D.ROUND} & \text{be.big.SUB-CL:2D.ROUND} \\
kù:mar-hi & \text{turtle-CL:2D.ROUND}
\end{array}
\]

“That big turtle”

\[ a^{3}ko^{3}\text{-}ba^{4} \quad i^{3}\text{-}ba^{2}\text{-}dom^{3} \quad ko^{4}be^{3} \quad be^{3} \]

banana-CL:LONG 3SG.POSS-CL:LONG-stay.FUT canoe LOC

“A banana will remain in the canoe.”

(6.19) Ngan’gityemerri (cf. Reid 1997: 204)

\[ wa= \quad wedimuy \quad ngagadi\text{-}tye \quad wa= \]

M little.boy 1SGS:go-PST M

\[ [deme-nga\text{-}pi\text{-}yi\text{-}tye\text{-}yedi] \]

3SGS:AUX-1SG.DIRECT.OBJECT-head-numb-PST-3SG.S:go

“As a young boy, I was (a male who was) habitually shy.”

However, class marking on nominal modifiers in Ngan’gityemerri is optional in that either all targets take markers simultaneously or none of them does (cf. 5.27 A and B = 6.20).

(6.20) Ngan’gityemerri (cf. Reid 1997: 174)

A \[ wa= \quad tyerrmusye \quad (wa=) \quad mirrisyarrn \quad perreyt \quad meny \]

M old.man M blind die do:3SG.S

“The old blind man has died.”

B \[ wur= \quad wedimuy \quad (wur=) \quad ngayt \quad tyusuy \quad -m \]

F child F mine sick do:3SG.S

“My little girl is sick.”

The optionality of class marking on modifiers is unexpected and might represent an instance of emerging obligatory target marking. However, this and other conspicuous traits such as a seeming ‘disagreement’ due to a classificatory variability e.g. in Bora (cf. 5.12 = 4, where the marking on the targets varies and does not coincide with the noun’s class marking in all cases) calls for a closer inspection of the hybrid systems’ way of class marking.

(6.21) Bora (cf. Seifart 2005: 169)

A \[ c\text{-}hi \quad ma\text{hau\text{-}hi} \quad k\text{\text{"u}\text{-}mu\text{-}hi} \quad turtle\text{-CL:2D.ROUND} \quad be\text{-}big\text{-}SUB\text{-}CL:2D.ROUND \]

“that big turtle”

B \[ c\text{-}hi \quad ma\text{hau\text{-}hc} \quad k\text{\text{"u}\text{-}mu\text{-}hi} \quad turtle\text{-CL:2D.ROUND} \quad be\text{-}big\text{-}SUB\text{-}CL:MG \]

“that big turtle”

C \[ aj\text{-}di \quad ma\text{hau\text{-}hi} \quad k\text{\text{"u}\text{-}mu\text{-}hi} \quad turtle\text{-CL:2D.ROUND} \quad be\text{-}big\text{-}SUB\text{-}CL:2D.ROUND \]

“that big turtle”
Multiple marking by means of agreement is one of the properties that clearly distinguishes classifier systems from concordial class systems in that classifiers do not occur as multiple agreement markers, whereas concordial classes are marked multiply by means of agreement. Agreement markers copy information from the head noun and occur on a range of nominal satellites and possibly on the noun itself. Classifiers “have meaning, in the sense that a classifier denotes some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer)” (Allan 1977: 285). Classifiers thus do not copy semantic information from the head noun, but carry this information themselves. It is hardly possible to decide unambiguously for any given language whether or not this is always the case, especially if there is not much data available in the respective language descriptions. Cf. 2.31 A – 6.22 where the classifier seemingly mirrors information that is already available in the head noun and thus redundant in the classifier; examples like this are highly frequent in classifier languages. In contrast, the concordial class marker in 31 B provides diminutive information that is not present in the head noun.


yi zhi luohu, yi zhi niao, yi zhi pangzie
one CL:ANIMAL tiger one CL:ANIMAL bird one CL:ANIMAL crab
“one tiger, one bird, one crab”

(6.23) Chichewa (cf. Mchombo 2004: 4f.)

A M-
CLASS1.SG lenje w-
SUBJ:CLASS1.SG ánú w-
ja w-
á nthábwa
that SUBJ:CLASS1.SG ASSOCIATIVE humor:CLASS10.PL w-
SUBJ:CLASS1.SG PF break FINAL.VOWEL CLASS4.PL spear
“That humorous hunter of yours has broken the spears.”

B Ku-
CLASS12.SG lenje k-
SUBJ:CLASS12.SG your
ka-
ja k-
á SUBJ:CLASS12.SG that CLASS12.SG ASSOCIATIVE nthábwa
humor:CLASS10.PL SUBJ:CLASS12.SG PF break FINAL.VOWEL t-
CLASS13.PL CLASS4.PL spear
“That small humorous hunter of yours has broken the tiny spears.”
For this reason, a more obvious distinctive criterion for the identification of the presence of agreement in the sample was adopted: The possible multiple occurrence of classificatory markers. Classifiers are commonly defined as occurring once and once only within the classificatory context (Dixon 1982 and 1986, Grinevald 2000: 62), which is the phrase that is headed by the categorised noun. Also in multiple classifier languages such as Mandarin, the classifier occurs once, even in the presence of multiple potential classifier hosts (cf. 6.24).

(6.24) Mandarin (cf. Sun 2006: 159f.)

A  naï yǐ-pí hēi mă
DEM-REMOTE one-CL:HORSE black horse
“That black horse.”

B  naï pí hēi mă
DEM-REMOTE CL:HORSE black horse
“That black horse.”

C  măi pī mă
buy CL:HORSE black horse
“Buy one black horse.”

The only ‘legitimate’ deviation from the single occurrence of classifiers within one and the same classificatory context is ‘double classification’. In those cases, more than one classifier may occur, but they differ in their classificatory semantics and thus each classify the nominal referent from their own perspective (cf. 6.25, where pă “fish” is categorised by both the general and the animate classifier).


căh tū pă du chăng āhn hō
CL:GENERAL CL:ANIMATE fish at insiden CL:INANIMATE house
ni, mưhn vă
FOC he say
“The fish inside the house said…”

According to the existing definitions, the occurrence of one and the same classificatory marker in a given classificatory context such as in 4A must result in the analysis of this marker as a non-classifier. Cases of multiple marking by the same marker like in 4A can be found not only in Bora, but also in the other hybrid systems (cf. 5.10 = 6.26, 167A = 6.27, and 5.25 = 6.28), which is why these systems were analysed as employing multiple marking.

(6.26) Nasioi (cf. Hurd 1977: 140)

A-u marasni-n-u bore’
DEM-CL:TABLET medicine-DERIV-CL:TABLET head
svipa-koo-n-u, sick-part-DERIV-CL:TABLET
“This pill is for headaches.”
Multiple vs. Single Marking

(6.27) Mundurukú (cf. Derbyshire and Payne 1990: 261)

\[
xepxep-'a \quad wezik-\dot{a} \\
two-\text{CL:ROUND} \quad \text{potato-CL:ROUND} \\
\text{“two potatoes”}
\]

(6.28) Ngan’gityemerri (cf. Reid 1997: 169)

\[
\begin{align*}
wa &= \text{yedi} & wa &= \text{wunu} & wa &= \text{lenggir} \\
M &= \text{man} & M &= \text{bad} \\
& & & & \\
\text{“that bad man”}
\end{align*}
\]

This straightforward judgement is however distorted by phenomena that do not fit this study’s understanding of multiple marking by means of agreement. Most prominently, class ‘disagreement’ may occur in the systems of Bora and Ngan’gityemerri. Example [4] exemplifies this phenomenon for Bora: [6.29] features two examples from Ngan’gityemerri, where *pope* “popë” originally belongs to the masculine class (cf. again 6.28), but is categorised in the animal class due to his position as a spiritual being. Similarly, *tyulut* “hookspear” can be categorised as either *yeli*= “bamboo spear” or *yerr*= “tree”, *match* “match” as *-yenggi* “fire” or *-yawurr* “tree”, etc. (Reid 1997: 202).

(6.29) Ngan’gityemerri (cf. Reid 1997: 184)

\[
gagu \quad pope \quad dini-ganu-tye \\
\text{ANIMAL} \quad \text{pope} \quad \text{3SGS:sit-while-PST} \\
\text{Alice Springs} \ldots \\
\text{“While the pope was in Alice Springs[…”}
\]

It is commonly accepted that nouns in a concordial class language typically and inherently belong to one class and one class only. A smaller number of exceptions from this rule is allowed (cf. e.g. Aihenvald 2000: 21, Corbett and Fedden 2015: 9, Dixon 1986: 106). This entails that a noun’s class membership is part of this noun’s lexical entry, and is read and copied by means of agreement from the lexical entry information on the class membership. This cannot be the case in [4] and [6.29] since the classificatory markers do not correspond to the head noun’s inherent class. What does that mean for the analysis of the hybrid languages? Examples like [4] and [6.29] strongly suggest that each occurrence of a classificatory marker does not represent a case of agreement with the head noun’s class, but a new instance of categorisation, and that the nominal modifiers cooccurring with the head noun are of appositional character rather than an NP-internal one (e.g. *that (one)*, *the big (one)*, *the turtle*). This also explains the optionality of the Ngan’gityemerri ‘agreement’, where “all targets show agreement, or none of them do” (Reid 1997: 168), while the head noun is obligatorily marked for class: If the Ngan’gityemerri system is analysed from the perspective of a noun classifier system, it is natural that NP-internal modifiers
do not host classifiers, while they are supposed to, if they are noun phrases in their own right. This analysis is confirmed at least for Bora by Seifart (2005: 5): “[T]wo coreferential nominal expressions in a clause are typically in a relation of loose apposition, rather than forming a tightly integrated constituent, as the literal translation of example 6.30 indicates”.

(6.30) Bora (cf. Seifart 2005: 5)

\[
\begin{align*}
\text{ó} & \quad \text{-?di} & \text{íhka} & \quad \text{-ko} & \quad \text{tsa} & \quad \text{-ko} \\
\text{1SG} & \quad \text{POSS} & \text{COP} & \quad \text{CL:1D.POINTED} & \text{one} & \quad \text{CL:1D.POINTED} \\
\text{pihhu} & \quad -ko \\
\text{fish.NOMINALISED} & \quad \text{CL:1D.POINTED} \\
\text{“I have one fishing rod.” (literally: “What [pointed] is to me, one [pointed], a fishing rod”)}
\end{align*}
\]

Next to class ‘disagreement’, the common and frequent involvement of class markers in the creation of nouns in the hybrid systems suggests that they are not mere mirrors of class information, but that a given class marker possesses a referential character itself and thus has “meaning, in the sense that [it; MP] denotes some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer)” (Allan 1977: 285). This phenomenon is present in all four hybrid systems; cf. 6.31, 6.32, 5.15 and 167 B = 6.33 and 6.34 for examples.

(6.31) Nasioi (cf. Hurd 1977: 147f.)

\[
\begin{align*}
\text{A} & \quad \text{na-} & \text{va} & \quad \text{CL:HOUSE} \\
& \quad \text{one} & \quad \text{“one house”} \\
\text{B} & \quad \text{ke} & \text{-veku} & \quad \text{-ka, ke} & \text{-vin} & \text{-ta} \\
& \quad \text{two} & \quad \text{CL:BASKET DL} & \text{two} & \quad \text{CL:VINE DL} \\
& \quad \text{“two baskets, two vines/ropes”} \\
\text{C} & \quad \text{bee} & \text{-veku} & \quad \text{-pi} \\
& \quad \text{three} & \quad \text{CL:BASKET PL} \\
& \quad \text{“three baskets”} \\
\text{D} & \quad \text{kare} & \text{-vin} & \quad \text{-tu, kare} & \text{-mii} & \text{-ri} \\
& \quad \text{four} & \quad \text{CL:VINE PL} & \text{four} & \quad \text{CL:HILL PL} \\
& \quad \text{“four vines/ropes, four hills/mountains/ridges”}
\end{align*}
\]


\[
\begin{align*}
\text{A} & \quad \text{tú?a} & \text{-pá:hi} & \quad \text{CL:HOLE} \\
& \quad \text{foot} & \quad \text{“shoe”} \\
\text{B} & \quad \text{sáa} & \text{-ha} & \quad \text{CL:COVER} \\
& \quad \text{one} & \quad \text{“one (house, etc.)”}
\end{align*}
\]
6.3. Multiple vs. Single Marking

C mĩ́hṹ -be
big  CL:M.SG
“a big one (male animate)”

D mĩ́hṹ -hi
big  CL:DISC
“a big one (hat, turtle, etc.)”

E mĩ́hṹ -ne
big  CL:INAN
“a big one (inanimate)”

(6.33) Mundurukú (cf. Derbyshire and Payne 1990: 260f.)

A wo- pi- men -na
FORMATIVE  CL:STICK  red  ADJ
“red (stick, pencil, etc.)”

B kape -di
coffee  CL:LIQUID
“coffee (drink)”

C kape -da
coffee  CL:SEED
“coffee (seed)”

(6.34) Ngan’gityemerri (cf. Reid 1997: 169)

A wur= yubu
F  good
“a good woman”

B wur= kinyi
F  this
“this woman”

C wur= ngayi
F  mine
“my woman”

D wur= wungume
F  one
“a certain woman”

E wur= nyin
F  anaphor
“that (previously mentioned) woman”

All four examples display a clear compositional character, which means that the morphemes all contribute equally to the semantics of the coined word. This also means that there cannot be agreement involved, since there is no class information copied from a noun. Due to the absence of a noun that could be copied from, the categorisation present in the above examples is a categorisation of the nominal referent itself (cf. Hengeveld 2012: 471ff. for a detailed discussion of agreement and referentiality). Consequently, the hybrid systems must be reanalysed as systems of nominal classification that lack multiple marking by
means of agreement due to their referential function. Table 6.12 provides a
revised overview over referentiality and agreement in those sample systems
that were initially analysed as multiple marking systems; the criteria used for
the identification of referentiality are the presence of class ‘disagreement’ and
nominalisation by means of classificatory markers; agreement was identified by.

Table 6.12: Multiple marking revised: Agreement vs. referential marking in
the sample.

<table>
<thead>
<tr>
<th>System</th>
<th>Bora</th>
<th>Ewe</th>
<th>Ganda</th>
<th>Hausa</th>
<th>Kwa</th>
<th>Kirghiz</th>
<th>Korean</th>
<th>Masarik</th>
<th>Sumerian</th>
<th>Tamil</th>
<th>Tainae</th>
<th>Burushaski</th>
<th>Maasai</th>
<th>Nasioi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreement</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>Referential Marking</td>
<td>a</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>a</td>
<td>r</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>r</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

A: agreement, r: referential marking

After the differentiation of agreement and referentiality, the most obvious
criterion for the clear identification of agreement in those systems that allow
multiple marking lies in the employment of portmanteau morphemes, which
convey not only information on a noun’s class, but in addition e.g. on case and
number. Since case and number marking is commonly and uncontroversially
analysed as non-referential agreement marking, portmanteau class markers that
carry information on case and/or number are analysed as non-referential agree-
ment markers, too. Note that almost all non-hybrid systems from table 6.12
employ portmanteau morphemes for their class marking; the only exception is
Tainae (19), where only the masculine and feminine class distinguish a singular
and a plural form (cf. Carlson 1991: 38). Are Tainae class markers agreeing
or referential? The specific class markers of Tainae “are used mainly when a
specific person or thing is in mind, and not when just talking about something
generall” (Carlson 1991: 27); in case of a general reference, the residue marker
-č is attached to the noun. The occurrence of the Tainae class marking only
with specific discourse referents argues for a certain degree of referentiality
of the markers. The markers may serve as nominalisers, yet the available examples
do not allow a judgement on the referentiality of the markers in those cases; cf.
6.35, where stua hamteatīo “shopkeeper” m. does not occur on its own, but is
accompanied by the lexical noun afao “man” m., which might represent either
a case of multiple classification, like in the case of the hybrid systems, or a case
of multiple marking by means of agreement.


afa -a  stua  hamte -ati -a  nipī -mtīti.
man 3SGM. store watch DUR 3SG.M come 3SG.FUT

“The storekeeper will come.”
6.3. Multiple vs. Single Marking

The clearest indication of the referential character of the Tainae class markers is provided in (5.32) where the ‘real life’ shape of a leaf is encoded in the class marking.

(6.36) Tainae (Carlson 1991: 27f.)

A  *ikyipw-*ina
   leaf-FLAT
   “leaf”
B  *ikyipw-*ina
   leaf-CYLINDRICAL
   “leaf (roll)”
C  *ikyipw-*iti
   leaf-THIN
   “leaf (strip)”
D  *ikyipw-*ai
   leaf-LONG
   “(folded) leaf (rod)”
E  *ikyipw-*ini
   leaf-FLUID
   “leaf (ball)”

The Tainae system is thus analysed as a referential one in this study and consequently stands out from the other concordial class systems. Table 6.13 provides an overview over the revised properties of the language sample. Note again that the hybrid systems have been reanalysed as markers involved in referentiality and thus each creating an own context of classification, but not multiply occurring agreement markers. The hybrid systems are thus in accordance with the grammaticalisation hypothesis.

Table 6.13: The revised sample’s form-based properties.
6.4 Locus Operandi

The locus operandi of the sample systems almost fully overlaps with their multiple marking behaviour (cf. table 6.14): All classifier systems and the system Dyirbal (18) categorise their nouns NP-internally, and all concordial class systems except the mentioned ones allow their markers to transcend the NP-borders. Only the ‘hybrid’ systems, which were reanalysed in the previous sub-section as classifier systems lacking multiple marking and instead employing multiple classification, allow their markers to occur outside of the NP borders.

The ‘hybrids’ attach their class markers among others to the verb itself, which is clearly an instance of NP-external marking (cf. 5.10 B = 6.37 6.38)

Teni tore’ tareuri-ma-∅-i
ART.F children care.for-PRS.HAB-DERIV-CL:F
bau’uri-ma-∅-i
feed-PRS.PROGR-DERIV-CL:F
“The lady who cares for the children (is) the one who is feeding them.”

(6.38) Bora (cf. Seifart 2007: 424)

ká:tùže -hi (bo:tá -hi)
fall.down CL:DISC button CL:DISC
“(The button,) it fell down.”


Ti dojot puye, o’-ti-mog ip basya’a be.
water bring when they-CL:WATER-place they basin in
“When they brought water, they placed it in a basin.”

(6.40) Ngan’gityemerri (cf. Reid 1997: 194)

ganbi yerr= buy- mem -garri kide
didgeridoo CL:TREE light 3SGS.do leg where
“Where is that white didgeridoo?”

Classifier systems do not employ multiple marking, and the classifiers occur in the categorised noun’s phrase (cf. 3.44 = 119 A). Once introduced to discourse, they may be used as carriers of anaphora without a cooccurring head noun. Yet they are embedded in an NP, where the head noun would occur, if it was not omitted (cf. 119 B).

(6.41) Jakaltek (cf. Craig 1986a: 264)

A zil [noj xuwan]NP [no’ lab’a]NP.
saw CL:MAN Juan CL:ANIMAL snake
“Juan saw the snake.”

B zil [noj ]NP [no’ ]NP
he CL:MAN CL:ANIMAL
“He saw an ANIMAL.”

Many classifier systems allow this type of reference tracking by classificatory means. It may be avoided though, if an ambiguous reading results from the ‘classifier only’ construction (cf. 6.42).

A mac ka' chawoche [naj pel]NP iz malin]NP mato
who more like:2SG CL:MAN Peter CL:WOMAN Mary or
[ya' cap]NP.
CL:RESPECTED.NONDEITY Gabriel

"Who do you like best: Peter, Mary, or Gabriel?"

B [naj pel]NP.
CL:MAN Peter

"Peter."

C [*naj ∅]NP
CL:MAN

"Him MALE."

There are exceptions, where classifier systems allow a ‘classifier only’ use without an overt noun at all. However, the consulted sources hardly ever describe this phenomenon, which leaves a considerable ambiguous reading potential and thus the correct identification of the intended referent depends on the context (cf. 6.43).

(6.43) Mandarin (cf. Tang 2015: 19)

A wo mai e yi Ben
1SG buy PFV one CL:VOLUME

"I bought one BOOK-SHAPED THING."

B wo na le yi zhang
1SG take PFV one CL:2D

"I took one TWO-DIMENSIONAL FLAT THING."

In concordial class systems, which feature multiple marking (cf. 2.67 B = 6.44 A), the class marking is allowed to transgress the NP-borders (e.g. in predicative constructions; cf. 6.44 B) and even clausal borders (cf. 6.44 C).

(6.44) Latin

A me-a maxim-a culp-a
POSS:1SG-ABL.F.SG biggest-ABL.F.SG fault-ABL.F.SG

"through my most grievous fault"

B aue [Maria]NP [grati-a plen-∅]NP [...]
ha1 Mary:VOC.F.SG grace-ABL.F.SG full-VOC.F.SG

"Hail Mary, full of grace..."

C [... ] [benedict-a tu in multieribus]VP
blessed-VOC.F.SG 2SG in woman:ABL.F.PL

"... blessed art thou among women."

Class marking may also occur on the full verb in cases where the controller noun does not overtly occur in the direct or prior context (cf. 6.45 and 6.46).
6.4. Locus Operandi

(6.45) Burushaski (cf. Munshi 2006: 164)

\[ k\hat{h}: \quad \text{menan} \quad e\text{-sqan-uman} \]
\[ \text{DEM.CLOSE} \quad \text{who.SG.ANIMATE} \quad 3\text{SG.M-kill-PST3PL} \]

"Whom did these people kill?"

(6.46) German

\[ \text{Wußtest} \quad \text{Du, daß die new-e} \]
\[ \text{know:IPF2SG} \quad 2\text{SG} \quad \text{that ARTDEF-NOM.F.SG} \quad \text{new-NOM.F.SG} \]
\[ \text{(Kollegin)} \quad \text{zum Boxen geht?} \]
\[ \text{colleague:NOM.F.SG to boxing:DAT.N.SG go:PRS3SG} \]

"Did you know that the new (colleague) takes boxing lessons?"

Class markers may also function as nominalisers, which leads in rare cases to a situation where even whole clauses may be marked for class and refer to an entity themselves (cf. 178 for examples from the ‘hybrid’ Ngan’gityemerri; 6.19 = 178 A).

(6.47) Ngan’gityemerri (cf. Reid 1997: 205f.)

A \[ w\hat{a}= \quad \text{wedimuy ngagadi-tye} \quad w\hat{a}= \]
\[ M \quad \text{little.boy} \quad 1\text{SG.S:go-PST} \quad M \]
\[ \text{[deme-ngi-pi-yiri-tye-ge-di]} \]
\[ 3\text{SGS:AUX-1SG.DIRECT.OBJECT-head-numb-PST-3SG.S:go} \]

“As a young boy, I was ‘shy-was-he’ (i.e. a male who was habitually shy).”

B \[ w\hat{u}r= \quad \text{[gagu wudumbun-ngirr-bul]} \]
\[ 3\text{SG.S:AUX-1PL.EX-cook} \]

“‘meat-cooked-she’ (i.e. the woman who cooked meat for us).”

C \[ mityity \quad \text{wagarri pey-endi} \quad widdibeng-gu \quad yerr=} \]
\[ \text{white.woman} \quad \text{two in.place-SAME} \quad 3\text{PL.S:stand-DL} \quad \text{TREE} \]
\[ \text{[tuyan-gi-merrendi gentyerrmu-gi-baty knife]} \]
\[ \text{something-LEST 2\text{PL.S:AUX-DL-hold knife}} \]

“There are two white women (security guards) both standing there (with) a thing lest you have something like a knife (i.e. a metal detector).”

The available data on Tainae actually does not allow a firm statement on whether or not Tainae marks NP-external modifiers which depend on the controller noun. Among all elements in question, the most obvious one is the predicative adjective, as even the only example provided for a ‘head adjective’ in the consulted source “is not a proper head at all” (cf. Carlson 1991: 31). The only example for a predicative adjective in Carlson (1991) is given in 6.48.

\[ ^5 \text{A marker’s ability to coin nominalisations entails that it is not a mere agreement marker: An agreement marker copies information from its controller noun and marks it on a satellite of the controller noun. In the case of the nominalisations, there is no controller noun from which information could be copied. This means that the marker itself in question must be referential to some degree. Cf. the discussion of agreement in section 6.3 for a more detailed argumentation.} \]
Carlson’s own interpretation of the so-called predicative adjective in (6.48) as “a bad person” instead of a mere interpretation as “bad” raises doubts about the element in question being a true predicative adjective, especially since the presumed adjective is glossed as a noun or verb “sin”. Tainae allows for straightforward nominalisation of e.g. verbs by the affixation of a class marker (Carlson 1991: 118); since the form *finifi* “sin” is not attested a second time in Carlson (1991), I interpret it in (6.48) as a nominalisation and consequently translate the example as “you are a bad person (literally: sinner)” rather than “you are bad”. In the case of class markers combining with verbs, the class markers function as nominalisers and thus create an NP. This is a common cross-linguistic function of class markers (cf. 205 for an example for deverbal nominalisation from Tainae and 6.50 for an example of deadjectival nominalisation from Mayali).

(6.49) Tainae (cf. Carlson 1991: 118)

\[
\begin{array}{ll}
F-ai & popti-ai \\
road-LONG & enter-LONG \hbox{ gone}
\end{array}
\]

“The way to enter was gone.”


A \[\begin{array}{ll}
\text{ngodo} \\
\text{leprous}
\end{array}\]

“A leper”

B \[\begin{array}{ll}
\text{na-} \\
\text{ngodo} \\
\hbox{M leprous}
\end{array}\]

“a leper”

C \[\begin{array}{ll}
\text{ngal-} \\
\text{ngodo} \\
\hbox{F leprous}
\end{array}\]

“A leper woman”

D \[\begin{array}{ll}
\text{gun-} \\
\text{ngodo} \\
\hbox{N leprous}
\end{array}\]

“leprosy”

Even pronouns, which are the most obvious indicators for class marking outside the NP and even clause boundaries, do not constitute a sound example in Tainae, since they can cooccur with an overt head noun that marks the referent’s class instead of the pronoun (cf. 6.51).
6.5. Flexibility

The clearest available example of anaphoric and transclausal class marking in Tainae is found in the translation of the books of Moses and given in (6.51) Tainae (cf. Carlson 1991: 133).

(6.51) Tainae (cf. Carlson 1991: 133)

\[
\text{do-REMOTE.PST-SG.DIFFERENTREFERENT} \quad \text{that-SUBJ} \quad \text{ki} \quad \text{afa-o-we}
\]
\[
\text{kiap-epi} \quad \text{a-po} \quad \text{ntaimau-aniki}
\]

wife.POSS-F ANAPHORA-near arrive-PST

“While he was doing that his wife arrived.”

The clearest available example of anaphoric and transclausal class marking in Tainae is found in the translation of the books of Moses and given in (6.52) Tainae (cf. Carlson 1991: 134).

(6.52) Tainae (cf. Carlson 1991: 134)

\[
\text{nīmti-aka} \quad \text{g} \quad \text{te-mti-ki,\ldots}
\]

“When he said that to him, he said this, (‘If 40 righteous people exist, I will not destroy it.’)”

Even if the available data on Tainae suggests that it is not a frequent phenomenon, transclausal class marking in Tainae appears to be possible, but only within the borders of a noun phrase. The sample thus reflects the predictions of the grammaticalisation hypothesis well.

6.5 Flexibility

The sample shows variation of the flexibility of both classifier and concordial class systems. In the following, the outliers are presented and discussed in the context of the full sample. The sample systems display preferences with respect to their classificatory flexibility: A noun can belong to either one invariant class (the typical case for concordial class systems) or to more than one class (typical for classifier systems). If it belongs to more than one class, this can be lexicalised or a noun can be productively reassigned to the multiple possible classes. Take e.g. German, where the noun *Schild* “shield” is either masculine and denotes a battleship, or neuter and denotes e.g. a traffic sign; the class membership and the related semantics are lexicalised. Another possibility is that a noun belongs to multiple classes, like German *Butter* “butter”, which is either feminine or masculine; both variants have the same semantics. However, the standard High German variant is always feminine, while the Bavarian and Austrian variant may be masculine. A speaker of High German typically refuses the masculine form of *Butter* as unacceptable. Multiple class nouns like German *Butter* thus represent locally restricted variants, which are mutually exclusive and thus do not represent a case of a truly flexible class membership (cf. Spitta-Bey 1880: 123f, for a discussion of the same phenomenon in Masri). A productive reassignment is typically found with classifiers, which may be variably assigned to the same noun in order to highlight certain properties of the nominal referent (cf. 152).
Grammaticalising Properties


A yi zhang hua
   “a painting”
B yi fu hua
   “a painting”
C yi zhen hua
   “a painting”

Epicene nouns work on a similar basis; these are nouns that change their concordial class with respect to the nominal referent (e.g. French un enfant m. “a male child” vs. une enfant f. “a female child”). However, the reassignment there is not free and mostly affects small semantic niches such as the distinction of biological sex in humans and higher animates or the distinction of plants and their fruits and/or products. In cases where only smaller semantic niches such as biological sex and plant–fruit relations are affected by a flexible classification, the system as a whole has been analysed as employing both flexible and rigid class membership. Concordial class systems are generally considered rigid systems, i.e. systems that do not allow their nouns ad hoc class shift, while classifier systems are generally considered to allow a variable classification of nouns. Table 6.15 presents the flexibility of class membership of the sample systems.

Table 6.15: Flexibility the language sample.

<table>
<thead>
<tr>
<th>Korean</th>
<th>Mandarin</th>
<th>Nisors</th>
<th>Bora</th>
<th>Mundurukú</th>
<th>Ngan’gityemerri</th>
<th>Basque</th>
<th>Filian</th>
<th>Hungarian</th>
<th>Jukapik</th>
<th>North Ambrym</th>
<th>Dongo-Ko (CLS)</th>
<th>Boumaa Fijian</th>
<th>Hungarian</th>
<th>Great Abasharesse</th>
<th>German</th>
<th>Hadza</th>
<th>Hausa</th>
<th>Masri</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
</tbody>
</table>

blank: flexible assignment, •: rigid assignment, n: partly flexible assignment, partly rigid assignment
6.5. Flexibility

The grammaticalisation cline is again mirrored well in Table 6.15 in that the classifier systems show a flexible class membership, whereas concordial class systems show a flexible membership to the left and rigid class membership to the right. There are five systems that blur the picture in that they show a (partly) rigid class membership, whereas their neighboring systems show a flexible one; these are the classifier systems of Mandarin (2), Jakaltek (10), and North Ambrym (11) as well as the concordial class systems of Dyirbal (18) and Burushaski (20). The system of Tohono O’odham (14) discriminates an animate and an inanimate class (Saxton 1983: 186), which prevents a flexible classification.

The Mandarin system is a case where both flexible and rigid class membership are found; the rigid cases occur in higher linguistic registers, where the ‘correct’ and invariant classifier of a given noun has to be learned (cf. again 152 where the different classifiers are supposed to be inserted in different registers), whereas in regular spoken language, the usage of classifiers tends to be more free (cf. Zhang 2007: 54ff.). Given the age of the Chinese classifier system, which is documented for some 3400 years (cf. Aikhenvald 2000: 409), it does not come as a surprise that classifier-noun pairs conventionalised and have to be learned. Considering the large inventory and the related complexity of classification, it is conceivable that this highly complex system has not found its way into spoken language fully, but restricts itself to the higher registers.

Jakaltek distinguishes social status classifiers which render the classified referent’s relation towards the speaker; this part of the classifier set can be flexibly assigned. The other part of the inventory denotes physical properties of the classified, and these categories are semantically mutually exclusive. Therefore, this part of the inventory is assigned rigidly. North Ambrym’s classifier system displays a largely rigid class membership; however, a number of nouns can take a variety of classifiers (cf. 6.54 and Franjieh 2016: 8ff.).

(6.54) North Ambrym (cf. Franjieh 2012: 201)

A Mwene-ng ol
   CL:RESIDUE-1SG coconut
   “my (coconut as) copra”

B Ye-ng ol
   CL:EDIBLE-1SG coconut
   “my coconut for eating”

C Mwe-ng ol
   CL:DRINKABLE-1SG coconut
   “my coconut for drinking”

There are a number concordial class systems which employ semantically flexible niches: Dyirbal (18), Burushaski (20), Ingush (28), Chichewa (29), Hadza (31), and Hausa (32). Dyirbal, Ingush, Hadza, and Hausa all have a number of common class nouns, which adjust their concordial class with respect to the sex of the animate referent (cf. Dixon 1968: 110ff., Nichols 2011: 145, Edenmyr 2004:...
5 and 9, and Newman 2000: 200). Cf. \[6.55\] and \[2.7\] = \[6.56\] for examples of flexibility from Dyirbal and Hausa.


A  
\[\text{yala-∅∅ diban-∅ ba-ŋgu-∅ yibi-ŋgu buran}\]
there-ABS-IV stone-ABS there-ERG-I person-ERG look.at
“The man there is looking at the stone here.”

B  
\[\text{bala-∅∅ diban-∅ ya-ŋgu-n yibi-ŋgu buran}\]
there-ABS-IV stone-ABS here-ERG-II person-ERG look.at
“The woman here is looking at the stone there.”


\[\text{w¯a¯d¯a,}\]
dwarf:MF.SG
\[\text{w¯a¯w¯a,}\]
fool:MF.SG
\[\text{zàbiy¯a,}\]
albino:MF.SG
\[\text{kùm¯a}\]
feeble.person:MF.SG

“dwarf, fool, albino, feeble person” (male or female depending on class agreement)

Hadza furthermore has some forty pairs of homophonous nouns which are distinguished in class and denote semantically related, yet distinct entities. The bulk of these appears to differentiate shape (cf. Edeyry 2004: 15). The Burushaski system allows for a classificational flexibility in order to individuate a nominal referent: Depending on the real world referent’s state (continuous vs. individuated), object nouns generally may take either the “concrete” or the “residue” class marker and thus the same noun stem shares a mass notion and individuated instances of masses (cf. Berger 1998: 36ff. and Munshi 2006: 163).

This also applies to trees and their fruits, such as \[\text{báalt}\] (concrete class) “apple” vs. \[\text{báalt}\] (residue class) “apple tree” (cf. Berger 1998: 36).

The Chichewa system features a slightly different kind of flexibility in its concordial class system, which is fairly common for African nominal classification systems: It employs 18 distinct class markers, which are counted as individual singular and plural classes by some (cf. Mchombo 2004: 6), but can be boiled down to ten concordial classes (Watkins 1937: 24). The classes 15–18, respectively 8–10, host infinitives and locatives, and the classes 12 and 13, respectively 7, host diminutives; these class markers may attach to practically all nouns in order to e.g. add a diminutive reading (cf. 31 = 6.57 and Mchombo 2004: 5).

(6.57) Chichewa (cf. Mchombo 2004: 4f.)

A  
\[\text{M-CLASS1.SG lenje ur-CLASS1.SG ánú ur-}
\]
\[\text{SUBJ:CLASS1.SG POSS:2SG SUBJ:CLASS1.SG}\]
\[\text{ja w-ú nithábwala that SUBJ:CLASS1.SG ASSOCIATIVE humor:CLASS10.PL}\]
\[\text{w- a thyal -a mi- kondo.}\]
\[\text{SUBJ:CLASS1.SG PF break FINAL.VOWEL CLASS4.PL spear}\]

“That humorous hunter of yours has broken the spears.”
6.5. Flexibility

Two concordial class systems, the one of Tainae (19) and Maasai (21), lack rigid assignment. Tainae employs a number of eleven semantically specific classes and an additional residue class. The specific classes are discriminating animacy, biological sex, shape features ("cylindrical", "flat", "long", and "thin and flexible"), and state/inherent nature features ("fluid", "tool", "indeterminate (frozen class)", "girl, rain (frozen class)"); cf. Carlson 1991: 23). A special property of the Tainae class marking indicates its use as a pragmatic tool for focussing discourse referents: Nouns are typically only marked for their class, when "a specific person or thing is in mind, and not when just talking about something generally" (Carlson 1991: 27). If reference to a thing in general is made, the so-called ‘plural’ marker -i is attached (cf. 212).


B  
\[ \text{Ku-}\text{nai-} \text{Wi-} \text{nai-e} \text{ne-ate} \text{naon} \]  
CLASS12.SG nai 1PL work-RESIDUE do-SG about  
“This is about our garden work.”

This means that the so-called ‘plural’ marker refers to entities of all numbers, which suggests a reanalysis of this marker as indeterminate for the category of number, but as a residue class marker. As 213A illustrates, nouns that are marked for a certain class can also be interpreted as a generic form which refers to more than just one real life referent. 213B demonstrates the same noun *apifa* “troll” being marked for the masculine class.

(6.59) Tainae; cf. Carlson 1991: 40 and 45)

A  
\[ \text{Mi} \text{apifa-e-ye} \text{Wi-nk-i} \text{wink-e} \text{te-ye} \]  
SR.SEQ troll-RESIDUE-SUBJ 3DL/PL-POSS-EP work-RESIDUE this-SUBJ  
“This is what trolls do.”

B  
\[ \text{Mi} \text{apifa-3mwa-mi-aksi} \text{wink-e} \text{te-ye} \]  
SR.SEQ troll-3SG.PL-3.POSS work-PL this-SUBJ  
“This is what trolls do.”

In Maasai, there is a small number of nouns that can take only the feminine gender. These comprise *enk-áí “god”, en-kiriŋó “meat”, en-kímná “fire”, and

---

6 The only noun that is listed as exclusively masculine is *3l-aláshè “brother”, while not all speakers object to its the feminine equivalent *enk-aláshè (cf. Payne 1998: 165f. and 169).
kulé “milks” (which is in its collective function unmarked for gender, but takes feminine agreement; Payne 1998: 169). The “vast majority of common nouns […] can choose either the masculine or the feminine gender prefix” (Payne 1998: 159f.). There are several semantic criteria involved in this flexibility: The biological sex of the real life referent is mirrored by the respective assignment to the masculine or feminine gender (cf. [6.60] and Payne 1998: 165f.).

(6.60) Maasai (cf. Payne 1998: 166)

\[
\begin{align*}
\text{\texttt{enk-apatáni}} & \quad \text{\texttt{\textendash-apatáni}} \\
\text{F.SG-wife.parent} & \quad \text{M.SG-wife.parent} \\
\text{“wife's mother, wife's father”}
\end{align*}
\]

A noun’s concordial class can also be related to size or strength (Payne 1998: 166). The masculine class is hereby related to bigger size or strength, the feminine class to smaller size and weakness (cf. [6.61] and Tucker and Tompo Ole Mpaayei 1955: 3).


\[
\begin{align*}
\text{\texttt{enk-abááni}} & \quad \text{\texttt{\textendash-abááni}} \\
\text{F.SG-doctor} & \quad \text{M.SG-doctor} \\
\text{\texttt{en-\textendash-anásh`}} & \quad \text{\texttt{\textendash-anásh`}} \\
\text{F.SG-sister} & \quad \text{M.SG-sister} \\
\text{\texttt{en-títo}} & \quad \text{\texttt{\textendash-títo}} \\
\text{F.SG-girl} & \quad \text{M.SG-girl} \\
\text{\texttt{enk-\textendash-anásh`}} & \quad \text{\texttt{\textendash-anásh`}} \\
\text{F.SG-doctor} & \quad \text{M.SG-doctor} \\
\text{\texttt{en-\textendash-anásh`}} & \quad \text{\texttt{\textendash-anásh`}} \\
\text{F.SG-sister} & \quad \text{M.SG-sister} \\
\text{\texttt{en-títo}} & \quad \text{\texttt{\textendash-títo}} \\
\text{F.SG-girl} & \quad \text{M.SG-girl} \\
\text{\texttt{en-kitók}} & \quad \text{\texttt{\textendash-kitók}} \\
\text{F.SG-respected} & \quad \text{M.SG-respected} \\
\text{“female/small doctor (or quak) vs. male doctor, sister vs. very large sister, girl vs. ‘large shapeless hulk of a woman’, weak brother vs. brother, wimpy male donkey vs. male donkey; woman vs. very respected man”}
\end{align*}
\]

The feminine appears to be the default class, which is e.g. assigned when a noun’s original class is unknown or if a noun is borrowed (Payne 1998: 172). Nominalisations are also usually feminine (Payne 1998: 169f.), except for deverbal abstracts that denote negative concepts (while the concept of negativity can also occur with feminine abstracts; cf. [144]).


\[
\begin{align*}
\text{\texttt{en-m-\textendash-búátá}} & \quad \text{\texttt{\textendash-m-\textendash-búátá}} \\
\text{\texttt{\textendash-m-\textendash-míshírár`}} & \quad \text{\texttt{\textendash-m-\textendash-míshírár`}} \\
\text{\texttt{\textendash-m-\textendash-rokán}} & \quad \text{\texttt{\textendash-m-\textendash-rokán}} \\
\text{\texttt{\textendash-m-kíbá}} & \quad \text{\texttt{\textendash-m-kíbá}} \\
\text{F.SG-extraction} & \quad \text{F.SG-extraction} \\
\text{F.SG-branding} & \quad \text{F.SG-branding} \\
\text{F.SG-blackness} & \quad \text{F.SG-blackness} \\
\text{F.SG-hate} & \quad \text{F.SG-hate} \\
\text{“extraction, branding, blackness, hate”}
\end{align*}
\]

\[
\begin{align*}
\text{\texttt{en-\textendash-lôm}} & \quad \text{\texttt{\textendash-lôm}} \\
\text{\texttt{\textendash-m-ená}} & \quad \text{\texttt{\textendash-m-ená}} \\
\text{M.SG-envy} & \quad \text{M.SG-envy} \\
\text{M.SG-contempt} & \quad \text{M.SG-contempt} \\
\text{“jealousy, contempt”}
\end{align*}
\]

Nominalisations with reference to concrete entities are more likely to occur with either gender (e.g. \texttt{\textendash-løm “big eraser”, en-\textendash-løm “eraser”}). As a rule, most nouns with an unmarked, fixed gender can be shifted, if the abovementioned
semantic mechanisms are adopted (Payne 1998: 171f.). There is also a number of nouns which occur in either gender without any addition of pejorative semantics (Payne 1998: 173).

The sample in general shows a preference for classifier systems to employ flexible classification and for concordial class systems to employ inflexible classification. It is however not uncommon for a system to deviate from this preference.

6.6 Assignment

The manners of assignment in the sample largely correspond to the grammaticalisation cline. Classifier systems prevailingly display a speaker-sided assignment and concordial class systems a system-sided assignment. However, a number of systems deviates from this preference (cf. also table 6.16). The classifier system of Mandarin (2) shows a rigid and lexicalised classifier-noun pairing in higher registers, which coexists with a free and speaker-sided assignment in lower registers; the classifier system of Great Andamanese (9) also shows a partially lexicalised and partially productive assignment, whereas the classifier systems of Jakaltek (10), North Ambyrm (11), and Tohono O’odham (14) display a system-sided assignment. The concordial class systems of Dyirbal (18), Tainae (19), Burushaski (20), and Sumerian (22) allow a speaker-sided assignment. The concordial class systems of Maasai (21) and Tamil (24) employ both a speaker- and a system-sided assignment.

<table>
<thead>
<tr>
<th>Language</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>Naso</td>
<td></td>
</tr>
<tr>
<td>Bora</td>
<td></td>
</tr>
<tr>
<td>Mundurukú</td>
<td></td>
</tr>
<tr>
<td>Ngan’gityemerri</td>
<td></td>
</tr>
<tr>
<td>Boumaa Fijian</td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td></td>
</tr>
<tr>
<td>Great-Alandese</td>
<td></td>
</tr>
<tr>
<td>Jakaltek</td>
<td></td>
</tr>
<tr>
<td>North-ambrym</td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
</tr>
<tr>
<td>Moro</td>
<td></td>
</tr>
<tr>
<td>Japnok</td>
<td></td>
</tr>
<tr>
<td>Chichewa</td>
<td></td>
</tr>
<tr>
<td>German</td>
<td></td>
</tr>
<tr>
<td>Hadza</td>
<td></td>
</tr>
<tr>
<td>Hausa</td>
<td></td>
</tr>
<tr>
<td>Masri</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.16: Assignment in the language sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyirbal</td>
<td></td>
</tr>
<tr>
<td>Tainae</td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td></td>
</tr>
<tr>
<td>Maasai</td>
<td></td>
</tr>
<tr>
<td>Sumerian</td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
</tr>
<tr>
<td>Moro</td>
<td></td>
</tr>
<tr>
<td>Japnok</td>
<td></td>
</tr>
<tr>
<td>Chichewa</td>
<td></td>
</tr>
<tr>
<td>German</td>
<td></td>
</tr>
<tr>
<td>Hausa</td>
<td></td>
</tr>
<tr>
<td>Masri</td>
<td></td>
</tr>
</tbody>
</table>

blank: speaker-sided, •: system-sided, b: both
The decision whether a system allows a free and speaker-sided assignment or not demands a certain degree of interpretation in many cases, which has to be based on the available data and description. The division which kind of assignment is classifier-like and which is concordial class-like cannot be drawn between semantic assignment on the one hand and form-based assignment on the other hand, as even “gender always has a semantic core: there are no gender systems in which the genders are purely formal categories” (Corbett 1991: 307). While a morphological and phonological assignment thus can be attributed a concordial class-like character without problems, lexical and semantic assignment needs a more finegrained approach, since both classifier and concordial class systems employ a semantic one. The literature does typically not subcategorise semantic assignment and attribute the diverse subtypes to the diverse nominal classification devices. Two semantic assignment types can be postulated on basis of the last decades’ notion of nominal classification systems: Free semantic assignment, i.e. an assignment that is limited only by the semantics underlying the system and the speaker’s own will, and ‘bound’ semantic assignment, which renders an assignment system in which the speaker is responsible for the assignment, while there is not much potential for interactive variation left. I elaborate on this line of conjecture in the following by presenting the systems that deviate from their expected behaviour.

There are classifier systems that do not allow a flexible assignment in form of a speaker-sided manipulability. Take e.g. the Tlingit (12) case, where a binary semantic distinction is made between humans and non-humans. The assignment here is carried out by the speaker due to the utter simplicity and transparency of the classification, yet the narrow semantics of the class “human” triggers a mechanic assignment which is ‘bound’ to the system’s semantic specifications. In contrast, Mandarin (2) allows the speaker to choose from a range of classifiers, the number of which is also delimited due to their semantics, at least in its lower and spoken linguistic registers. Nevertheless, the speaker has a free choice over the classifier he or she considers the most appropriate for insertion in the given context. For instance, “[t]he Chinese noun for ‘snake’ can either be classified by tiáo (one-dimensional) or by zhí (animal)” (Bisang 1999: 23; emphasis in the original) depending on whether the speaker wants to emphasise the snake’s property of being of a long shape or an animal. Cases like this thus are instances of free and speaker-sided assignment in the ‘classifier-like’ sense.

Even flexible classifier-noun pairings can display a preference for one of the possible variants. in the case of Mandarin “snake”, “the classifier tiáo (one-dimensional) is replaced [...] in 19 out of 35 cases by zhí (animals)” (Bisang 1999: 23; emphasis in the original). For reasons of an economic language process, it seems natural that systems tend to conventionalise frequent classifier-noun pairs, which they indeed often do and which may lead to an overall semantic bleaching of the referential class to the degree of idiosyncracy. Again, Mandarin constitutes an example for this phenomenon; cf. 2.38 = 6.63, where the classifier tiao for long and 1-dimensional objects has lexicalised for the noun ku(zi) “pants”, which results in a classification by tiao of all compounds headed
by *ku*(zi), even if the denoted nominal referent is not long, as in the case of briefs and shorts (cf. 6.63 B and D).


<table>
<thead>
<tr>
<th>A</th>
<th>yi</th>
<th>biao</th>
<th>nei</th>
<th>underwater</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>yi</td>
<td>biao</td>
<td>sanji</td>
<td>brief</td>
</tr>
<tr>
<td>C</td>
<td>yi</td>
<td>biao</td>
<td>youyong</td>
<td>swimming.trunk</td>
</tr>
<tr>
<td>D</td>
<td>yi</td>
<td>biao</td>
<td>duan</td>
<td>shorts</td>
</tr>
</tbody>
</table>

Great Andamanese (9) also employs both a speaker- and a system-sided assignment, which is semantics-based. The set is divided into seven categories, which relate to body partitions and their semantic extensions. These are semantically largely not overlapping (cf. table 41).

Great Andamanese (9) also employs both a speaker- and a system-sided assignment, which is semantics-based. The set is divided into seven categories, which relate to body partitions and their semantic extensions. These are semantically largely not overlapping (cf. table 41).


<table>
<thead>
<tr>
<th>class</th>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a-</td>
<td>“mouth and its semantic extensions” (9%)</td>
</tr>
<tr>
<td>2</td>
<td>er-</td>
<td>“major external body parts” (36%)</td>
</tr>
<tr>
<td>3</td>
<td>oŋ-</td>
<td>“extreme ends of the body (e.g. toes, fingernails)” (6%)</td>
</tr>
<tr>
<td>4</td>
<td>ut-</td>
<td>“bodily products, part-whole relationships” (13%)</td>
</tr>
<tr>
<td>5</td>
<td>e-</td>
<td>“inner organs” (17%)</td>
</tr>
<tr>
<td>6</td>
<td>ara-</td>
<td>“parts designating round shape and sexual organs” (12%)</td>
</tr>
<tr>
<td>7</td>
<td>o- ɔ-</td>
<td>“parts for legs and related terms” (7%)</td>
</tr>
</tbody>
</table>

Idiomatic categorisation occurs for instance on adverbs, which “do not give near transparency of original body division semantics” (Abbi 2011: 776). This argues for a lexicalised categorisation in the adverbial domain, while the Great Andamanese data in general includes “a substantial amount of variation of forms noted among speakers” (Abbi 2011: 761). The general assignment in Great Andamanese is thus speaker-sided, and complemented by system-sided assignment in specific domains such as adverbs.

The noun classifiers of Jakaltek (10) are semantically highly transparent and constitute a smallish set, which can be divided into two subsets of twelve classifiers respectively. One set accounts for inherent physical properties (e.g. stones or objects made from stone), the other set for the social relation of a
person to the speaker. Taking into account its small size and straightforward semantics, it is standing to reason to postulate a linguistically more economic and speaker-sided assignment. While the speaker is the assigning agent, Jakaltek does not allow for a free assignment in the physical property set, since the system’s categories are in complementary distribution and mutually exclusive (cf. again table 6.9 on page 188, where the semantic categories are listed). Craig (1986a: 266f., 275ff., and 281) provides a number of examples that verify the lack of class variation in the physical property set: For instance, ch’ew “ice” and sañb’ut “hail” are assigned to the class ch’en “rock”, and not to the class ha’ “water”, since they share the physical properties of rocks; the theoretically overlapping categories metx’ “dog (domesticated)” and no’ “animal” are mutually exclusive. Craig (1986a: 276) suggests a list of four strategies that apply to speaker-sided assignment in Jakaltek:

1. natural classification by material substance (unmarked case).
2. natural classification by perceptual analogy (limited to rock and corn class).
3. totemic classification of the objects as members of the spiritual world in the social interaction subsystem (e.g. sun, moon, lightning, crosses, mountains, corn and wind)
4. absence of classification

The absence of classification occurs when the semantic principles are in conflict, e.g. with objects that are of a mixed consistency or are composed of several or unidentified substances (e.g. garbage, smoke, clouds, and beer), or when the noun’s semantics provide no ‘jacking point’ for their categorisation (e.g. abstracts; cf. Craig 1986a: 273 and 276).

In contrast, the classificatory subset that serves the categorisation for humans specifically allows a variable and speaker-sided classification of a given nominal referent: On the one hand, a given member of a social group is related differently to different members of this group – inferior to some, superior to others; kin to some, close or distanced towards others. This is reflected in how this person is categorised by these different sorts of people. On the other hand, it is also possible for the speaker of Jakaltek to compliment or insult somebody by categorising this person by means of an ‘inappropriate’ classifier, which either adds an expression of respect (or kinship, if the addressed is unrelated, but close to the speaker) or strips him or her off a respectful connotation (Craig 1986a: 269f.).

The possessive classifiers of North Ambrym (11) comprise a set of five categories, which account for alienable objects. They correspond well to the semantic structure of many other Polynesian possessive classifier systems, which are often referred to as ‘relational classifiers’ (cf. table 117).\footnote{Cf. the discussion and reanalysis of ‘relational classifiers’ on the pages 28ff. and in Passer (2015a).}
Table 6.18: The inventory of North Ambrym classifiers (cf. Franjieh 2012: 244 and 2014: 5).

<table>
<thead>
<tr>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>a(n)</td>
<td>animals, tools, units of time, some trees and kin terms</td>
</tr>
<tr>
<td>ma(n)</td>
<td>liquids, containers of liquids, buildings, holes, mats</td>
</tr>
<tr>
<td>bo(n)</td>
<td>fire, firewood</td>
</tr>
<tr>
<td>to(n)</td>
<td>baskets</td>
</tr>
<tr>
<td>muen(a)n</td>
<td>residue</td>
</tr>
</tbody>
</table>

The classifier assignment in North Ambrym is originally semantic, and classifier-noun pairs have conventionalised. This results in a rigid and overall inflexible categorisation, which cannot be manipulated by the speaker (Franjieh 2016: 24) by e.g. reassignment to the residue category.

The system of Tohono O’odham (14) displays straightforward and semantic assignment criteria in that it distinguishes an animate and an inanimate class (Saxton 1983: 186). The speaker can carry out the assignment, but is bound to the mutually exclusive class semantics.

Let us now turn to the concordial class systems: Both Maasai (21) and Tamil (24) allow for a both speaker- and a system-sided assignment. Maasai has a sex-based, binary concordial class system: The masculine class further carries a notion of size and strength, the feminine class one of smaller size and weakness and is further considered the default class for e.g. nominalisations and borrowings (Payne 1998: 166 and 172 and Tucker and Tompo Ole Mpaayei 1955: 3). Due to the size and strength connotation, “the vast majority of Maasai noun roots can occur with either a feminine or a masculine prefix” (Payne 1998: 165). A shift from the masculine to the feminine class or vice versa can also add a pejorative meaning to the referent (cf. 6.64). While the system imposes a certain rigidity on the natural class membership of Maasai nouns, the recategorisation can be carried out by the speaker according to his or her will.

(6.64) Maasai (cf. Payne 1998: 166)

A  en-  títo
   F3G  girl
   “girl”
B  ol-  títo
   M3G  girl
   “large shapeless hulk of a woman”

The concordial class assignment of Tamil distinguishes a rational class (also labeled “superior”, “high caste”, “human”, and “animate”) and a nonrational class (also labeled “inferior”, “low caste”, “nonhuman”, and “inanimate”; cf. Andronov

(6.65) Tamil (cf. Andronov 2004: 147)

A nall-ēn, nall-ēm
  good-1SG  good-1PL.
  “good I, good we”
B nall-ay, nall-ār
  good-2SG  good-2PL.
  “good thou, good you”
C nall-ān, nall-āl, nall-ār
  good-3SG.M  good-3SG.F  good-3PL.RATIONAL
  “good man, good woman, good people”
D nār-ru, nall-a
  good-3SG.NONRATIONAL, good-3SG.NONRATIONAL
  “good thing, good things”

The class assignment is entirely semantically based, carried out by the speaker, and determined by the nominal referent, which naturally results in a number of epicene nouns (cf. 221).

(6.66) Tamil (cf. Asher 1982: 136)

A veelekaar, paqkaark, pickekar
  worker:M  rich:M  beggar:M
  “male worker, rich man, beggarman”
B veelekar, paqkkaari, pickekkar
  worker:F  rich:F  beggar:F
  “female worker, rich woman, beggarwoman”

Since nouns referring to children or animals are usually members of the nonrational class, but can be reassigned with respect to their biological sex, the speaker has a certain influence on the class assignment. Therefore, both a speaker- and a system-sided assignment system are found in Tamil.

The system of Sumerian employs a distinction of animacy in the form of a common and a neuter class. The Sumerian class assignment is rigid: With terms for slaves and groups of people being the only exception, every noun displays no class variation and thus “does not betray its class” (Edzard 2003: 29 and Jagersma 2010: 102).

Dyirbal is probably one of the most well-known examples for semantic class assignment, since its system’s semantics has been extensively described in Dixon (1968) and served as a reference for many later works on nominal classification. Due to this and the limitations of this section, I refer to Dixon (1968: 119f.) in stating that the Dyirbal speakers at the time of the data collection (1963) had not to learn a noun’s class individually, but were able to assign new or unkown words intuitively and uniformly to a certain class. This semantic
and speaker-sided assignment was thus productive, yet bound. However, note that the stage of Dyirbal which was included in the sample is not the most recent documented one: Only one generation of speakers later (1982), it had been massively restructured in order to maintain its transparent semantics (Lakoff 1990: 97 and Schmidt 1985). This shift is studied in section 8.2.

The Tainae (19) system assigns semantically, both bound and free. Its eleven classes distinguish animacy (masculine, feminine, and animate), shape (cylindrical, flat, long, thin/flexible), two more specific classes (fluid, tool), and two frozen classes (one for the nouns for “girl” and “rain” and one indeterminate class; cf. Carslon 1991: 23). There are epicene nouns for humans (e.g. at-o in. “grandfather” vs. at-epi f. “grandmother”; Carlson 1991: 25), and lexicalised assignment (e.g. apak-2 “woman” (animate class) and the two frozen classes). However, the shape classes constitute the potential of flexibility for the assignment: For instance, ikyipw-ina (flat) “leaf” refers to a ‘normal’ leaf, which could be rolled up and referred to by ikyipw-iwa (cylindrical) “leaf (roll)”. A strip of that very leaf could be referred to by ikyipw-iit (flexible) “leaf (strip)”, a stiff, folded rod from that leaf could be referred to by ikyipw-ai (long) “(folded) leaf (rod)”, and if one is tired of playing with the leaf and crumbles it to a ball, one throws away an ikyipw-ipi (fluid) “leaf (ball)” (all examples from Carlson 1991: 27f.). Tainae thus possesses both free and bound semantic assignment.

The system of Burushaski (20) distinguishes four classes, one for human males, one for human or superhuman females, one for concrete entities (i.e. all non-human countable objects and animates), and one residue class (Berger 1998: 33ff.). This system is characterised by a high semanticity in that it allows for a speaker-sided reassignment of noun roots, which can be free or bound. In the case of nouns referring to humans, the speaker can add a pejorative connotation in assigning them to the concrete class. Furthermore, the concrete and residue class often share noun stems in order to express a mass-count distinction or a fruit-plant distinction: For instance, the noun bâalt refers to apples if the noun is assigned to the concrete class, but to the apple tree if it is assigned to the residue class. Similarly, the noun ˙gunō bears the meaning “seed” if assigned to the concrete class, but means “semen” if assigned to the concrete class (Berger 1998: 36). The basic distinction of the Burushaski is thus between a supercategory for humans and a supercategory for non-humans, the latter of which displays a lot of class fluidity.

This section has shown that flexibility in the sample largely coincides with the expectations from the grammaticalisation hypothesis: With only few exceptions, classifier systems allow for a largely flexible assignment, while concordial class systems prefer a rigid assignment.

6.7 Semantic Contribution

Classifiers are expected to semantically contribute to the noun phrase that hosts them. Concordial class systems are not expected to always contribute to
the noun phrase, though they virtually always employ semantically productive niches (such as the distinction of biological sex). While many systems correspond to this expectation, table 6.19 illustrates that this distinction cannot be drawn strictly between the two systemic types: There are both classifier systems which do not obviously semantically contribute to their host NP, as well as concordial class systems which show a high degree of semantic productivity.

Table 6.19: Semantic contribution to the NP in the language sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
</tr>
<tr>
<td>Nias</td>
<td></td>
</tr>
<tr>
<td>Bora</td>
<td></td>
</tr>
<tr>
<td>Mundurukú</td>
<td></td>
</tr>
<tr>
<td>Nga-giyamantu</td>
<td></td>
</tr>
<tr>
<td>Ramaoa Tiwan</td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td></td>
</tr>
<tr>
<td>Jeti</td>
<td></td>
</tr>
<tr>
<td>North Ambrym</td>
<td></td>
</tr>
<tr>
<td>Dyurgyo-ko (CAS)</td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
</tr>
<tr>
<td>Hmong Daw</td>
<td></td>
</tr>
<tr>
<td>Nùng</td>
<td></td>
</tr>
<tr>
<td>Dyirbal</td>
<td></td>
</tr>
<tr>
<td>Tainae</td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td></td>
</tr>
<tr>
<td>Maasai</td>
<td></td>
</tr>
<tr>
<td>Sumeiran</td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
</tr>
<tr>
<td>Moro</td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td></td>
</tr>
<tr>
<td>Ingush</td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
</tr>
<tr>
<td>German</td>
<td></td>
</tr>
<tr>
<td>Hadza</td>
<td></td>
</tr>
<tr>
<td>Masri</td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td></td>
</tr>
<tr>
<td>Chichewa</td>
<td></td>
</tr>
<tr>
<td>German</td>
<td></td>
</tr>
<tr>
<td>Homa</td>
<td></td>
</tr>
<tr>
<td>Hina</td>
<td></td>
</tr>
<tr>
<td>Maeli</td>
<td></td>
</tr>
</tbody>
</table>

The classifier systems from the sample which lack a clear semantic contribution to their host NP are the ones of Dongo-Ko (13), Tohono O’odham (14), Hmong Daw (15), and Nùng (16). In these cases, it is not a semantic feature of the referent which is chosen to be highlighted by classificatory means, since there is no choice of a semantic feature which is mirrored by the classificatory marker. Due to this lack of competition of classificatory marking, a clear semantic contribution cannot be assumed for these systems, which are discussed individually in the following.

The Dongo-Ko possessive classifier system distinguishes three more general categories ((m)bá “property”, dà “house”, and nɔ “animate”) from seven body part classifiers (kɔ “body”, mɔ “head”, lɔ “head”, etc.; Pasch 1986: 246). Variability is present in the Dongo-Ko possessive classification: For instance, nɔ “animate” may refer to any body part and replace the respective body part classifiers (cf. 6.67 and Pasch 1986: 249).
6.7. Semantic Contribution

(6.67) Dongo-Ko (cf. Pasch 1986: 249 and 251)

A  l-i l-a n-{}, r-{}
   eye-3:SG 3:SG-POSS  CL:ANIMATE  1SG
   “my eye”

B  l-i-s-{}, s-{}, l-{}
   “the dog’s eyes”

However, neither the available examples from Dongo-Ko nor the classificatory semantic structure suggests that a classifier adds semantic content to its host NP. Instead, the classifiers reflect semantic features that are present in the noun. The possessive classifiers of Tohono O’odham consist of two categories, an inanimate one, which is marked by șoi-, and an animate one, which is marked by șoi-. The markers occur either with the possessive marker for alienables (-g(a) or -ka) or with the bare noun (Saxton 1983: 183 and 186f.; cf. 5.7 = 6.68).

(6.68) Tohono O’odham (cf. Langacker 1977: 91f.)

A  huan  goys  șoi-ga
    Juan  dog  CL:ANIMATE-POSS.ALI
   “Juan’s dog”

B  huan  uns  iɾi-ga
    Juan  tree  CL:GENERAL-POSS.ALI
   “Juan’s tree”

The nominal referent determines the class membership, and class membership does not highlight a property of the referent. Therefore, the Tohono O’odham classifiers do not semantically contribute to the semantics of the noun phrase.

The numeral classifier systems of Hmong Daw and Nùng also do not semantically contribute to the noun phrase. Their inventories are fairly comparable in that both employ a number of more general categories and a few more specific, mostly shape-based classes (cf. table 6.20).
Table 6.20: The Nùng and Hmong Daw classifiers (cf. Saul and Wilson 1980: 25ff.)

<table>
<thead>
<tr>
<th>Nùng</th>
<th>Tohono O’odham</th>
</tr>
</thead>
<tbody>
<tr>
<td>cáh</td>
<td>lab</td>
</tr>
<tr>
<td>tū</td>
<td>tus</td>
</tr>
<tr>
<td>ánh</td>
<td>leej</td>
</tr>
<tr>
<td>dāhn</td>
<td>raḥ</td>
</tr>
<tr>
<td>tēu</td>
<td>tzhais</td>
</tr>
<tr>
<td>bāh</td>
<td>dam</td>
</tr>
<tr>
<td>cōn</td>
<td>troj</td>
</tr>
<tr>
<td>mōi</td>
<td></td>
</tr>
<tr>
<td>hōnh</td>
<td></td>
</tr>
<tr>
<td>phōhn</td>
<td></td>
</tr>
<tr>
<td>maht</td>
<td></td>
</tr>
</tbody>
</table>

Like in the cases of Dongo-Ko, and Tohono O’odham, a semantic contribution to the noun phrase is not present in the four more general classes ("general", “animate”, “human”, and “inanimate/artefacts”). In those cases where classifier variation is possible, it seems not to alter the semantics of the noun phrase as a whole (cf. 90 = 95); this does not mean however that the classes themselves are semantically empty, but that the classifier itself does not add semantics to the noun phrase.


A

<table>
<thead>
<tr>
<th>sb</th>
<th>bug</th>
<th>neeg</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>CL:ANIMATE</td>
<td>man/person</td>
</tr>
<tr>
<td>“a man/person”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>leej</th>
<th>neeg</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL:HUMAN</td>
<td>man/person</td>
</tr>
<tr>
<td>“the man/person”</td>
<td></td>
</tr>
</tbody>
</table>

The remaining non-shape based classifiers (“clothing”, “bedding”, “fabric”, and “one of a natural pair”) display semantics which also do not suggest an active semantic contribution in the context of classification, but to reflect a property of the nominal referent, which is invariably related to it. While the shape-based classes have the potential of contributing semantically to their host NP, there is no data available for the actual use and variability of these. Even if it cannot be excluded at this point that these classifiers semantically contribute to their morphosyntactic environment, the remaining classifiers of the systems of Nùng and Hmong Daw do not. Therefore, an overall semantic contribution is not present in these languages.

The concordial class systems that display a higher degree of semantic productivity are the ones of Dyirbal (18), Tainae (19), Burushaski (20), Maasai (21), and Tamil (24). Dyirbal allows the reassignment of whole parts of the nominal lexicon with respect to two conditions:
6.7. Semantic Contribution

“(1) if some noun has characteristic X (on the basis of which its class membership would be expected to be decided) but is, through belief or myth, connected with characteristic Y, then generally it will belong to the class corresponding to Y and not that corresponding to X.

(2) if a subset of nouns has some particular important property that the rest of the set does not have, then the members of the subset may be assigned to a different class from the rest of the set, to ‘mark’ this property; the important property is most often ‘harmfulness’.”

(Dixon 1968: 120 and 1982: 179; cf. also Lakoff 1990: 94)

As the reassignment serves the emphasis of certain inherent properties, a semantic contribution is involved in this class shift. The Tainae system allows its speaker to assign nouns to classes even more freely, e.g. by varying shape categorisation of an object, the shape of which is manipulated (cf. 5.32 = 6.84); the respective classifier adds information on the shape of the object referred to and thus contributes semantically to the noun phrase.

(6.70) Tainae (Carlson 1991: 27f.)

A  skewp-we-ina
    leaf-FLAT
    “leaf”

B  skewp-we-iv
    leaf-CYLINDRICAL
    “leaf (roll)”

C  skewp-iti
    leaf-THIN
    “leaf (strip)”

D  skewp-ai
    leaf-LONG
    “(folded) leaf (rod)”

E  skewp-ip
    leaf-FLUID
    “leaf (ball)”

In addition to a distinction of biological sex, the Burushaski system allows for a productive mass-count distinction by class assignment (cf. Berger 1998: 36ff. and Munshi 2006: 162ff.) and thus also contributes semantically to the noun phrase. Since this distinction affects a large range of object nouns, the Burushaski system proves a great degree of semanticity.

In Maasai, class membership does not only serve the distinction of biological sex, but also encodes size or strength (cf. Payne 1986: 166). Class assignment also serves the function of pejorisation (cf. again 6.64 on page 220). Therefore, the Maasai concordial class system clearly contributes to the noun phrase’s semantics.
Tamil finally distinguishes a masculine, a feminine, and a neuter (or ‘nonrational’) class. Masculine and feminine are summarised in a ‘rational’ superclass. ‘Nonrational’ nouns referring to a ‘rational’ entity are freely reassigned to the ‘rational’ class and vice versa. The new class adds information on the animacy and sex of the referent (cf. 220), which allows the conclusion that the Tamil concordial class system contributes to the noun phrase’s semantics.

(6.71) Tamil (cf. Andronov 2004: 59)

A  člai
   poor.man:M
   “poor man”

B  člai
   poor.woman:F
   “poor woman”

C  člai
   stupidity:N
   “stupidity”

In conclusion, classifier systems often display a semantic contribution to the noun phrase, while concordial class systems most often lack a semantic contribution to the noun phrase or restrict their contribution to semantic niches. This preference is however blurred by a few deviations for each systemic type respectively, and the kind of semantic distribution is different for the individual systems. This suggests that the notion of this property is too wide to be usefully applied, and that its various forms ought to be distinguished for future studies. After the discussion of the individual properties in this and the prior sections, section 6.8 addresses correlations between the properties in the sample.

6.8 Correlations Between Properties

The previous sections have shown that the individual properties of the language sample mirror a clear grammaticalisation cline, and that the sample includes systems at all kinds of intermediate steps between prototypical classifier and concordial class systems. There is a number of deviations with respect to the preferred order of grammaticalising properties, yet these deviations can be analysed and explained. The semantic contribution in its wide notion has proven not to possess an expressive value for the study of the grammaticalisation of nominal classification devices and thus plays no role in the further course of this study. This subsection examines the correlations between of grammaticalising properties in the language sample and their order. The main correlations occur between multiple marking and the locus operandi, a noun’s classificational flexibility and the assigning agent, and the obligatoriness of classification and multiple marking.


Marking and the locus operandi

Multiple marking and the locus operandi, which distinguishes between NP-internal and NP-external classificatory marking, show a correlating pattern: All non-'hybrid' languages that employ multiple marking also mark class membership beyond the borders of the noun phrase. The ‘hybrids’ are reanalysed as classifier systems lacking multiple marking and instead employing multiple classification; they attach their class markers to NP-external elements such as the verb and thus stand out among the other classifier systems. Table 6.21 displays the correlation of the locus operandi and systemic marking in the language sample.

Table 6.21: Correlation between multiple marking and locus operandi.

<table>
<thead>
<tr>
<th>Language</th>
<th>Korean</th>
<th>Mandarin</th>
<th>Niuafoʻou</th>
<th>Bora</th>
<th>Munduruku</th>
<th>Ngan’gityemerri</th>
<th>Boumaa Fijian</th>
<th>Hungarian</th>
<th>good-Andamanese</th>
<th>Anikey</th>
<th>North Ambrym</th>
<th>Dongo-Ko (CLS)</th>
<th>Tchenu Ochhina</th>
<th>Hinuwe Diwe</th>
<th>Ningaloo</th>
</tr>
</thead>
<tbody>
<tr>
<td>F3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>F4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Multiple markers may occur inside or outside the borders of the noun phrase where the classified noun occurs. Tamil (24) for instance marks for the subject’s class on the verb (cf. 6.72), and Ngan’gityemerri (6) references objects’ classes on the verb (cf. 184 = 6.73).


A anta kugante kaarile ceran-atu
that child:N car:LOC climb:PST-3SG.N
“That child climbed into the car.”

B iran-ta naay (-kal) vant-atu
two dog:N PL come:PST-3SG.N
“Two dogs came.”
A lack of multiple marking in the sample correlates with NP-internal marking, except for the ‘hybrid’ systems. It is however common for classifiers to function as carriers of anaphora, and thus to occur outside of a categorised head noun’s phrase in discourse (cf. 2.1 – 6.74).

In these cases, classifiers enter the discourse together with an overt noun that is classified; in later discourse stretches, the classifier itself is used anaphorically and lacks an overt noun. I analyse these cases as instances of ellipsis, where the head noun is deleted for a more economic speech. This phenomenon can also be found in concordial class systems (cf. 6.75, where ellipsis occurs with the marking of adjectives in headless noun phrases, but marking occurs outside the NP with pronouns).

The ‘hybrid’ systems are the only sample-internal systems that do not employ multiple marking, but attach their classificatory markers to NP-external elements. Cf. 6.37 – 6.40 – 6.76 – 6.79 for examples from all four hybrid systems.


Teni toire' tareuri-ma∅-∅
ART.F children care.for-PRS.HAB-DERIV-CL:F
bua'uri-ma∅-∅
feed-PRS.PROGR-DERIV-CL:F

“The lady who cares for the children (is) the one who is feeding them.”
6.8. Correlations Between Properties

(6.77) Bora (cf. Seifart 2007: 424)

\[ \textit{ká:túfše -hi} \quad \text{(bo:tá -hi)} \]
fall.down CL:DISC button CL:DISC

“(The button,) it fell down.”

(6.78) Mundurukú (cf. Mithun 1986: 381)

\[ \textit{Ti dojot puye, o'-ti-moq ip baseya'a be.} \]
water bring when they-CL:WATER-place they basin in

“When they brought water, they placed it in a basin.”

(6.79) Ngun’gityemerri (cf. Reid 1997: 194)

\[ \textit{ganbi yerr= buy-mem -garri ksde} \]
didgeridoo CL:TREE light 3SGS.do leg where

“Where is that white didgeridoo?”

In summary, all concordial class systems allow their markers to occur multiply and to transcend the noun phrase’s borders, while all classifier systems allow only a single and NP-internal marking. Only the ‘hybrids’ employ a single marking that may transcend the noun phrase’s borders.

**Assignment and flexibility** The assignment and classificatory flexibility show partial correlations in that a free and flexible assignment typically involves a speaker-sided act of assigning, whereas a lexicalised and possibly idiosyncratic assignment is not carried out by the speaker. Table 6.22 shows their distribution over the sample.

**Table 6.22: Correlation between assignment and flexibility.**

<table>
<thead>
<tr>
<th>Language</th>
<th>F6</th>
<th>F7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Mandarin</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Nasioi</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Bora</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Mundurukú</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Ngun’gityemerri</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Bouma Fijian</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Hungarian</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Great Andamanese</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>North Ambrym</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Dong é-Ko (CLS)</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Tahoro O’odham</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Dyirbal</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Tainae</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Burushaski</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Maasai</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Tamil</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Dong é-Ko (CCS)</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Ket</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Ingushí</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Chichewa</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>German</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Hausa</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>Marí</td>
<td>b</td>
<td>b</td>
</tr>
</tbody>
</table>

blank: classifier-like, ●: concordial class-like, n: both;
F6: assignment, F7: semantic contribution
Most languages with inflexible class membership employ a system-sided assignment, and most languages with a flexible assignment employ a speaker-sided assignment. German nouns for instance inherently belong to an invariable concordial class, which has to be learned together with the noun; Korean nouns can take a variety of classifiers, which are selected by the speaker. Mandarin even combines both principles language-internally in that higher linguistic registers demand a fixed classifier-noun pairing, which has to be learned by the speaker, whereas it displays a flexible and speaker-sided assignment in its lower registers (Zhang 2007: 54ff.).

Sumerian (22) does not correspond to this pattern and displays an inflexible, speaker-sided assignment. Its nouns (except for saḫ “male slave” and gemen “female slave”; Jagersma 2010: 103) are inflexible, but can be assigned by the speaker due to the minimalistic and semantically straightforward class inventory, which comprises a common and a neuter class. Dyirbal (18) and Burushaski (20) deviate from the pattern in either employing a speaker-sided semantic assignment, while the class membership is flexible only for parts of the nominal lexicon. Great Andamanese (9), Jakaltek (10), and Maasai (21) deviate in allowing a generally flexible class membership, which is only partly manipulable by the speaker. Nouns denoting animates and humans are variable with respect to their real life referent’s biological sex. Burushaski further distinguishes the state of a real world referent’s individuation by class assignment. The remaining nominal lexicon in these languages displays a rigid class membership. Great Andamanese classifiers and the ‘physical’ classifier set of Jakaltek constitute mutually exclusive classes, which are assigned by the speaker according to the nature of the real world referent (cf. (6.80) and (6.81)); in these cases, the speaker has no freedom of choice of class, but is bound to the assignment of the ‘correct’ classifier.

(6.80) Great Andamanese (cf. Abbi 2011: 764)

A
\[
\begin{align*}
  a\=\text{cal\=a}, & \quad e\=\text{cal\=a}, & \quad o\=\text{cal\=a} \\
  \text{CL:4\=scar} & \quad \text{CL:2\=scar} & \quad \text{CL:3\=scar} \\
  \text{“scar left by arrow-head, scar on the head, scar on the limbs”}
\end{align*}
\]

B
\[
\begin{align*}
  e\=\text{tei}, & \quad a\=\text{tei}, & \quad o\=\text{tei} \\
  \text{CL:5\=blood} & \quad \text{CL:4\=blood} & \quad \text{CL:3\=blood} \\
  \text{“blood inside the body, blood outside the body, blood on/from finger”}
\end{align*}
\]

(6.81) Jakaltek (cf. Craig 1986a: 265, B compiled after Craig 1986a)

A
\[
\begin{align*}
  a\=\text{ca\=b’two-PL\=INANIMATE} & \quad (o)\quad \text{no\=ANIMAL} & \quad \text{\= Carnival} \\
  \text{“two saddles” (inanimate object made of animal product)}
\end{align*}
\]

B
\[
\begin{align*}
  a\=\text{ca\=b’two-PL\=INANIMATE} & \quad (o)\quad \text{te\=ANIMAL} & \quad \text{\= Carnival} \\
  \text{“two chairs”}
\end{align*}
\]

Most Maasai nouns can take either concordial class and are assigned on basis of the speaker’s cognitive construal of a real world referent (Payne 1998: 159f.}
and 168). However, the Maasai nominal lexicon displays a partially lexicalised concordial class membership (ib.: 169) and thus is not assigned by the speaker in all cases.

**Obligatoriness and marking** The obligatoriness of classification and its indexing by multiple markers strongly correlate, which has its reasons in both properties being most crucial and characteristic for the definition of concordial class systems. Concordial class systems are defined as systems that obligatorily classify every noun in every context and mark this class membership by means of concordial marking. Classifier systems on the other hand are defined as systems that classify nouns only in given contexts and mark class membership by an individual, morphosyntactically independent or clitic element. Multiple marking and the locus operandi largely overlap in their distribution over the sample; since multiple marking and obligatoriness correlate in the sample, the locus operandi also shows a correspondence to obligatoriness. Table 6.23 displays their distribution.

Table 6.23: Correlation between obligatoriness, agreement, and the locus operandi.

<table>
<thead>
<tr>
<th>Language</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarín</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Náuáá</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boru</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mundurukú</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ngan’gityemerri</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bóuma Fijian</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Great Andamanese</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Ambrym</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyirbal</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tainae</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maasai</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sumerian</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>German</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhojpur</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chichewa</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hausa</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masri</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingush</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmong</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinh</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmung</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tais</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’odham</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ing</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skolt</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haska</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hmung</td>
<td>33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to predict that a system classifies the nominal lexicon obligatorily, multiple marking is the most reliable criterion, since it is an obligatory morphosyntactic means overtly reflecting an act of classification. This does not exclude a system lacking multiple marking (e.g., since it is of an isolating type)
from the option to obligatorily classify all nouns in all contexts, but predicts that in the case that a classifier system is grammaticalising into a concordial class system, class agreement precedes an obligatory classification (cf. the hybrid systems Nasioi (3), Bora (4), Mundurukú (5), and Ngan’gityemerri (6)). Following this line of argumentation, one would expect a classifier system to expand the range of morphosyntactic classifier hosts, where classifier marking may multiply occur in the extended range of morphosyntactic contexts (~class concord), before classification in general is obligatorised across contexts. Unsurprisingly, all concordial class systems display an obligatory classification, and all except Dyirbal employ multiple marking ‘proper’.

The Dyirbal system lacks ‘proper’ multiple marking in that it marks class membership on the so-called ‘noun marker’, an article-like element that marks the noun’s case, class, and visibility (Dixon 1968: 106; cf. 5.1).

(6.82) Dyirbal (cf. Dixon 1972: 93)

\[
\begin{align*}
\text{bala} & \quad \text{yagumbil} & \quad \text{baygu} & \quad \text{yaranggu} & \quad \text{baygu} \\
\text{there} & \quad \text{woman} & \quad \text{there} & \quad \text{man} & \quad \text{there} \\
\text{yataggu} & \quad \text{balgan} \\
\text{stick:INSTR} & \quad \text{hit:PRS/PST}
\end{align*}
\]

“man is hitting woman with stick [sic]”

Class marking also occurs on the demonstrative and interrogative form of the noun marker (Dixon 1968: 107). However, these do not cooccur in discourse. The class marking thus rather resembles the one of a multiple classifier system such as the one of Mandarin, where a classifier can occur with different morphosyntactic hosts, but only once per context (cf. 147 = 6.83).

(6.83) Mandarin (cf. Sun 2006: 159f.)

\[
\begin{align*}
\text{A} & \quad \text{na} & \quad \text{yi-pi} & \quad \text{he}i & \quad \text{ma} \\
\text{DEM:REMOTE} & \quad \text{one-CL:HORSE} & \quad \text{black} & \quad \text{horse} \\
\text{“That black horse.”}
\end{align*}
\]

\[
\begin{align*}
\text{B} & \quad \text{na} & \quad \text{pi} & \quad \text{he}i & \quad \text{ma} \\
\text{DEM:REMOTE} & \quad \text{CL:HORSE} & \quad \text{black} & \quad \text{horse} \\
\text{“That black horse.”}
\end{align*}
\]

\[
\begin{align*}
\text{C} & \quad \text{mai} & \quad \text{pi} & \quad \text{ma} \\
\text{buy} & \quad \text{CL:HORSE} & \quad \text{black} & \quad \text{horse} \\
\text{“Buy one black horse.”}
\end{align*}
\]

Therefore, and because of the lack of a noun marker in the ablative, allative, and locative case (Dixon 1982: 163), Dyirbal is reanalysed as employing neither obligatory classification nor multiple marking. Similarly, Tainae markers have been shown to be of a referential nature (cf. 5.32 = 6.84) and to occur non-obligatorily (cf. 6.12 = 6.85), which is why it was reanalysed as employing neither agreement nor an obligatory classification.
6.8. Correlations Between Properties

(6.84) Tainae (Carlson 1991: 27f.)

A  
\textbf{skypw-ina}  
leaf-FLAT  
“leaf”

B  
\textbf{skypw-ive}  
leaf-CYLINDRICAL  
“leaf (roll)”

C  
\textbf{skypw-it}  
leaf-THIN  
“leaf (strip)”

D  
\textbf{skypw-ai}  
leaf-LENGTH  
“(folded) leaf (rod)”

E  
\textbf{skypw-ini}  
leaf-FLUID  
“leaf (ball)”


A  
\texttt{to-nai ankwi mante-y sakun-e}  
1SG.M-1PL today school-RESIDUE  
\texttt{ne-at-akwina[...]}  
do-DUR-1PL.DIFFERENT.REFERENT  
“We were doing schoolwork today, monday, when[... ]”

B  
\texttt{as-i yai niwi-ntai-i}  
male.cross.cousin-EPENTHETIC 1DL go-1DL.FUT.UNQUOTE  
“His cousin replied, ‘Let’s both go’.”

C  
\texttt{mi tumpunu stoti at-iwa-mi sta[...]}  
QUOTE ancestor story grandparent-3PL.M-OBJ story  
“The ancestral story, (about how people scared the bushknife which customarily did work by itself)[... ]”

The ‘hybrid’ systems of Nasioi (3), Bora (4), Mundurukú (5), and Ngan’git-yemerri (6) do not obligatorily classify their nouns. If class marking occurs, it resembles multiple marking and is not bound to the borders of the noun phrase; I argue in section 6.3 that each instance represents a classificatory context of its own. The supposed ‘hybrid’ systems thus have shown to correspond well to the notion of classifier systems, without the classifier-like restriction to the noun phrase’s borders.

Order of grammaticalising properties Grammaticalisation is the assumed vehicle for change phenomena in nominal classification. Even though grammaticalisation does not follow a strict script of mechanic changes, grammaticalisation theory entails the prevalence of a preferred order of processes of change, as was laid out in section 3.2. The sample confirms independently from grammaticalisation theory that there is a preferred order of grammaticalising properties.
This becomes evident if one orders the sample languages with respect to an increase of concordial class-like properties. The order of properties that mirrors a grammaticalisation cline most obviously is the one given in table 5.1 on page 120 and is repeated in figure 6.3.

However, the order of changing processes that may be expected from grammaticalisation theory is only partly met in the sample, which allows conclusions about the impact of the individual grammaticalisation processes (semantic bleaching, decategorisation, phonological erosion, cliticisation, and morphologisation). The processes and the related properties of nominal classification systems are addressed individually below.

Semantic bleaching involves a reduction of the class inventory, since individual classes lose their semantic specificity and a smaller number of larger, more abstract classes outcompetes categories which are semantically highly specific and have fewer members. This expectation is met well in the sample, since there are only two classifier and three hybrid systems that employ a large class set, while all other systems employ a system of two dozen classes or less. In contrast, a loss of semantic specificity does not necessarily entail a loss of semantic contribution to the noun phrase, since also an abstract class can semantically contribute to it. Recall, for instance, the Maasai concordial class system, where the speaker is able to shift a noun to the masculine class for the emphasis of strength or size and to the feminine class for the emphasis of a lack of strength or size. The Maasai system also exemplifies that a lower semantic specificity does not necessarily correlate with a more system-sided assignment. Not only do fewer and more abstract categories provide a higher degree of transparency and identifiability of a system’s inventory than a great many semantically specific classes, but also do cues for assignment not have to be semantic, but can also be morphological or phonological. For instance, the system of Bukiyip mainly relies on a phonological assignment, the only regular exception being male and female human beings (cf. Fortune 1942: 6f. and 11). Both these expectations are also met well in the sample, since it is the semantic contribution and the assignment which ranks lowest with respect to the grammaticalisation cline and is most evenly distributed over the sample.

Decategorisation renders the loss of characteristics of the original category of the grammaticalising element, for instance the incapability to take inflec-
6.9 Discussion

What has the sample data taught us about grammaticalisation? This section discusses the synchronic study’s outcomes for the ‘hybrid’ systems, addresses the research questions, hypotheses, and predictions, and concludes the syn-

The emergence of multiple concordial marking is inevitably related to a loss of phonological substance and morphosyntactic independence and, as was argued in the previous section, is the most reliable indicator for the existence of an obligatory classification. Systems lacking obligatory marking may obligatorily classify the nominal lexicon, but one cannot predict them to do so. A classifier system undergoing the shift postulated by the grammaticalisation hypothesis however is expected to develop into a multiple classifier system first and then to introduce multiple marking. Optional multiple marking then develops into an obligatorised class marking across morphosyntactic contexts. This conjecture of a tendency towards ‘multiple contexts before obligatoriness’ is reflected in the sample by the ‘hybrid’ systems, which all employ multiple classification (an optional one in the case of Ngan’gityemerri (6)), but do not obligatorily classify the nominal lexicon, whereas all systems that feature multiple marking obligatorily classify the nominal lexicon.

The sample study thus generally confirms the predicted preferred order of grammaticalising properties, though semantic bleaching and decategorisation seem to operate independently from each other in the case of a grammaticalising system. Before a separate account for these observations is presented in the next chapter, section 6.9 summarises the findings of the sample study so far.
chronic study by an overview over the observed developmental pathways in the sample. Grammaticalisation is present throughout the sample, but is not the only dimension that is involved in nominal classification.

The reanalysis of the ‘hybrids’ as actual classifier-like systems results from instances of class ‘disagreement’, where nominal modifiers show markers that are different from the head noun’s class and thus cannot represent instances of class agreement (cf. e.g. 5.12 = 6.86), and a productive nominalisation by means of class markers, where also no copying process can be verified, since any reference to an overt noun is lacking (cf. 6.32 = 6.87).

(6.86) Bora (cf. Seifart 2005: 169)

A  
$v$-hi  $m\=\text{\textit{h}}\=\text{\textit{h}}$  
DISTANT-CL:2D.ROUND be.big.SUB-CL:2D.ROUND  
k\=\text{\textit{u}}:m\=\text{\textit{u}}:hi  
turtle-CL:2D.ROUND 
“that big turtle”

B  
$v$-hi  $m\=\text{\textit{h}}\=\text{\textit{h}}$  
DISTANT-CL:2D.ROUND be.big.SUB-CL:MG  
k\=\text{\textit{u}}:m\=\text{\textit{u}}:hi  
turtle-CL:2D.ROUND 
“that big turtle”

C  
$aj$-$di$  $m\=\text{\textit{h}}\=\text{\textit{h}}$  
DISTANT-CL:2D.ROUND be.big.SUB-CL:2D.ROUND 
k\=\text{\textit{u}}:m\=\text{\textit{u}}:hi  
turtle-CL:2D.ROUND 
“that big turtle”

D  
$aj$-$di$  $m\=\text{\textit{h}}\=\text{\textit{h}}$  
DISTANT-CL:2D.ROUND be.big.SUB-CL:MG  
k\=\text{\textit{u}}:m\=\text{\textit{u}}:hi  
turtle-CL:2D.ROUND 
“that big turtle”


A  
t\=\text{\textit{u}}\=\text{\textit{u}}\=\text{\textit{a}}$  $p$-hi  
foot  CL:HOLE  
“shoe”

B  
s\=\text{\textit{a}}  $sa$  
one  CL:COVER  
“one (house, etc.)”

C  
m\=\text{\textit{u}}\=\text{\textit{u}}\=\text{\textit{u}}  $-\text{\textit{h}}\=\text{\textit{e}}$  
big  CL:M.SG  
“a big one (male animate)”

D  
m\=\text{\textit{u}}\=\text{\textit{u}}\=\text{\textit{u}}  $-\text{\textit{h}}$  
big  CL:DISC  
“a big one (hat, turtle, etc.)”

E  
m\=\text{\text{\textit{u}}\=\text{\textit{u}}\=\text{\textit{u}}}  $-\text{\textit{n}}\=\text{\textit{e}}$  
big  CL:INAN  
“a big one (inanimate)”


\footnote{In order to account for nominal classification in a way that yields more precise results, I introduce an additional account for the dimension of transparency in chapter 7; the synchronic study serves as a basis for the development of an instrument that serves to map nominal classification devices with respect to both their degree of grammaticalisation and transparency.}
This renders the supposed hybrid systems actual classifier-like systems. Their most obvious deviation from ‘ordinary’ classifier-like systems lies in the density of classificatory contexts, as every element carrying a classificatory marker can be identified with a context of classification. Multiple classification in the same morphosyntactic environment is also occasionally found in other classifier-like systems (cf. 6.25 – 6.88 where pá “fish” is categorised by the general classifier and additionally the animate classifier), but it is the exception rather than the rule.


cánh  tù  pá  du  chăng  ánh  hóa
CL:GENERAL  CL:ANIMATE  fish  at  inside  CL:INANIMATE  house
nǐ,  mạnh  vọ
FOC  he  say
“The fish inside the house said…”

Another visible difference of the former hybrid systems and the other classifier systems is that the hybrids allow their classificatory markers to occur outside of the noun phrase, e.g. on the verb (cf. 6.37 – 6.40 = 6.89 – 6.92), while ‘ordinary’ classifier systems operate NP-internally.


Teni  tore’re  tareuri-ma-∅-i
ART.F  children  care.for-PRS.HAB-DERIV-CL:F
bau’uri-ma-∅-i.
feed-PRS.PROGR-DERIV-CL:F
“The lady who cares for the children (is) the one who is feeding them.”

(6.90) Bora (cf. Seifart 2007: 424)

kà:tûfe  -hí  (bo:tá  -hí)
fall.down  CL:DISC  button  CL:DISC
“(The button,) it fell down.”

(6.91) Mundurukú (Mithun 1986: 381)

Ti  dojot  puye,  o’ti-mog  ip  baseya’a  be.
water  bring  when  they-CL:WATER-place  they  be.
“When they brought water, they placed it in a basin.”

(6.92) Ngan’gityemerri (cf. Reid 1997: 194)

ganbi  yerr=  buy-  mem  -garri  kide
didgeridoo  CL:TREE  light  3SGS.do  leg  where
“Where is that white didgeridoo?”
Figure 6.4 displays the taxonomy of nominal classification devices after the reanalysis of the hybrid systems as classifier systems.

![Diagram of systems of nominal classification]

Figure 6.4: Final taxonomy of systems of nominal classification.

Different morphosyntactic classifier hosts do not entail a different classificatory behaviour. Given the large range of classifier hosts in the ‘hybrid’ systems, presumably originating from one single host, it stands to reason that the classificatory material transgresses the borders of the noun phrase at a given point of host expansion, while its classificatory function is maintained. In summary, the former hybrid systems can be reanalysed as classifier-like systems, from which they marginally divert in that they allow their classifiers to occur NP-externally.

Table 6.24 presents the revised systems and their properties in the sample. The semantic contribution to the noun phrase is removed as a property due to its lack of reliability, while the inventory size is maintained as an indirect criterion for a system’s analysis.

Table 6.24: The revised form-based properties of the language sample.
In the following, the research questions (Q), hypotheses (H), and predictions (P) from section 4.1 are addressed. Since the hypotheses are based on the outcomes of the predictions, and since the research questions build on the hypotheses, the predictions are addressed before the hypotheses, which precede the research questions.

**P1** The different systemic (sub-)types of nominal classification cover different areas on the functional continuum.

The sample study confirms P1 in that classifier-like systems cover the less grammaticalised area of the continuum and concordial class systems cover the more grammaticalised area of the continuum (cf. table 6.24, where (1)–(19) are classifier systems and (20)–(33) are concordial class systems). There is no noticeable difference between the individual subtypes of classifiers, i.e. between noun, numeral, and possessive ones. The ‘hybrid’ systems group together towards the less grammaticalised side of the continuum, since they are the only classifier-like systems that allow their markers to transgress the borders of the noun phrase. Their remaining form-based properties do not further set them apart from other classifier systems. Beyond that, it is common sense that noun class systems and gender systems are two labels for the same phenomenon. The individual subtypes of nominal classification systems can thus not be said to cover distinct areas in the grammaticalisation continuum (except for the ‘hybrids’, if those are defined as a distinct subtype), while classifier-like systems and concordial class-like systems do.

**P2** The development of nominal classification devices consists in (at least) three steps: Lexical nouns, which grammaticalise into classifiers, which may further grammaticalise into concordial class markers.

The sample study confirms P2. Classifiers mainly derive from lexical nouns, as has been shown by earlier studies for numerous systems such as Korean (Hwang et al. 2008: 155f.), Hungarian (Csirmaz and Dékány 2010a: 10f.), and Jakaltek (cf. Craig 1986b: 253 and Day 1973: 60 and 132). Concordial class systems may stem from a variety of sources such as demonstratives (cf. Greenberg...
1978a), case marking patterns (cf. Luraghi 2011, following Fodor 1959), or, according to the grammaticalisation hypothesis, classifiers. The sample study was able to elicit three types of deviations (first, second, and third order deviations) from the definition of a prototypical classifier system and concordial class systems. The first order deviations include the ‘hybrids’; these share core properties of classifier-like and concordial class-like systems (non-obligatoriness, single marking, and a locus operandi that stretches beyond the NP borders). They represent the candidates that are the most likely sources for a potential future concordial class system.

\[P_3\] There are intermediate systemic types which display properties of both classifier and concordial class systems, ... 

\[P_4\] ... and these intermediate types form a coherent and so far not further specified grey area in the functional continuum.

The sample study confirms the existence of intermediate systemic types, which display properties of both classifier-like and concordial class-like systems and thus form a grey area between prototypical classifier and concordial class systems. The bulk of systems deviates from the two systemic prototypes in at least one form-based property. The study thus explored to what degree these deviations could be subcategorised and ranked. It proves fruitful to structure the underdefined notion of the grey area, which comprises all systems that deviate in one or more properties. As a result, the deviations were grouped into first, second, and third order deviations. First order deviations are those systems that show a mixed classifier-like and concordial class-like property setup in the otherwise highly stable ‘core properties’ (obligatoriness, marking, and locus operandi), and comprise the hybrid systems of Nasiioi, Bora, Mundurukù, and Ngan’gityemerri as well as the systems of Dyirbal and Tainae. Second order deviations are those systems that show a mixed classifier-like and concordial class-like property setup in the remaining properties (set size, flexibility, assignment) and comprise for instance the classifier systems of Jakaltek and North Ambrym and the concordial class systems of Maasai and Sumerian. Third order deviations are those systems that show a consistent classifier-like or concordial class-like property setup, which features individual properties that are both classifier-like and concordial class-like, such as the concordial class systems of Ingush, Chichewa, Hadza, and Hausa, which allow a flexible class membership for larger parts of the nominal lexicon. Depending on the problem looked into, a study of intermediate systems of nominal classification may delimit the grey area to the form of first and second order deviations or first order deviations only, since it may not be fruitful to include ‘light’ deviations, which are fairly close to their respective prototype and can be identified as either a classifier or a concordial class-like system without greater problems.

Both \[P_3\] and \[P_4\] are thus confirmed. In addition, it has been shown that the grey area can be specified; table 7.39 provides a revised overview of the distribution of the first, second, and third order deviations.
Table 6.25: The revised grey area.

<table>
<thead>
<tr>
<th>Language</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td>2</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasioi</td>
<td></td>
<td>4</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bora</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mundurukú</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern Japanese</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panamanian Fiji</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Great Andamanese</td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jakaltek</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Ambrym</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hangang Dow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nung</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyirbal</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tainae</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maasai</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amor</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruk</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingush</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingush</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chichewa</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hadza</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hausa</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bawa</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1: first order deviation, 2: second order deviation, 3: third order deviation, blank: prototypical system.

P5.1 By the course of CLS>CCS, the classificatory elements (partially or fully) lose their potential of semantic contribution to the noun phrase.

The sample study could not confirm this prediction. While it is true that the bulk of classifier-like systems semantically contributes to the noun phrase, and that concordial class-like systems tend not to semantically contribute to the noun phrase, there are a number of classifier systems that lack a semantic contribution (e.g. Dongo-Ko and Tohono O’odham) and concordial class systems that exert a strong semantic contribution (e.g. Burushaski and Maasai). The data from the sample suggests that a general tendency towards semantically loaded classifiers and semantically vain concordial classes holds, but that the semantic productivity of a system is independent from its type.

P5.2 By the course of CLS>CCS, the set of classificatory elements decreases in size.

A system’s inventory size has turned out to be only indirectly related to its type, since all systems except the ones of Korean, Mandarin, Nasioi, Bora, and Mundurukú employ an inventory of less than 25 classes. The classifier-like systems tend to employ larger inventories than concordial class-like systems, and there are no concordial class-like systems that employ a large inventory in the definition of this study (cf. table 6.26).
The synchronic data from the sample thus fits P₅.₂, but cannot be considered as evidence in favour of it, or even as its confirmation. Since concordial class systems are among others defined by a small class inventory, a classifier-like system with a large inventory undergoing CLS>CCS must reduce its inventory in order to grammaticalise into a concordial class-like system ‘proper’. Yet there is a fair number of classifier systems that already employ a small class inventory, and the synchronic sample data cannot verify that these inventories would decrease in size when grammaticalising into a concordial class system.

P₅.₃ By the course of CLS>CCS, the classificatory elements cliticise and later morphologise to agreement markers.

The boundedness of the classificatory markers is indirectly linked to the type of nominal classification system in that the sample’s concordial class systems do not employ independent markers, but clitic markers (e.g. Sumerian), fused forms with articles or demonstratives (e.g. Burushaski, German), or affixes (e.g. Burushaski, Maasai, Sumerian, Tamil, Dongo-Ko, Moro, Ket, Hadza, Hausa). The classificatory markers within a given system can be of a mixed nature: For instance, German employs both affixed markers and ones that are fused with the article, and Sumerian possesses both clitic and affixed markers. Classifier systems also may employed mixed markers; e.g. Ngan’gityemerri and Korean employ both independent and bound markers.

It is however not possible in most cases to reliably evaluate the degree of boundedness of a marker, since often the only criterion that such a judgement could be based on is a given author’s usage of either ⟨−⟩ or ⟨−⟩ in the glossing of his or her examples. This usage is often determined on an intuitive basis by the field researcher who collects the data (Swintha Danielsen, p.c.). Since a clear distinction of clitic and affixed marking cannot be carried out on a reliable basis, the general boundedness (bound vs. unbound) of the classificatory markers must suffice as a distinctive criterion for this study.

The boundedness of marking in the sample is distributed in the same way as the systemic inventory size: Classifier systems display bound or unbound

---

Table 6.26: Inventory size in the revised sample.

<table>
<thead>
<tr>
<th>Language</th>
<th>Inventory Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>1</td>
</tr>
<tr>
<td>Mandarin</td>
<td>&gt;187</td>
</tr>
<tr>
<td>Nasioi</td>
<td>2</td>
</tr>
<tr>
<td>Bora</td>
<td>3</td>
</tr>
<tr>
<td>Mundurukú</td>
<td>4</td>
</tr>
<tr>
<td>Ngan’gityemerri</td>
<td>5</td>
</tr>
<tr>
<td>Boumaa Fijian</td>
<td>6</td>
</tr>
<tr>
<td>Hungarian</td>
<td>7</td>
</tr>
<tr>
<td>Great Andamanese</td>
<td>8</td>
</tr>
<tr>
<td>North Ambrym</td>
<td>9</td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td>10</td>
</tr>
<tr>
<td>Dongo-Ko (CCS)</td>
<td>11</td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td>12</td>
</tr>
<tr>
<td>Hmong Daw</td>
<td>13</td>
</tr>
<tr>
<td>Nùng</td>
<td>14</td>
</tr>
<tr>
<td>Dyirbal</td>
<td>15</td>
</tr>
<tr>
<td>Tainae</td>
<td>16</td>
</tr>
<tr>
<td>Burushaski</td>
<td>17</td>
</tr>
<tr>
<td>Maasai</td>
<td>18</td>
</tr>
<tr>
<td>Sumerian</td>
<td>19</td>
</tr>
<tr>
<td>Tamil</td>
<td>20</td>
</tr>
<tr>
<td>Sumerian (CLS)</td>
<td>21</td>
</tr>
<tr>
<td>Meru</td>
<td>22</td>
</tr>
<tr>
<td>Ket</td>
<td>23</td>
</tr>
<tr>
<td>English</td>
<td>24</td>
</tr>
<tr>
<td>Chichewa</td>
<td>25</td>
</tr>
<tr>
<td>German</td>
<td>26</td>
</tr>
<tr>
<td>Hadza</td>
<td>27</td>
</tr>
<tr>
<td>Hausa</td>
<td>28</td>
</tr>
<tr>
<td>Masri</td>
<td>29</td>
</tr>
<tr>
<td>Ket</td>
<td>30</td>
</tr>
<tr>
<td>Hausa</td>
<td>31</td>
</tr>
<tr>
<td>Masri</td>
<td>32</td>
</tr>
<tr>
<td>Ket</td>
<td>33</td>
</tr>
</tbody>
</table>

---

The boundedness of marking in the sample is distributed in the same way as the systemic inventory size: Classifier systems display bound or unbound...
markers, both system-internally and across systems, while concordial class systems display only bound markers. The presence of bound markers only thus does not tell much about the system’s type, whereas a system that employs unbound markers is likely to be a classifier system.

P5.4 By the course of CLS->CCS, a classificatory element’s loss in semantic contributive force precedes its functional reanalysis into an agreement marker.

This prediction builds on the same assumption as P5.1, namely that a grammaticalising classifier system loses its potential to semantically contribute to the noun phrase. The synchronic sample data cannot verify this prediction, since there are semantically productive and semantically unproductive classifier systems, just as there are semantically productive and semantically unproductive concordial class systems. P5.4 can thus not be confirmed by the sample study.

The predictions that came from this study’s hypotheses could be partly confirmed by the sample study. In the following, the hypotheses are addressed.

H1.1 Classifier and concordial class systems constitute a functional continuum.

This hypothesis has been confirmed by the sample study. A number of properties such as the inventory size, the flexibility of classification, and the assigning agent are not tied to a given systemic type in their manifestation and can thus be expressed ‘classifier-like’ or ‘concordial class-like’ in either classifier systems or concordial class systems. This proves that we are dealing with a true continuum of nominal classification, which comprises classifier systems and concordial class systems.

H1.2 Classifier and concordial class systems are different systemic types, which display formal and/or functional differences.

The sample study confirms this hypothesis. Means of morphosyntactic expression such as the obligatoriness of classification or a lack thereof and multiple vs. single marking actually are tied to concordial class systems and classifier systems respectively. This proves that classifier systems and concordial class systems are different phenomena within the same continuum and thus not two dichotomous categories.

H2 Classifier systems may grammaticalise into concordial class systems.

The sample study indirectly confirms H2 in that the sample data displays a clear distribution of less grammaticalised classifier-like systems and more grammaticalised concordial class-like systems. The two form-based properties that all classifier systems in the sample lack are the obligatoriness of classification and the marking by means of agreement. All it takes for a classifier system to grammaticalise into a concordial class system is thus the development of these
two properties, which involves all processes that take part in a grammaticali-
sation processes (semantic bleaching, decategorisation, loss in morphosyntactic
independence and phonological substance). The most likely candidates for a
suchlike process are the ‘hybrids’. These however possess an inventory that
is too large for concordial class systems, and are semantically too productive
in that they do not copy class information from the head noun, but typically
add semantic information to it. The fairly intermediate systems of Dyirbal and
Tainae are commonly interpreted as concordial class systems; yet they have
been reanalysed as classifier systems in this study, as their modus operandi is
closer to the state of a classifier system than a concordial class system. Their
reanalysis demonstrates that it does not take much of formal changes to turn
these systems into a concordial class system. For instance, Dyirbal had to ex-
pand an overt class marking over the whole case paradigm and potentially also
mark other parts of speech than the noun marker for a noun’s class in order to
gain the nature of a ‘truly agreeing’ concordial class system.

The hypotheses of this study have been partly confirmed by the sample
study. This enables us to address the research questions.

Q1 To what degree are classifier systems and concordial class systems inher-
ently similar or different?

The sample study has shown that classifier systems and concordial class sys-
tems share many commonalities: Both types may be semantically productive or
unproductive, both may allow for a more or less flexible class membership, both
may assign nouns to classes either via the speaker or the system, both prefer
a small inventory size, and both may operate within and beyond the borders
of the noun phrase (though it is only the ‘hybrid’ systems among the classifier
systems that allow an NP-external occurrence of the classificatory markers).
The most salient difference between the two types lies in the obligatoriness of
classification and the multiple marking of a noun’s class membership in the
same classificational context. Both these properties are exclusive to the sam-
ple’s concordial class systems, while classifier systems consistently display a
non-obligatory classification and single marking of class membership in the
same classificational context. In summary, both systemic types can display full
semanticity and transparency or a lack thereof, while it is the manner and
obligatoriness of their marking that distinguishes them.

Q2 Can we observe spontaneous classification and conventionalised classifi-
cation in both systemic types?

The sample study has shown that spontaneous classification occurs in both
classifier systems and concordial class systems, just as conventionalised classifi-
cation does. Recall the Mandarin classifier system for instance, where a flexible
ad hoc classification is common in lower linguistic registers, while higher regist-
iers show classifier-noun pairs prevalingly conventionalised and inflexible (cf.
Bisang 1999: 23 and Zhang 2007: 54). In concordial class systems, the class
membership of a given noun is part of its lexical entry and thus inherent and inflexible. This means that ad hoc classification is not necessarily found in concordial class systems. However, a number of concordial class systems display a spontaneous, flexible, and speaker-sided classification, the most obvious case being the system of Maasai. The two Maasai concordial classes carry a notion of size and strength, which allows the speaker to shift nouns from their ‘natural’ class to the respective other class in order to impose the size or strength connotation of this class on the nominal referent.

The remaining two research questions are directed at the diachrony and raison d'être of concordial class systems:

Q3 Where do concordial class systems come from?

Q4 Why do concordial class systems exist, if they are functionwise seemingly vain?

Q3 and Q4 rely on the study of a system’s diachrony and cannot be answered on basis of mere synchronic data. At this point, the sample study confirms that there are no formal obstacles towards concordial class systems that derive from classifier systems by grammaticalisation. A more elaborated answer to Q3 and Q4 is given after the investigation of the diachrony of selected systems in chapter 8.

In conclusion, the synchronic study has shown that there is not one common grammaticalisation pathway in the development of classifier systems, but that there are two independent possible lines of development: Conventionalisation, i.e. the decrease of a system’s semantic productivity and accessibility, and formalisation, i.e. the grammaticalisation of a system’s formal properties. Conventionalisation represents the by far more common process in the sample and consists in the conventionalisation or lexicalisation of classifier-noun pairs without an accompanying grammaticalisation of the formal properties of the system. Both conventionalisation and formalisation correspond to the concept of grammaticalisation: Conventionalisation corresponds to a system’s grammaticalisation in the semantic domain, formalisation to a system’s grammaticalisation in the formal domain. The implications of this finding for the grammaticalisation hypothesis are studied in chapter 7 while the following paragraphs describe all possible pathways of development that were observed in the sample, i.e. conventionalisation and formalisation as well as the less frequently observed functional expansion or reanalysis of a system or its decay.

Virtually all classifier systems display conventionalised classifier-noun pairs, which can be explained from the higher frequency of given nouns in discourse and semantically incompatible alternative classes that a noun can be assigned to. Why is conventionalisation so much more common than formalisation? A system’s way of formal expression depends among others on the language type: Take e.g. Mandarin, which is an isolating language with a highly aged classifier system. While continuous changes in the composition of the classifier inventory and its semantics can be observed over the documented period of this system,
the isolating language type of Mandarin remained the same. Mandarin simply has not the formal means available to develop a concordial class system, which is among others characterised by class marking by means of affixes that agree with the head noun’s class. Those system-internal changes that are possible in Mandarin can also be observed, i.e. individual classifiers semantically expand and/or bleach to a degree where the classifier drops out of the system, and new classifiers enter the inventory. Furthermore, the higher linguistic registers make ample use of conventionalised classifier-noun pairs, so that the speaker of Mandarin has to learn the ‘correct’ classifier for a given noun (Zhang 2007: 54). A classifier system can thus develop into a state where it is to be analysed as a prototypical classifier system with respect to its form, while its semantics and transparency have eroded to such a degree that an intuitive access to the underlying structure of classification is not available to the speaker. Such a system can be surpassed by concordial class systems in terms of semanticity and transparency. A system’s conventionalisation does however not mean the end of its development: In theory, all a conventionalised system needs to gain in order to turn into a concordial class system is agreement marking and, if it does not possess it already, a small class inventory. Due to the conventionalised class membership of the nominal lexicon, the original class semantics fade away. Yet the development of first conventionalisation and a subsequent formalisation into the direction of a concordial class system seems a detour for the development of an individual system, or a ‘last resort’ for a system that has conventionalised to an idiosyncratic state and can be remotivated by its structural ‘renovation’ to a concordial class system.

Great Andamanese shows another possible development from the stage of a conventionalised nominal classification system. Its nominal classification system displays a high degree of semantically not necessarily transparent lexicalisations (cf. 6.93), which proves its conventionalised character.


A ara=mik h u-tei, ot=tei, er=e=ęŋe
   CL:6=middle-blood CL:4=blood CL:2=CL:5=measles
   “stomach ache, splitting headache, measles”

B firo ter=likh u(s), firo tara=creel
   sea CL:2=lap sea CL:6=green/blue
   “deep sea, open sea”

Great Andamanese employs a number of morphosyntactic hosts including the noun, the possessive, the adjective, the adverb, and the verb. Instead of developing a concordial class-like modus operandi, the classificatory markers have adopted an additional functionality that falls outside the scope of nominal classification: If the markers attach to the verb or adverb, their function is the specification of the concept expressed by the predicate (cf. 5.3 = 6.94).

A comparable functional extension can be found in the gender system of Mawng, where
6.9. Discussion

(6.94) Great Andamanese (cf. Abbi 2011: 770)

A  \textit{ut=fle}, \textit{e=fle}, \textit{ek=fle}  \\
\textit{CL:4=aim} \textit{CL:5=aim} \textit{OBJ=aim}  \\
“aim from above, aim to pierce, aim at”

B  \textit{ara=p^b^o}, \textit{er=p^b^o}, \textit{ut=p^b^o}  \\
\textit{CL:6=cut} \textit{CL:2=cut} \textit{CL:4=cut}  \\
“cut down/fell, hit with a stick (in the front), separate from the source”

C  \textit{er=bate}, \textit{ek=bate}, \textit{ut=bate}  \\
\textit{CL:2=slap} \textit{OBJ=slap} \textit{CL:4=slap}  \\
“slap on the face, slap suddenly, slap hard”

The less frequent case consists in the formalisation of a classifier system, which is for instance present in the sample’s classifier systems of Nasioi, Bora, Mundurukú, Ngan’gityemerri, and Tainae. These systems feature multiple morphosyntactic hosts for the classificatory markers (which indicates decategorisation), bound markers (which indicates the loss of morphosyntactic independence and phonological substance) and visually resemble agreeing concordial class systems due to the potential multiple occurrence of classificatory markers. Note though that three out of the five mentioned systems employ a large inventory of more than one hundred classes, and that these classes are semantically highly specific and also contribute semantically to the noun phrase. This indicates an incomplete formalisation process, where not all mechanisms related to grammaticalisation occur. Figure 6.5 summarises the potential pathways of development of a classifier system: The semantic conventionalisation and remotivation of both classifier and concordial class systems, the formalisation of classifier systems into the direction of concordial class systems, and the decay or functional remotivation of both classifier and concordial class systems.

gender markers occur on the verb in order to mediate selectional restriction of the predicate’s semantics, while the gender system itself maintains its functions as a nominal classification system (for instance reference tracking and referent differentiation; cf. Singer 2006, 2010, 2012). These cases demonstrate that a nominal classification device can be physically maintained, while its usage and operational nature may be partly or even fully amended by functions that are alien to nominal classification.
The sample study has shown that nominal classification, as varied and diverse as it presents itself in real life systems, shares a number of common traits. These motivate the assumption of a nominal classification continuum and comparative measures of individual systems. It has further shown that there are only two systemic types in the sample, classifier systems and concordial class systems. Even supposed ‘hybrid’ systems prove to be a classifier-like phenomenon on closer investigation. The properties that are used to describe, define, and measure nominal classification devices can be ranked into core properties, which are found in both concordial class systems or in classifier systems; these are the obligatoriness of classification or lack thereof and the potential for multiple classificatory marking or lack thereof. The locus operandi is closely tied to the employment of multiple marking, and only the hybrid systems employ NP-external marking without a parallel multiple marking. This means that a system that restricts its marking to the noun phrase is most likely not a concordial class system.\footnote{This is a strong tendency observed in the sample, yet not a strict rule: The concordial class system of German for instance confines its class marking to the borders of the noun phrase. Also, the markers transcending the noun phrase’s borders have to occur NP-externally in a function serving nominal classification. Recall e.g. the Great Andamanese system, which employs its classificatory markers inside and outside of the noun phrase, but serves the purpose of nominal classification only NP-internally.} The available ‘classifier-like’ and ‘concordial class-like’ values of the non-core properties occur in both classifier-like and concordial class-like systems, regardless of the systemic type at hand. These properties comprise the flexibility in class membership, the assigning agent, the semantic contribution, and the locus operandi. Based of these findings, chapter 7 introduces an instrument for mapping nominal classification more comprehensively.

Figure 6.5: Possible developmental pathways of systems of nominal classification.
This chapter introduces a novel way to map and graphically represent nominal classification devices with respect not only to their grammaticalisation of formal properties, but also with respect to their degree of transparency. The sample study so far has shown that it is possible to depict a grammaticalisation cline in nominal classification devices, which corresponds to the distribution of systemic types in that classifier systems form the less grammaticalised devices, and concordial class systems the more grammaticalised ones. The grammaticalisation cline is blurred by a number of outliers; these deviations from the cline can be explained by a number of reasons, most prominently the conventionalisation of a system. The sample data also suggests that the degree of grammaticalisation of a system’s formal appearance (“formalisation”) is not the only criterion that shapes a system, which is the reason why there is a considerable number of deviations that has to be explained.

In order to represent all types of nominal classification systems and their degree of grammaticalisation in a more appropriate way, another dimension has to be considered. I argue that this additional dimension is the one of a system’s transparency or accessibility, and that established and tested criteria can serve as measures for this dimension. Let us recapitulate what grammaticalisation in nominal classification devices actually is, namely a shift in its formal properties towards a more dependently and frequently marked, more rigid, and less manipulable means of category marking. The underlying structure of these categories and their interaction with phenomena different from nominal classification (e.g. case marking) is captured by the dimension of transparency, which accounts for a given system’s assignment and accessibility of its class inventory. Section 7.1 discusses the developmental pathways that are observed
in the sample and relates them to the two dimensions of grammaticalisation in nominal classification devices, transparency and form. Section 7.2 provides the basic arguments for the mapping method and the theory that was considered for its composition and elaborates the criteria and measurements at use. Section 7.3 exemplifies the model and discusses its account for the full sample.

7.1 Developmental Pathways in the Sample

This subsection identifies and discusses the sample systems at their respective developmental state. The synchronic study has shown that there are conventionalised, semantically idiosyncratic representatives of both classifier and concordial class systems (e.g. North Ambrym, Great Andamanese, German, and Masri) as well as semantically transparent ones (e.g. Nùng, Dongo-Ko, Maasai, and Sumerian). A semantically transparent system may conventionalise, such as in the case of the higher linguistic registers of Mandarin. Since the process of formalisation is a gradual one and consists of a growing number of conventionalised or lexicalised class-noun pairings, most of the systems are neither fully formalised nor fully transparent, but at an intermediate stage (e.g. Jakaltek and Chichewa). A conventionalised or idiosyncratic system may be semantically remotivated, as it has occurred in Dyirbal between the 1960s and 1980s (cf. Schmidt 1985: 151ff.). Since there are more and less transparent representatives of classifier systems and concordial class systems and since both systemic types may entertain the same functional purposes, the remaining main distinctive trait between those two is their formal appearance and its degree of grammaticalisation. This means that the grammaticalisation hypothesis holds, and the candidates from the sample which qualify for a potential ongoing transition from classifier to concordial class systems are found in the former hybrid systems of Nasioi, Bora, Mundurukú, and Ngan’gityemerri. These are the possible developmental changes that can occur within the realm of nominal classification; a system can however develop into a direction that transcends nominal classification. In this case, a functional expansion may occur, i.e. the nominal classification function is maintained, and additional functions are adopted (as in Great Andamanese, where an original system of nominal classifiers adopted the function of verbal classification). The other possibility is that the system is reanalysed and loses its nominal classification function; while there is no clear candidate in the sample for such a functional conversion, it is conceivable that e.g. a possessive classifier system with a reduced class set is reanalysed as a system for possessive marking, where the former classifiers distinguish alienability from inalienability. Finally, a system may be given up by its speakers and decay. This has been argued to be the case for the systems of Tlingit and Chukchi in the previous chapter. Figure 7.1 provides an overview over the distribution of the sample systems over the possible pathways of change that were elaborated in the synchronic study.
The conventionalisation or semantic remotivation of nominal classification devices appears to be a more common phenomenon. It does not necessarily affect a given system as a whole, but individual classifiers or groups of classifiers. Nonetheless, cases like Dyirbal, where the whole classificational system is reanalysed, restructured, and potentially reduced by the speakers, are documented (cf. e.g. the changes in the Indo-European concordial class system and its daughters as they are described in Irslinger 2009 and Luraghi 2011). The
following subsections are therefore focusing on those changes that are less well-attested, i.e. the grammaticalisation of classifiers into concordial class systems, the functional expansion or reanalysis of a system, and a system’s decay.

**Grammaticalisation** The following lines serve the alignment of the grammaticalisation hypothesis and the findings from the synchronic study by reference to the most likely candidate systems for undergoing grammaticalisation of classifiers into concordial classes, which are the sample’s first order deviations (Nasioi, Bora, Mundurukú, and Ngan’gityemerrí). The synchronic sample study suggests a distinction of conventionalisation and formalisation as the two main vehicles for change for nominal classification systems. This has an important consequence for the changes that are expected in the case of a system’s formalisation: A full grammaticalisation process involves semantic bleaching and the reduction of a system’s class inventory; a system’s path down the grammaticalisation cline is thus theoretically closely linked to a loss in its semanticity and transparency. The sample data however shows that the bulk of systems employs a smallish inventory, which includes the majority of classifier systems. It also shows that a system’s semanticity and transparency are not fully correlating with its type. Concordial class systems are thus not necessarily ‘less semantic’ than classifier systems, despite their more grammaticalised form of morphosyntactic expression. In the following, those formalisation processes that distinguish the former hybrid systems from classifier systems are recapitulated. After that, the systems are evaluated with respect to the formalisation processes they lack in order to develop into a concordial class system.

Prototypical classifier systems can be distinguished from conventionalised classifier systems, which display a more rigid and less manipulable class assignment and are represented by the sample’s second and third order deviations, and from formalised classifier systems, which display an increase in their targets’ class marking. Formalised systems are represented by the first order deviations in the sample; they primarily develop a different manner of formal expression, while conventionalising systems primarily develop a different manner of underlying assignment and classification. Within the form-based package of properties, this means that a formalising system corresponds to the expected order of grammaticalising properties and thus start to look like a concordial class system, while their underlying classification remains classifier-like. In contrast, conventionalising systems display the opposite pathway by first shifting their assignment and flexibility from a classifier-like to a concordial class-like state and thus maintain a classifier-like formal appearance, while they start to assign and classify the nominal lexicon in a concordial class-like manner.

What distinguishes the grammaticalising systems of Nasioi, Bora, Mundurukú, and Ngan’gityemerrí from the other classifier systems in the sample? With respect to the form-based package of properties, it is only the locus operandi which separates these systems from prototypical classifier systems: All four systems allow their classificatory markers to attach to the verb, which is an
NP-external constituent. A prototypical classifier system restricts its classificatory marking to the noun phrase. In addition, there are two more properties beyond the locus operandi that are not adopted as a measure for this study and that distinguish the grammaticalising classifier systems from the other classifier systems in the sample: First, all of the grammaticalising systems employ bound markers. Note that classifier systems may employ bound markers, but are not required to, whereas concordial class systems are defined by their agreement marking, which occurs in form of bound morphemes. Therefore, the development of bound marking is a precondition for a classifier system that grammaticalises into a concordial class system. Second, the grammaticalising systems all employ a large number of morphosyntactic hosts for their classificatory markers, which is the norm for concordial class systems, but not classifier systems. Table 7.1 lists the individual systems and the range of nominal satellites that carry class marking.

<table>
<thead>
<tr>
<th>language</th>
<th>N</th>
<th>ADJ</th>
<th>NUM</th>
<th>V</th>
<th>DEM</th>
<th>P</th>
<th>INT</th>
<th>ART</th>
<th>POSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasiö</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Bora</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ngan’gityemerri</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mundurukú</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

What distinguishes the grammaticalising systems from the concordial class systems in the sample? All of the grammaticalising systems allow the multiple occurrence of classificatory markers, which gives them a visual appearance that closely resembles agreement marking. As has been argued in section 6.3, this multiple occurrence is different from agreement in that it is an instance of multiple classification of appositional noun phrases (*that*^[M] [one], *big*^[M] [one], *turtle*^[M] [one]) and not an instance of multiple agreement marking of elements of the same noun phrase (*that*^[M] *big*^[M] *turtle*^[M]). Furthermore, no concordial class system employs a largish inventory, while Nasiö, Bora, and Mundurukú employ an inventory of more than one hundred classificatory markers. At the same time, the ‘hybrids’ classify the nominal lexicon in a flexible way, which means that a common noun in these systems has not an inherent class value. Consequently, the assignment is carried out by the speaker. Finally, the grammaticalising systems’ classification is non-obligatory, as all of them feature unclassified nouns in discourse. A concordial class system classifies nouns inherently, and hence each noun in every context carries a class value. Table 7.2 summarises the properties that separate the grammaticalising classifier systems from prototypical classifier systems on the one hand and from concordial class systems on the other hand.
### 7.1. Developmental Pathways in the Sample

Table 7.2: The properties of prototypical classifier systems, the sample’s formalising classifier systems, and prototypical concordial class systems (*: except Ngan’gityemerri).

<table>
<thead>
<tr>
<th>prototypical &amp; formalising CLSs</th>
<th>prototypical CCSs</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-obligatory classification</td>
<td>obligatory classification</td>
</tr>
<tr>
<td>single marking</td>
<td>multiple marking</td>
</tr>
<tr>
<td>flexible classification</td>
<td>rigid classification</td>
</tr>
<tr>
<td>speaker-sided assignment</td>
<td>system-sided assignment</td>
</tr>
<tr>
<td>large inventory*</td>
<td>small inventory</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>prototypical CLSs</th>
<th>formalising CLSs &amp; prototypical CCSs</th>
</tr>
</thead>
<tbody>
<tr>
<td>independent/bound marking</td>
<td>bound marking*</td>
</tr>
<tr>
<td>NP-internal marking</td>
<td>NP-internal and -external marking</td>
</tr>
<tr>
<td>few targets</td>
<td>many targets</td>
</tr>
</tbody>
</table>

The formalising classifier systems are first order deviations, which means that their ‘core criteria’ show a mixed setup of classifier-like properties (non-obligatory classification and single marking) and concordial class-like properties (NP-internal and -external marking). In order to develop into a second order deviation of a concordial class system, it takes those systems the introduction of an obligatory classification and agreement marking. Given the already existing occurrence of markers with multiple modifiers in these systems, the adoption of agreement mainly requires a reanalysis of these markers on adjacent, independent noun phrases as the multiple marking of one head noun’s class value. Though it is not a precondition for achieving the state of a second order deviation, the systems of Nasioi, Bora, and Mundurukú are expected to reduce their class inventory to some two dozen classes or less, as no concordial class system in the sample employs a large inventory. Two factors may facilitate this process: First, speakers of a language with a large classifier system are reported to make use of only a fraction of the available classifier inventory. For instance, speakers of Japanese use only some 30 classifiers from a full set of about 150 classifiers (cf. Downing 1996: 17 and 21). Second, there are large classifier systems that may not be used by all speakers, but only by certain societal groups, which are typically higher ranking than others (e.g. in Kilivila; Senft p.c.). Since the majority of speakers in these languages use a reduced systemic variety, a reduction of the whole system is a matter of the potential with which the higher societal group can maintain their system’s richness.

The step from a second order system to a prototypical or third order system respectively requires the full or partial shift from a flexible to a rigid classification. Even flexible classifier systems tend to display preferred classifier-noun
pairings, which override a potential classificatory variation (cf. Bisang 1999: 23). A conventionalisation or lexicalisation of class-noun pairings from a small class inventory that lacks the potential for a greater classificatory flexibility thus seems natural. On the same basis, i.e. a reduced class inventory with less potential for classificatory variation and a rigid class membership, speakers lose their ability to carry out the assignment at their own will, which is the last property that separates the grammaticalising classifier systems from a concordial class system.

In summary, the grammaticalising classifier systems from the sample possess the potential to turn into concordial class systems. The necessary steps for that are at least the obligatorisation of classification, the adoption of agreement marking, and the reduction of the class inventory. These require an in-depth reanalysis of the system, both morphosyntactically and semantically. Since all of the systems are characterised by a high degree of semanticity and semantic productivity, a drastic bleaching of these two attributes is the most obvious factor that motivates such an in-depth reanalysis. All of the involved systems need further observation and documentation in order to verify the grammaticalisation hypothesis. Note that the grammaticalising classifier systems from the sample are all fairly transparent with respect to their semantics and do not display a higher degree of formalisation. Given the frequently observed different degrees of grammaticalisation of individual classifiers within one and the same system (e.g. in Mandarin and Great Andamanese), the more common strategy to make an idiosyncratic classifier system productive again appears to be its semantic remotivation by the adoption of new and transparent categories. This study does however not suggest that only transparent classifier systems may grammaticalise into concordial class systems: A highly conventionalised and idiosyncratic classifier system is not excluded from a grammaticalisation of its means of formal expression; such a system just happens not to be present in the sample.

In the following, I discuss another possible developmental pathway of classifier systems: The expansion or reanalysis of a system’s functionality.

**Functional Expansion or Reanalysis** The sample contains one example of a system’s functional expansion beyond the scope of nominal classification, which is described in the following. The classifier system of Great Andamanese is described as a primary possessive classifier system (Abbi 2011: 747), but this primary function is complemented by others, which relate to the morphosyntactic host of the class marker. The part of the Great Andamanese system that serves nominal classification was investigated in the sample study; it applies to the classificatory markers in possessive constructions (cf. 7.1) and in combination with the adjective and noun (cf. 7.2).

A tat
1SG 1CL:1 tongue
“my tongue”

B Licho
Ut= thire
Licho 4CL:4 child
“Licho’s child”

C 2SG 4CL:4 child PL much COPULA
“You have many children.”

(7.2) Great Andamanese (cf. Abbi 2011: 774)

A Lephai
Lep= cok nol cai tabol =o
Lephai 2CL:2 face good 5CL:5 bad REFL naughty PST
“Lephai was goodlooking, but naughty.”

The classifiers in (7.1) and (7.2) account for inherent characteristics of the nominal referent, as it is common for nominal classifiers: The referent of tat “tongue” is categorised as a part of the mouth in (7.1) A, children as a body product in (7.1) B and C, the face as a part of the head in (7.2) and the naughty character of the referent as a body-internal property. (7.2) also demonstrates the occurrence of multiple classification in Great Andamanese: Both the adjective and the noun carry a classificatory marker, which independently assign their respective host to different categories. The system is involved in the denotation of a referent’s discourse specificity, as general and unspecified referents remain uncategorised, while specific discourse referents carry a class marker (cf. 3.31 = 7.3).

(7.3) Great Andamanese (cf. Abbi 2011: 762)

A cokbi
Cokbi omu
turtle flesh
“turtle meat (general, non-specific)”

B cokbi
Ot= cok omu
turtle 4CL:4 flesh
“turtle meat (still on its body)”

C cokbi
Ot= cok omu
turtle INANIMATE- 4CL:4 flesh
“turtle meat (cut up for consumption)”

This part of the Great Andamanese classifier system is unobtrusive and classifier-like, even if the extended range of morphosyntactic classifier hosts and the phenomenon of multiple classification indicate that the system is not a fully prototypical one. If the classifiers however combine with an adverb or verb, it is not a noun or its referent which is categorised and modified, but the
concept expressed by the predicate. This phenomenon does not fall under the scope of nominal classification, but serves the function of verbal classification. Cf. table 42 for the individual semantic categories and examples.


<table>
<thead>
<tr>
<th>class</th>
<th>classifier</th>
<th>verbal semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a-</td>
<td>“mouth-related activity, origin”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. a=jire “abuse”, a=kabpa “sprout”</td>
</tr>
<tr>
<td>2</td>
<td>eγ-</td>
<td>“action involving the front part of the body”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. era=luk “weigh”</td>
</tr>
<tr>
<td>3</td>
<td>oγ-</td>
<td>“hand-related activity”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. oγ=cβa “stitch”, un=tupuro “trembling of hands”</td>
</tr>
<tr>
<td>4</td>
<td>ul-</td>
<td>“directional, experiential”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. ot=cone “leave”, tβ e=bon “be hungry”</td>
</tr>
<tr>
<td>5</td>
<td>ε-</td>
<td>“action involving the interior of an object”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. ε=leko “suck”, ε=reno “tear”</td>
</tr>
<tr>
<td>6</td>
<td>ara-</td>
<td>“action involving the middle portion of the body”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. ara-del “be pregnant”</td>
</tr>
<tr>
<td>7</td>
<td>o- a-</td>
<td>“resultative state”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. o=crono “make nest”, o=reas “sting”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>class</th>
<th>classifier</th>
<th>verbal semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a-</td>
<td>“deixis front/back, anteriority of action”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. a-karap “behind”, a-kau “prior to”</td>
</tr>
<tr>
<td>2</td>
<td>eγ-</td>
<td>“deixis of adjacency”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. er-be[tloa “adjacent/near to X”</td>
</tr>
<tr>
<td>3</td>
<td>oγ-</td>
<td>“haste or hurriedly done action”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. oγ-koč “fast, hurriedly”</td>
</tr>
<tr>
<td>4</td>
<td>ul-</td>
<td>“directional deixis”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. ot-le “seaward”, ot-bo “backwards”</td>
</tr>
<tr>
<td>5</td>
<td>ε-</td>
<td>“deixis of internal space”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. te-kil “in the middle”, e-kotra “inside”</td>
</tr>
<tr>
<td>6</td>
<td>ara-</td>
<td>“deixis of immediate vertical/horizontal space”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. ara-balo “behind X”, tara-tal “right under X”</td>
</tr>
<tr>
<td>7</td>
<td>o- a-</td>
<td>“temporal deixis”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e.g. o-pó “day break”, o-kara “sunset”</td>
</tr>
</tbody>
</table>

Considering the original nature of the classifiers as nominal classifiers in possessive constructions, the Great Andamanese system in its current state appears to have undergone two separate change phenomena: First, the possessive classifier system underwent a grammaticalisation process, which extended the range of morphosyntactic classifier hosts to nouns and adjectives; this multiple classifier system remained a nominal classifier system. At a later stage, the system further grammaticalised and extended its range of classifier hosts to verbs and adverbs. The result of this grammaticalisation process is the verbal classification system described above, which lies outside the scope of nominal classification. The multiple classifier ‘source system’ was maintained as a nom-
7.1. Developmental Pathways in the Sample

inal classification device, as the semantics of the classifiers in their respective contexts demonstrate (cf. again table 5.7 on page 130). Figure 7.2 summarises the development of the Great Andamanese system at the three postulated stages.

![Diagram of the development of the Great Andamanese classifiers]

As the sample contains only productive nominal classification devices, there is no example for a classifier system that was fully reanalysed and lost its nominal classification function. It is however conceivable that a given nominal classification system undergoes a functional reanalysis such as Great Andamanese without maintaining it as a nominal classification device. In such a case, it seems more likely that the original nominal classification device has grown idiosyncratic to a certain degree, and demands a remotivation or reanalysis. The Great Andamanese system in contrast appears to be fairly accessible in all its components, which may be the reason for broad functionality and stable productivity. Less transparent, productive, and stable systems may also be driven out of grammar instead of being functionally reanalysed or remotivated. This phenomenon is discussed in the following.

**Decay** The decay of the nominal classification systems of Tlingit and Chukchi is described in the following. Numerals in Tlingit distinguish two sets, one regular set for counting objects and animals (tléíx “one”, deíx “two”, nás’k “three”, etc.), and one set that serves the quantification of human referents and features an additional infix -ná- (tléí-ná-x, dáx-ná-x, nás’gi-ná-x; Edwards 2009: 17). Note that -ná- and the word-final -x are described as a postposition ná.x in Swanton (1911: 198). The classification in Tlingit shows striking parallels with Athabaskan systems (e.g. Hupa; cf. Aikhenvald 2000: 123), and diachronic data for Tlingit is not available. The genetic affiliation of Tlingit varies between the sources; while Ruhlen (1987: s.v.) treats it as a member of the Na-Dene
family, Lewis et al. (2013: s.v.) treat it as an isolate. Since Tlingit and its surrounding (and genetically unrelated) languages are known for a number of areally shared linguistic features (cf. Mithun 1999: 314ff.), the Tlingit system is contrasted here against a richer Athabaskan system, the one of Carrier (ISO 639-3: crx). The Carrier system shows a better preserved state from a common protosystem, which has been given up in Tlingit. Mithun (1999: 104) reports Carrier to employ a fivefold set of numerals and quantifiers, which includes a human set that is marked by the suffix -ne (cf. table 7.4).

Table 7.4: Quantificational categories in Carrier (cf. Mithun 1999: 104).

<table>
<thead>
<tr>
<th></th>
<th>human</th>
<th>general</th>
<th>abstract</th>
<th>areal</th>
<th>multiplicative</th>
</tr>
</thead>
<tbody>
<tr>
<td>“two”</td>
<td>nane</td>
<td>nanki</td>
<td>nazw</td>
<td>nadon</td>
<td>nat</td>
</tr>
<tr>
<td>“three”</td>
<td>tane</td>
<td>ta</td>
<td>tazw</td>
<td>tandon</td>
<td>tat</td>
</tr>
<tr>
<td>“four”</td>
<td>dine</td>
<td>dmyi</td>
<td>dizw</td>
<td>dibon</td>
<td>dit</td>
</tr>
</tbody>
</table>

Table 7.4 shows that one of the categories is unusual for classification systems, namely the multiplicative category. The other classes are however cross-linguistically well-attested. The individual categories have a straightforward semantic basis and are marked word-finally on the quantifier, which is commonly observed in numeral classifier systems. The fusion of the numeral and classificatory marker indicates an advanced grammaticalisation process, just as the wide semantic scope of the individual categories does. Considering the existence of a human and a non-human category in Hupa numerals (Aikhenvald 2000: 123), and other, by far larger systems that include a human category (e.g. Kwakwala, ISO 639-3: kwk; Yurok (a neighbouring language of Hupa), ISO 639-3: yur; Mithun 1999: 105 and 109f.), nominal classification in quantificational contexts appears to be common for the linguistic area in question. The Tlingit and Hupa system appear to be impoverished remnants of an earlier productive, richer classification system. If the classificatory distinction of the Tlingit system was marked elsewhere than the numeral, one could make a case for its conversion into a concordial class system. Its current state however gives little reason to advocate its productivity as a system of nominal classification, as only one of the supposed categories is overtly marked. Instead, it suggests -ná- to be a petrified relic of an earlier productive classification system. This relic results in the distinction of two numeral paradigms rather than the distinction of a classificatory set of numerals.

The supposed nominal classification of Chukchi displays a similar state: A number of affixed case markers (ergative/instrumental, locative, equative, and allative) are regularly preceded by an additional suffix -ne- in the singular and -r- or -c- in the plural, if the noun is a personal name or kin term (Dunn 1999: 100f.). Quantitative pronouns and demonstratives referring to such nouns can also carry this additional marking. One could argue from a synchronic point
of view that this suffix is a remnant of an impoverished, earlier and richer classification system like Tlingit, and that it marks a ‘high animate’ class, which is opposed to an unmarked residue class. However, as diachronic data is lacking, all conjectures about the diachrony of the Chukchi system are nothing more but guesses, which leaves two options for its synchronic state: Either it is a remnant of an earlier productive system, or it is a system that allows for emphatic marking of ‘high animates’. Either way, the Chukchi system at its contemporary state is not a productive system of nominal classification, yet let us assume for the moment that it is the remnant of an earlier productive system. Why is it not productive anymore? First, it does not distinguish noun classes in the form of declensional classes, as it is common for concordial class systems. Second, it restricts itself to overtly mark only the small group of proper names and kinship terms, while the overwhelming majority of nouns, which constitutes the residue class, remains unmarked. This results in the third argument against the Chukchi system being a productive means of nominal classification: The only clearly identifiable class is considerably small, as it comprises proper names and kinship terms. A productive system would be expected to include at least more common nouns for higher animates, such as human, man, or woman. Fourth, even within the overtly marked ‘class’, only four out of thirteen grammatical cases display the animate suffix -ne- (Dunn 1999: 101). All in all, Chukchi does not display evidence that lets the researcher conclude to deal with a productive and ‘proper’ system of nominal classification.

After the different potential developmental pathways in the sample have been addressed, section 7.2 lays the theoretical basis for mapping systems in a more efficient way by the distinction of two subtypes of grammaticalisation in nominal classification devices: Conventionalisation and formalisation.

### 7.2 Theoretical Considerations

This section conveys the aims, theoretical arguments and reflections that lay the foundations for the proposed mapping instrument. Most studies on nominal classification focus on aspects that relate to a system’s formal expression; ultimately, the types of nominal classification systems are defined on basis of their morphosyntactic properties, while their semantic and functional criteria so far play an inferior role in research. Actual systems often do not fully correspond to the definition of a prototypically or canonically ideal system and are thus automatically branded as deviations. The presented mapping aims at a representation of all nominal classification systems as systems in their own right rather than defining an ideal which is not often met in the real world. Nominal classification devices are better understood and can be better represented, if the dimension of systemic transparency is considered next to a given system’s mere formal expression. A system’s dimension of formal expression and its dimension of transparency constitute two components, which together fall under the scope of grammaticalisation in its most general notion; in nomi-
nal classification devices however, they display developmental shifts that occur independent of each other and together constitute ‘full’ grammaticalisation (cf. figure 7.3).

Figure 7.3: Grammaticalisation in nominal classification devices: Formal expression vs. transparency.

This section demonstrates that a comparability along the lines of both dimensions, formal grammaticalisation and bleaching of transparency, allows for a much better and clearer comparability of systems. Beyond that, the adoption of transparency as a second measure eliminates those exceptional cases from the sample that had to be separately analysed and explained, as their setup of properties proves to differ from the predicted grammaticalisation cline. Finally, concordial class systems seem to be much better understood than classifier systems; classifier systems “differ substantially, and we do not find systems which could be usefully analysed as genuine opposites of gender systems” (Corbett and Fedden 2015: 2). As a result, current studies prevailingly define classifiers as all types of nominal classification that are not concordial class systems (cf. e.g. Corbett and Fedden 2015 and Senft 2007, cited in Contini-Morava and Kilarski 2013: 266), which indeed results in a highly varied category of classifier systems.

Manifestations of nominal classification are in general of a “gradient rather than categorical” nature (Aikhenvald 2000: 14; cf. also Contini-Morava and Kilarski 2013: 266). It seems not fruitful for our understanding of this phenomenon to define a single category of concordial class systems well and treat those systems that do not correspond to this category as a residue. In contrast, it seems more promising to ignore existing categorial pigeonholes for the intended mapping. Systems such as Nasioi, Bora, Mundurukú, Ngan’gityemerri, Dyirbal, and Tainae are examples for systems that profit from avoiding a subcategorisation of nominal classification devices. Note that the labels “classifier system” and “concordial class system” are used as descriptive reference points in the following (e.g. to refer to the sample systems), but are not meant to pinpoint systemic types or assign labels to systems.

By avoiding a categorial division based on the definition of concordial class systems, a better understanding of nominal classification in general is gained. The established definitions of types of nominal classification systems serve as a mere delimitation of the nominal classification continuum in the presented mapping: Systems that meet the definition of a prototypical classifier system (e.g. Korean) represent the lowly grammaticalised form of nominal classification, and systems that meet the definition of a prototypical concordial class system
264  7.2. Theoretical Considerations

(e.g. German) represent the more grammaticalised form of nominal classification. For the rest, a labelling of systemic types is avoided. Another aim of this mapping method is the creation of a graphically represented theoretical space which allows to display and compare nominal classification systems. The representation of systems in this model gives the viewer a more complete idea of the mapped systems in that the model captures not only formal properties, but also the additional dimension of a system’s transparency. In the following, I discuss their implementation in the presented model.

Nominal classification systems are characterised by three main dimensions: Their formal expression, their degree of transparency or intuitive accessibility, and their functionality. As has been argued in chapter 2, the most promising account for the functionality of nominal classification devices stems from Contini-Morava and Kilaraki (2013), where a typology of four semantic and three discourse functions is presented. A functional account demands a certain extent of documentation and especially discourse data, which is not available for most languages. In addition, functionality is not reliably quantifiable, which both stands in contrast with the aim for comparability in the mapping method. Therefore, this study adopts the functional typology as a filter for the sample composition, i.e. a candidate system for the sample is only selected and incorporated if it serves at least one of the functions proposed by Contini-Morava and Kilaraki (2013); if it does not, the system is discarded as a device that does not serve nominal classification.

The mapping accounts for the two remaining dimensions, a system’s formal expression (subsection 7.2.1) and its transparency (subsection 7.2.2). Beyond the insights from this study, the model is based on the ‘prototype account’ of nominal classification by Dixon (1982, 1986) and others (e.g. Allan 1977, Grinevald 2000, etc.), the canonical typology of nominal classification by Corbett (cf. Corbett and Fedden 2015), and the studies of complexity of concordial class systems by Audring (2014, 2015).1

Note that the mapping as it is presented here represents an approximation model, and not an accuracy model. It is an instrument to visualise a system’s overall transparency and degree of formalisation in form of a diagram. For this purpose, properties are distinguished with respect to their low, medium, or high presence in a given system; this low, medium, and high degree is translated into a numerical value for practical purposes. However, since the individual measures for the calculation of the transparency value and the formalisation value vary with respect to their degree of precision, it is important not to presuppose an absolute accuracy of the values, but to take a margin of tolerance into account.

1Note that the term “transparency” as it is used here partly overlaps with Audring’s (2014, in print) use of “complexity”. I decided to use the label “transparency” due to the fact that Audring’s notion of “complexity” also includes formal, morphosyntactic criteria, which I intend to analyse as a separate dimension.
7.2.1 The Dimension of Form

Formal properties of a system's morphosyntactic expression are its most tangible ones and generally serve as defining criteria for nominal classification devices. The main formal difference between systemic types is their degree of grammaticalisation. The sample study has shown that it is common for nominal classification systems to conventionalise rather than formalise, which has an important consequence for the account of the formal properties. Grammaticalisation theory assumes a preferred order of properties (cf. again figure 3.6 on page 70), which is not always met throughout the sample due to the conventionalisation of systems. If the mapping is intended to account for all kinds of nominal classification devices, it is thus standing to reason that it is not so much the order or type of grammaticalised properties that determines a system's position on the formal scale, but rather the 'grammaticalisation package' as a whole: The more grammaticalised properties a system has (independent from their type), the more concordial class-like it is.

There are seven properties that serve the quantification of a system's ‘formalisation value’: Its inventory size, number of morphosyntactic hosts, locus operandi, obligatoriness of classification, boundedness of markers, presence or lack of multiple marking, and the exhaustivity of classification. These can be employed to a high, medium, or low degree by an individual system. In order to calculate a mean value of a system’s formalisation value, each property’s submeasures are translated into numerical subvalues.

Inventory size The first criterion relates to a system’s inventory size: Larger systems are less grammaticalised than smaller systems. In order to capture the inventory size in a more finegrained way, the inventories of 140 cross-linguistic systems (cf. table 7.5) were consulted, before the actual measurements below were determined.

Table 7.5: Selected languages’ inventory size.

<table>
<thead>
<tr>
<th>Language 1</th>
<th>Language 2</th>
<th>Inventory Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chukchi</td>
<td>Mawung</td>
<td>7–13 Njangumarta</td>
</tr>
<tr>
<td>Emmi</td>
<td>North Ambrym</td>
<td>7–13 Kayardild</td>
</tr>
<tr>
<td>Hadza</td>
<td>Palikūr</td>
<td>7–13 Wati</td>
</tr>
<tr>
<td>Hausa</td>
<td>Persian</td>
<td>7–13 Diyari</td>
</tr>
<tr>
<td>Kilivila</td>
<td>Persian</td>
<td>7–13 Koko Bera</td>
</tr>
<tr>
<td>Kusubian</td>
<td>Tsafiki</td>
<td>7–13 Kugu-Muninh</td>
</tr>
<tr>
<td>Manam</td>
<td>Wogamusin</td>
<td>7–13 Wik-Mungkh</td>
</tr>
<tr>
<td>Masai</td>
<td>Anejom</td>
<td>14 Akatek</td>
</tr>
<tr>
<td>Sumerian</td>
<td>Baining</td>
<td>14 Babungo</td>
</tr>
<tr>
<td>Teligu</td>
<td>Ingush</td>
<td>14 Ndal</td>
</tr>
<tr>
<td>Tlingit</td>
<td>Jarawa</td>
<td>14 Tashkent Uzbek</td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td>Lonwolwol</td>
<td>15 Kanjobal Mayan</td>
</tr>
<tr>
<td>German</td>
<td>Motuna (CCS)</td>
<td>15 Tibetan</td>
</tr>
<tr>
<td>Ket</td>
<td>Ndunga-Le</td>
<td>16 Ngan’gityemerri</td>
</tr>
<tr>
<td>Maasai</td>
<td>Rikvani Andi</td>
<td>17 Itonama</td>
</tr>
<tr>
<td>Paxtijamalj</td>
<td>Sakao</td>
<td>18 Hungarian</td>
</tr>
</tbody>
</table>
### 7.2. Theoretical Considerations

<table>
<thead>
<tr>
<th>Language</th>
<th>Inventory Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamil</td>
<td>6</td>
</tr>
<tr>
<td>Ata</td>
<td>6</td>
</tr>
<tr>
<td>Bounaa Fijian</td>
<td>7</td>
</tr>
<tr>
<td>Burushaski</td>
<td>7</td>
</tr>
<tr>
<td>Dyirbal</td>
<td>7</td>
</tr>
<tr>
<td>Gaagudju</td>
<td>7</td>
</tr>
<tr>
<td>Godié</td>
<td>7</td>
</tr>
<tr>
<td>Gurr-Goni</td>
<td>7</td>
</tr>
<tr>
<td>Khinalug</td>
<td>7</td>
</tr>
<tr>
<td>Lolovoli</td>
<td>8</td>
</tr>
<tr>
<td>Marind</td>
<td>8</td>
</tr>
<tr>
<td>Mian</td>
<td>8</td>
</tr>
<tr>
<td>Omani Arabic</td>
<td>8</td>
</tr>
<tr>
<td>Paamese</td>
<td>8</td>
</tr>
<tr>
<td>Paumarí</td>
<td>9</td>
</tr>
<tr>
<td>Rotokas</td>
<td>10</td>
</tr>
<tr>
<td>Tape</td>
<td>10</td>
</tr>
<tr>
<td>Wambaya</td>
<td>10</td>
</tr>
<tr>
<td>Zande</td>
<td>10</td>
</tr>
<tr>
<td>!Xóó</td>
<td>10</td>
</tr>
<tr>
<td>Anindilyakwa</td>
<td>10</td>
</tr>
<tr>
<td>Araki</td>
<td>11</td>
</tr>
<tr>
<td>Bengali</td>
<td>11</td>
</tr>
<tr>
<td>Chamalal</td>
<td>11</td>
</tr>
<tr>
<td>Chambri</td>
<td>11</td>
</tr>
<tr>
<td>Chechen</td>
<td>11</td>
</tr>
<tr>
<td>Chenapian</td>
<td>11</td>
</tr>
<tr>
<td>Ejagham (CLS)</td>
<td>12</td>
</tr>
<tr>
<td>Gunbarlang</td>
<td>12</td>
</tr>
<tr>
<td>Iwam</td>
<td>12</td>
</tr>
<tr>
<td>Lenakel</td>
<td>13</td>
</tr>
</tbody>
</table>

Note that this criterion only tackles the actual size of the inventory and not the semanticity of the individual classes. If a system's transparency decreases by a large inventory, this is typically compensated by straightforward and thus easily interpretable class semantics, whereas small inventories often are less straightforward with respect to their semantics and thus less transparent. For this reason, class semanticity serves as a determination criterion of systemic transparency, and not its grammaticalisation. The measurements for the value of the inventory size are based on the sample study and are listed in table 7.6.
Table 7.6: Formalisation measures: Inventory size.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;130 classes</td>
<td>low</td>
</tr>
<tr>
<td>91–130 classes</td>
<td>0.125</td>
</tr>
<tr>
<td>71–90 classes</td>
<td>0.25</td>
</tr>
<tr>
<td>51–70 classes</td>
<td>0.375</td>
</tr>
<tr>
<td>31–50 classes</td>
<td>medium</td>
</tr>
<tr>
<td>21–30 classes</td>
<td>0.625</td>
</tr>
<tr>
<td>11–20 classes</td>
<td>0.75</td>
</tr>
<tr>
<td>5–10 classes</td>
<td>0.875</td>
</tr>
<tr>
<td>&lt;5 classes</td>
<td>high</td>
</tr>
</tbody>
</table>

**Host number**  The second criterion relates to the number of morphosyntactic hosts for classificatory markers: It is low if only one morphosyntactic host accepts classificatory marking, which comes down to a numerical value of 0 (e.g. the classifiers of Boumaa Fijian, which occur only together with the possessive).

It is high, if five or more morphosyntactic hosts for classificatory marking are available (e.g. the German class marking, which occurs on articles, adjectives, possessive, interrogative, demonstrative, relative, and personal pronouns etc.). If a system possesses three potential morphosyntactic hosts for classificatory marking, it lies in the medium range of this property. Mandarin classifiers for instance may occur together with quantifiers, demonstratives, or the bare noun.

The number of hosts for classificatory markers increases by the grammaticalisation routine of decategorialisation: Classifiers are reported to originally occur in a single morphosyntactic context (e.g. in Hungarian or Boumaa Fijian); by means of grammaticalisation, this context can be extended, which results in a multiple classifier system (e.g. in Mandarin and Nùng). Concordial class systems are renowned for marking a noun’s class membership on a wider range of nominal satellites: “Clearly the more targets that mark gender, the more canonical the system” (Corbett and Fedden 2015: 16). Therefore, the value of the criterion of a system’s number of hosts is calculated as presented in table 7.7.

Table 7.7: Formalisation measures: Host number.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>one host</td>
<td>low</td>
</tr>
<tr>
<td>two hosts</td>
<td>0.25</td>
</tr>
<tr>
<td>three hosts</td>
<td>medium</td>
</tr>
<tr>
<td>four hosts</td>
<td>0.75</td>
</tr>
<tr>
<td>five or more hosts</td>
<td>high</td>
</tr>
</tbody>
</table>

**Locus operandi**  The third property relates to the locus operandi of a system: A given system can mark class membership inside the borders of the classified
noun’s phrase, outside the borders of the classified noun’s phrase, and even beyond clausal borders. Again, “the more evidence there is for gender […] the more canonical the system[…]”, hence with respect to a system’s locus operandi “the more domains the better” (Corbett and Fedden 2015: 23). A highly grammaticalised system is expected to entertain more domains than a less grammaticalised one. Most systems allow their markers to function as carriers of anaphora irrespective of their type, yet all classifier systems from the sample except the ones of Nasioi, Bora, Mundurukú, and Ngan’gityemerri display classifiers only in NPs, even if the marking occurs across clausal borders. A distinction of clause-internal and -external marking thus seems not a valid measure for the locus operandi. Instead, the locus operandi distinguishes NP-internal marking within the NP that hosts the classified noun (cf. 2.19 A = 7.4 A) from NP-internal classifier occurrence without the classified noun (cf. 2.19 B = 7.4 B, where the class marker serves as a carrier of anaphora and thus occurs outside the borders of the NP that hosts the classified noun, but is embedded in another NP), and NP-in- and -external marking (cf. 2.19 C = 7.4 C, where a class marker occurs on the verb and thus NP-externally).

(7.4) Bora (cf. Seifart 2005: 259 and 2007: 426)

A \[\text{í:huː-pé} \quad [\text{te-ne} \quad \text{tòdhí-hu}]_{\text{NP}} \quad \text{yesterday-PST} \quad 3\text{SG}-\text{CL:INANIMATE} \quad \text{blowgun-CL: TUBE} \quad \text{o} \quad \text{pákiguá:jhku-k} \quad […] \quad 1\text{SG.SUBORDINATE} \quad \text{rasp-PURPOSE} \]

“Yesterday I rasped the [inanimate] blowgun, […].”

B \[ […] \quad \text{a-ne} \quad \text{nuh} \text{tsókâ-\dot{\text{n}}} \quad \text{ú\ddot{\text{b}}\ddot{\text{ı}}\ddot{\text{o}}\ddot{\text{ʔ}}\ddot{\text{ʔ}}-kó} \quad […] \quad \text{CON-CL:INANIMATE} \quad 1\text{SG} \quad \text{try.out-PRD} \quad \text{good} \]

“[…] and I tried it [inanimate] out: good […].”

C \[ […] \quad \text{tëtsi:t-\dot{\text{u}}} \quad \text{á:bačá-hpi-ké} \quad \text{\ddot{\text{a}}\ddot{\text{h}}\ddot{\text{k}}\ddot{\text{u}}\ddot{\text{k}}-\dot{\text{b}}\ddot{\text{u}}} \quad \text{then-TENSE.ASPECT.MOOD} \quad \text{owner-CL:M.SG-ACC} \quad 1\text{SG} \quad \text{give-PF:FUT.PREDICATIVE} \quad 3\text{SG}-\text{CL:TUBE-ALLATIVE} \]

“[…] and then, indeed, I will give it [tube-shaped] to its owner.”

The locus operandi of a given system translates into the numerical values listed in table 7.8.

Table 7.8: Formalisation measures: Locus operandi.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>marking within classified noun’s phrase</td>
<td>low</td>
</tr>
<tr>
<td>NP-internal marking outside classified noun’s phrase</td>
<td>0.5</td>
</tr>
<tr>
<td>marking outside the noun phrase</td>
<td>high</td>
</tr>
</tbody>
</table>

Note that this criterion requires the absolute value of the individual measure found in a given system. It does not allow for the calculation of a mean value, e.g. if a system allows for classificatory marking within the classified noun’s phrase and outside the noun phrase, but not for NP-internal marking outside the classified noun’s phrase.
Obligatoriness The fourth property is the obligatoriness of classification: A system may classify its nouns obligatorily, i.e. in all morphosyntactic contexts, or only in a restricted set of contexts or a single one. The sample study demonstrates that these instances of obligatoriness relate to a system’s degree of grammaticalisation in that concordial class systems classify their nominal lexicon obligatorily, and classifier systems non-obligatorily. The most extreme case of a lack of obligatoriness from the sample is the Hungarian numeral classifier system, where classification is optional even in quantificational contexts, where other numeral classifier systems employ an obligatory classification. The measurements for a system’s obligatoriness are given in table 7.9.

Table 7.9: Formalisation measures: Obligatoriness.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>fully optional classification</td>
<td>0</td>
</tr>
<tr>
<td>obligatory classification (single morphosyntactic context)</td>
<td>0.333</td>
</tr>
<tr>
<td>obligatory classification (restricted morphosyntactic contexts)</td>
<td>0.667</td>
</tr>
<tr>
<td>obligatory classification (across contexts)</td>
<td>1</td>
</tr>
</tbody>
</table>

Boundedness The fifth property relates to the boundedness of the classificatory markers: Systems may mark a noun’s class membership by morphosyntactically independent morphemes, clitics, or affixes, which translates to a rise in the marker’s degree of grammaticalisation. The most extreme form of ‘boundedness’ is the fusion and phonological erosion of an affix with its host to the point where both become unseparable; this is for instance the case in German, where the article *der m.*, *die f.*, *das n.* are not segmentable anymore. These instances are treated as equally bound as affixes, since they are not a requirement for nominal classification devices. Only if they are reanalysed by the speaker community as a separate class marker, this has an effect on its boundedness, since the reanalysed elements are free markers. The measures that determine the value of a system’s boundedness are listed in table 7.10.

Table 7.10: Formalisation measures: Boundedness.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>independent markers</td>
<td>0</td>
</tr>
<tr>
<td>clitic markers</td>
<td>0.5</td>
</tr>
<tr>
<td>affixed/infixed/fused markers</td>
<td>1</td>
</tr>
</tbody>
</table>

Systems may employ both bound and unbound markers at the same time; in this case, both numerical values are added and their mean value is determined. Ngan’ghtmerri and Korean for instance both employ free and bound markers in their inventory (cf. Reid 1997: 173 and Martin 1992: 179ff.), which results
in a boundedness value of 0.5: 0 (free markers) + 1 (affixed markers) = 1 is divided by the number of values met in the given system, which is 2 in this case.

**Marking** The occurrence of single or multiple classificatory marking is the sixth property: A system may feature a single marker per classificatory context or multiple markers. This is irrespective of the number of the nominal satellites that are able to carry classificatory marking. In multiple classifier systems for instance, a single classifier occurs, even if there are two or more modifiers that are capable to host a classificatory marker (cf. 147 A = 7.5, where only one of the available classificatory hosts carries class marking).

(7.5) Mandarin (cf. Sun 2006: 160)

\[ \text{nà yí-pǐ hēi mā} \]

DEM:REMOTE one-CL:HORSE black horse

“That black horse.”

In contrast, the more grammaticalised concordial class systems commonly mark all available hosts for a noun’s class, which results in the multiple occurrence of classificatory markers in the same context (cf. 2.68 = 7.6, where all available targets are marked for the nominal referent’s class).

(7.6) German

\[ \text{Sylvia Maria Kristel} \quad \text{war eine} \quad \text{niederländische} \]
\[ \text{Schauspielerin.} \quad \text{Sie spielte in mehr als fünfzig internationalen Spielfilmen.} \]

“Sylvia Maria Kristel was a Dutch actress. She played in more than fifty international movies.”

The measures for a system’s marking distinguishes thus single from multiple marking in the same context. Their values are provided in table 7.11.

Table 7.11: Formalisation measures: Single vs. multiple marking.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>single marking</td>
<td>low</td>
</tr>
<tr>
<td>multiple marking</td>
<td>high</td>
</tr>
</tbody>
</table>
Exhaustivity The exhaustivity of classification is the seventh and last property determining a system’s formalisation value: A system of nominal classification may be able to classify the full nominal lexicon, or only a part of it. For instance, abstract nouns and nouns denoting objects of mixed ingredients in Jakaltek cannot be classified (cf. Craig 1986a: 273 and 276). In Tohono O’odham, there are only two referential classes, animate and inanimate; these allow a classification of the full nominal lexicon. At an advanced degree of grammaticalisation, the system is expected to classify the full nominal lexicon. The last measure therefore distinguishes the exhaustivity of a classificatory system according to the measures in table 7.12.

Table 7.12: Formalisation measures: Exhaustivity.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>partial classification of the nominal lexicon</td>
<td>low</td>
</tr>
<tr>
<td>exhaustive classification of the nominal lexicon</td>
<td>high</td>
</tr>
</tbody>
</table>

Note that the measures are at times represent tendencies rather than bijective criteria. This issue is also acknowledged by Audring (in print: 19) and also concerns the transparency measurements. Nevertheless, they represent valid measurements, which can be refined and specified with help of the insights gained by future studies. If the dimension of the formal criteria that are described above are all accounted for, they are added up and divided by seven, which is the number of criteria at use. The result is the mean ‘formalisation value’, which is between 0 and 1 and can be represented on a linear graph. In the following subsection, the dimension of transparency is addressed; table 7.13 summarises the criteria for the determination of the transparency value.

Table 7.13: Formal criteria: Summary.

| inventory size | low 0 | >130 classes |
|               | 0.125 | 91–130 classes |
|               | 0.25  | 71–90 classes |
|               | 0.375 | 51–70 classes |
|               | 0.5   | 31–50 classes |
|               | 0.625 | 21–30 classes |
|               | 0.75  | 11–20 classes |
|               | 0.875 | 5–10 classes |
|               | high 1 | <5 classes |

| host number | low 0 | one host |
|            | 0.25  | two hosts |
|            | 0.5   | three hosts |
|            | 0.75  | four hosts |
|            | high 1 | five or more hosts |

| locus operandi | low 0 | marking within the classified noun’s phrase |
|               | 0.5   | general NP-internal marking |
|               | high 1 | marking outside the noun phrase |
7.2. Theoretical Considerations

<table>
<thead>
<tr>
<th>obligatoriness</th>
<th>low</th>
<th>0</th>
<th>fully optional classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.333</td>
<td>obligatory classification (single context)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.667</td>
<td>obligatory classification (restricted set of contexts)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>obligatory classification (across contexts)</td>
</tr>
<tr>
<td>boundedness</td>
<td>low</td>
<td>0</td>
<td>independent markers</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>clitic markers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>affixed/infixed/fused markers</td>
</tr>
<tr>
<td>marking</td>
<td>low</td>
<td>0</td>
<td>single marking</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>multiple marking</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>low</td>
<td>0</td>
<td>partial classification of the nominal lexicon</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>exhaustive classification of the nominal lexicon</td>
</tr>
</tbody>
</table>

7.2.2 The Dimension of Transparency

The sample systems show different degrees of transparency or accessibility irrespective of their systemic types: There are highly transparent classifier systems (e.g. the ones of Hmong Daw and Nùng) and concordial class systems (e.g. the one of Maasai and Sumerian), and there are idiosyncratic classifier systems (e.g. the higher registers of Mandarin) and concordial class systems (e.g. the ones of Masri and German). As different degrees of transparency are found throughout the sample, transparency appears to determine a system’s nature as much as its morphosyntactic expression and thus is introduced to the mapping of nominal classification devices. The calculation of a system’s mean ‘transparency value’ requires the isolation of the criteria that make a system’s transparency and that allow for quantification. The established criteria that serve the determination of the formalisation value fall under the scope of a system’s inventory, its assignment, and its semantic contribution.

Semantic structure  The first criterion relates to the semantic structure of the classes: Referential classes may possess a wider or narrower semantic range, both intrasystemically and intersystemically. The extreme ends of this range are classes that host one member only (e.g. the ‘place class’ in Maasai and the ‘basket class’ in North Ambrym) and classes that have no recognisable underlying semantics (e.g. in the Ingush system, where only two of the six classes have recognisable semantics; cf. Nichols 2011: 144). semantically bijective classes such as “non-flesh food”, “human female”, “inanimate”, or “round object” are more transparent than classes that host different semantic areas, such as the Dyirbal class for females, water, fire, fighting, and attributes that are related to the latter three. The most intraparent kind of class is naturally the one that (seemingly) lacks underlying semantics (e.g. most Chichewa classes; cf.

---

3Note that the notion of “transparency” as it is presented here holds for a system’s actual structure and does not imply consequences for e.g. first language acquirers. While the age of a speaker’s full acquisition of a nominal classification system may vary interlinguistically, it typically is acquired to a full extent by unimpaired first language learners. The mapping method considers transparency an abstract, linguistic dimension, which any speaker may be unaware of, both in the process of acquisition and use.

Table 7.14: Formalisation measures: Semantic class structure.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>opaque classes</td>
<td>low</td>
</tr>
<tr>
<td>mixed classes</td>
<td>0.5</td>
</tr>
<tr>
<td>bijective classes</td>
<td>high</td>
</tr>
</tbody>
</table>

Rule types and productivity  
The transparency of a system’s assignment comprises two criteria, the first being its rule types and their productivity: Systems may employ semantic assignment rules, morphological ones, or phonological ones. Beyond that, nouns may belong to a given referential class due to opaque rules, i.e. rules that are not productive anymore at the system’s synchronic state. In this case, a noun’s class membership is inherited, while the respective assignment rule is not. German for instance employs all four types of assignment rules (cf. Audring in print: 23); in contrast, Sumerian nouns display a straightforward semantic assignment only. Though virtually all systems have semantic assignment rules at their command, I do not consider the type of rule to be of influence on a system’s transparency. Instead, the number of a system’s productive assignment rule types is taken as a measure – the more rule types a system employs, the less transparent it is (cf. table 7.15).

Table 7.15: Formalisation measures: Rule types and productivity.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>four rule types</td>
<td>low</td>
</tr>
<tr>
<td>three rule types</td>
<td>0.333</td>
</tr>
<tr>
<td>two rule types</td>
<td>0.667</td>
</tr>
<tr>
<td>one rule type</td>
<td>high</td>
</tr>
</tbody>
</table>

Rule number  
The third criterion is the number of rules: A system’s assignment is more transparent if it possess a small number of assignment rules (independent of a system’s number of rule types). The more assignment rules it has, the more intransparent it becomes. Sumerian for instance has a straightforward human and a straightforward nonhuman class, which results in only two assignment rules: Nouns denoting humans are assigned to the human class and nouns denoting nonhumans are assigned to the nonhuman class. German on the other hand possesses countless assignment rules: The feminine class for instance comprises among others females, ships, motorcycles, most trees, nouns ending in -ung, -heit, -keit, -schaft, and loans ending in -euse, -age, and -ion. This short list comprises eleven assignment rules, which cover only a fraction
of the total number of German class assignment rules. The high number of rules awarded German the reputation to possess a semantically unmotivated, intransparent, and ‘unlearnable’ classification system. The distinction of rule numbers thus is adopted as a measure for a system’s transparency; its measures are listed in table 7.16.

Table 7.16: Formalisation measures: Rule number.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>high number of rules</td>
<td>low</td>
</tr>
<tr>
<td>low number of rules</td>
<td>high</td>
</tr>
</tbody>
</table>

A system’s semantic contribution to the noun phrase has an effect on its transparency: A system that productively contributes to a language’s discourse semantics is more accessible than a semantically opaque system, whose underlying structure is not readable from linguistic input. As there is no reliable way to measure the degree of a given system’s semantic contribution, four subcriteria that relate to the identifiability of its semantic contribution are adopted for the calculation of the transparency value:

**Categorial independence** Class markers may either exclusively convey information on a noun’s class or be fused with other grammatical categories such as person, number, or case. The sample study has confirmed that lowly grammaticalised systems typically convey class information only (e.g. Tohono O’odham, Hungarian, or Korean), whereas highly grammaticalised systems typically are marked by portmanteau markers, which are fused with at least one other grammatical category (e.g. case and number in German). A system with categorially independent class markers is more transparent than a system with fused class markers, as there is less linguistic information to be distinguished and decoded. Similarly, a system fused with more categories is less accessible than a system fused with fewer categories. This results in the measurements provided in table 7.17.

Note that this criterion accounts only for grammatical categories and not for other dimensions such as specificity or definiteness in discourse. The latter dimensions are frequently involved in all kinds of nominal classification systems (cf. e.g. Contini-Morava and Kilarski 2013: 283ff.), but are different from grammatical categories. Beyond that, discoursal properties can only be accounted for properly, if there is sufficient discourse data available. Furthermore, the highly frequent phenomenon of systems that distinguish class membership only in the third person are not considered as fusing the category of person and class: In these systems, nominal classification is limited to the third person, but is not interacting with the category of person in the sense that the class marking consistently applies to the paradigm of the category of person.
Table 7.17: Formalisation measures: Categorial independence.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>portmanteau markers (fused with three or more categories)</td>
<td>low</td>
</tr>
<tr>
<td>portmanteau markers (fused with two categories)</td>
<td></td>
</tr>
<tr>
<td>portmanteau markers (fused with one category)</td>
<td></td>
</tr>
<tr>
<td>categorially independent markers</td>
<td>high</td>
</tr>
</tbody>
</table>

Discreteness of the classificatory markers  A highly transparent system employs discrete, bijective markers per class, i.e. it follows a one-to-one mapping of noun-class-relationships (cf. e.g. the inventory of the Nùng classifier system, which has one and only one distinct classifier per class). In contrast, a system that employs either several markers for the same class or uses the same marker to index different classes is less bijective and therefore less transparent (cf. e.g. the German definite articles in table 38 on page 422 which feature forms that occur in all three classes). There are also mixed systems that employ both markers that index a single class and markers that are employed by several classes (e.g. Chichewa). The measures for the discreteness of a given system’s markers therefore comprise the values listed in table 7.18.

Table 7.18: Formalisation measures: Discreteness of the markers.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>markers encode more than one class</td>
<td>low</td>
</tr>
<tr>
<td>mixed set of markers</td>
<td></td>
</tr>
<tr>
<td>distinct markers per class</td>
<td>high</td>
</tr>
</tbody>
</table>

Redundancy  A system’s markers may carry semantics, and these semantics can either contribute to the noun phrase where the markers occur or mirror redundant semantic content that is already present in the classified noun. Beyond that, markers can lack semantic content. It is not uncommon to find class markers partly contributing and partly not: Cf. e.g. German *der Gläubige* m. “the (male or unspecified) believer” vs. *die Gläubige* f. “the (female) believer”, where the class marker conveys information on the biological sex of the nominal referent, whereas no information is conveyed in the class marking of *der Löffel* m. “the spoon” vs. *die Gabel* f. “the fork”. There are also instances of redundant class marking: Cf. e.g. *der Hengst* m. “the stallion” and *die Stute* f. “the mare”, where the information on the referent’s biological sex is present in the lexical noun and repeated in the class marking. The measure for a system’s semantic redundancy thus comprises the values given in table 7.19.
7.2. Theoretical Considerations

Table 7.19: Formalisation measures: Redundancy.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantically vain class marking</td>
<td>low</td>
</tr>
<tr>
<td>semantically redundant class marking</td>
<td>0.5</td>
</tr>
<tr>
<td>semantically informative class marking</td>
<td>high</td>
</tr>
</tbody>
</table>

Flexibility  Systems may allow a noun to belong to more than one class, or they may allow a noun to have a single class value. For instance, the noun for turtle in German, *Schildkröte*, is rigidly classified as a member of the feminine class, whereas in Bora, it may be classified as e.g. masculine or a two-dimensional, disc-like object (cf. 5.12 A and D = 7.7).

(7.7) Bora (cf. Seifart 2005: 169)

A  
\[ \text{māhu-hi} \]
DISTANT-CL:2D.ROUND be.big.SUB-CL:2D.ROUND
kū:mā-hi
turtle-CL:2D.ROUND
“that big turtle”

B  
\[ \text{māhu-be} \]
“that big turtle”

A noun’s flexibility in class membership is not necessarily related to a formalisation process, though grammaticalisation theory predicts a loss in classificatory flexibility in the case of a system’s formalisation. Instead, a flexible classification allows a much better speaker-sided analysis of the system’s underlying semantics and its contribution to discourse. The measurements for a system’s flexibility are provided in table 7.20.

Table 7.20: Formalisation measures: Flexibility.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>fully rigid classification</td>
<td>low</td>
</tr>
<tr>
<td>largely rigid classification</td>
<td>0.333</td>
</tr>
<tr>
<td>largely flexible classification</td>
<td>0.667</td>
</tr>
<tr>
<td>fully flexible classification</td>
<td>high</td>
</tr>
</tbody>
</table>

Note that some of the criteria stand in a mutual correlation, namely the inventory size, the semantic class structure, the number of rules, and the redundancy of conveyed semantic categories. In the following, I argue why they are to be treated as separate criteria nonetheless. First, the inventory size may correlate with the semantic class structure in that a small number of classes may lack semantic specificity of classes, whereas a large number of classes has
a higher probability to employ semantically more specific classes. German, for instance, has three classes, masculine, feminine, and neuter. Only a fraction of the respective class members such as nouns denoting humans and higher animates can be synchronically motivated for their class membership, while the classes themselves host nouns from all semantic areas. Mandarin Chinese on the other hand possesses about 200 classes, which are semantically highly specific. Nevertheless, small inventories that are semantically straightforward exist, for instance in Sumerian, where there is one class for humans and one for nonhumans. This justifies a separation of the inventory size and the semantic structure of the referential classes.

Second, the number of assignment rules may depend on the inventory size and the semantic structure. Each class has at least one assignment rule; mixed classes may have more than one assignment rule. Systems with larger inventories therefore automatically possess a large number of assignment rules, whereas smaller systems may either have a small number (cf. again Sumerian, where there are two assignment rules), or a large number in case that they are semantically underspecified (cf. again German, where there are 44 rules only for monosyllabic nouns; cf. Köpcke 1982, cited in Audring in print: 20). Since there is no full correlation between the number of rules, the inventory size, and/or the semantic specificity of the classes, a distinction of these criteria is justified.

Third, the redundancy of conveyed class semantics may correlate with the semantic structure of the classes in one respect: If a system possesses opaque classes, these classes are semantically vain, and both criteria score as transparent. However, if a system’s classes are not opaque, but mixed or even bijective, the results for the same system’s semantic redundancy may differ. Maasai for instance has two mixed classes, masculine and feminine, and a specific class, which only hosts the noun wuéjí “place”. While logic suggests that semantically specific classes are more likely to be semantically productive (as it is often the case in classifier systems) than mixed and less specified classes, the opposite is the case in Maasai: The ‘place class’ does not contribute semantically to the noun or its referent, since this notion is already conveyed by the lexical noun wuéjí “place”. The mixed classes contain all other nouns, including ones denoting humans (e.g. en-k-apútání f. “wife’s mother”, òl-apútání m. “wife’s father”), objects (e.g. en-jutét f. “eraser”, òl-pánkà m. “machete”), and abstractions (e.g. e-ksbá f. “hate”, ò-löm m. “envy”). Despite this mixed nature of the masculine and feminine class, they are semantically highly productive and informative due to an inherent connotation to size and strength. For instance, if the feminine noun en-jutét f. “eraser” is shifted to the masculine class, the noun phrase’s semantics change, and the result is òl-jutét “big eraser” (cf. Payne 1998: 171f.). Therefore, the semantic class structure and the redundancy of the classes’ conveyed semantics have to be treated as separate criteria.

After all the abovementioned transparency criteria have been accounted for within a given system, the values are added up and divided by the number of criteria, which is seven. The result of the division is this system’s transparency value, which can be represented on a linear graph. Table 7.21 summarises the
criteria for the determination of the transparency value, before subsection 7.2.3 elaborates the representation of the transparency value and the formalisation value in a twodimensional graph.

Table 7.21: Transparency criteria: Summary.

<table>
<thead>
<tr>
<th></th>
<th>low</th>
<th>0</th>
<th>opaque classes</th>
<th>0.5</th>
<th>mixed classes</th>
<th>high</th>
<th>1</th>
<th>bijective classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>class structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rule types</td>
<td></td>
<td>0</td>
<td>four rule types</td>
<td></td>
<td></td>
<td>high</td>
<td>1</td>
<td>one rule type</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.333</td>
<td>three rule types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.667</td>
<td>two rule types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rule number</td>
<td></td>
<td>0</td>
<td>high number of rules</td>
<td></td>
<td></td>
<td>high</td>
<td>1</td>
<td>low number of rules</td>
</tr>
<tr>
<td>categorical independence</td>
<td>low</td>
<td>0</td>
<td>portmanteau markers (3+ more categories)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.333</td>
<td>portmanteau markers (2 more categories)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.667</td>
<td>portmanteau markers (1 more category)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>categorially independent markers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discreteness</td>
<td>low</td>
<td>0</td>
<td>markers encode more than one class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>mixed set of markers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>distinct markers per class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>semantic redundancy</td>
<td>low</td>
<td>0</td>
<td>semantically vain class marking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>semantically redundant class marking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>semantically informative class marking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flexibility</td>
<td>low</td>
<td>0</td>
<td>fully rigid classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.333</td>
<td>largely rigid classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.667</td>
<td>largely flexible classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>1</td>
<td>fully flexible classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7.2.3 Graphic Representation of Form and Transparency

This subsection introduces the graphic representation of the formalisation value and transparency value and projects it on the core characteristics of nominal classification devices as they presented themselves in the sample study. The formalisation value and the transparency value can be combined and represented in a graph, where the formalisation value is represented on the x-axis and the transparency value is represented on the y-axis (cf. figure 7.4).

Before the mapping is applied to actual systems, I elaborate on its areas and the systems that are represented in them. I do not expect fully intransparent systems to occur, which would be in contradiction with Corbett’s (1991) observation that concordial class systems always have a semantic core. Beyond that, a fully idiosyncratic system violates the principles of linguistic economy and expressivity and would either be expected to drop out of a grammar or to experience a semantic motivation. Figure 7.5 shows all predicted possible phenomena; the boxes marked by small caps host possibly occurring types of systems, and the arrows marked by Roman numerals indicate the possible developmental shifts of systems.
Classifier systems in their prototypical definition and systems that are close to the status of a prototypical classifier system correspond to the top left corner of the graph ($x$: ca. 0–0.4, $y$: ca. 0.6–1): They are fairly ungrammaticalised.
in their formal expression and semantically transparent and productive (cf. area A in figure 7.5). As the sample study has shown, classifier systems may conventionalise, which results in a loss of semanticity and transparency (cf. arrow i in figure 7.5); this loss is however not necessarily related to changes in the systems formal appearance. Such a conventionalised classifier system (e.g. the one of the higher registers of Mandarin) therefore corresponds to the lower left corner (x: ca. 0–0.4; y: ca. 0.2–0.6; cf. area B in figure 7.5). In contrast, a conventionalised system which is not transparent to the speakers anymore may be semantically remotivated and thus gain in its y-value again (cf. arrow ii in figure 7.5).

The definition of prototypical concordial class systems and systems that are close to the status of a prototypical concordial class system correspond to the lower right corner in the graph (x: ca. 0.8–1; y: ca. 0.2–0.6); they are defined as the products of a grammaticalisation process and consequently lack a strong semanticity (cf. area C in figure 7.5). Canonical concordial class systems in contrast correspond to the upper right corner (x: ca. 0.8–1; y: ca. 0.6–1), as they are defined as devices that are characterised by unambiguous semantics (cf. Corbett and Fedden 2015: 26ff.; cf. area D in figure 7.5). Like classifier systems, transparent concordial class systems may become subject to conventionalisation and thus decrease in their y-value (cf. arrow III in figure 7.5), or conventionalised concordial class systems may be semantically remotivated and thus increase in their y-value (cf. arrow IV in figure 7.5).

Less clear-cut cases are expected to occur in the transitional area between those two layers that correspond to the notion of classifiers and concordial classes (x: ca. 0.3–0.7; y: ca. 0.2–1; cf. area E in figure 7.5). These display a formal expression that is involving formalisation processes which are not fully processed (yet) or partial formalisation in that not all systemic properties are subject to change (cf. Nasiö, Bora, Munduruku, and Ngani’gityemmeri). Recall however that these systems are typically characterised by a high semanticity and thus are likely to possess a higher y-value. Considering the different degrees of transparency that classifier systems display and the continuous availability of grammaticalisation for any system, there are several theoretical options available for a classifier system to undergo the grammaticalisation hypothesis: The ‘ideal case’ originates from the state of a prototypical classifier system, which jointly formalises and loses in transparency. The logical endpoint of this development is a prototypical concordial class system (cf. arrow V in figure 7.5). At the same time, a transparent classifier system’s semanticity may be maintained, while its formal properties are formalising (cf. arrow VI in figure 7.5). Beyond that, a conventionalised classifier system may formalise and either remain conventionalised (cf. arrow VII in figure 7.5) or be semantically remotivated in the

---

^Note that the formalisation span of the diagram covered by systems corresponding to the notion of classifiers is larger than the one of concordial class systems. This is due to the observation from the sample that classifier systems are more varied with respect to their degree of grammaticalisation. In addition, fully optional systems were included in the calculation of the x-value, which increases the range of less grammaticalised systems.
process of formalisation or thereafter (cf. arrow VIII in figure 7.5). The model does not permit a system to decrease in its formalisation value. However, since “[g]rammaticalization is unidirectional in the sense that changes that go in the opposite direction of grammaticalization are observed very rarely” (Lehmann 2004: 15; cf. also Aikhenvald 2000: 374), such a reverse shift is not expected to occur.

Not all shifting pathways are equally frequent, as the sample study has shown. Beyond that, not all are in this study’s focus, since the main subject of research is the grammaticalisation hypothesis (arrow V and arrow VI in figure 7.5). Due to their relative frequency in the sample, the conventionalising shifts I and III are of a general interest to this study, just like their opposite shifts II and IV. After the description of the mapping model and its criteria and their values, the model and the shifting pathways that are found in the sample are exemplified in section 7.3.

7.3 Application of the Mapping Instrument

This section maps the sample in order to exemplify the languages’ distribution over the nominal classification continuum. Their distribution is characterised by the systems’ degree of grammaticalisation as well as their transparency. Since a detailed exemplification of all individual sample systems would exceed the spatial limits of this subsection, a respective representative system of transparent and conventionalised systems with a low degree of grammaticalisation (Hungarian and the higher linguistic registers of Mandarin), currently formalising systems (Bora), and conventionalised and transparent systems with a high degree of grammaticalisation (German and Burushaski) are analysed in detail. These were hypothesised to correspond to the areas A–E in figure 7.5. After the detailed account, the dimensions of the model are applied to the sample as a whole. The two dimensions are discussed separately (first the formalisation value, then the transparency value), before the values are combined and presented in form of the full model.

7.3.1 The Formalisation Value

The formalisation value is determined on basis of six criteria, which are presented in tabular form and then discussed for each language individually. Table 7.22 lists the criteria and their values for Hungarian.
Table 7.22: The formalisation value of Hungarian.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>12</td>
</tr>
<tr>
<td>host number</td>
<td>0.5</td>
</tr>
<tr>
<td>locus operandi</td>
<td>0.5</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>0</td>
</tr>
<tr>
<td>boundedness</td>
<td>0</td>
</tr>
<tr>
<td>marking</td>
<td>0</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>0</td>
</tr>
<tr>
<td>mean</td>
<td>0.214</td>
</tr>
</tbody>
</table>

As predicted, Hungarian displays a low degree of formalisation: It attaches its classifiers to numerals, quantifiers, and demonstratives (cf. 7.8 A–C) and thus within the borders of the NP only. The classifiers may function as carriers of anaphora and thus appear transclausal (cf. 7.8 D and E).

(7.8) Hungarian (cf. Csirmaz and Dékány 2010a: 5f. and 11)

A három fej saláta
three CL:HEAD lettuce
“three heads of lettuce”

B sok fej saláta
many CL:HEAD lettuce
“many lettuces”

C az a fej saláta
that ART CL:HEAD lettuce
“that [head of] lettuce”

D az a szem gyöngy
that ART CL:EYE pearl
“that pearl”

E az a szem
that ART CL:EYE
“that one”

They occur as free forms, optionally, and once per classificatory context (cf. 5.1 = 7.9 A–C). The same noun can take various classifiers and thus is largely flexible with respect to its classification (cf. 7.9 D–E). The classification is not fully flexible though: For instance, the ‘general’ classifier darab “piece” can only be applied to segmentable objects according to my informant, not to e.g. humans. Other classifiers such as kötet “volume” find no cognate category that would allow for a classificatory variation.

(7.9) Hungarian (cf. Csirmaz and Dékány 2010a: 12 and 2010b: 8)

A hét (fej) saláta
seven CL:HEAD lettuce
“seven lettuces”
Next, Mandarin is accounted for as a representative of a lowly formalised system. Since Mandarin shows a distinction between its lower registers, which are not conventionalised, and its higher registers, which are conventionalised, this account is only directed at the higher registers of Mandarin for the sake of exemplifying the mapping. The account for the full sample includes again the full system of Mandarin, while the following description focuses on its conventionalised parts. Table 7.23 gives an overview over the criteria that the formalisation value of Mandarin is based on.

Table 7.23: The formalisation value of Mandarin.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>187</td>
</tr>
<tr>
<td>host number</td>
<td>3 (numerals/quantifiers, demonstratives, nouns)</td>
</tr>
<tr>
<td>locus operandi</td>
<td>NP-internal (free of classificatory context)</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>obligatory in restricted contexts</td>
</tr>
<tr>
<td>boundedness</td>
<td>free and clitic markers</td>
</tr>
<tr>
<td>single marking</td>
<td>0</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>exhaustive classification</td>
</tr>
<tr>
<td>mean</td>
<td>0.417</td>
</tr>
</tbody>
</table>

Mandarin attaches its classifiers to numerals and other quantifiers, demonstratives, and also the bare noun. They can occur in free and bound form (cf. \[147 = 7.10\,\text{A–C}\]) and are an obligatory device in quantificational and demonstrative contexts (Zhang 2007: 47). All instances of nominal classifiers occur NP-internally, also if the classifier is used as a carrier of anaphora (cf. \[153 = 7.10\,\text{D}\]).

Note that the ‘boundedness value’ in table 7.23 amounts to 0.25, which is not an individual measure in the model. In case that a system employs more than one measure at the same time, all met values are added up and their mean value is calculated. In the case of Mandarin, we find free markers with a value of 0 and clitic markers with a value of 0.5. The sum of both divided by the number of met measures, which is two, results in a value of 0.25.

A
\[ \text{na} \quad \text{yì-pí} \quad \text{hēi} \quad \text{mă} \]
DEMON:REMOTE one CL:horse black horse
“That black horse.”

B
\[ \text{na} \quad \text{pì} \quad \text{hēi} \quad \text{mă} \]
DEMON:REMOTE CL:horse black horse
“That black horse.”

C
\[ \text{măi} \quad \text{pì} \quad \text{hēi} \quad \text{mă} \]
buy CL:horse black horse
“Buy one black horse.”

D
\[ \text{Zhe} \quad \text{jian} \quad \text{wuzi} \quad \text{you} \quad \text{liang} \quad \text{zhang} \quad \text{shuzhuo}, \text{na} \quad \text{zhang} \quad \text{shi} \quad \text{wo} \quad \text{de}. \]
this CL:room room have two CL:flat desk that CL:flat be mine
“This room has two desks, that (one) is mine.”

Furthermore, the classifiers occur once per classificatory context (cf. again A, where both a demonstrative and a numeral occur, but only the numeral carries a classifier). Finally, “traditional and teaching descriptions of Mandarin require an invariant, specific classifier in an enormous variety of cases” (Erbaugh 1986: 404ff., cited in Zhang 2007: 54), which comes down to a largely rigid classification, but not a full one.

Turning over to the more formalised end of the spectrum, Burushaski represents a more transparent system. Table 7.24 summarises its formal criteria and their values.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>4</td>
</tr>
<tr>
<td>host number</td>
<td>5+ (numerals, pronouns, adjectives, nouns, verbs)</td>
</tr>
<tr>
<td>locus operandi</td>
<td>NP-internal (free of classificatory context) and -external</td>
</tr>
<tr>
<td>obligatory</td>
<td>obligatory across contexts</td>
</tr>
<tr>
<td>boundedness</td>
<td>affixed/fused markers</td>
</tr>
<tr>
<td>marking</td>
<td>multiple marking</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>exhaustive classification</td>
</tr>
<tr>
<td>mean</td>
<td>1</td>
</tr>
</tbody>
</table>

As expected, Burushaski displays a high degree of formalisation: Class distinctions are found in verbal affixes and auxiliary forms, the nominal plural endings, numerals, adjectives, and a range of pronouns (demonstratives, interrogatives, possessives), where they occur in affixed or fused form. The occurrence of classificatory markers on the verb proves the ability of the markers to occur NP-externally. Beyond that, classificatory marking on e.g. pronouns allows the markers to occur NP-internally, but outside of the immediate classificatory context (cf. 7.11, where reference is made to an earlier established discourse referent).
Mapping Nominal Classification

(7.11) Burushaski (cf. Berger 1998: 80)

\[ \text{máímo} \quad \text{báalar} \quad \text{s-chá-m-a.} \]
POSS3SG.F house-GEN.RESIDUE 3SG.ANIMATE-brought-IPF-3SG.F

“She brought him to her (own) house.”

Since every Burushaski noun belongs to a class (Berger 1998: 33), the classification is obligatory. Its markers occur obligatorily and multiply across contexts (cf. 7.12).

(7.12) Burushaski (cf. Tiffou and Pesot 1989: 54)

\[ \text{we} \quad \text{gušiga} \quad \text{kančéni(-tiŋ)} \quad \text{ban}. \]
DEM:3PL.ANIMATE woman:PL.ANIMATE volatile(-PL.ANIMATE)

“These women are volatile.”

Finally, there is a great amount of classificatory flexibility involved in the concordial class assignment of the Burushaski nominals: Depending on the real world referent’s consistency (continuous vs. individuated), object nouns generally may take either the “concrete” or the “residue” class marker (cf. Berger 1998: 36ff. and Munshi 2006: 163). Furthermore, the masculine and feminine class allow for the distinction of the referent’s biological sex, which results in a largely flexible classification in Burushaski.

Turning over to German, we find a system that is a highly formalised one and lacks an overall transparent structure. Its properties are listed in table 7.25.

Table 7.25: The formalisation value of German.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>3</td>
</tr>
<tr>
<td>host number</td>
<td>5+ (nouns, adjectives, articles, pronouns, numerals,...)</td>
</tr>
<tr>
<td>locus operandi</td>
<td>NP-internal (free of classificatory context)</td>
</tr>
<tr>
<td>obligatoryness</td>
<td>obligatory across contexts</td>
</tr>
<tr>
<td>boundedness</td>
<td>affixed/fused markers</td>
</tr>
<tr>
<td>marking</td>
<td>multiple marking</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>exhaustive classification</td>
</tr>
<tr>
<td>mean</td>
<td>0.929</td>
</tr>
</tbody>
</table>

German employs a large number of morphosyntactic hosts for the classificatory markers: A noun’s class membership can be marked on the noun, the article (definite and indefinite), the adjective, the numeral, and various pronouns (demonstratives, interrogatives, possessives, relatives). Since there is no marking on the verb, adverb, or other NP-external constituents, the marking
occurs only NP-externally, but is not bound to the noun phrase where the
classified noun occurs (cf. 61 where class marking occurs transclausally, yet
NP-externally).

(7.13) German

\[
\begin{array}{lll}
\text{A} & \text{Hast} & \text{Du} \\
& \text{have:PRS2SG} & \text{2SG} \\
& \text{der} & \text{ARTdot:F.SG} \\
& \text{Gesichtstäutorung} & \text{seen} \\
& \text{Frau} & \text{woman:ACC.F.SG} \\
& \text{mit} & \text{with} \\
& \text{der} & \text{ARTdot:DAT.F.SG} \\
& \text{Gesichtstäutorung} & \text{seen} \\
& \text{gesehen?} & \\
\end{array}
\]

“Have you seen the woman with the facial tattoo?”

B

\[
\begin{array}{lll}
\text{Welche?} & \\
\text{INTERR2SG:ACC.F.SG} & \\
\text{“Which [one]?”} & \\
\end{array}
\]

C

\[
\begin{array}{lll}
\text{Die} & \text{da} & \text{verne.} \\
\text{ARTdot:ACC.F.SG} & \text{there} & \text{ahead} \\
\text{“The [one] over there.”} & \\
\end{array}
\]

A noun’s class membership is obligatorily marked across contexts and by
means of affixes or fused forms. The markers may occur multiply in the same
class (cf. 7.14). Except for a small number of nouns that distinguish e.g. the
referred’s sex by classificatory means (e.g. der Gläubige m. “the (male) believer”
vs. die Gläubige f. “the (female) believer”), a noun’s class membership is rigid
and cannot be altered.

(7.14) German

\[
\begin{array}{llllll}
\text{dies-} & \text{nett-} & \text{gesichtstäutoriert-} & \text{Frau} & \text{woman:NOM.F.SG} & \\
\text{this-NOM.F.SG} & \text{nice-NOM.F.SG} & \text{facially.tattooed-NOM.F.SG} & \\
\text{“this nice woman with the facial tattoo”} & \\
\end{array}
\]

Finally, the Bora system is accounted for as a representative of the systems
that are expected to correspond to the intermediate area ε in figure 7.5. Table
7.26 presents Bora’s properties.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>125</td>
</tr>
<tr>
<td>host number</td>
<td>5+</td>
</tr>
<tr>
<td>locus operandi</td>
<td>NP-internal (free of classificatory context) and -external</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>obligatory in restricted contexts</td>
</tr>
<tr>
<td>boundedness</td>
<td>affixed/fused markers</td>
</tr>
<tr>
<td>marking</td>
<td>multiple marking</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>partial classification</td>
</tr>
<tr>
<td>mean</td>
<td>0.470</td>
</tr>
</tbody>
</table>
As expected, Bora displays a highly intermediate formalisation value: It knows a large number of targets of classificatory marking, including nouns, adjectives, pronouns (demonstratives, interrogatives possessives, relatives), numerals, and the verb. The classification is non-obligatory in that not all nouns in all contexts are classified (cf. 7.15 where classificatory marking is omitted altogether) and in that not all nouns can be classified: Non-countables for instance generally remain unclassified (Seifart 2005: 113). However, a restricted set of hosts is obligatorily marked for a noun’s class in discourse, for instance demonstratives (cf. Seifart 2005: 3).

(7.15) Bora (cf. Seifart 2005: 57)

\[
\text{táj } -\text{náʔbe } -\text{úʔú } =\text{pe}
\]
POSS.1SG brother DECEASED PST
“my late brother”

The classificatory markers occur as affixes and are not bound to the borders of the noun phrase. They can be used as carriers of anaphora (cf. 2.19 A and C = 7.16, where the class marker attaches to the verb across clausal boundaries).

(7.16) Bora (cf. Seifart 2005: 259 and 2007: 426)

\[
\text{A c:huʔ-} \text{te-}\text{-ne } \text{tččiː-} \text{hnum}
\]
\[
\text{yesterday-PST 3SG-CL:INANIMATE blowgum-CL:TUBE}
\]
\[
o \text{pákiqviʔhmu-k}:t \quad [\ldots]
\]
1SG.SUPERORDINATE rasp-PURPOSE
“Yesterday I rasped the [manimate] blowgun, [\ldots]”

\[
\text{B } [\ldots] \quad \text{tččitá-t} \text{ʔdum}
\]
\[
\text{ʔ:} \text{bává-hpi-k}:t \quad \text{ʔó}
\]
then-TENSE.ASPECT.MOOD owner-CL:M.SG-ACC 1SG
\[
\text{ʔhku} \text{-} \text{lam}:\text{ʔt}:n \quad \text{te-}\text{-hun-} \text{ʔmu}
\]
give-PF-FUT.PREDICATIVE 3SG-CL:TUBE-ALLATIVE
“[\ldots] and then, indeed, I will give it [tube-shaped] to its owner.”

As has been argued in section 6.3, Bora does not allow multiple marking. Instead, the referentiality of the markers suggests that seeming instances of multiple marking actually represent appositional contexts of multiple classification. This is further confirmed by the classificational variability in these contexts (cf. 5.12 = 7.17).

(7.17) Bora (cf. Seifart 2005: 169)

\[
\text{A c:hi } \text{máʔha} \text{-hi}
\]
\[
\text{DISTANT-CL:2D.ROUND be.big.SUPER-CL:2D.ROUND}
\]

\[
\text{kú:mu} \text{-hi}
\]
turtle-CL:2D.ROUND

“that big turtle”

\[
\text{B c:hi } \text{máʔha} \text{-be}
\]
\[
\text{DISTANT-CL:2D.ROUND be.big.SUB-CL:M.SG}
\]
turtle-CL:2D.ROUND

“that big turtle”
3. Application of the Mapping Instrument

C  aj-di
DISTANT-CL:M.SG
mάhu-hi
be.big.SUB-CL:2D.ROUND
turtle-CL:2D.ROUND
“that big turtle”

D  aj-di
DISTANT-CL:M.SG
mάhu-be
be.big.SUB-CL:M.SG
turtle-CL:2D.ROUND
“that big turtle”

After having accounted for the representative systems individually, table 7.27 gives an overview over their values.

Table 7.27: Overview over the representative systems’ formalisation criteria and value.

<table>
<thead>
<tr>
<th>System</th>
<th>Inventory size</th>
<th>Host number</th>
<th>Locus operandi</th>
<th>Obligatoriness</th>
<th>Boundedness</th>
<th>Marking</th>
<th>Exhaustivity</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungarian</td>
<td>0.75</td>
<td>0.25</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.214</td>
</tr>
<tr>
<td>Mandarin</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.667</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0.417</td>
</tr>
<tr>
<td>Bora</td>
<td>0.125</td>
<td>1</td>
<td>0.5</td>
<td>0.667</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0.470</td>
</tr>
<tr>
<td>German</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.929</td>
</tr>
<tr>
<td>Burushaski</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The predicted results are met well in the evaluation of the systems’ formalisation value. Only the higher linguistic registers of Mandarin exceed the estimated value due to the extended set of contexts in which classifiers occur. Nonetheless, the distribution of systems along the grammaticalisation scale bears no contradictory surprises towards the established notions of the systems themselves (and their commonly assumed types). Also note that there is a strong order in the rising degree of the individual grammaticalising properties, which is only disturbed by the mixed inventory sizes.

After having elaborated the calculation of the formalisation value of the representative systems, the values of the sample as a whole are presented. Table 7.28 lists the individual measures and values that determine the individual systems’ formalisation value. Note that the values of the Mandarin system have been reevaluated, as the full Mandarin system is subject to this study, and not only its higher registers.
Table 7.28: Overview over the sample systems’ formalisation criteria and value.

<table>
<thead>
<tr>
<th>System</th>
<th>Inventory Size</th>
<th>Nominal Number</th>
<th>Locus Operandi</th>
<th>Obligatory</th>
<th>Boundedness</th>
<th>Marking</th>
<th>Exhaustivity</th>
<th>&amp;-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dongo-Ko (13)</td>
<td>0.875</td>
<td>0</td>
<td>0</td>
<td>0.333</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0.208</td>
</tr>
<tr>
<td>Hungarian (08)</td>
<td>0.75</td>
<td>0.25</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.214</td>
</tr>
<tr>
<td>Korean (01)</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.333</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0.226</td>
</tr>
<tr>
<td>Jakaltek (10)</td>
<td>0.625</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.304</td>
</tr>
<tr>
<td>North Ambrym (11)</td>
<td>0.875</td>
<td>0</td>
<td>0</td>
<td>0.333</td>
<td>0.25</td>
<td>1</td>
<td>0</td>
<td>0.351</td>
</tr>
<tr>
<td>Nung (16)</td>
<td>0.75</td>
<td>0.25</td>
<td>0.5</td>
<td>0.333</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.408</td>
</tr>
<tr>
<td>Mandarin (02)</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.667</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
<td>0.417</td>
</tr>
<tr>
<td>Mandarin (formal)</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.667</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
<td>0.417</td>
</tr>
<tr>
<td>Tohono O’odham (14)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.429</td>
</tr>
<tr>
<td>Bouman Fijian (07)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.333</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.476</td>
</tr>
<tr>
<td>Ngun’gityemerri (06)</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>0.333</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0.512</td>
</tr>
<tr>
<td>Bora (04)</td>
<td>0.125</td>
<td>1</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.542</td>
</tr>
<tr>
<td>Mundurukú (05)</td>
<td>0.125</td>
<td>1</td>
<td>1.0000</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.542</td>
</tr>
<tr>
<td>Hmong Daw (15)</td>
<td>0.75</td>
<td>1</td>
<td>0.5</td>
<td>0.667</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.559</td>
</tr>
<tr>
<td>Great Andamanese (09)</td>
<td>0.875</td>
<td>1</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.649</td>
</tr>
<tr>
<td>Dyirbal (18)</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.667</td>
</tr>
<tr>
<td>Nasioi (03)</td>
<td>0.125</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.732</td>
<td></td>
</tr>
<tr>
<td>Taiaue (19)</td>
<td>0.75</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.893</td>
</tr>
<tr>
<td>Ingush (28)</td>
<td>0.875</td>
<td>0.75</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.899</td>
</tr>
<tr>
<td>Sumerian (22)</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.929</td>
</tr>
<tr>
<td>German (30)</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.929</td>
</tr>
<tr>
<td>Tamil (24)</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.964</td>
</tr>
<tr>
<td>Moro (26)</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.964</td>
</tr>
<tr>
<td>Ket (27)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.964</td>
</tr>
<tr>
<td>Dongo-Ko (25)</td>
<td>0.875</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.982</td>
</tr>
<tr>
<td>Chichewa (29)</td>
<td>0.875</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.982</td>
</tr>
<tr>
<td>Burushaski (20)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Maasai (21)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hadza (31)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hausa (32)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Masri (33)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

A first observation in table 7.28 is that those systems described as concordial class systems form a coherent group; they all have a formalisation value above 0.8. The other systems display a broad range of formalisation, spanning from 0.208 (Dongo-Ko (13)) to 0.732 (Nasioi (03)). As the systems group themselves with respect to the systemic labels used in the literature, there is one property that is divided between concordial class and classifier systems, namely the presence or absence of multiple marking respectively. There is no salient pattern in the individual systems’ setup of grammaticalised properties; on the contrary, they display fairly mixed constellations. The classificatory flexibility seems to be able to be maintained up to a very high degree of classification, whereas a system’s number of hosts and its locus expand from an early stage on. In addition, bound markers are common from an early developmental stage on. Finally, the contexts in which a system’s classification is obligatory displays a fairly steady rise in grammaticalisation. In subsection 7.3.2 the transparency
value of the representative and sample systems is accounted for.

### 7.3.2 The Transparency Value

The transparency value is determined on basis of seven criteria, which are presented in tabular form and then discussed for each representative language individually. Table 7.29 gives an overview over the properties of Hungarian.

Table 7.29: The transparency value of Hungarian.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantic structure</td>
<td>bijective classes</td>
</tr>
<tr>
<td>rule types</td>
<td>semantic</td>
</tr>
<tr>
<td>rule number</td>
<td>one per class</td>
</tr>
<tr>
<td>categorial independence</td>
<td>categorially independent</td>
</tr>
<tr>
<td>discreteness of markers</td>
<td>distinct markers per class</td>
</tr>
<tr>
<td>redundancy</td>
<td>mostly redundant</td>
</tr>
<tr>
<td>flexibility</td>
<td>largely flexible</td>
</tr>
<tr>
<td><strong>mean</strong></td>
<td><strong>0.917</strong></td>
</tr>
</tbody>
</table>

Hungarian possesses a high transparency value: Its medium inventory consists of 18 classes, which are semantically straightforward and assigned by semantic principles only. The categories and their members reveal a simplistic assignment system with one assignment rule per class, which are organised by an entity’s shape or configuration. The markers are not fused with other categories and distinct per class. The semantic information conveyed by the markers is generally redundant, as the Hungarian classifiers are generally optional, which means that a given construction has the same semantics with and without a classifier. However, certain categories serve the specification of the nominal referent (cf. again 7.9 D and E on page 283), which raises the semantic redundancy of Hungarian. This is achieved by generating the mean value of both the redundant measure (0.5) and the informative measure (1). Hungarian’s transparency value comes down to 0.917.

The properties determining the transparency value of the Mandarin system’s higher registers is summarised in table 7.30.
Table 7.30: The transparency value of Mandarin.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantic structure</td>
<td>opaque/mixed classes 0.25</td>
</tr>
<tr>
<td>rule types</td>
<td>semantic, redundant 0.667</td>
</tr>
<tr>
<td>rule number</td>
<td>high (one per conventionalised pair) 0.667</td>
</tr>
<tr>
<td>categorial independence</td>
<td>categorially independent 1.0</td>
</tr>
<tr>
<td>discreteness of markers</td>
<td>mixed set, predominantly distinct markers 0.75</td>
</tr>
<tr>
<td>redundancy</td>
<td>redundant 0.5</td>
</tr>
<tr>
<td>flexibility</td>
<td>largely rigid 0.333</td>
</tr>
<tr>
<td>mean</td>
<td>0.536</td>
</tr>
</tbody>
</table>

As expected, Mandarin’s higher registers possess a relatively low transparency value. The categorial independence of the Mandarin markers is the only fully transparent criterion, whereas the number of classes and rules both are high and thus not transparent. The system is characterised by a largely conventionalised and rigid classifier-noun pairing (Erbaugh 1986: 404ff. and Zhang 2007: 54), which replaces the original semantic assignment for large parts of the nominal lexicon. Most of the Mandarin classes are semantically specific, yet there are mixed classes such as the one marked by the general classifier *ge*. Due to the weight of the specific classes over the mixed ones, their mean value applies to the Mandarin higher registers. A similar situation is present for the distinctness of the markers per class: While the majority of markers indeed is distinct, there are classifiers referring to more than one class (e.g. *zhi* “animal (mammals, birds, insects, frogs); one of a pair; three-dimensional, container-/ball-like object”; cf. 156; 2.31 A = 156 A).

(7.18) Mandarin (cf. Zhang 2007: 56)

A yi zhi laohu, yi zhi niao, yi zhi *pangxie* crab
“one tiger, one bird, one crab” (Mandarin; cf. Zhang 2007: 55)

B yi zhi yanjing, yi zhi xie
“one eye, one shoe”

C yi zhi lanzi, yi zhi qiqiu
“one basket, one balloon”

Therefore, the mean value of distinct and mixed markers applies to the system. Finally, the fixed class membership of the majority of nouns reveals that there are no productive semantics involved in the assignment of those cases. Therefore, potential semantic information carried by the markers is redundant...
in the conventionalised cases. For this reason, the higher registers of the Mandarin system are at an intermediate position between being semantically redundant and informative. The combination of the individual criteria results in a transparency value of 0.536 for the higher registers of the Mandarin system.

The transparency value of the Burushaski system and its individual criteria are given in table 7.30.

Table 7.31: The transparency value of Burushaski.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantic structure</td>
<td>bijective classes</td>
</tr>
<tr>
<td>rule types</td>
<td>semantic and redundant</td>
</tr>
<tr>
<td>rule number</td>
<td>low</td>
</tr>
<tr>
<td>categorial independence</td>
<td>fused (number)</td>
</tr>
<tr>
<td>discreteness of markers</td>
<td>mixed markers</td>
</tr>
<tr>
<td>redundancy</td>
<td>informative</td>
</tr>
<tr>
<td>flexibility</td>
<td>largely flexible</td>
</tr>
<tr>
<td>mean</td>
<td></td>
</tr>
</tbody>
</table>

The Burushaski system is a highly transparent one, as it is equipped with a small number of semantically straightforward classes ([+human male], [+human female], [+concrete], [residue]; Berger 1998: 144). Except for some redundantly assigned “exceptions which either fall under the class of [+concrete] or [-concrete] [i.e. the residue class; MP]” (Munshi 2006: 166), these classes allow for a productive semantic reassignment of large parts of the nominal lexicon based on four rules, which require only two steps in a decision tree (cf. figure 7.6).

The class markers occur only in the third person singular and are fused with the category of number. Due to a series of formal syncretisms, a number of targets for class markers fuse either the masculine and feminine class to a superclass [+human], the concrete and residue class to a superclass [-human], or both (cf. Berger 1998: 34 for a more detailed description and table 13 for the third person forms of the auxiliary b(ā)-, where the plural displays the fusion
of the masculine and feminine class).

Table 7.32: The third person forms of the Burushaski auxiliary $b(á)$- (Berger 1998: 144).

<table>
<thead>
<tr>
<th>class</th>
<th>singular form</th>
<th>plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>masculine</td>
<td>bái</td>
<td>báan</td>
</tr>
<tr>
<td>feminine</td>
<td>bo</td>
<td>báan</td>
</tr>
<tr>
<td>concrete</td>
<td>bi</td>
<td>běi</td>
</tr>
<tr>
<td>residue</td>
<td>běá</td>
<td>běčá</td>
</tr>
</tbody>
</table>

These syncretisms, which apply differently to different parts of speech, obscure the otherwise very transparent structure of the Burushaski system, as the same marker may refer to different underlying classes. However, the Burushaski markers are informative in that they semantically contribute to the noun phrase, where they specify the referent’s animacy, sex, degree of individuation, concreteness, or countability. The combination of these criteria gives Burushaski a transparency value of 0.786.

Next, the German system is investigated; its properties are listed in table 7.33.

Table 7.33: The transparency value of German.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantic structure</td>
<td>opaque/mixed classes</td>
</tr>
<tr>
<td>rule types</td>
<td>semantic, morphological, phonological, redundant</td>
</tr>
<tr>
<td>rule number</td>
<td>high</td>
</tr>
<tr>
<td>categorial independence</td>
<td>fused (number, case)</td>
</tr>
<tr>
<td>discreteness of markers</td>
<td>markers encode more than one class</td>
</tr>
<tr>
<td>redundancy</td>
<td>opaque/redundant/informative</td>
</tr>
<tr>
<td>flexibility</td>
<td>largely rigid</td>
</tr>
<tr>
<td>mean</td>
<td>0.202</td>
</tr>
</tbody>
</table>

As expected, German has a very low transparency value of 0.202. Despite its small inventory of only three classes, there is a high number of assignment rules, which are based on semantics, morphology, and phonology. Beyond this, there are many redundant assignment rules. This results in classes that are inbetween opaque and mixed. The classification applies to the third person, and the markers are fused with the category of case and number. The classes feature syncretic markers, which occur in different paradigmatic slots across the individual classes (cf. e.g. table 38 on page 422 for an overview over the German articles). Finally, the German class marking can be opaque (cf. 7.19 A, where no information on the referent is derivable from the class marking), redundant (cf. 7.19 B, where information on the referent is encoded in both
class marking and the noun itself), or informative (cf. 7.19 C and D, where the class marking semantically contributes to the noun phrase).

(7.19) German

A. der Löffel, Messer
   die Gabel, Messer
   "the spoon, the fork, the knife"

B. der Müllmann, Krankenschwester
   die Müllmann, Krankenschwester
   "the dustman, the nurse"

C. der Gläubige, Gläubige
   die Gläubige, Gläubige
   "the (male or unspecified) believer, the (female) believer"

D. der BMW, BMW
   die BMW, BMW
   "the BMW (car), the BMW (motorcycle)"

Finally, the transparency of the system of Bora is accounted for. Table 7.34 lists the individual properties contributing to its calculation.

Table 7.34: The transparency value of Bora.

<table>
<thead>
<tr>
<th>criterion</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantic structure</td>
<td>bijective</td>
</tr>
<tr>
<td>rule types</td>
<td>semantic, redundant</td>
</tr>
<tr>
<td>rule number</td>
<td>high (one per class)</td>
</tr>
<tr>
<td>categorial independence</td>
<td>independent</td>
</tr>
<tr>
<td>discreteness of markers</td>
<td>distinct markers per class</td>
</tr>
<tr>
<td>redundancy</td>
<td>redundant/informative</td>
</tr>
<tr>
<td>flexibility</td>
<td>largely flexible</td>
</tr>
<tr>
<td><strong>mean</strong></td>
<td><strong>0.726</strong></td>
</tr>
</tbody>
</table>

Despite its productive semantics, the Bora system is fairly balanced with respect to its transparency: Its large inventory and the related high number of assignment rules lower the overall transparency of the system. In contrast, the bijective semantic structure of the individual classes and the categorial independence and classificatory distinctness of the markers raise it. Next to a number of nouns the assignment of which is opaque, the assignment rules are semantic. In many cases, the markers serve as compositional building blocks of the noun phrase and thus are informative, but non-compositional instances of classification can be found, too (cf. 7.20, where A–D exemplifies compositional classification and E–F a non-compositional one).

Note that informative class marking, where crucial semantic information is conveyed by the class markers only, is fairly rare in German.
(7.20) Bora (cf. Seifart 2005: 195, 203 and 214)

A  `a`hi  -ʔa
    `banana`  CL:3D.LONG.OBJ
    "banana fruit"

B  mɛ:me  -ʔa
    `palm`  CL:3D.LONG.OBJ
    "palm fruit"

C  t`ub`ó
    shoot.NOMINALISED -ɛʔo
    CL:LITTLE.STICK
    "arrow"

D  adó
    drink.NOMINALISED -ra
    CL:BOTTLE
    "drinking bottle"

E  e:`he` -ko
    `flue`  CL:1D.POINTED
    "flue"

F  p`ó`ka
    `year`  CL:3D
    "year"

After having accounted for the representative systems’ transparency value individually, table 7.35 gives an overview over their values.

Table 7.35: Overview over the representative systems’ transparency criteria and value.

<table>
<thead>
<tr>
<th></th>
<th>category Independence</th>
<th>reversibility</th>
<th>flexibility</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungarian</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>Burushaski</td>
<td>1</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Bora</td>
<td>1</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Mandarin</td>
<td>0.25</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>German</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0.333</td>
</tr>
</tbody>
</table>

Table 7.35 shows that the transparency of the individual system widely varies, and that the reasons for a system to be transparent are varied, too: Those systems that are characterised by a larger inventory tend to balance this out by a more productive and straightforward semantic substructure of their class system. Systems that possess a smaller inventory tend to be semantically less straightforward and thus less transparent in this respect. Nevertheless, a system’s inventory size in and by itself is a formalisation property. It only tends towards a counterbalance by its semanticity and is thus not a strong indicator
of a system's overall transparency, as both large and small inventories display large differences in their transparency value (cf. e.g., the smaller systems of Burushaski and German). The representative systems display a transparency value that was predicted on basis of their description and earlier studies, German and the higher levels of Mandarin scoring relatively low, Bora displaying a medium value, and Hungarian and Burushaski scoring relatively high.

After having elaborated the calculation of the formalisation transparency value, the sample's values are presented. Table 7.36 provides an overview over the transparency measures and values of the individual systems. Note that the values of the Mandarin system have been reevaluated, as the full Mandarin system is subject to this study, and not only its higher registers.

Table 7.36: Overview over the sample systems’ transparency criteria and value.

<table>
<thead>
<tr>
<th>Language</th>
<th>Rule Type</th>
<th>Rule Number</th>
<th>Categorial Independence</th>
<th>Discreteness of Markers</th>
<th>Redundancy</th>
<th>Flexibility</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>German (30)</td>
<td>0.25</td>
<td>0</td>
<td>0.333</td>
<td>0.5</td>
<td>0.333</td>
<td>0</td>
<td>0.202</td>
</tr>
<tr>
<td>Ket (27)</td>
<td>0.5</td>
<td>0.667</td>
<td>0</td>
<td>0.667</td>
<td>0.5</td>
<td>0</td>
<td>0.298</td>
</tr>
<tr>
<td>Dongo-Ko (25)</td>
<td>0.25</td>
<td>1</td>
<td>0.667</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0.345</td>
</tr>
<tr>
<td>Ingush (28)</td>
<td>0.25</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>0.333</td>
<td>0.357</td>
</tr>
<tr>
<td>Hausa (32)</td>
<td>0.25</td>
<td>0.333</td>
<td>0</td>
<td>0.333</td>
<td>1</td>
<td>0.25</td>
<td>0.357</td>
</tr>
<tr>
<td>Moro (26)</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>0.667</td>
<td>0</td>
<td>0</td>
<td>0.404</td>
</tr>
<tr>
<td>Chichewa (29)</td>
<td>0.5</td>
<td>0.667</td>
<td>0</td>
<td>0.667</td>
<td>0</td>
<td>0.75</td>
<td>0.464</td>
</tr>
<tr>
<td>Mandarin (formal)</td>
<td>0.25</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
<td>0.75</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Masri (33)</td>
<td>0.5</td>
<td>0.333</td>
<td>0</td>
<td>0.667</td>
<td>1</td>
<td>0.75</td>
<td>0.512</td>
</tr>
<tr>
<td>Hadza (31)</td>
<td>0.5</td>
<td>0.667</td>
<td>0.5</td>
<td>0.333</td>
<td>1</td>
<td>0.75</td>
<td>0.583</td>
</tr>
<tr>
<td>Massai (21)</td>
<td>0.5</td>
<td>0.667</td>
<td>0</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
<td>0.690</td>
</tr>
<tr>
<td>Korean (01)</td>
<td>0.75</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.75</td>
<td>0.667</td>
<td>0.702</td>
</tr>
<tr>
<td>Mandarin (02)</td>
<td>0.75</td>
<td>0.667</td>
<td>1</td>
<td>0.75</td>
<td>0.75</td>
<td>1</td>
<td>0.702</td>
</tr>
<tr>
<td>Nasi (03)</td>
<td>0.75</td>
<td>1</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
<td>0.667</td>
<td>0.702</td>
</tr>
<tr>
<td>Great Andamanese (09)</td>
<td>0.5</td>
<td>0.667</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.714</td>
</tr>
<tr>
<td>North Ambrym (11)</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.714</td>
</tr>
<tr>
<td>Bora (04)</td>
<td>1</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>0.726</td>
</tr>
<tr>
<td>Mundurukú (05)</td>
<td>0.75</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.75</td>
<td>0.738</td>
</tr>
<tr>
<td>Sumerian (22)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>0.5</td>
<td>0.738</td>
</tr>
<tr>
<td>Dyirbal (18)</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>Tohono O’odham (14)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.786</td>
</tr>
<tr>
<td>Burushaski (20)</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>0.667</td>
<td>0.5</td>
<td>1</td>
<td>0.786</td>
</tr>
<tr>
<td>Hmong Daw (15)</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.798</td>
</tr>
<tr>
<td>Jakaltek (10)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>0.5</td>
<td>0.893</td>
</tr>
<tr>
<td>Nùng (16)</td>
<td>0.75</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.893</td>
</tr>
<tr>
<td>Ngan’gityemerrri (06)</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.667</td>
<td>0.905</td>
</tr>
<tr>
<td>Tamil (24)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>0.75</td>
<td>0.917</td>
</tr>
<tr>
<td>Hungarian (08)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>0.917</td>
</tr>
<tr>
<td>Dongo-Ko (13)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.929</td>
</tr>
<tr>
<td>Tohono O’odham (14)</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.929</td>
</tr>
<tr>
<td>Boumaa Fijian (07)</td>
<td>1</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.952</td>
</tr>
</tbody>
</table>
As predicted, the transparency value is much more varied in the sample than the formalisation value. The constellation of the individual criteria per system varies greatly, which means that a system may be intransparent due to varied reasons (which are represented by the criteria themselves). The lowest transparency value in the sample is the one of German (30) with 0.202. This is in accordance with the earlier made conjecture that a fully intransparent system is highly unlikely, as it would experience either a semantic remotivation process in order to lift its transparency value, or to drop out of the grammar. After having calculated the transparency value of the sample, the next section combines both the transparency and formalisation value.

7.3.3 Combining the Values

This section presents the graphic representation of the two-dimensional account for nominal classification devices. After having calculated the values for both formalisation and transparency for the sample, they can be combined and represented in a two-dimensional graph. The formalisation value is represented on the $x$-axis in figure 7.7, the transparency value on the $y$-axis. Figure 7.7 displays the distribution of all sample systems in the model.

![Figure 7.7: Comparative mapping of the full language sample.](image_url)

Without imposing a systemic typology on the mapped sample ex ante, figure 7.7 shows a clear horizontal distribution along the $x$-axis in the higher transparent layers, and a clear vertical distribution along the $y$-axis in the higher formalised layers. Figure 7.8 shows the distribution of the sample and the location of the individual sample systems in the graph. Note again that the full Mandarin system is accounted for here, not only its higher registers.
7.3. Application of the Mapping Instrument

Figure 7.8: Comparative mapping of the individual sample systems (index cf. below).

<table>
<thead>
<tr>
<th>language</th>
<th>№</th>
<th>x-value</th>
<th>y-value</th>
<th>type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>1</td>
<td>0.226</td>
<td>0.702</td>
<td>CLS</td>
</tr>
<tr>
<td>Mandarin</td>
<td>2</td>
<td>0.417</td>
<td>0.702</td>
<td>CLS</td>
</tr>
<tr>
<td>Nasiol</td>
<td>3</td>
<td>0.732</td>
<td>0.702</td>
<td>CLS</td>
</tr>
<tr>
<td>Bora</td>
<td>4</td>
<td>0.542</td>
<td>0.726</td>
<td>CLS</td>
</tr>
<tr>
<td>Mundurukù</td>
<td>5</td>
<td>0.542</td>
<td>0.738</td>
<td>CLS</td>
</tr>
<tr>
<td>Ngan’gityemerri</td>
<td>6</td>
<td>0.512</td>
<td>0.905</td>
<td>CLS</td>
</tr>
<tr>
<td>Boumaa Fijian</td>
<td>7</td>
<td>0.476</td>
<td>0.952</td>
<td>CLS</td>
</tr>
<tr>
<td>Hungarian</td>
<td>8</td>
<td>0.214</td>
<td>0.917</td>
<td>CLS</td>
</tr>
<tr>
<td>Great Andamanese</td>
<td>9</td>
<td>0.649</td>
<td>0.714</td>
<td>CLS</td>
</tr>
<tr>
<td>Jakaltek</td>
<td>10</td>
<td>0.304</td>
<td>0.893</td>
<td>CLS</td>
</tr>
<tr>
<td>North Ambyr</td>
<td>11</td>
<td>0.351</td>
<td>0.714</td>
<td>AMI</td>
</tr>
<tr>
<td>Dongo-Ko</td>
<td>13</td>
<td>0.208</td>
<td>0.929</td>
<td>CLS</td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td>14</td>
<td>0.429</td>
<td>0.786</td>
<td>CLS</td>
</tr>
<tr>
<td>Hmong Daw</td>
<td>15</td>
<td>0.559</td>
<td>0.798</td>
<td>CLS</td>
</tr>
<tr>
<td>Nùng</td>
<td>16</td>
<td>0.406</td>
<td>0.893</td>
<td>CLS</td>
</tr>
<tr>
<td>Dyirbal</td>
<td>18</td>
<td>0.667</td>
<td>0.75</td>
<td>CLS</td>
</tr>
<tr>
<td>Tainae</td>
<td>19</td>
<td>0.893</td>
<td>0.929</td>
<td>CLS</td>
</tr>
<tr>
<td>Burushaski</td>
<td>20</td>
<td>1</td>
<td>0.786</td>
<td>CCS</td>
</tr>
<tr>
<td>Maasai</td>
<td>21</td>
<td>1</td>
<td>0.690</td>
<td>CCS</td>
</tr>
<tr>
<td>Sumerian</td>
<td>22</td>
<td>0.929</td>
<td>0.738</td>
<td>CCS</td>
</tr>
<tr>
<td>Tamil</td>
<td>24</td>
<td>0.964</td>
<td>0.917</td>
<td>CCS</td>
</tr>
<tr>
<td>Dongo-Ko</td>
<td>25</td>
<td>0.982</td>
<td>0.345</td>
<td>CCS</td>
</tr>
<tr>
<td>Moro</td>
<td>26</td>
<td>0.964</td>
<td>0.404</td>
<td>CCS</td>
</tr>
<tr>
<td>Ket</td>
<td>27</td>
<td>0.964</td>
<td>0.298</td>
<td>CCS</td>
</tr>
<tr>
<td>Ingush</td>
<td>28</td>
<td>0.899</td>
<td>0.357</td>
<td>CCS</td>
</tr>
<tr>
<td>Chichewa</td>
<td>29</td>
<td>0.982</td>
<td>0.464</td>
<td>CCS</td>
</tr>
<tr>
<td>German</td>
<td>30</td>
<td>0.929</td>
<td>0.202</td>
<td>CCS</td>
</tr>
<tr>
<td>Hadza</td>
<td>31</td>
<td>1</td>
<td>0.583</td>
<td>CCS</td>
</tr>
<tr>
<td>Hausa</td>
<td>32</td>
<td>1</td>
<td>0.357</td>
<td>CCS</td>
</tr>
<tr>
<td>Masri</td>
<td>33</td>
<td>1</td>
<td>0.512</td>
<td>CCS</td>
</tr>
</tbody>
</table>

formalisation value

transparency value
The systems in figure 7.8 show a clear organisation in groups that correspond to the areas sketched out in figure 7.5 – 7.9, which displays the possibly populated areas in the mapping graph.

![Diagram showing the mapping of nominal classification](image)

**Figure 7.9:** Corresponding areas and systems in the model.

However, there is no full correspondence to the possible areas in figure 7.9; area B is not met in the sample, and area E needs to be readjusted in size, as the following paragraphs demonstrate. There are no lowly formalised or intermediate systems with a transparency value below 0.7. All other systems with a transparency value below 0.6 are systems that are typically referred to as concordial class systems. This means that a system’s ability to transcend a certain transparency threshold correlates with its degree of formal grammaticalisation, while a higher degree of formalisation does not require a loss in transparency. In other words: A system first has to formalise to a greater degree, before it gains the ability to grow fairly intransparent. A possible explanation for this is the visibility of the classificatory marking: In contrast to concordial class systems, which attach class markers to a typically wider range of morphosyntactic hosts and thus are visible ‘all over the place’, classifiers occur once and overtly in discourse. If such a comparatively rare element is productive part of a grammar, it is likely that the speaker community prefers this productive element to be transparent. The data from figure 7.8 suggests the existence of a transparency threshold of 0.7, as there are no systems with a lower transparency value. What happens if a system falls below that threshold? The most likely two options are that these comparatively rarely visible elements, which are idiosyncratic to a certain degree, either are omitted altogether and the classification system is dropped. Alternatively, the rarely occurring element can be reanalysed and given a new and transparent semantic underpinning. This was
the case in Dyirbal, as the next chapter shows. The omnipresent formal marking of concordial class systems on the other hand establishes a strong visible presence of classification, be it productive and flexible or conventionalised and opaque. This greater degree of visibility strengthens the maintenance of an existing classification and reduces the acceptability of a reanalysed classificatory basis (cf. also Audring 2014: 13–16). Take the concordial class-like systems of Dongo-Ko, Ket, and German for instance, which have a very low transparency value, and the class membership of nouns has to be learned (cf. Pasch 1986: 194 for Dongo-Ko and Georg 2007: 89). This state is not a new one, as section 8.3 shows for the German system. Why are these classification systems maintained in this state generation after generation? First, they mark a broad range of nominal modifiers for a noun’s class membership, which guarantees a high visibility in discourse. If a noun is not correctly marked on the agreement targets, this class disagreement is easily recognisable for the recipient. At the same time, a first language acquirer typically masters a language’s grammatical phenomena regardless of their complexity without troubles. Both these factors may contribute to the maintenance of highly visible yet opaque classification systems.

As there are no sample systems corresponding to area b for lowly formalised and rather intransparent systems, it is ignored for the rest of this analysis. Area e is reduced to a transparency value of 0.7 or more. The areas a and e further suggest a common coherent grouping of systems rather than a categorial distinction. Note though that area a is by far less densely populated than area e. Table 7.38 summarises the areas and their corresponding systems.

Table 7.38: The mapped sample systems and their correspondences to systemic areas.

<table>
<thead>
<tr>
<th>area</th>
<th>language</th>
<th>type</th>
<th>area</th>
<th>language</th>
<th>type</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Korean (01)</td>
<td>CLS</td>
<td>c</td>
<td>Tainae (19)</td>
<td>CLS</td>
</tr>
<tr>
<td></td>
<td>Hungarian (08)</td>
<td>CLS</td>
<td></td>
<td>Burushaski (20)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Jakaltek (10)</td>
<td>CLS</td>
<td></td>
<td>Maasai (21)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>North Ambrym (11)</td>
<td>CLS</td>
<td></td>
<td>Sumerian (22)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Dongo-Ko (13)</td>
<td>CLS</td>
<td></td>
<td>Tamil (24)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Mandarin (02)</td>
<td>CLS</td>
<td>d</td>
<td>Dongo-Ko (25)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Bora (4)</td>
<td>CLS</td>
<td></td>
<td>Moro (26)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Mundurukú (5)</td>
<td>CLS</td>
<td></td>
<td>Ket (27)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Ngan’gityemerri (06)</td>
<td>CLS</td>
<td></td>
<td>Ingush (28)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Boumaa Fijian (07)</td>
<td>CLS</td>
<td></td>
<td>Chichewa (29)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Great Andamanese (09)</td>
<td>CLS</td>
<td></td>
<td>German (30)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Tohono O’odham (14)</td>
<td>CLS</td>
<td></td>
<td>Hausa (32)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Hmong Daw (15)</td>
<td>CLS</td>
<td></td>
<td>Masri (33)</td>
<td>CCS</td>
</tr>
<tr>
<td></td>
<td>Nùng (16)</td>
<td>CLS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dyirbal (18)</td>
<td>CLS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7.10 visualises the distribution of systemic types in the graph. Note
that there is a noticeable population gap between the areas E and C.

Figure 7.10: Comparative mapping of the sample systems: Distribution of systemic types (+: CLS; •: CCS).

The mapping of the sample systems immediately confirms a conjecture: If one accepts the reanalysis of the Dyirbal (18) system as a non-concordial class system, the concept of concordial class systems corresponds to the systems that are generally referred to as concordial class systems in that only concordial class systems group in the areas C and D. The Tainae (19) system, which was also reanalysed as a classifier-like system, groups together with the concordial classifier systems.

The notion of classifier systems in contrast is highly varied and comprises systems of a wide range of formalisation and transparency. All systems with a formalisation value below 0.4 (area A) are more or less corresponding well to the established notion of classifiers, while those systems in area E comprise among others all the sample’s ‘misfits’, namely Nasioi (03), Bora (04), Mundurukú (05), Ngan’gityemerri (06), Great Andamanese (09), and Dyirbal (18). In addition, the classifier systems of Mandarin (2), Boumaa

---

8The Tainae system poses a few general problems: It is described in only one grammar (Carlson 1991) and interacts with the discoursal domain. It employs its specific class markers only if “a specific person or thing is in mind, and not when just talking about something generally” (Carlson 1991: 27); general and unspecified discourse referents are marked by the ‘plural’ marker -i/-e. This ‘plural’ marker can denote singular entities and the specific markers can refer to plural entities, which is why -i/-e is reanalysed as a generic class markers (cf. again the discussion on page 577). By the availability of a generic marker, the Tainae system applies to the full nominal lexicon, and the classification becomes an exhaustive and obligator one. The exhaustiveness and obligatoriness value are those measures that attribute a higher formalisation value to the Tainae system, which otherwise would group together with the other classifier-like systems. The case of Tainae suggests a further review and refinement of the mapping instrument on the basis of more ‘misfits’ such as Tainae.
302 7.3. Application of the Mapping Instrument

Fijian (07), Tohono O’odham (14), and Hmong Daw (15) fall into the scope of ε due to their advanced degree of formalisation, which allows its markers to attach to a large range of nominal modifiers, or due to its small class inventory’s lack of classificatory flexibility. This suggests to drop the distinction of different types of nominal classification devices altogether. Those four systems with a formalisation value between 0.4 and 0.5 (Mandarin (2), Boumaa Fijian (07), Tohono O’odham (14), and Nùng (16)) seem to form a ‘hinge area’ between fairly prototypical classifier-like systems (x-value <0.4) and more intermediate classifier-like systems (x-value >0.5).

Grammaticalisation does not apply in its full range of possibly involved processes to the developmental states that are found in the sample, as a system’s semanticity and transparency typically is not affected by drastic bleaching (i.e. by a bleaching that pushes the system’s transparency value below 0.7) until it displays concordial class-like properties. Therefore, those properties that contribute to a system’s transparency and those properties that shape a system’s formal appearance have been separated in the mapping; both together constitute grammaticalisation. The mapped sample confirms that it is indeed grammaticalisation that is the main vehicle for developmental shifts in nominal classification systems: Only sample systems with a formalisation value of 0.8 or more show the potential to distinctly lack an accessible and transparent underlying structure, i.e. to have a transparency value of considerably less than 0.6. In contrast, the variation of transparency in less grammaticalised systems spans from just above 0.7 to 0.929, which suggests that conventionalisation and bleaching processes of these systems are less pervasive in their effect than formalisation processes, and supposedly reversible by the speaker community. Formalisation processes in contrast are known for their lacking potential for reversibility. In the light of these findings, the number of developmental pathways decreases: The only developments that can be traced in the sample are III, IV, and VI (cf. figure 7.11).

This observation from the sample does not mean that other pathways do not occur at all; it however suggests that the circumstances necessary for other shifts are specific to such a degree that these other developments are infrequent exceptions. The pathways III, IV, and VI are studied in the documented diachrony of individual systems in chapter 8.

The distribution of the prototypical systems and the first, second, and third order deviations from the sample shows a distribution similar to those of the systemic types. First order deviations are those systems that show a mixed classifier-like and concordial class-like pattern within the ‘core criteria’ (obligatoriness, marking, and locus operandi). In the sample, these are the systems of Nasioi, Bora, Mundurukú, and Ngan’gityemerri, which display classifier-like single marking and an optional classification together with a concordial class-like NP-external occurrence of classificatory markers. Second order deviations are those systems that show homogeneous classifier-like or concordial class-like ‘core properties’, but an overall mixed composition of properties. The system of North Ambrym is a second order deviation in that it displays classifier-like
core properties, but at the same time a concordial class-like rigid classification and class inventory. Third order deviations are those systems that display an overall homogeneous set of classifier-like or concordial class-like properties, but employ individual properties that are both classifier-like or concordial class like. The Mandarin system for instance employs a fully classifier-like setup of properties, with the assignment, which is partly classifier-like in that it is free and speaker-sided and partly concordial class-like in that classifier-noun pairs have conventionalised and have to be learned, as the only exception. Together, these deviations constitute the grey area (cf. figure 5.23 = 7.39).

Table 7.39: The grey area.

<table>
<thead>
<tr>
<th>Language</th>
<th>Core Properties</th>
<th>Third Order Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mandarin</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Nasiq</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Bora</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Mundurukú</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Ngan'gityemerri</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Burea Fijian</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Hungarian</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Jakaltek</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>North Ambrym</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Dongo-Ko (CLS)</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Tohono O’odham</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Hmong Daw</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>Nùng</td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>

Figure 7.11: Systemic areas and developmental pathways in the sample.
7.3. Application of the Mapping Instrument

The location of these deviations and the prototypes from the sample is provided in figure 7.12.

![Figure 7.12: The grey area: Deviations and prototypes (●: prototypical system; 1: first order deviation; 2: second order deviation; 3: third order deviation).](image)

A first observation from figure 7.12 is the concentration of deviations in the top right corner of the graph: Thirteen out of fifteen systems with a transparency value and a formalisation value of more than 0.5 respectively are deviations. There are only two exceptions, the system of Hmong Daw, which has a formalisation value of just above 0.5 (0.559), and the one of Masri, which has a transparency value of just above 0.5 (0.512). The order of deviation shows a correlation with the systems’ variation in transparency value and formalisation value, but does not strictly depend on a system’s type: The bulk of the third order deviations is located among the highly formalised systems. Only Mandarin represents a case of a lowly grammaticalised third order deviation. Like the mapped concordial class systems, the third order deviations display a variation along the y-axis, reaching from fairly intransparent systems (Ingush; y-value: 0.357) to highly transparent ones (Tamil; y-value: 0.917).
Second order deviations show the opposite distribution: Like the sample’s classifier systems, their variation occurs along the $x$-axis, while their transparency value is not varied much and not lower than 0.69 (Maasai). They include both classifier and concordial class systems. The lowest formalisation value of a second order deviation is 0.304 (Jakaltek), and the highest one 1 (Burushaski, Maasai). The first order deviations are centered around the gap that separates the highly formalised systems from the lowly formalised ones. They comprise the highly intermediate and typologically problematic systems of Nasioi, Bora, Mundurukú, Nganjmirra, Dyirbal, and Tainae. The synchronic study has shown that especially these systems demand an intensive and careful analysis in order enable a typologically unproblematic account.

Summing up the observations from the mapping of the sample, the most eyecatching finding is that classifier-like systems allow a great variation in their degree of formalisation. At the same time, they display a delimited variation in their transparency value, which is always higher than 0.7. Second order deviations show a similar distribution, but are not tied to a systemic type. Concordial class-like systems show only little variation in their formalisation value, which is always higher than 0.8. They display however a great variation in their transparency value, which lies between 0.2 and 1. Third order deviations correspond to this distribution in a looser way. Chapter 8 investigates why this converse distribution occurs, and further exemplifies the diachronic applicability of the mapping by carrying out studies of the history of individual systems that have undergone the observed developmental pathways.
This chapter studies the diachrony of nominal classification devices from the perspective of the insights from the synchronic sample study. I have argued for the segregation of two distinct dimensions of grammaticalisation in the previous chapter, as the sample systems have shown to independently develop along the lines of formal expression and semantic transparency. The mapping presented in chapter 7 accounts for these dimensions separately and thus is used as a comparative instrument for the diachrony of selected sample systems in this chapter.

The synchronic study isolated one main developmental pathway for lowly formalised systems: A system may formalise into the direction of a concordial class system without a greater gain or loss in its transparency (cf. arrow VI in figure 7.1). Beyond this possible development for lowly formalised systems, the sample’s highly formalised systems show a large variation of their transparency, which suggests that a highly formalised system may either lose (arrow III) or gain (arrow IV) in its underlying semanticity (cf. figure 8.1).
These three pathways are studied in this chapter on basis of diachronic data from the sample languages: Mandarin and Dyirbal both provide detailed documentations of their respective systems at different developmental states. Mandarin offers an example for a documented formalisation of its system from its beginning on; spoken Mandarin represents a case of ‘mere’ formalisation without a greater change in its transparency value. Its higher registers with their conventionalised classifier-noun pairings represent an example for systems that both conventionalised and formalised notably. Both cases are studied in section 8.1 Section 8.2 accounts for the development of the Dyirbal system. Furthermore, the West Germanic languages, which include German and English, are well-documented and allow for a presentation of their respective systems’ diachrony. Section 8.3 presents the case studies of the systems of Dyirbal and the West Germanic languages, which display an observable loss or gain in their transparency without an accompanying grammaticalisation process.

Cf. again section 7.2 for the detailed description of the calculation of the grammaticalisation value of the two domains of formalisation and transparency of nominal classification systems. Note that the values are represented by means of descriptive labels (low, medium, and high degree of grammaticalisation) in the running text. Section 8.4 concludes the diachronic part of this study. The case studies are ordered with respect to their degree of formalisation, beginning with the lowly formalised system of Mandarin, and ending with the highly formalised West Germanic systems.
8.1 Grammaticalisation in Mandarin

This section studies the observable formalisation of the Mandarin classifier system from its beginnings in the first centuries BC up to its contemporary status. The system in general first formalised without a larger shift in transparency. Only the higher linguistic registers display a stronger loss in their transparency value; these are presented after the discussion of the Mandarin system in general.


Prior to the classifier system, Mandarin possessed a productive system of measure words, which were inserted in quantificational expressions. Its original structure displays the order [noun + numeral + measure word] (cf. 8.1). According to Peyraube (1991: 106 and 120), the classifiers entered this construction with a parallel change in word order: [numeral + classifier/measure word + noun] (cf. 8.2).

(8.1) Pre-Han Chinese (cf. Wang 1994: 102)

Qihuangong qi zhi, you [ma er shi cheng].
duke.Huan wife him have horse twenty MEASURE.WORD:4.HORSES
“Duke Huan gave him a lady to wife, and he had 80 horses (in 20 teams of four).”


Zheng pan si zhao hua, jing xi [san tiao shu].
struggle break four radiant flower compete play three CL:LONG road
“Contend in breaking, compete in playing on the three roads.”

There are about fifty classifiers that can be identified for the Han period, which represents a medium inventory size. This number grew to about 110 in Early Medieval Chinese and further to its current state (cf. Peyraube 1991: 118 and Wang 1994: 107), which indicates a low degree of formalisation of the class inventory. These classifiers combined with numerals as the only classificatory host (cf. 8.3). Later, the classifiers started to expand their range of morphosyntactic hosts and combine with demonstratives in Late Medieval Chinese (cf. 8.4) at the same time, they additionally can occur together

---

1Peyraube (1991: 119f.) motivates the development of classifiers by the semantic bleaching of the postnominal measure words. Once having lost the categorial characteristics of a lexical noun, a change in word order reorganises the system and allows classifiers to emerge.
with the bare noun (cf. Peyraube 1991: 120 and 8.5). This raises the value for the classifiers’ number of hosts to a medium one.

(8.3) Han Chinese (cf. Wang 1994: 112)

\[
\begin{array}{llll}
\text{chu} & \text{yi} & \text{pian} & \text{shu.} \\
\text{take out} & \text{one} & \text{CL:WRITTEN.PIECE book} \\
\end{array}
\]

“Took out a book.”

(8.4) Mandarin (cf. Sun 2006: 160)

\[
\begin{array}{llll}
\text{na} & \text{p˘ı} & \text{h¯ei} & \text{mă} \\
\text{DEM:REMOTE CL:HORSE black horse} \\
\end{array}
\]

“That black horse.”

(8.5) Song Chinese (cf. Wang 1994: 141)

\[
\begin{array}{llllllllll}
\text{you} & \text{ge} & \text{Guanxi Yanzhou Yannan fu} & \text{de ren} \\
\text{have CL:GENERAL Gunaxi Yanzhou Yannan prefecture one person} \\
\end{array}
\]

“There is one person from the Yannan prefecture of Yanzhou in Guanxi.”

They can be used as carriers of anaphora, which means that they occur NP-internally, but independent of the presence of the classified noun (cf. 8.6, where the classifier adopts a pronominal function). This translates to a medium formalisation value for the system’s locus operandi.

(8.6) Han Chinese (cf. Wang 1994: 137)

\[
\begin{array}{llllllllll}
\text{[...]} & \text{ma} & \text{san} & \text{pi,} & \text{pi} & \text{er} & \text{li.} \\
\text{horse three CL:HORSE CL:HORSE two official} \\
\end{array}
\]

“[… ]there are three horses, each serving two officials.”

Erbaugh (1986: 403) however states that her data suggests that “classifiers were, in fact, seldom used as pro-forms in spoken Mandarin”. According to Erbaugh (1986: 427), classification in quantificational contexts became obligatory in the Song period (960–1127 AD). Obligatory classification in a single morphosyntactic context corresponds to a lower medium value. Classification in the remaining contexts is becoming obligatorily marked only in the Qing period (1644–1912), which raises the lower medium value to a higher medium value. The consulted sources feature the classifiers up until Mandarin as independent elements. The value for the markers’ boundedness is thus low. Mandarin however may feature classifiers additionally as clitics (cf. 149), which lifts the boundedness value to the lower medium.
(8.7) Mandarin (cf. Wu 2004: 78)

\[ \text{tamen duì zheì -ge wents de taolun} \]

they towards this **CL** problem POSS discussion

“their discussion of this problem”

The classifiers occur once and overtly, and multiple marking is not present in Mandarin (marking: low value; cf. 147 A = 8.8 where one classifier occurs despite the presence of two potential hosts, the demonstrative and the numeral).

(8.8) Mandarin (cf. Sun 2006: 160)

\[ \text{nà yē-pǐ hēì mà} \]

DEM:REMOTE one-**CL:**HORSE black horse

“That black horse.”

Finally, the presence of a general classifier, which obviously is not present in the classifierless Pre-Han period, allows an exhaustive classification of the nominal lexicon. This translates to a high value with respect to the system’s exhaustivity. Table 8.1 displays the rise from a Pre-Han low value to a medium Mandarin value in the system’s degree of formalisation. Only the growing inventory size goes against the trend of formalisation. Note that the Pre-Han period represents the predecessor of the classifier system and is not interpreted as a system of nominal classification in this study: “The number of proto-classifiers is quite small, and the function of those classifiers is more for quantification than categorization” (Wang 1994: 104).

<table>
<thead>
<tr>
<th>System</th>
<th>Inventory size</th>
<th>Loss number</th>
<th>Focus operand</th>
<th>Obligatoriness</th>
<th>Boundariness</th>
<th>Marking</th>
<th>Exhaustivity</th>
<th>( r )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Han Chinese</td>
<td>0.875</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.125</td>
</tr>
<tr>
<td>Han Chinese</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.214</td>
<td></td>
</tr>
<tr>
<td>Early Medieval Chinese</td>
<td>0.125</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.160</td>
<td></td>
</tr>
<tr>
<td>Late Medieval Chinese</td>
<td>0</td>
<td>0.5</td>
<td>0.333</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.262</td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td>0</td>
<td>0.5</td>
<td>0.667</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
<td>0.417</td>
<td></td>
</tr>
</tbody>
</table>

While displaying a constant rise in formalisation, the Mandarin system slightly loses its transparency. Only two properties remain the same throughout the Chinese history: First, the classificatory markers are not fused with other categories at any time, which attributes the markers’ categorial independence
a high value; second, nouns are generally flexible with respect to their classification (cf. Erbaugh 1986: 406), which also results in a high value. In the Han period for instance, “books were classified by several different classifiers one of which is pian” (Wang 1994: 111). Before the coining of the classifier system in the Han period, measure words could either add semantic content to the noun phrase (cf. again twenty [group of four] horses in 8.1) or mirror redundant semantic information (cf. 8.9), which results in a medium redundancy value.

(8.9) Jinwen Chinese (cf. Wang 1994: 80 and 83)

\[
\begin{array}{lll}
má & sì & pí \\
\text{horse} & \text{four} & \text{CL: HORSE} \\
\end{array}
\]

“four horses”

The same is the case for the later classifier system: The classifiers however can have a more thorough influence on the noun phrase in that they do not impose a temporary measure on the nominal referent, but serve e.g. referent distinction by classifying permanent inherent properties of the referent. At later stages, this may result in the lexicalisation of classifier-noun pairings and the loss of the classificatory quality of the classifier in those lexicalised constructions (cf. 2.36 B and C = 8.10, where the classifier-noun pairing is lexicalised and corresponds to our notion of a lexical noun; Jing Lin, p.c.). The redundancy value is thus raised to the higher medium.

(8.10) Mandarin (Zhang 2007: 48)

\[
\begin{array}{lll}
A & yì & lún \\
\text{one} & \text{CL: ROUND} & \text{moon} \\
\end{array}
\]

“a full moon”

\[
\begin{array}{lll}
B & yù & wán \\
\text{one} & \text{CL: CURVED} & \text{moon} \\
\end{array}
\]

“a crescent moon”

The assignment is solely carried out on semantic grounds. The number of involved rule types and their productivity is thus low, which results in a high transparency of this measure. The abovementioned conventionalised classifier-noun pairings are reported only for Mandarin, where the value is reduced to the higher medium. In the process of aging, two phenomena occur: First, the classifier inventory considerably grows, which results in a large number of assignment rules from the Han period on: “[M]any new classifiers appeared in the Han [period; MP] for the first time and their members are not as stable as they were to become in the later time” (Wang 1994: 111). With the number of classifiers increasing, the number of assignment rules also grows, which results in a low value. Second, individual classifier categories bleach and become semantically more general. At the same time, the use of general classifiers instead of specific ones becomes more frequent (Erbaugh 1986: 427 and 429; cf. 8.11).
(8.11) Chinese

zhāng “bows” (Han) > “zithers” (Han) > “paintings” (Tang) > “flat things, flags, tables” (Song)

tiáo “small branch” (Pre-Han) > “sticks” (Pre-Han) > “snakes, lengths of cloth, strings of gold ingots” (Post-Han) > “string, clothing” (Post-Han) > “long things” (Song)

gè “bamboo” (Pre-Han) > “lengths of bamboo” (Han) > “arrows” (Post-Han) > “arrows, candles, dogs, chicken, horses” (Tang) > “fruit, birds, people” (Song) > “general classifier” (Ming; cf. Erbaugh 1986: 429)

This results in a less straightforward semantic structure of individual classes from Late Medieval Chinese on, and consequently in a higher medium value. Cf. again 2.38 = 8.12 for an example of the classifier tiào “long”, which originally classifies pants, but is analogically extended to all sorts of compounds which feature ku “pants” as their head. This analogical classification does not look at the classified object’s actual shape, as it is also used for short objects.


A yi tiao CL:LONGISH nei ku underwear
“one pair of underwear”
B yi tiao CL:LONGISH sanjiao ku brief
“one pair of briefs”
C yi tiao CL:LONGISH youyongku ku swim.trunk
“a swimming trunk”
D yi tiao CL:LONGISH duanku ku shorts
“a pair of shorts”

Examples like 8.12 which are reported not before Mandarin, also indicate a loss of discreteness between the markers: A given marker may classify referents that are not falling into the marker’s original semantic scope. The measure of the markers’ discreteness thus receives a higher medium value. Finally, the classifiers may convey either redundant semantics (cf. again 7.10 on page 284 where the classifier pǐ “horse” classifies the noun mǎ “horse”), or they may contribute to a noun phrase’s meaning at all stages, which translates to a higher medium redundancy value. Cf. 2.8 = 8.13 where different classifiers may cause different readings).


A yi tong CL:COMMUNICATION dian-hua telephone
“one phone call”
Table 8.2 summarises the individual properties’ measures and the transparency values for the individual periods of Chinese, which display a steady drop from a high Pre-Han value to a higher medium Mandarin value.

Table 8.2: The transparency value in the diachrony of the Chinese system.

<table>
<thead>
<tr>
<th>Period</th>
<th>Semantic</th>
<th>Rule Types</th>
<th>Rule Number</th>
<th>Categorial</th>
<th>W. Markers</th>
<th>Redundancy</th>
<th>Flexibility</th>
<th>$y$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Han Chinese</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.929</td>
</tr>
<tr>
<td>Han Chinese</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.821</td>
</tr>
<tr>
<td>Early Medieval</td>
<td>0.75</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.786</td>
</tr>
<tr>
<td>Late Medieval</td>
<td>0.75</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.702</td>
<td></td>
</tr>
<tr>
<td>Mandarin</td>
<td>0.75</td>
<td>0.667</td>
<td>0</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.702</td>
<td></td>
</tr>
</tbody>
</table>

The conventionalisation and formalisation of the Mandarin system in general is mapped in figure 8.2, which displays a clear rise in formalisation and a decrease in transparency, which remains steady except for the transition from the Hand period to the Early Medieval period.

Figure 8.2: Developments in the Mandarin system (PH: Pre-Han Chinese, H: Han Chinese, EM: Early Medieval Chinese, LM: Late Medieval Chinese, M: Mandarin).
The higher registers of Mandarin display a more conventionalised character, which is discussed in the following as an example for a lowly formalised system’s stronger loss in systemic transparency. Erbaugh (1986: 413) states that “[l]evel of formality, discourse type, especially narrative and poetry; length of speech turn, presence of the referent, familiarity of the referent, and age of the hearer are all critical, synergistic, and variable determinants of special classifier use”. Lower registers display a more common use of general classifiers and flexible classification, and the more formal a discourse is, the more rigid and specific the classification becomes.

The formal properties of the higher registers of Mandarin are the same as described above and thus not presented again here. In short, the medium formalisation value of Mandarin is slightly higher than the low one of Late Medieval Chinese. The spoken variety of Mandarin shows a slightly more formalised setup with a medium value: According to Erbaugh (1986: 403), “22 core classifiers [...] cover virtually all the classifiers produced by both children and adults in this [Erbaugh’s; MP] study” (cf. table 8.3), which comes down to a higher medium value for the inventory size.

Table 8.3: The core classifiers of Mandarin (Erbaugh 1986: 405).

<table>
<thead>
<tr>
<th>single referent</th>
<th>shape</th>
<th>category</th>
</tr>
</thead>
<tbody>
<tr>
<td>dǐng “hats”</td>
<td>tǐao “long”</td>
<td>zhi “animals”</td>
</tr>
<tr>
<td>dāo “flowers”</td>
<td>zhāng “flat”</td>
<td>yuàn “clothing, situations”</td>
</tr>
<tr>
<td>běn “books”</td>
<td>kuài “square”</td>
<td>tōu “animals”</td>
</tr>
<tr>
<td>kē “trees”</td>
<td>pīan “slice”</td>
<td>bā “movable objects, with handles”</td>
</tr>
<tr>
<td>wěi “honored”</td>
<td>kē “small, round, hard”</td>
<td>yuàn “room”</td>
</tr>
<tr>
<td>yǔ “word, utterance”</td>
<td>gēn “stall, root-like”</td>
<td>yá “objects w. frame-like construction”</td>
</tr>
<tr>
<td>shǒu “verse”</td>
<td>zhǐ “small, slender”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lǐ “tiny, grain-like”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>duàn “length of X”</td>
<td></td>
</tr>
</tbody>
</table>

In contrast, the locus operandi in spoken Mandarin is more restricted than in its other varieties, as the possibility to use classifiers as carriers of anaphora is seldom exploited (Erbaugh 1986: 403). This results in the classifiers of spoken Mandarin mostly occurring together with the noun that they classify. However, since they have the potential to occur without the noun that they classify, a lower medium value comes closer to the real life situation in spoken Chinese than a low value. The higher registers of Mandarin allow the classifiers to serve the same function, and no restrictions on their usage as carriers of anaphora are reported. Therefore, the locus operandi of the higher registers of Mandarin displays a medium value. Table 8.4 lists the formalisation value of Late Medieval Chinese as the precursor of the system of Mandarin, which is presented both

---

2Note that Erbaugh’s (1986: 405) core classifiers do not include the general classifier ge, which is reported to be the most frequently used of all classifiers. This suggests that the set of core classifiers may be larger than the one identified in Erbaugh’s study.
in its higher registers and spoken variety.

Table 8.4: The formalisation value of the linguistic registers of Mandarin.

<table>
<thead>
<tr>
<th></th>
<th>inventory size</th>
<th>host number</th>
<th>loco-grammatical</th>
<th>obligatoriness</th>
<th>marking</th>
<th>exhaustivity</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Late Medieval Chinese</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Mandarin (higher registers)</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.667</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Spoken Mandarin</td>
<td>0.625</td>
<td>0.5</td>
<td>0.25</td>
<td>0.667</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

While the higher registers of Mandarin and its spoken variety do thus not vary much with respect to their medium formalisation value, their transparency value is medium and high respectively. The spoken variety of Mandarin displays a greater conformity with Late Medieval Chinese and thus is addressed first. The semantic structure of the individual classes in spoken Mandarin displays a lower medium value: There is a number of classes that contain a single referent (e.g. hats, flowers, books), classes encoding a referent’s shape, and classes that denote a referent’s hypernym category (e.g. animals, movable objects; cf. again table 8.3 and Erbaugh 1986: 405). The assignment is carried out partly on semantic grounds and partly consists of conventionalised classifier-noun pairings (cf. Bisang 1999: 23), which results in a higher medium value for the measure of rule types and their productivity. The number of assignment rules is consequently neither low nor high, but somewhere inbetween; the number of assignment rules is thus medium and lower than in Late Medieval Chinese, where the measure of the assignment rule number displays a low value. The classifiers are not fused with other categories in both Late Medieval Chinese and spoken Mandarin (cf. again table 8.3). The systemic markers’ categorial independence consequently displays a high value. Similarly, both varieties allow for a flexible classification and convey either redundant semantics or contribute to their host NP’s meaning. The flexibility value is thus high, and the redundancy value is in the higher medium. The list of spoken Mandarin ‘core classifiers’ provided by Erbaugh (1986: 405) contains two homophonous classifiers  ke, one denoting trees, the other one denoting small, round, hard, kernel-like objects. Similarly, it contains a classifier pair jiàn that can refer both to clothing and rooms (cf. again table 8.3). While those classifiers are graphically represented by different characters, this makes the set of spoken Mandarin classifiers a mixed one, and the discreteness value a higher medium one. In summary, the transparency value of spoken Mandarin is hardly higher than the one of Late Medieval Mandarin.

Turning over to the higher registers of Mandarin, a first observation is striking: Despite their low formalisation value, their transparency value is distinctly lower than the transparency threshold for lowly formalised system that was observed in the sample. This makes the higher registers of Mandarin an ex-
traordinary system within this study, as no other system with a formalisation value lower than 0.899 has a transparency value below the earlier postulated threshold of 0.7. However, the higher registers of Mandarin represent only a fraction of the Mandarin system, and an exceptional fraction indeed: If the assignment in these registers is as rigid as reported, this makes the Mandarin system fairly uneconomic and unlikely to survive in any environment that is less formal and regularised than the higher registers. Therefore, I interpret the low transparency value not as a contradiction to the earlier made observation that lowly formalised systems are fairly transparent. Instead, the higher registers are interpreted as an exception to this observation, which came into being and was maintained only in a highly specialised domain of speech.

The higher registers maintain only two properties of Late Medieval Chinese: First, the classifiers remain independent from other grammatical categories. This results in a high value for the markers’ categorial independence. Second, the number of assignment rules remains high, which results in a low value for the system’s number of assignment rules. This is due to the high number of opaquely assigned nouns — whenever a ‘correct’ class marker has to be learned for a given noun, this represents an assignment rule of its own (cf. again the lexicalised classifier-noun pairings in 2.36 B and C = 8.14 A and B).

(8.14) Mandarin (Zhang 2007: 48)

A  yi   lun yueliang
   one   CL-ROUND moon
“a full moon”
B  yu  wan yueliang
   one   CL-CURVED moon
“a crescent moon”

As the ‘correct’ classifier has to be learned, there seems to be a vast lack of semantic assignment rules, which could generalise the assignment for larger parts of the nominal lexicon and thus organise the assignment in a more economic manner. Therefore, the measure of assignment rule types and their productivity receives a higher medium value. Beyond that, this indicates that the classification is mostly rigid and semantically vain or opaque. The flexibility value is thus located in the lower medium of the scale. A semantic contribution of the classificatory markers to their host NP is however not excluded, which results in a medium redundancy value. Like in the more restricted set of spoken Mandarin, there are also homophonous classifiers in the larger set of Mandarin’s higher registers. This gives the system’s markers a higher medium discreteness value. Table 8.5 lists the transparency values of Late Medieval Chinese and the registers of formal and spoken Chinese.

---

3Similarly, the Kilivila classifier system features differences in its richness with respect to the speaker’s social ranking: The higher one’s social status is, the more classes are available. The full system is only available to the clan with the highest social status (Senft, p.c.).
Table 8.5: The transparency value in the diachrony of the Chinese system.

<table>
<thead>
<tr>
<th></th>
<th>semantic structure</th>
<th>rule types</th>
<th>role</th>
<th>number</th>
<th>categorial independence</th>
<th>discreteness of markers</th>
<th>redundancy</th>
<th>flexibility</th>
<th>y-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Late Medieval Chinese</td>
<td>0.75</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.786</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin (higher registers)</td>
<td>0.25</td>
<td>0.667</td>
<td>0</td>
<td>1.75</td>
<td>0.5</td>
<td>0.333</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoken Mandarin</td>
<td>1</td>
<td>0.667</td>
<td>0.5</td>
<td>1</td>
<td>0.75</td>
<td>0.75</td>
<td>1</td>
<td>0.809</td>
<td></td>
</tr>
</tbody>
</table>

Figure 8.3 illustrates the transition from Late Medieval Chinese to both the spoken variety and the higher registers of Mandarin.

The study of the diachrony of the Chinese classifiers have shown that a full system’s transparency does not undergo greater changes. In Chinese, we can observe a steady, but limited decline in systemic transparency. The overall high transparency is achieved by a continuous maintenance of the inventory: Individual categories bleach and may even drop out of the system; at the same time, new classifiers and thus new categories enter the system. In contrast, the formal properties grammaticalise to a higher degree. The higher registers of Mandarin constitute an exceptional case, as they display a severe loss in transparency. It is however unlikely that highly specified systems like this occur as independent systems due to their sociologically formal character and their violation of the principle of linguistic economy. In this study, the higher registers of Mandarin...
are the only lowly grammaticalised system that possesses a transparency value of less than 0.7. After grammaticalisation and formalisation processes in lowly formalised systems have been demonstrated with the example of Mandarin, the sections 8.2 and 8.3 deal with bleaching and conventionalisation processes in stronger formalised systems.

8.2 There and Back Again: Dyirbal Crossing the Transparency Threshold

The system of Dyirbal underwent a profound reanalysis and semantic remotivation of its system between the 1960s and 1980s. The fact that the semantic underpinning of the system was completely reanalysed within only one generation suggests that an idiosyncratic system calls for an immediate intervention by the speaker community and to be ‘renovated’ or dismissed. The Dyirbal language was studied in the field by Dixon in 1963 and again by Schmidt in 1982, at a point when it was approaching extinction. Within this period, the Traditional Dyirbal classification system faced a shift from an intact and transparent system, which was mainly based on semantics that relate to the cultural and mythological world of its speakers, to an idiosyncratic one for the Younger Dyirbal generation, as these speakers were not anymore familiar with the traditional assignment criteria. This idiosyncratic system was semantically fully restructured and stripped of all cultural connotations by speakers of Younger Dyirbal. Its formal properties remained unaltered and thus are touched on only briefly in the following. Since the Dyirbal system is documented in detail both in its original state and in its reanalysed state, its eliminated idiosyncratic aspects can be identified and presented. The intermediate, idiosyncratic state is not documented as such, but can be detected and reconstructed by filtering out those parts of the system that have been reanalysed. For this reason, the intermediate state of Dyirbal is marked by an asterisk. The assignment of Traditional Dyirbal is described first, before the later idiosyncracies in its assignment are addressed and presented together with the remotivated system. After that, the three developmental states of the Dyirbal system are mapped and commented on.

Dixon (1963: 120ff. and 1982: 178ff.) provides a detailed description of the semantic system underlying the Traditional Dyirbal concordial class system, which involves a large culture-specific and mythological knowledge. Dyirbal has four concordial classes, whose semantics and formal markers are listed in table 8.6.

<table>
<thead>
<tr>
<th>class</th>
<th>class marker</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-l</td>
<td>“masculine; animate”</td>
</tr>
<tr>
<td>II</td>
<td>-n</td>
<td>“feminine; water; fire; fighting”</td>
</tr>
<tr>
<td>III</td>
<td>-m</td>
<td>“non-flesh food”</td>
</tr>
<tr>
<td>IV</td>
<td>-∅</td>
<td>residue class</td>
</tr>
</tbody>
</table>

Since speakers of Traditional Dyirbal did not have to learn a noun’s concordial class individually (cf. Dixon 1968: 119ff.), but were able to assign new or unknown words intuitively and uniformly to a certain concordial class, this semantic system was productive at the time when Dixon collected his data. The classes I and II refer to the sex of human referents and (exceptionally) to the sex of animates in general (Dixon 1968: 123). Class I contains all human males and animates besides birds, whereas human females, birds, water, fire, and dangerous ‘things’ (fighting weaponry, toxic or otherwise harmful animals, etc.; cf. Dixon 1968: 120ff.) are hosted by class II. Class III comprises “all vegetable foodstuffs (and the trees and bushes that bear them” (Dixon 1968: 115). The fourth class in its function as the residue class finally gathers all nouns that do not fit into the categories of the other three classes (cf. Dixon 1968: 123). These basic semantic features are not carved in stone, but know a number of exceptions. These exceptions can be captured by two rules, which are:

1. if some noun has characteristic X (on the basis of which its class membership would be expected to be decided) but is, through belief or myth, connected with characteristic Y, then generally it will belong to the class corresponding to Y and not that corresponding to X.

2. if a subset of nouns has some particular important property that the rest of the set does not have, then the members of the subset may be assigned to a different class from the rest of the set, to ‘mark’ this property; the important property is most often ‘harmfulness’.

(Dixon 1968: 120 and 1982: 179; cf. also Lakoff 1990: 94)

This results in a reassignment of some parts of the nominal lexicon, which bear by mythological belief or actual properties characteristics of another concordial class than the ‘natural’ one. For instance, animals that are female beings in the Dyirbal mythology or that are harmful are assigned to class II instead of class I: Birds and crickets, which are mythological females, are assigned to class II, just like the firefly (due to its ‘fiery’ bite; cf. Dixon 1968: 120 and 122), and so forth. The moon, rainbows, and storms are assigned to class I due to their perceived masculinity in mythology, the sun and stars to class II due to their perceived femininity. Furthermore, fighting spears are assigned to class II,
which hosts dangerous items and beings, instead of the residue class IV. Fishing spears join the regular class of fish, which is I, and big short spears (supposedly for hunting) are assigned to class IV. As these examples show, larger parts of the Dyirbal nominal lexicon are not classified with respect to their referents’ natural inherent properties, but within the framework of the Dyirbal cultural conceptualisation of the world.

As the semantic assignment rules described above are deeply rooted in the culture and worldview of the Dyirbal tribe, the disintegration of its native customs and beliefs inevitably led to the loss of productive semantic assignment by the following generation of speakers. All assignment criteria that are not related to a referent’s animacy or sex bleached to an opaque degree and were given up. This notably includes the classification of non-flesh food, which is widespread in Australian languages, irrespective of the systemic type and genetic affiliation. These speakers remodelled the semantics of the assignment system severely: The ‘vegetable class’ III was given up completely, class I was maintained as the class for animates and males, and class II was reduced in its scope to human female referents only. Humans are thus always differentiated with respect to their sex, while animals are marked for their sex optionally. All other nouns are assigned to the unmarked residue class IV (Schmidt 1985: 153ff.). The reanalysed semantic system thus resembles the English system, the main difference being the assignment of non-human animates: These are assigned to the neuter class in English, but to the masculine class in Dyirbal. Table 8.7 provides an overview over selected nouns and their class membership in Traditional and Younger Dyirbal.

Table 8.7: Class membership of selected nouns in Traditional and Younger Dyirbal (cf. Schmidt 1985: 154).

<table>
<thead>
<tr>
<th>class I</th>
<th>class II</th>
<th>class III</th>
<th>class IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Dyirbal</td>
<td>man</td>
<td>woman</td>
<td>vegetable food</td>
</tr>
<tr>
<td></td>
<td>kangaroo</td>
<td>fire</td>
<td>black bean</td>
</tr>
<tr>
<td></td>
<td>fish</td>
<td>fighting spear</td>
<td>wild fig</td>
</tr>
<tr>
<td></td>
<td>moon</td>
<td>bird</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fish spear</td>
<td>stone fish</td>
<td></td>
</tr>
<tr>
<td>Younger Dyirbal</td>
<td>man</td>
<td>woman</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>kangaroo</td>
<td>—</td>
<td>moon</td>
</tr>
<tr>
<td></td>
<td>fish</td>
<td>—</td>
<td>fishing spear</td>
</tr>
<tr>
<td></td>
<td>bird</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td></td>
<td>stone fish</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

The changes in the semantic assignment allows an exact determination of the ‘inbetween’ state of the system, and consequently the mapping of this state. In the following, the Dyirbal system is accounted for in its traditional, opaque,
and remotivated state. The morphosyntactic expression remains unaltered at all stages and thus is presented only once for all three developmental stages. The Dyirbal system consists of four original classes, which are later reduced to a set of three classes by the deletion of the vegetable class (cf. again table 8.7). This attributes the measure of inventory size a high value. The classes are obligatorily marked on the plain, demonstrative, or interrogative form of the ‘noun marker’ by means of affixes (cf. again table 8.6 above and 53 = 8.15), which prevents a multiple class marking in the same context, but allows the markers to occur in noun phrases that do not contain the classified noun.


A  bala-∅∅  diban-∅∅  ya-ŋgu-∅  yibi-ŋgu  buran
  there-ABS-IV  stone-ABS  here-ERG-II  woman-ERG  look.at
  “The woman here is looking at the stone there.”

B  yala-∅∅  diban-∅∅  ba-ŋgu-∅  yibi-ŋgu  buran
  there-ABS-IV  stone-ABS  there-ERG-I  man-ERG  look.at
  “The man there is looking at the stone here.”

This results in a higher medium value for the system’s obligatoriness, a medium value for its host number, a high value for the markers’ boundedness, a low value for the measure of marking, and a medium value for its locus operandi.

Table 8.8 lists the individual measures and the higher medium formalisation value of the different stages of Dyirbal.

Table 8.8: The formalisation value in the diachrony of the Dyirbal system (*: reconstruction).

<table>
<thead>
<tr>
<th></th>
<th>inventory size</th>
<th>host number</th>
<th>focus operandi</th>
<th>obligatory classes</th>
<th>boundedness</th>
<th>marking</th>
<th>exhaustivity</th>
<th>x-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Dyirbal</td>
<td>1</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.667</td>
<td></td>
</tr>
<tr>
<td>*Intermediate Dyirbal</td>
<td>1</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.667</td>
<td></td>
</tr>
<tr>
<td>Younger Dyirbal</td>
<td>1</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.667</td>
<td></td>
</tr>
</tbody>
</table>

While the formal properties of the Dyirbal system remain the same, the transparency value varies between the three stages. Three of the transparency measures remain stable across the developmental stages. These are the categorial independence and the cross-class discreteness of the markers, both with a high value, and the flexibility of classification, which is largely rigid at all stages. The system’s flexibility thus displays a lower medium value. The other properties display changes: First, the semantic structure of the traditional classes is mixed and thus of a medium value. For instance, the female class also hosts
referents that are related to danger or animals and objects that correspond to females in mythology (like the sun and stars). This semantic classification fades away, and at the intermediate stage of Dyirbal, the speaker is confronted with many nouns whose assignment he or she cannot know due to a lack of understanding of the traditional assignment criteria. This lowers the value for the system’s semantic structure to the lower medium: For instance, there is no synchronic semantic motivation available why the sun and stars are assigned to class II, but the moon to class I. This results in not only mixed classes for intermediate Dyirbal, but mixed and opaque classes at the same time. After the reanalysis of the assignment, the semantic underpinning of the classes is bijective and thus again of a high value: Class I hosts biological males and other animates, class II hosts females, and class IV hosts all other nouns.

Second, there are two types of assignment rules in both traditional and intermediate Dyirbal, namely a semantic assignment and an opaque assignment, which attribute a higher medium value to the measure of assignment rule types and their productivity. The difference lies in the growing number of nouns that are assigned opaquely, which is addressed under the property of the number of assignment rules. In Younger Dyirbal, the reformed assignment knows only semantic rules, and the measure of rule types is thus rising to a high value.

Third, the number of assignment rules changes: The Traditional Dyirbal assignment rules as they are described above are few in number, which results in a high value for the measure of assignment rule numbers. Though “Dixon does not […] claim that all Dyirbal classification works by his principles” (Lakoff 1990: 95; emphasis in the original), the number of nouns that displays an unexplicable class membership and thus are assigned opaquely is small. Their number considerably grows on the way to intermediate Dyirbal, as all culturally motivated assignment rules become idiosyncratic. The value for the system’s assignment rule numbers thus decreases to a low measure. As has been shown above, a large number of nouns is affected by the culturally motivated assignment. After the system’s remotivation, the number of assignment rules is even lower than in Traditional Dyirbal and comprises only three rules. The system thus regains a high value for its number of assignment rules.

Fourth, the redundancy of conveyed semantics changes: In Traditional Dyirbal, the class semantics expresses either redundant information, which is already present in the lexical noun that is classified (e.g. in the case of non-flesh food), or are informative (e.g. in the case of sex distinction of humans; redundancy: higher medium). Since most of the assignment rules grow opaque on the way to intermediate Dyirbal, the class assignment of the bulk of nouns becomes idiosyncratic. This lowers the value for the system’s redundancy to a low measure. In Younger Dyirbal, the class assignment is remotivated, and the class semantics result in a fairly rigid classification. The information conveyed by the remotivated classes is basically redundant. Only in the case of the optional distinction of an animate referent’s biological sex by classificatory means, the classification is informative. The system’s redundancy measure thus corresponds to a higher medium value.
The account for the Dyirbal system demonstrates the considerable drop in its system’s transparency value: The original higher medium value, which is in the acceptable range of lowly grammaticalised systems, drops considerably to a medium one as a result of the bleaching of culture-specific knowledge of the younger speaker generation. These speakers obviously felt the need to remotivate the system’s idiosyncracies rather than eliminating the system, and remodelled it to a highly transparent system that closely resembles the English one. This remotivated system displays a transparency value of 0.869, which is higher than in Traditional Dyirbal. Table 8.9 lists the individual measures and transparency values of the different stages of Dyirbal.

Table 8.9: The transparency value in the diachrony of the Dyirbal system (*: reconstruction).

<table>
<thead>
<tr>
<th></th>
<th>semantic structure</th>
<th>role types</th>
<th>role number</th>
<th>categorial independence</th>
<th>discreteness of markers</th>
<th>redundancy</th>
<th>flexibility</th>
<th>y-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Dyirbal</td>
<td>0.5 0.667</td>
<td>1 1 1</td>
<td>0.75 0.333</td>
<td>0.75</td>
<td></td>
<td>0.464</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Intermediate Dyirbal</td>
<td>0.25 0.667</td>
<td>0 1 1</td>
<td>0</td>
<td>0.333</td>
<td>0.333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger Dyirbal</td>
<td>1 1 1</td>
<td>1 1 1</td>
<td>0.75 0.333</td>
<td>0.869</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is in order to address the question why the system was changed to such a great degree in such a short span of time. As has been argued in subsection 7.3.3, two main factors may influence the fundamental reanalysis of a system: The ‘visibility’ of the system in discourse, and the speaker’s own proficiency. Dyirbal employs only a small range of morphosyntactic hosts for its classificatory markers, only one of which can occur in the same construction at the same time. This makes it less visible in discourse than e.g. concordial class systems that employ five or more morphosyntactic hosts; concordial class markers are possible to cooccur in the same construction. The more important factor in the Dyirbal case however is the proficiency of the speakers: According to Schmidt (1985: 23), “[n]o individual in the 0–15 age group could construct a Dyirbal sentence”, and speakers in the age group 15-35 spoke an impoverished version of Dyirbal (Younger Dyirbal), which had a low reputation both in the community and the speakers themselves. Older speakers are reported to show a low acceptancy of Younger Dyirbal and to often correct younger speakers (ib.: 18). At the same time, English is the only language used at schools and also replaced Dyirbal as the primary language spoken at home, as children are intended to be well prepared for school by learning and practicing their English (ib.: 19ff.). Together with a lack of Dyirbal media, this results in a situation where Younger Dyirbal is a second language for every speaker. If a language is inherited to a group of second language acquirers, and its grammar brings with it a phenomenon that is opaque for this group of speakers, it might either be
dropped or remotivated rather than maintained and inaccurately passed on to the next generation of speakers. Either option requires immediate action, and the Dyirbal case shows that the reanalysis of the system can happen within a time span that short that one cannot even speak of a transitional period, where the phenomenon in question gradually changes, but rather of a sudden point of reanalysis. The only transitional period that can be theoretically postulated is the period where both speakers of Traditional and Younger Dyirbal coexist and use their respective variety. This is the case for Schmidt’s (1985) study, where both varieties are documented, but used without interfering with each other. Figure 8.4 summarises the developmental pathways of the Dyirbal system.

Figure 8.4: Developments in the Dyirbal system (T: traditional state, I: intermediate state, Y: younger state).

The study of Dyirbal has no direct consequence for the original grammaticalisation hypothesis, as the system’s formal properties do not shift into the direction of a concordial class system. Dyirbal however provides an example for the potential suddenness of systematic change in nominal classification devices, as the reanalysis of the system occurred within only one generation of speakers. The case study of West Germanic in section 8.3 provides the opposite example, where a change in both formal and semantic properties occurred over centuries.

8.3 Loss and Gain in Transparency in West Germanic

This subsection discusses the formalisation and bleaching of the West Germanic classification systems. Since the most eye-catching changes in the documented history of the Germanic languages occurred in Old and Middle English, this case is presented in depth. The non-documented, reconstructed earlier changes
and other systemic shifts in West Germanic sister languages of English (German and Dutch) are presented in short. The Proto-Germanic concordial class system was inherited from Proto-Indo-European. Proto-Indo-European originally distinguished two fairly transparent classes, an animate one and an inanimate one. The introduction of a third feminine class in late Proto-Indo-European, which attracted among female referent nouns from the animate class also nouns such as abstracts from the neuter class, caused a severe loss in systemic transparency. The Indo-European daughter languages, among them Proto-Germanic and its daughters English, German, and Dutch, applied different strategies to make the system effective again. Since the documented ‘Germanic solutions’ mostly did not result in greater changes of the formalisation value or transparency value, they are referred to only where it is necessary. Beyond that, the reconstructed state of the Proto-Indo-European and Proto-Germanic system is included in the account for illustrative reasons only. Since they are not documented, they are not intended to serve as evidence for a system’s loss in transparency. The sources consulted for the calculation of the formalisation value and transparency value of the West Germanic languages are Braune and Reiffenstein 2004, Burrow and Turville-Petre 2005, Duke 2009, Fodor 1959, Iglesias-Rábade 2003, Luraghi 2011, Meier-Brügger 2002, Paul et al. 1989, Ringe 2006, and Sievers 1903.

The Indo-European tripartite class system is remarkably stable throughout its history, reaching from the protolanguage and the Germanic diachrony to the current synchronic systems. The main difference between (early) Proto-Indo-European and its daughter languages is the system’s more restricted number of morphosyntactic hosts: Originally, only lower numerals were marked for class, and adjectives were originally marked for case and number, but not class (host number: lower medium value; cf. Luraghi 2011: 452 and Meier-Brügger 2002: 236). Since the formalisation value of all West Germanic languages except Modern English is the same, the following elaborations refer to the state of Old English exemplarily. Old English inherited the late Indo-European and Proto-Germanic tripartite concordial class system, which patterned the full nominal lexicon into classes. This results in a high value for the system’s inventory size and exhaustivity. Class membership was marked multiply, attributing a high value to the measure of marking, and NP-internally, independent of the presence of a classified noun, which results in a medium value for the locus operandi. The classificatory marking occurred obligatorily on the noun and the adjective, the numeral, and the different pronominal forms by means of affixes or fused forms (cf. 8.16). The measures of the system’s obligatoriness and boundedness therefore both display a high value.

For a more detailed presentation of the changes occurring in the classification systems of the individual Germanic languages, consult Duke (2009).
(8.16) Old English

A [pu... ]  
þu  
2SG beor-e  
3SG.M.GEN about  
spece[... ]  
say:IPF2SG from  
Breca-n  
speak:IPF2SG from  
“Drunk on beer, you [...] spoke about Breca[...], told of his journey.”

B [... ]  
*hildedeor-e  
battle.brave-NOM.M.PL nobleman-GEN.M.PL  
bearn caf-ra  
offspring:NOM.N.SG/PL all-GEN.PL twelve-NOM.M.PL wish:IPF3PL  
cear(q) e  
cwihan[... ]  
sorrow:ACC.F.PL mourn:INF  
“[…]the battle-brave sons of noblemen, twelve altogether, wished to bewail  
their sorrow[...].”

C [... ]  
faðh-e  
hostility-GEN.F.SG and  
fyrn-e  
guilt-GEN.F.SG be:IPF3SG too firmly  
on pam  
on DEM3SG.M.DAT  
“[…] violence and viciousness was too entrenched in him.”

D sele  
sele  
hall:NOM.M.SG tower:IPF3SG high-NOM.M.SG and  
and  
*horngeap[... ]  
wide.horned-NOM.M.SG  
“The hall towered high and horn-gabled[...].”

Old English underwent drastic changes in its grammar, including comprehensive phonological weakening and erosion of word-final segments. Since class membership and case are expressed by means of suffixes, the system grew quickly irrecognisable with respect to its formal expression. Though phonological changes contributed much to the development of the English system, the exact constellation and number of influential factors is debated until today: “Although the starting and end points of this development are generally agreed upon among scholars, little else about the change is clear” (Duke 2009: 237). Since most of the changes only finished in the Middle English period, this period represents the transition between the traditional West Germanic system and the current system of English. The calculation in table 8.10 is based on early Middle English, which is closer to the status of the Old English system than to the resulting status of Early Modern English.

The formal properties that were maintained in the Middle English period are the inventory size and the exhaustive classification of the nominal lexicon; both have a high value. Furthermore, the (possibly multiple; cf. 8.16 D, where the same noun’s class is marked on multiple elements) marking by means of affixes and fused forms was initially maintained, though the nominal affixes eroded considerably in the course of Middle English. The system’s measures of marking and boundedness both display a high value within this timespan. The survival of the traditional system into early Middle English is mostly “indicated by forms of the article, or by pronouns or adjectival inflexion” (Burrow and Turville-Petre
2005: 38), which indicates the maintenance of the original set of classificatory hosts (including numerals, which “were declined as adjectives”; Iglesias-Rábade 2003: 63) and locus operandi. The larger number of hosts translates to a high value, the NP-internal marking in absence or presence of the classified noun to a medium value of the locus operandi. The obligatoriness of class marking lies in the suffixes’ paradigms: Nouns and their modifiers are obliged to take the ending that corresponds to the nominal referent’s appropriate class, number, and case. As long as the suffixes survived, they were attached to nouns and their modifiers, which is mirrored in a high value for the system’s obligatoriness.

The reduction and restructuring of the system was finished “by the end of the [Middle English] period”, when “one essentially has the same system found in English today” (Duke 2009: 234). This system’s formal properties are described in the following. The properties that maintain a high value are the inventory size, the exhaustive classification of the nominal lexicon, and the expression by means of bound markers, which are fused with the third person singular forms of the personal pronoun. Since Modern English lacks affixal marking of case and class, class marking occurs only on the personal pronoun (he/she/it), interrogative pronoun (who/what), and possessive pronoun (his/her/its). The measure of the number of classificatory hosts thus corresponds to a medium value. The locus operandi of the English system diverts from all other sample systems in that the class marking occurs only in NPs that do not feature the classified noun, which results in a low value for the locus operandi. Since class information is expressed only on the mentioned pronominal forms, multiple marking in the same construction is not possible. The measure of marking thus also displays a low value. These pronouns inherently express class information (for instance, she refers to females and females only), which is why the classification is obligatory. The range of marked elements is however restricted to the mentioned pronominal expressions, which is reflected in a higher medium value for the system’s measure of obligatoriness. The most noticeable formal change towards the modern system is the loss of class marking on all non-pronominal hosts. This is however not an instance of degrammaticalisation, but rather the continuation of a formalisation process. Its result is the decay of formal marking on all non-pronominal parts of speech. Table 8.10 provides an overview over the formalisation value of German, Dutch, and English in their synchronic state as well as historical states of German and English. Note that the high degree

5This results in an interesting temporary reanalysis in Middle English, where a class mismatch can be often observed. This mismatch does not correspond to the traditional assignment, nor to the later classification according to a referent’s biological sex and animacy. Instead, the masculine accusative marker -ne was used for all classes as a surrogate accusative marker and sometimes as a dative marker, while the feminine genitive and dative marker -re was used as a dative marker for all classes (Jones 1967 and 1988, cited in Duke 2009: 236). This indicates that speakers give priority to the maintenance of the category of case above nominal classification, if both categories are endangered.

6Note that examples such as *He who pays the piper calls the tune*, where pronominal expressions occur in direct adjacency, are interpreted as featuring two separate phrases, each of which carries a pronominal: [[[He]$_1$]$_{DP}$ [who pays the piper]$_{CP}$ [ti calls the tune]$_{VP}$]
of formalisation is maintained through all stages and languages except Modern English, which features a reduced higher medium value.

Table 8.10: The formalisation value in the diachrony of the West Germanic system.

<table>
<thead>
<tr>
<th></th>
<th>inventory size</th>
<th>host number</th>
<th>locus operandi</th>
<th>obligatory marking</th>
<th>exhaustivity</th>
<th>*value</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Proto-Indo-European</td>
<td>1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.667</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>*Proto-Germanic</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Old High German</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Middle High German</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>New High German</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Modern Dutch</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Old English</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Middle English</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Modern English</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>0.667</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

After the changes in the Germanic and English system’s formal setup have been addressed, the shifts in its transparency are discussed. The value for Proto-Indo-European is based on more recent studies of its system wherever possible. Since the specifications of the Proto-Indo-European system are subject to lively debate and are not in the focus of this study, the following outline from Duke (2009: 83f.) may suffice as a description:

“One may assume a relatively high degree of semantic and formal transparency. We have little indication that the system was particularly opaque: There may well have been cases, however, in which the speaker of late Indo-European simply had to know the gender of the noun by rote because for that particular noun neither semantic nor formal criteria applied. All in all, there appears to be reason to believe the system was neither particularly opaque nor completely transparent.”

The original early Indo-European two-class system was based on the distinction of a referent’s animacy, thus possessing an animate and an inanimate class. This semantic system has been confused by the introduction of the feminine gender, which attracted nouns denoting female referents from the animate class and abstracts from the inanimate class among others. The succeeding systems of the individual Indo-European and even Germanic daughter languages show a great variation with respect to all kinds of properties, which indicates the efforts of Post-Indo-European speakers to remotivate their respective system. It is thus in order to assume that the Post-Indo-European languages inherited
a system that was by far less transparent and manageable than the proto-
system outlined by Duke above. Note though that the transparency value of
Proto-Indo-European is already in the medium range.

The systemic confusion is mirrored by the low transparency value of Proto-
Germanic. The semantic structure of the classes is confused and mainly re-
dundant or opaque, which results in a lower medium value. The assignment
to a given class may occur on a morphological or semantic basis, and in many
cases is opaque; this leads to a high number of assignment rules. The measure
of assignment rule types and their productivity thus is located in the lower
medium, and the value of the number of assignment rules is a low one. The
markers are fused with the categories of number and case, which results in a
lower medium value for the system’s measure of the categorial independence
of its markers, and the same class marker can refer to more than one class.
The markers’ discreteness value hence is low. The information conveyed by the
class markers can be opaque, redundant, or informative, and a noun’s class
membership is mostly rigid. The system thus displays a medium redundancy
value and a lower medium value for its classificatory flexibility. This results in
a dropping transparency value from word to word, and comes down to a lower
medium value.

The Germanic daughter languages applied minor changes and attempts for
restructuring their systems both formally and semantically, yet these changes
did not result in greater variations of both the grammaticalisation and trans-
parency value. Table 8.11 illustrates that German, Dutch, and Old English do
not exceed a transparency value of 0.333.

Only the diachrony of the English
language shows a significant drop in its formalisation value and a significant
rise in its transparency value, which is described in the following.

The transparency value of Old English is 0.333 and thus slightly higher
than the Proto-Germanic one. This is due to the fusion of the Old English
markers with case only, and due to the partial employment of markers that
denote one class only. The value for the markers’ categorial independence is
thus a higher medium one, and the one for its markers’ discreteness a medium
one. In contrast, the semantic information conveyed by the markers is redu-
dant or opaque in most cases, which means a further reduction of the medium
redundancy value in the course from Proto-Germanic to Old English, where
the value corresponds to the lower medium. From the Old English period on,
the system’s semantic underpinning is getting lost due to the loss of distinc-
tive inflectional endings and the great number of opaquely assigned nouns.
Together with the inflectional morphology, the instances of morphological as-
signment crumble away. This results in the reduction of assignment rule types

\[7\]

---

An account for the North Germanic systems of Danish, Swedish, and Nynorsk (additional
sources consulted: Kilarski 2003; Matthias Krause, p.c.) shows the same trend: The Danish
and Swedish system retain a high formalisation value of 0.929 and drop in their transparency
value down to 0.202. The Nynorsk system decreased its number of classificatory hosts and
thus shows a slightly lower formalisation value of 0.893. Beyond that, it employs discrete
markers for its classes and thus has a higher transparency value of 0.345.
to semantic and opaque ones and translates to a higher medium value. As both the categories of case and class are collapsing, the Middle English markers are growing categorically more and more independent, which is mirrored in a high value for the markers’ categorial independence. Like in Proto-Germanic, a part of the class markers may refer to more than one class (the relative pronoun þe for instance refers to both the masculine and the feminine class (Iglesias-Rábade 2003: 100); in the further course of time, it is extended to the neuter class and thus does not distinguish class anymore). Except in the cases where a referent’s biological sex is indicated by class marking, the semantic information conveyed is opaque, which is mirrored in the great number of class mismatches and general flexibility in class membership (cf. again footnote 5 on page 328).

The system’s redundancy value at this stage is thus low. The Middle English system with a transparency value of 0.440 is thus more flexible and less complex in its assignment than the Old English one. The loss of understanding of the system’s underlying semantics, however, demands a further revision of the system. This revision occurred in the further progress of Middle English and led to the fully transparent classification system of Modern English by an in-depth semantic remotivation.

The Modern English assignment rules are strictly semantic and low in number: Human males are assigned to the masculine class, human females to the feminine class, and all other nouns to the neuter class. This results in a high value for both the measure of the assignment rule types and their number. The structure of the reanalysed classes is thus bijective and its value high; nonhuman animates can be optionally assigned to the human animate classes according to their biological sex. Beyond that, all inanimate nouns can be optionally anthropomorphised and are thus flexible in their class membership. The system’s classificatory flexibility has a higher medium value. For instance, ships, motorcycles, or instruments can be assigned to the feminine class, which expresses a certain degree of intimacy or personal importance of the denoted referent. The pronominal markers are not fused with categories such as case, which attributes a high value to the markers’ categorial independence and are discrete per class, except for the interrogative pronoun who, which refers to both the masculine and the feminine class. The discreteness of the markers corresponds to a medium value. The semantic information conveyed by the classificatory markers is either redundant, if the information on a referent’s sex or humanness is part of its lexical entry (cf. king m., queen f., and crown n.), or it is informative, if the information on a referent’s sex is not part of the classified noun’s lexical entry (cf. 8.17 A and B, where the pronominal forms convey the information on the referent’s sex) or if a neuter noun’s referent is anthropomorphised (cf. 8.17 C). The measure of semantic redundancy of the markers translates to a higher medium value.
8.3. Loss and Gain in Transparency in West Germanic

A I’ll introduce you to my friend. She’s a gem.
B I’ll introduce you to my friend. He’s a gem.
C My guitar needs a repair. Her neck broke. (English)

These properties render the Modern English system a highly transparent one. Its transparency value amounts to 0.845 and thus exceeds the transparency value of its (West) Germanic siblings by a factor of up to four. Table 8.11 lists the individual values of the West Germanic systems.

Table 8.11: The transparency value in the diachrony of the Germanic system.

<table>
<thead>
<tr>
<th></th>
<th>semantic structure</th>
<th>rule types</th>
<th>rule number</th>
<th>categorial independence</th>
<th>discreteness of markers</th>
<th>redundancy</th>
<th>flexibility</th>
<th>y-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Proto-Indo-European</td>
<td>0.5 0.667</td>
<td>1</td>
<td>0.333</td>
<td>0</td>
<td>0.5</td>
<td>0.333</td>
<td>0.476</td>
<td></td>
</tr>
<tr>
<td>*Proto-Germanic</td>
<td>0.25 0.333</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0.5</td>
<td>0.333</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Old High German</td>
<td>0.25 0.333</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0.5</td>
<td>0.667</td>
<td>0.298</td>
<td></td>
</tr>
<tr>
<td>Middle High German</td>
<td>0.25 0</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0.5</td>
<td>0.667</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>New High German</td>
<td>0.25 0</td>
<td>0</td>
<td>0.333</td>
<td>0</td>
<td>0.5</td>
<td>0.333</td>
<td>0.202</td>
<td></td>
</tr>
<tr>
<td>Modern Dutch</td>
<td>0.25 0.333</td>
<td>0</td>
<td>0.667</td>
<td>0.5</td>
<td>0.25</td>
<td>0.333</td>
<td>0.333</td>
<td></td>
</tr>
<tr>
<td>Old English</td>
<td>0.25 0.333</td>
<td>0</td>
<td>0.667</td>
<td>0.5</td>
<td>0.25</td>
<td>0.333</td>
<td>0.333</td>
<td></td>
</tr>
<tr>
<td>Middle English</td>
<td>0.25 0.667</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
<td>0</td>
<td>0.667</td>
<td>0.440</td>
<td></td>
</tr>
<tr>
<td>Modern English</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.75</td>
<td>0.667</td>
<td>0.845</td>
</tr>
</tbody>
</table>

In contrast to the Dyirbal system, which changed within one generation, the English system displays a very slow and gradual transformation, which stretches over centuries. This may be due to the a grammatical restructuring that is by far more comprehensive than in the Dyirbal case: In English, not only nominal classification is affected, but also the category of case, the whole phonological system, the word order, and other parts of grammar. Such in-depth changes are not likely to unfold ‘over night’, but demand a gradual and consensual development. Figure 8.5 illustrates the shifts of the English classification system, including its prior stages.
Diachronic Study

8.4 Discussion

The case studies of Mandarin, Dyirbal, and English in this chapter have two outcomes: First, they exemplify the possibilities of the diachronic application of the mapping presented in chapter 7. Second, they provide a number of cues about the mechanisms and motivations behind systemic shifts in nominal classification devices. These cues and motivations are summarised in this section. They comprise a system’s visibility in discourse, the speaker’s proficiency of the language, potential ‘degrammaticalisation’ phenomena, and the extent or restriction of grammatical change. Finally, the research questions Q3 and Q4, which require the outcomes of both the synchronic and the diachronic study, are addressed at the end of this section.

Visibility  The diachronic case studies suggest that the lower a system’s visibility in discourse is, the higher the chances for a reanalysis of its underlying semantics are. Visibility refers to a given system’s potential of marking class membership: A typical classifier system, which allows only a single classificatory marker to occur, is less visible than a system that marks class membership on a range of nominal satellites at the same time (cf. 8.18 and 8.19).
(8.18) Nùng (cf. Saul and Wilson 1980: 26)

He then threw a stone into the river.

(8.19) Italian

“My girl is clever and educated.”

The more visible a noun’s class membership in discourse is, the more cues a speaker has available to access the underlying classification mechanisms. This is obviously a tendency that is met in this chapter’s case studies and not a strict rule that applies to all systems, as all Romance languages for instance altered the gender system of Latin, which is known for its richness in classificatory marking and represents a ‘highly visible’ classification system. Nevertheless, the development of the Dyirbal and English system suggests that a low degree of visibility supports a system’s extensive reanalysis. In contrast, highly visible systems such as the one of German do not display extensive changes of their underlying classification (cf. Audring 2014: 13–16 for a discussion of the same phenomenon).

Another factor that influences a system’s visibility is the frequency of occurrence of given class-noun pairs: Certain nouns occur more often than others in discourse, and certain classificatory contexts feature certain groups of nouns more often than others. For instance, alienable possessive constructions most likely feature more often non-human entities than human entities, which means that nouns referring to non-human entities are classified more frequently in a possessive classifier system than nouns referring to human entities. Frequency effects typically stabilise grammatical phenomena, even if they display more severe irregularities than more regular, yet less frequent representatives of the same phenomenon. For instance, the existential verb to be is highly frequent in its present tense form in discourse, and it displays a suppletive and thus more irregular paradigm than its less frequent preterite tense paradigm or the verb to exist (cf. table 8.12).
Table 8.12: Frequency stabilising irregular grammatical phenomena: to be and to exist.

<table>
<thead>
<tr>
<th>PRS</th>
<th>IPF</th>
<th>PRS</th>
<th>IPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>am</td>
<td>was</td>
<td>exist</td>
</tr>
<tr>
<td>2SG</td>
<td>are</td>
<td>were</td>
<td>exist</td>
</tr>
<tr>
<td>3SG</td>
<td>is</td>
<td>was</td>
<td>exists</td>
</tr>
<tr>
<td>1PL</td>
<td>are</td>
<td>were</td>
<td>exist</td>
</tr>
<tr>
<td>2PL</td>
<td>are</td>
<td>were</td>
<td>exist</td>
</tr>
<tr>
<td>3PL</td>
<td>are</td>
<td>were</td>
<td>exist</td>
</tr>
</tbody>
</table>

Nouns that refer to basic concepts such as body parts, humans, foodstuff, or housekeeping, are likely to occur more often in discourse than more specific ones or abstract ones. This means that the speaker is more often dealing with the classification of those basic ‘core concepts’, which may lead to a situation where e.g. abstracts do not fall under the scope of nominal classification at all (cf. Jakaltek) or where the more concrete basic referents possess a greater categorial specificity (cf. the inventory of Nùng in table [121] where the classification of shape allows a greater specificity than the classification of inanimate non-concrete objects).

Table 8.13: The Nùng classifiers (cf. Saul and Wilson 1980: 25ff.)

<table>
<thead>
<tr>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>áhn</td>
<td>“inanimate”</td>
</tr>
<tr>
<td>tû</td>
<td>“animate”</td>
</tr>
<tr>
<td>ỏhng</td>
<td>“human”</td>
</tr>
<tr>
<td>côh</td>
<td>general classifier</td>
</tr>
<tr>
<td>báb</td>
<td>“leaf-like object”</td>
</tr>
<tr>
<td>côn</td>
<td>“lump-shaped object”</td>
</tr>
<tr>
<td>hông</td>
<td>“clothing”</td>
</tr>
<tr>
<td>môí</td>
<td>“bean-like object without shell”</td>
</tr>
<tr>
<td>phôn</td>
<td>“bedding”</td>
</tr>
<tr>
<td>léu</td>
<td>“stick or string-like object”</td>
</tr>
<tr>
<td>maht</td>
<td>“fabric, board, brick, tile, or raindrop”</td>
</tr>
</tbody>
</table>

At the same time, nouns that occur more frequently than others in discourse are also likely to occur more often in classificatory contexts than others, which means that the speaker receives more cues to the classification of these basic terms than less frequent terms. Similarly, certain classificatory contexts such as possessive classifiers may prefer or disprefer certain groups of nouns and thus create a disproportional representation of nouns in classification. The conjecture that can be drawn at this point is that less visible classification as well as less frequent nouns are more prone to an either flexible classification.
or to a reanalysis of their underlying classification in case that the system as a whole is not transparent to the speaker (cf. again Dyirbal, where only the distinction of animacy and sex remained after the system’s reanalysis).

**Proficiency**  The Dyirbal case study brought another trigger of change to light: The overall linguistic and cultural proficiency of a speaker community. According to Schmidt (1985: 23), there were three groups of speakers at the time of her data collection, which showed a correlation between their age and proficiency in Dyirbal. The youngest age group (0–15) was not able to phrase a Dyirbal sentence at all; the age group of 15–35 spoke a variety (Younger Dyirbal) that showed severe differences towards Traditional Dyirbal, and the speakers above the age of 35 spoke ‘proper’ Traditional Dyirbal. The attribute ‘proper’ also fits the reputation of the language: The younger variety is perceived as an impoverished and imperfect distortion of Traditional Dyirbal, even by speakers of Younger Dyirbal themselves. Despite the successful and comprehensive grammatical reanalysis of the Dyirbal language, the speakers of Younger Dyirbal thus perceivedly lack proficiency in their own language.

Next to the perceived lack of proficiency between the groups of speakers of Traditional and Younger Dyirbal, there is a difference in the context of acquisition of the language, which has an effect on the speakers’ actual linguistic proficiency: While speakers of Traditional Dyirbal acquired it as their mother language, the role of Dyirbal within the speaking community decreased with the growing influence of English. English is present in media, schools, and daily discourse, which suppresses the role of Dyirbal as a means of communication. Beyond that, English is more prestigious than Dyirbal, and thus replaces Dyirbal as the communication language even at home (cf. Schmidt 1985: 16–26). In a context like the one described, the preconditions for language acquisition are less favourable than in a situation of ‘natural’ first language acquisition, as the acquirer receives less input and possesses less chances to practice and establish the grammatical input. On top of that, the more prestigious and present language (English) enters in competition with the less prestigious and present language (Dyirbal), which facilitates eventual language attrition. If a language is passed on to the next generation under such unconducive circumstances, the next generation is likely to consist of second language speakers rather than to show a proficiency like native language speakers. Native language speakers are typically able to acquire the most complex and intransparent grammatical phenomena without bigger problems. The comprehensive restructuring of Younger Dyirbal as compared to Traditional Dyirbal shows that this is indeed the case for Dyirbal, where together with the grammatical changes in the language a loss in traditional cultural knowledge occurred.

‘Degrammaticalisation’  The case study of the English system and its diachronic development showed a phenomenon that could easily be confused as an instance of degrammaticalisation. Degrammaticalisation constitutes the re-
verse process of grammaticalisation; it renders a linguistic element’s shift from a more grammatical function to a more lexical one. Since the English system is not degrammaticalising, the representation of its development in the proposed mapping graph is commented on in the following; figure 8.6 displays the developmental shifts of the English system.

Figure 8.6 shows a notable decrease in the x-value between Middle and Modern English. Since the x-axis represents the degree of grammaticalisation of formal properties, a reduction of formal grammaticalisation might be misinterpreted as degrammaticalisation. In fact, the opposite is the case, but let us first recapitulate what the x-value exactly denotes: The correspondence of a number of formal properties to the expected outcomes of a grammaticalisation process. For instance, a grammaticalisation process is prone to result in the grammaticalising element’s loss of morphosyntactic indepence and first cliticise, then morphologise. In the case of Old and Middle English, we are dealing with affixed class markers, the logical endpoint of a linguistically productive result of a grammaticalisation process. However, if the mechanisms of grammaticalisation remain operating, the grammaticalising element can further erode and ultimately disappear. This is the case in Middle English, where the inflectional markers carrying case and class information fully dissolve.

If the formal means of nominal classification in the Middle English system undergo a further grammaticalisation process, how can its x-value decrease? The loss of formal marking marks the decay of the grammaticalising phenomenon; in contrast, the grammaticalising phenomenon may be assumed to stay a productive linguistic unit up until its morphologisation. If marking on a range of nominal modifiers in a decaying system like the English one disappears due to phonological erosion, this correlates with an earlier, less gram-
maticalised state of the system, when classificatory marking occurs on fewer elements. In the state of formal decay, a system can thus formally resemble a less grammaticalised system again, which results in a lower $x$-value. This also means that a loss in a system’s $x$-value does not equal an instance of degrammaticalisation, as degrammaticalisation is a highly infrequent process, during which the degrammaticalising element is developing a more lexical nature.

**Extent of grammatical change**  The English case study has shown that in-depth changes of a system’s formal properties may be related to a grammatical shift that alters more than just nominal classification. The English case includes among others the breakdown of the whole system of nominal inflection by phonological weakening, and the development of a more rigid word order. If those factors were not present, it is highly unlikely that the classification system of English would have undergone a development as drastic as it was the case.

Similarly, the Mandarin classification system came into being due to a change in word order, which in and by itself is not related to nominal classification at all: The classifier system derives from an earlier postnominal construction featuring measure words (word order [noun – numeral – measure word]; cf. 8.1 = 8.20), and emerged only after the measure word construction developed a prenominal structure (word order [numeral – measure word/classifier – noun]; cf. 8.2 = 8.21).

(8.20) Pre-Han Chinese (cf. Wang 1994: 102)

\[Qihuangong qi zhi, you [ma er shi cheng].\]

“Duke Huan gave him a lady to wife, and he had 80 horses (in 20 teams of four).”


\[Zhen pan si zha hua, jing xi [san tiao shu].\]

“Contend in breaking, compete in playing on the three roads.”

The third case study of Dyirbal does not involve formal changes in the classification system, but serious changes in the sociolinguistic and acquisition situation of Dyirbal. Consequently, all three diachronic case studies involve an extent of developmental shifts that goes far beyond the scope of nominal classification and certainly cannot be reduced to a mere economisation of the respective system by means of grammaticalisation. What does that mean for the grammaticalisation hypothesis? Recall that the sample study does not contradict a classifier system grammaticalising into a concordial class system. Systems that display an $x$-value of more than 0.8 and lower than 0.8 are however separated by a considerable gap (cf. figure 7.10 = 8.7) instead of a more densely populated transitional area.
Taking these observations into consideration, an answer to the initial question why the grammaticalisation hypothesis is not documented for a single language can be provided: It would require not only a grammaticalisation of the nominal classification system’s formal properties, but also accompanying changes in other domains such as word order or paradigmatic qualities. As for the paradigmatic qualities, recall that most classifier systems are located in isolating languages in South East Asia, which typically do not feature affixes. If those languages were to develop concordial class systems, which inevitably employ agreement, they were obliged to introduce the potential for agreement marking by means of affixes. While this is not impossible, it is highly unlikely and demands such specific and comprehensive additional changes in the grammar and language type, that an actual ‘switch’ from a classifier to a concordial class system has a very low probability to ever occur. The chance to document a language in the process of switching from classifiers to concordial classes is thus close to zero.

**Research questions**  After having discussed the main observations from the diachronic study, it is in order to return to the original research questions. As indicated in section 6.9 on page 246, the questions Q3 (Where do concordial class systems come from?) and Q4 (Why do concordial class systems exist, if they are functionwise seemingly vain?) can now be addressed, as the diachronic study has been conducted.

As for Q3, this study has shown that there are no theoretical or typological obstacles for concordial class systems to derive from classifier systems; the probability of the origin of concordial classes from classifiers has however shown...
to be very low. Earlier studies (e.g. Fodor 1959, Greenberg 1978a, Luraghi 2011) have convincingly argued that concordial class systems such as the Indo-European one most likely derived from the reanalysis of other, existing grammatical phenomena such as case marking patterns or demonstratives. This mechanism presupposes that certain groups of nouns occur more often with certain case markers (or demonstratives) than others. Due to their frequent cooccurrence, they are reanalysed as classificatory markers rather than case markers or demonstratives, and a concordial class system is the result of this reanalysis. A concordial class system that emerges out of a suchlike reanalysis thus possesses the formal preconditions of a concordial class system from the very beginning on. It gains its classificatory semantics indirectly, and not by the step-by-step bleaching and reorganisation of a classifier inventory.

Question Q4 can be answered more in detail than Q3: Figure 8.7 demonstrates that concordial class systems display very little formal variation, but a great variation with respect to their transparency and semanticity. The existence of semantically highly productive concordial class systems disproves the seeming functional vanity of concordial class systems, as those systems are highly productive in functions such as referent specification and differentiation, individuation, and attribution of speaker-sidedly perceived properties (cf. Contini-Morava and Kilarski 2013: 268–279 and 6.64 = 8.22, where class reassignment in Maasai may serve among others the attribution of properties).

(8.22) Maasai (cf. Payne 1998: 166)

<table>
<thead>
<tr>
<th>A</th>
<th>en-</th>
<th>tito</th>
</tr>
</thead>
<tbody>
<tr>
<td>F.SG</td>
<td>girl</td>
<td>“girl”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>ol-</th>
<th>tito</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.SG</td>
<td>girl</td>
<td>“large shapeless hulk of a woman”</td>
</tr>
</tbody>
</table>

Contini-Morava and Kilarski (2013: 279ff.) furthermore show that concordial class systems are functionally productive in the domain of discourse in a variety of ways. Therefore, the alleged functional vanity attributed to concordial class systems cannot be backed by a broad range of systems. Nevertheless, there are individual systems that seem to be more of an obstacle to mastering a given language than to functionally contribute to it. Beyond that, the commonly known functions of nominal classification, despite being present, seem to play only a minor role in these systems. Take e.g. the Maasai system, which can be shown to serve among others the functions of referent differentiation (e.g. ol-anásh’ m. “brother” vs. enk-anásh’ f. “sister”), the attribution of properties (e.g. ol-abááni m. “(male) doctor” vs. enk-abááni f. “female doctor, quak”; cf. Payne 1998: 166), and reference identification (cf. 146, where the class information allows an identification of the referent).

---

8 For instance, reference tracking was long thought to be a “major function of gender” (Corbett 1991: 322), yet systems of all kinds recently have been shown to be involved in reference tracking to a smaller degree than previously assumed (Feist 2016).
These functions are also involved in the German system, but they are not exploited to a greater degree: Referents are typically distinguished not by classificatory means, but directly by lexemes (e.g. \textit{der Eber} m. “boar” vs. \textit{die Sau} f. “sow”), while examples of referent distinction by means of class marking such as in \textit{der/die Gl"aubige} m./f. “the (male/female) believer” are not very frequent. Class reassignment in German is not available to the speaker on a productive basis. The German system thus exploits its functional options to a small degree only. Instead of being functionally vain, the system could thus be labeled functionally underexploited. Considering the German system’s lack of transparency and accessibility and the enormous efforts that a (second language) learner has to invest in its acquisition due to its high complexity, it is indeed in order to ask why the system as such still exists and, as has been argued in subsection 8.3, not changed to a great degree since Proto-Germanic times.

The sample study has shown that only systems with highly formalised properties are able to develop a considerable degree of intransparency. Taking into account the abovementioned arguments of visibility and proficiency, the passing on of such a system over many generations can be explained by its firm embedding in discourse and by the fact that any child acquiring the German language as its mother language will have no difficulties to acquire its classification system. Second language acquirers in contrast are facing enormous difficulties in the system’s acquisition, yet those speakers did not exert influence on systemic developments until today. This may be related to their comparatively low number or the low acceptance of an incompletely acquired or mastered grammar in the speaker community. Experimental research supports this conjecture; Irslinger (2009: 4, referring to studies by Plemmenou et al. 2002 and Monpiou et al. 1995) reports that “[…]while gender slows down language procession, if one retrieves it or has to learn it as an L2-acquirer, that is in case of conscious access. This is however not the common situation of speech production, in which gender accelerates processes in unconscious and automaticised application” (own translation).

Coming back to Q$_1$, the answer lies in the manner of passing on of the system: Concordial class systems such as the German one are not “a useless fossil” (Bonfante 1946: 849, cited in Kilariski 2013: 329), but deeply embedded in grammar and the speakers’ consciousness. As native speakers of a concordial class language have no reason to question the use of the classification system in question, but are simply acquiring and mastering it, the presence and stability of concordial classes in a large number of languages does not come as a surprise, despite their functionality being largely unrecognisable. This concludes the discussion of the diachronic part of this investigation. Chapter 9 summarises and discusses the results of this study as a whole.
Conclusions

Since nominal classification is nothing else than the grouping of nominals, if and when it is expressed linguistically; since, furthermore, the grouping of nominals may change again and again under influences of a magical-religious or cultural kind so as to finally erode to a mere formal matter, classification represents a question which entails copious quantities of further questions.

(Royen 1929: IV; own translation)

This study has shown that the above statement holds until today. Nevertheless, it offers an answer to various questions that are related to nominal classification. This chapter summarises the findings and outcomes of the investigation of grammaticalising properties in nominal classification devices. The central subject of this study is the grammaticalisation hypothesis, i.e. the assumption that classifiers may grammaticalise into concordial class systems. Section 9.1 therefore presents the arguments and findings on the likeliness of the grammaticalisation hypothesis from a typological perspective. Section 9.2 puts the outcomes of the present sample study into a more general context and discusses its consequences for theory of nominal classification in general. Section 9.3 concludes the study with an overview of questions and phenomena that are relevant and promising for future research on nominal classification.

9.1 The Grammaticalisation Hypothesis

Grammaticalisation relates to nominal classification on all levels and hence contributes a lot to our understanding of the typology and functionality of nominal
classification devices. Nominal classification constitutes a continuum of linguistic phenomena that serve the classification of nouns and their referents. This continuum is internally organised with respect to an increasing degree of grammaticalisation, which means that it features one pole that is characterised by a low degree of grammaticalisation, where classifier-like systems are found, and an opposite pole that is characterised by a high degree of grammaticalisation, where concordial class-like systems are found.

The grammaticalisation hypothesis assumes classifier systems to potentially grammaticalise into concordial class systems. While this assumption is supported by our theoretical knowledge of both nominal classification and grammaticalisation, as I have argued in chapter 2 and 3, an actual transition of a classifier-like system into a concordial class-like system is not documented. Therefore, this study analysed a sample of forty languages with respect to their nominal classification systems' degree of grammaticalisation. If the scenario that is assumed by the grammaticalisation hypothesis occurs, the systems, which are at different stages of grammaticalisation, should be able to project onto a 'grammaticalisation scale' without a gap.

This grammaticalisation scale derives from existing theory and language descriptions: The majority of descriptive properties that apply to both classifier-like and concordial class-like systems distinguishes different stages of grammaticalisation. For instance, classifier-like systems are typically described as optional systems, which feature morphosyntactically independent or clitic markers, and concordial class-like systems are typically described as obligatory systems, which feature clitic or affixed markers. Both these properties correspond to a low degree of grammaticalisation in the case of classifier-like systems and to a high degree of grammaticalisation in the case of concordial class-like systems. In section 3.2, I have isolated a number of frequently used descriptive properties of nominal classification devices which apply to both systemic types; this package of properties serves as a measure for a given system's degree of grammaticalisation. More specifically, a system that employs only classifier-like properties is a prototypical classifier system, one that employs only concordial class-like properties is a prototypical concordial class system, and one that employs both classifier-like and concordial class-like properties is of an intermediate nature; depending on the number of either classifier-like or concordial class-like properties, such an intermediate system is either more classifier-like or concordial class-like. Grammaticalisation theory furthermore predicts a preferred order of individually grammaticalising properties to occur, which allows the formulation of a grammaticalisation scale on theoretical grounds. This predicted internal order of the scale is also sketched in section 3.2.

The sample study has produced a number of insights, which either confirm predictions and existing conceptualisations or call for their revision. The sample systems confirm the existence of a continuum of nominal classification, which hosts systems that are located at the poles of the continuum as well as a broad range of systems at various intermediate stages of grammaticalisation. The majority of systems shows a mixed setup of classifier-like and concordial
class-like properties, while only six systems display a fully concordial class-like setup of properties, and only one system displays a fully classifier-like setup of properties. Considering that grammaticalisation is assumed to be the main vehicle for the development of nominal classification systems and that prototypical classifier systems and concordial class systems are the extreme points of the continuum, a low correspondence of ‘real world’ systems to the linguistic ideal may be expected. The sample displays a clear grammaticalisation cline, if one orders the systems with respect to an increase in grammaticalised properties.

While this observation is clearly in favour of the grammaticalisation hypothesis, two issues require further consideration: First, the predicted preferred order of grammaticalising properties is not found back in the sample; for instance, agreement is predicted to occur at a later developmental stage, as it depends in most cases on the development of affixal markers. The loss of morphosyntactic independence and subsequent morphologisation of grammaticalising linguistic elements is generally observed at a later stage than e.g. its semantic bleaching and decategorisation. In contrast, the sample study shows the systems’ morphosyntactic independence to decrease at an earlier stage in many cases. Second, ten less grammaticalised classifier-like systems deviate from the order of grammaticalising properties that is found within the sample. These deviations form two coherent groups within the sample; they show a concordial class-like character of their assignment, locus operandi, and classificatory flexibility, while their remaining properties, which are the first ones to grammaticalise in the rest of the sample, are classifier-like.

It is possible to describe these deviations and to provide a rationale for their occurrence, yet such a large portion of deviations allows to question the accuracy of the conceptualisation of the grammaticalisation hypothesis. If one looks at the data, two separate developmental trends become apparent: Those systems that show a ‘regular’ grammaticalisation cline correspond to a grammaticalisation of their means of formal expression, i.e. their class inventory decreases in size, classification becomes obligatory, and the class markers occur multiply and outside of the borders of the noun phrase. I use the term “formalisation” when referring to this form of grammaticalisation. Those systems that deviate from this grammaticalisation cline correspond to a grammaticalisation in their semantic domain, i.e. in their class assignment, which becomes rigid, their classificatory flexibility, which decreases, and their decreasing potential to semantically contribute to the noun phrase. I use the term “conventionalisation” when referring to this form of grammaticalisation.

Formalisation and conventionalisation together constitute a ‘full’ or ‘unitary’ grammaticalisation process. A ‘unitary’ grammaticalisation approach assumes the parallel loss of morphosyntactic independence, semanticity, transparency, and involvement of the speaker. A ‘split’ grammaticalisation approach as it is introduced in the chapters 7 and 8 is novel to the study of nominal classification. A separate account for formalisation and conventionalisation in nominal classification systems leads to much more homogenous results than a ‘unitary’ one: Chapter 7 shows a strong preference for classifier-like systems
9.1. The Grammaticalisation Hypothesis

to not conventionalise to a greater degree, while varying a lot with respect to their degree of formalisation, and for concordial class-like systems to uniformly display a highly formalised setup, while greatly varying with respect to their degree of conventionalisation. This means that classifier-like systems are prone to shifts and cross-linguistic variation in their formal expression, while their semanticity and interactivity is largely maintained, and concordial class-like systems are prone to shifts and cross-linguistic variation in their semanticity, while their formal expression is fairly homogenous.

While classifier-like systems and concordial class-like systems are fairly evenly spread along the full formalisation scale, there is a notable gap on the formalisation axis between the two types. This gap spans over more than 15 per cent of the scale (cf. section 7.3.3). What consequences does this have for the grammaticalisation hypothesis? In a nutshell, the conceptualisation of grammaticalising systems can be revised, and the scenario that can be derived from the sample study suggests that systems have to formalise to a high degree before they are able to conventionalise. It is indeed technically possible for a classifier-like system to develop into a concordial class-like system, yet the probability of this actually happening is fairly low. The reason for the improbability lies in the grammatical requirements that come with a concordial class-like system: It is not sufficient for a classifier-like system to semantically bleach and develop bound markers, but the system must turn obligatory in all contexts and mark class membership by means of agreement. This entails grammatical changes that reach far beyond nominal classification. For instance, many classifier-like systems are found in isolating languages of Asia; it is unlikely that these systems turn into concordial class-like systems, since the grammar would have to introduce affixed agreement markers and thus change the language type. While this is not impossible, a grammar is more likely to alter or abandon structures that it already possesses instead of altering its most basic parameters. What can be observed instead in these languages is a continuous process of conventionalisation and semantic remotivation, where individual classifiers undergo grammaticalisation, eventually drop out of the system, and give way to newly coined classifiers.

In summary, the grammaticalisation hypothesis is highly unlikely to occur, and a ‘missing link’ between classifier-like and concordial class-like systems could not be found. A typological account for nominal classification gains in precision and accuracy of the description of nominal classification devices if it distinguishes the two dimensions of formalisation and conventionalisation, which are introduced in the chapters 7 and 8. Section 9.2 presents the implications of the two-dimensional account for formalisation and conventionalisation in nominal classification in general.
9.2 Nominal Classification

This section summarises the findings of this study that do not specifically relate to the grammaticalisation hypothesis, but to nominal classification in general. Recall that nominal classification in the notion of this study does not include the concept of ‘verb classifiers’, ‘relational classifiers’, deictic classifiers, and locative classifiers, as has been laid out in section 2.3 and Passer (2015, 2016).

The brief survey of the functional typology of nominal classification systems in section 4.3 has shown that there are no system-specific patterns of functions except for numeral classifier systems being involved in individuation. Furthermore, there is a strong tendency for all types of systems to be involved in reference tracking and referent differentiation by classificatory means. The majority of systems employs both functions simultaneously, and there is no sample system that serves neither function. Functions other than reference tracking and referent differentiation may be entertained additionally.

I have shown in this study that it is feasible to maintain the notion of both classifier systems and concordial class systems and to define a prototypical ideal for both types, despite the large formal diversity of classifier-like systems. Furthermore, the study of the supposed ‘hybrid systems’ of Nasioi, Bora, Mundurukú, and Ngan’gityemerri has shown that these actually are multiple classifier systems, and that they indeed represent the most likely candidate systems to develop into the direction of concordial class systems.

The six criteria underlying the definition of prototypical classifier systems and concordial class systems are well-established in earlier descriptive investigations of nominal classification and apply to both systemic types in their classifier-like and concordial class-like form respectively. Not all criteria are equally expressive; the sample data has shown that there are three core criteria, which consistently correlate with the respective systemic prototype almost without exception. These are the obligatoriness of classification, the potential to feature a classificatory marker multiple times in the same context of classification, and the potential to feature a classificatory marker outside of the borders of the noun phrase. In their classifier-like form, these properties are found exclusively in classifier-like systems, and in their concordial class-like form only in concordial class-like systems (cf. table 9.1). The only exception to this observation is found in the overall classifier-like systems of Nasioi, Bora, Mundurukú, and Ngan’gityemerri, which allow their markers to attach e.g. to the verb, and thus operate both NP-internally and -externally.

Non-core criteria are the flexibility of classification and the either systemsided or speaker-sided class assignment. They apply well to the respective prototype in their classifier-like and concordial class-like form, but do not show a correlation as strong as the criteria and thus are more prone to exceptions (cf. table 9.1). The last criterion is the size of a system’s class inventory. It does not apply to the notion of core criteria or non-core criteria, but can be used as
an indirect indication for a system’s type: All concordial class-like systems and the majority of classifier-like systems employ a small class inventory. Only the systems of Mandarin, Korean, Nasiol, Bora, and Mundurukú display a large class inventory (cf. table 9.1). Therefore, a given system’s inventory size is only an explanatory property for the identification of this system’s type, if it is a large one.

The definition of core criteria and non-core criteria makes it possible to answer the question about the quality of the grey area of intermediate systems, which are located within the continuum between clearly classifier-like and clearly concordial class-like systems. In its widest notion, the grey area consists of all systems that deviate in one or more properties from either prototypical system. This is certainly not a fruitful definition, as the bulk of systems falls under this notion, and yet the majority of these systems displays a recognisable classifier-like or concordial class-like quality. Depending on one’s research aims and interests, the grey area can be defined in a wider or narrower sense on basis of the distinction of core and non-core properties: Systems that deviate in one or more core properties from their respective systemic prototype are first order deviations. First order deviations represent the narrowest form of
the grey area. Systems that deviate in their assignment and/or classificatory flexibility from their respective systemic prototype are second order deviations. First and second order deviations comprise a wider form of the grey area. Third order deviations are those systems which display an overall prototypical setup of properties, but employ one or more criteria that is neither classifier-like nor concordial class-like, but both. This is e.g. the case in systems that allow a speaker-sided class assignment for parts of its nominal lexicon and a systems-sided class assignment for other parts of the nominal lexicon. First, second, and third order deviations together constitute the largest form of the grey area.

The sample study in chapter 7 and 8 has furthermore shown that a two-dimensional quantificational account for nominal classification, which distinguishes the dimensions of a system’s formal expression and its transparency, yields much more precise and expressive results than a ‘unitary’ account. Its application to the sample clearly demonstrates that both systemic prototypes show a mutual distribution along both dimensions’ axes: Classifier-like systems allow a great deal of variation with respect to their formal expression’s degree of grammaticalisation, but show little variation with respect to their transparency, which is consistently high. Concordial class-like systems allow a great deal of variation with respect to their systemic transparency, while their formal properties are consistently highly grammaticalised. This distribution of variation in the two dimensions confirms the feasibility of a distinction of two basic types of nominal classification systems.

There are four possible developmental pathways for nominal classification systems: First, a system may formalise, i.e. its means of formal expression may undergo grammaticalisation. This is the case for the systems of Nasioi, Bora, Mundurukú, and Ngan’gityemerri. Second, a system may alter its degree of transparency, i.e. it may conventionalise and lose in semantic accessibility and interactivity, or the other way round, i.e. it may gain in semantic accessibility and interactivity. Both these pathways may occur independent of each other or in combination. Section 8.2 exemplifies both the conventionalisation and remotivation of the system of Dyirbal. Third, a system may acquire additional functions outside of nominal classification. This is the case in Great Andamanese, where the class markers serve a nominal classification function (i.e. the manipulation or specification of nominal semantics) in combination with nouns and adjectives, whereas they serve a verbal classification function (i.e. the manipulation or specification of verbal semantics) in combination with verbs and adverbs. A system that experiences a functional extension may maintain its original nominal classification function, or it may abandon it in favour of the newly acquired functionality. Finally, a system may decay and lose its functionality. In this case, unproductive remnants of the earlier functional system may remain visible. This is the case in Tlingit and Chukchi, where modifiers of nouns with human referents carry an additional marker; the marking itself however does not display a recognisable function of nominal classification.

In conclusion, this study has shown that nominal classification, as diverse and rich a phenomenon it is, shows a binary structure of formally diverse
classifier-like systems and concordial class-like systems that are highly varied with respect to their semantic transparency. The phenomenon as a whole can be segmented into these two types, and that the so far unspecified notion of a grey area, which consists of systems of an intermediate type, distinguishes three orders of deviations. Moreover, a two-dimensional account for nominal classification, as it is put forward in chapters 7 and 8, captures all sorts of nominal classification systems comprehensively and allows a broad synchronic and diachronic comparability. Section 9.3 concludes this study by addressing aspects and questions that are promising for future research.

9.3 Outlook

The present study proposes an answer to a variety of questions that aim at a clarification of the internal structure of the continuum of nominal classification. There are various aspects of nominal classification which remain promising subjects for further study, and which may deepen our knowledge and understanding of this linguistic phenomenon. This section sketches the areas that appear fruitful for future research.

The two-dimensional mapping of nominal classification has shown that systems are generally fairly transparent, unless they are formalised to a high degree. In this case, they may conventionalise and grow largely idiosyncratic. I have hypothesised that this dependency on a semantically transparent underpinning of the majority of system is due to a system’s visibility in discourse: Only highly formalised systems allow classificatory marking on multiple elements, which makes these systems more visible than ones that feature a single marker per context. This higher visibility exposes enough cues to maintain the formal side of a classification system, and allows the semantic underpinning to erode. This explanation is a conjecture made on basis of observations from the sample data. A comparative in-depth study of the discourse behaviour of nominal classification systems not only promises to shed more light onto the relation of systems’ means of formal expression and their underlying classification apparatus, but is also vital for a better understanding of the yet understudied discourse-functional richness of nominal classification systems.

The functionality of systems in general is a largely understudied aspect of nominal classification. A broad comparative functional analysis of systems of nominal classification may gain us valuable insights into the range of functions that are employed by individual systems and cross-linguistically. Furthermore, nominal classification systems are reported to consistently employ multiple functions, which are not all exploited equally. Functional shifts within a system may be due to semantic reanalysis or formal developments, and can push a system beyond the scope of nominal classification (as in Great Andamanese). An investigation of the shifts and variation of functionality in nominal classification devices may thus give us more insight into the nature of nominal classification as well as its boundaries.
Furthermore, factors such as the speakers’ proficiency and the prestige of a given language variety have been shown to play an important role in the re-analysis and development of the diachrony of the Dyirbal system. These factors may apply in individual cases of language change, and are to my knowledge not studied in detail, except for Dyirbal. The study of further systems' reorganisation under the influence of a genetically unrelated superstratum may isolate and identify systemic aspects that are crucial and maintained in a system even throughout comprehensive changes, and ones that are prone to drop out or alter.

Beyond the investigation of specific phenomena such as the visibility in discourse, the speaker proficiency, and the prestige of a variety that undergoes change, the two-dimensional mapping instrument itself can be further studied and refined. The properties that are suggested in the chapters 7 and 8 to render the dimension of formalisation and transparency are by no means exhaustive or exclusive. They were selected here, because earlier research has shown that they play a crucial role in nominal classification, and that they can be used reliably in both descriptive and quantitative studies of nominal classification. However, the individual properties as well as their measure points as they are put forward here may be further tested to confirm their validity. The mapping has proved to generate precise and expressive results and to make a fine-grained cross-linguistic comparative study of nominal classification possible, yet I do not exclude the possibility that both the properties and their measure points are in need of further refinement, expansion, and/or weighing.
Appendix

The appendix contains the dossiers of the individual sample languages. It is available on the data carrier that is attached to the print version of this thesis and under [http://www.lotpublications.nl](http://www.lotpublications.nl). Each dossier describes the nominal classification system of the respective language and provides an account for its functional typology, its formal properties, and its dimensions of formalisation and transparency.

Introduction

This appendix contains the individual dossiers on the classification systems from the language sample. The sampling method and composition as well as the structure of the dossiers is described in detail in the study that this document is attached to. This introduction summarises the basic information about the sample and the dossiers.

The sampling method at use in this study is extensively described in Rijkhoff et al. (1993); it aims at the composition of a variety sample, i.e. a language sample that considers both a maximum genetic divergence and a maximum typological divergence. The ideal sample size of 40 languages is reduced to an actual sample size of 32 languages, since a number of basic phyla or language isolates (e.g. Basque, Etruscan, and Hurrian) are either too sparsely documented or do not feature a nominal classification system in their grammar. As Dongo-Ko employs both a possessive classifier system and a concordial class system, the number of sample systems is 33. Figure [] lists the sample languages and their respective systemic type according to this study’s analysis.

The dossiers are compiled on basis of language descriptions, grammars, and other secondary literature. Wherever possible, native speakers and/or specialists on the respective language were consulted to confirm the validity and accuracy of the presented data and analyses. Information that turned out to be irrelevant for this study was removed before publication. Each dossier is
organised in the same way in four sections in order to guarantee an easy comparability: Section 1 provides a general background for the language in question (its genetic affiliation according to Ruhlen (1987) and Lewis et al. (2013), its geographic location, its type of classification system, and the main sources at use). Section 2 addresses the possibly employed systemic types; it excludes those that are not present and exemplifies those that are present. Section 3 contains the analyses of the system(s) in question: First, their functionality is accounted for by a functional typology (adopted from Contini-Morava and Kilarski 2013 and adjusted for the purposes of this study). Second, their formal properties are accounted for by a package of properties that is adopted from Dixon (1982, 1986). Third, the mapping instrument that is presented in this study and based on the studies of Dixon (1982, 1986), Corbett and Fedden (e.g. 2015), and Audring (under review), is applied to the system(s) in question. Section 4 concludes with a brief summary of the observations made and the respective system’s degree of grammaticalisation.

Figure 1: Sample coverage of the terminal subtypes of nominal classification systems.
1 Bora

1.1 Background

Bora (or Miraña) is spoken in the North West Amazon area, which comprises areas in Southern Colombia, Peru, and Brazil (cf. Lewis et al. 2013: s.v.). It displays a “fairly polysynthetic, agglutinating, and suffixing profile” (Seifart 2007: 414). The following dossier is mainly based on Payne (1990), Seifart (2005 and 2007), Seifart and Payne (2007), and Weber (2002). The Bora system of nominal categorisation is an instance of a supposed ‘hybrid’ system that displays an overall classifier-like character, while it seemingly employs agreement. This study analyses the supposed agreement in Bora as an instance of multiple, adjacent contexts of classification. As the classificatory markers may combine with a wide range of nominal modifiers, this renders Bora a multiple classifier system.

nominal classification devices

CLSs

CCSs

noun

numeral

possession

multiple

CLS

Figure 2: Classification of the Bora system of nominal classification.

1.2 Classification Typology

1.2.1 Noun Classifiers

Bora employs a nominal classification system. Its markers attach among others to the noun (cf. [1]). The system however is a multiple classifier system, as the marking does not occur in the form of agreement. The Bora classifiers are thus multiple classifiers.

(1) A ńki
banana
“banana(s) (fruits, plants, etc.)”

B ńki
banana

CT:LONG.OBJ/CL:POINTED

“a banana fruit, a banana plant”
1.2.2 Numeral Classifiers

Bora employs a nominal classification system. Its markers attach among others to numerals (cf. 2). The system however is a multiple classifier system, as the marking does not occur in the form of agreement. The Bora classifiers are thus multiple classifiers.

(2) A tsá -连接 (áhi -连接) one CL:BUNCH banana CL:BUNCH
“one (banana) bunch”

B mìi -连接 (áhi -连接 -ku) two CL:BUNCH DL banana CL:BUNCH DL
“two (banana) bunches”

C ma:kíní (-连接 -ßa) (áhi -连接 -ne) three CL:BUNCH PL banana CL:BUNCH PL
“three (bunches) (of banana)”

D tsan´ndééjííkátst (-连接 -ßa) (áhi -连接 -ne) four CL:BUNCH PL banana CL:BUNCH PL
“four (bunches) (of banana)”

E tsá -连接 -ßa) (áhi -连接 -ne) one REPEATER.hand CL:BUNCH PL banana CL:BUNCH PL
“five (bunches) (of banana)” (literally: “one hand of banana”; Seifart 2005: 130f.)

1.2.3 Possessive Classifiers

Bora employs a nominal classification system. Its markers attach among others to the possessive (cf. 3). The system however is a multiple classifier system, as the marking does not occur in the form of agreement. The Bora classifiers are thus multiple classifiers.

(3) A táj -na?be -muy kábó?kó -hkatsí POSS.1SG brother CL:ANIM.PL beat RECIPROCAL -连接 PREDICATE
“My brothers hit each other.” (Bora; cf. Seifart 2005: 59)

B táy?ne (-gwa), dí?n?e (-gwa), my CL:2D.STRAIGHT your CL:2D.STRAIGHT his/her/their (-gwa), n?e?n?e (-gwa) CL:2D.STRAIGHT our CL:2D.STRAIGHT
“mine, yours, his/her/their/s, ours (respectively machete, plank, etc.)” (Bora; cf. Seifart 2005: 128)
1.2.4 Multiple Classifier Systems

Bora employs a nominal classification system. Its markers attach to a variety of morphosyntactic hosts (cf. 1, 2, and 3). The system however is a multiple classifier system, as the marking does not occur in the form of agreement. The Bora classifiers are thus multiple classifiers. It is analysed in section 1.3.

1.2.5 Concordial Class Systems

Bora employs a nominal classification system. Its markers attach to a variety of morphosyntactic hosts (cf. 1, 2, and 3). The system however is a multiple classifier system, as the marking does not occur in the form of agreement; cf. 4 for an example of variable class ‘disagreement’. The Bora classifiers are thus multiple classifiers.

(4) A e:-hi māhu:-hi
DISTANT-CL:2D.ROUND be.big.SUB-CL:2D.ROUND
kū:mmu:-hi
turtle-CL:2D.ROUND
“that big turtle”

B e:-hi māhu:-be kū:mmu:-hi
“that big turtle”

C aj:-di māhu:-hi kū:mmu:-hi
“that big turtle”

D aj:-di māhu:-be kū:mmu:-hi
“that big turtle” (Bora; cf. Seifart 2005: 169)

1.3 Analysis

1.3.1 Functional Setup

Table 2: The semantic and discourse properties of the Bora system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Classifier</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating</td>
<td>C₁</td>
<td>productive derivation, variable classification</td>
</tr>
<tr>
<td>discriminative synonyms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>individuation</td>
<td>C₂</td>
<td>unitisation for countability</td>
</tr>
<tr>
<td>attribution of properties</td>
<td>C₃</td>
<td>speaker-sided values</td>
</tr>
<tr>
<td>reference identification</td>
<td>C₄</td>
<td>reference tracking via AGR, anaphora</td>
</tr>
</tbody>
</table>

**ad C₁**: The interpretation of classifiers being employed in forming noun stems goes as far as to treat them as structurally bound nouns (cf. Weber 2002 and 5).
It is further possible to mark animate and inanimate nouns and their satellites with various class markers for the purpose of referent differentiation (cf. 6).

\begin{itemize}
\item A \textit{muuts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts \textit{-pá} pear.apple \text{CL:FRUIT} “pear apple fruit”
\item B \textit{muuts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts \textit{-pá} pear.apple \text{CL:TREE} “pear apple tree”
\item C \textit{muuts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts \textit{-páh} pear.apple \text{CL:GROVE} “pear apple grove” (Bora; cf. Weber 2002: 3)
\item D \textit{muuts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts\textsuperscript{b}ts \textit{-pá} \textit{P} pear.apple \text{CL:TREE-\textit{P}á:m} \text{CL:LEAF} “leaf of pear apple tree” (Bora; cf. Weber 2002: 5)
\end{itemize}

\textbf{ad C\textsubscript{2}}: The Bora classificatory markers serves the function of individualisation as we find it in numeral classifier systems (cf. Seifart 2005: 233 and 1 and 2).

\textbf{ad C\textsubscript{3}}: There is no information available that the Bora classificatory markers are used to attribute affective values on a referent.

\textbf{ad C\textsubscript{4}}: Speakers of Bora “use class markers both as their primary means to identify the task objects in first mentions and to anaphorically refer back to them in later mentions” (Seifart 2005: 190; cf. 7 for an example of reference tracking, which includes variable classification). The class markers occur in both predicative and referential function (ib.: 193).

\begin{itemize}
\item A \textit{ixi\textsuperscript{h}ú-pé te-ne tágci-hau yesterday-PST 3SG-CL:INANIMATE blowgun-CL:TUBE 1SG.SUBORDINATE rasp-PURPOSE “Yesterday I rasped the [inanimate] blowgun, […]”
\item B […] \textit{a-ne} \textit{ó} nuhts\textsuperscript{a}kau-\textit{ʔ} ni u\textit{b}\textit{u}\textit{ʔ}k\textit{á} good […]” “[…] and I tried it [inanimate] out: good […]”
\item C […] \textit{tétsi\textsuperscript{t}á:s\textsuperscript{d}áu} \textit{á:báká-hpi:-ké} \textit{ó} then-TENSE.ASPECT.MOOD owner-CL:M.SG-ACC 1SG \textit{áhkú\textsuperscript{a}:ko-\textit{ʔ} t\textit{é}:-\textit{hun-\textit{b}au} \textit{give-PF:FUT.PREDICATIVE} 3SG-CL:TUBE-ALLATIVE “[…] and then, indeed, I will give it [tube-shaped] to its owner.” (Bora; cf. Seifart 2005: 259 and 2007: 426)
\end{itemize}
1.3.2 Formal Setup

Table 3: The formal properties of the nominal classification system of Bora.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong> $(F_1)$</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility</strong> $(F_2)$</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td><strong>inventory</strong> $(F_3)$</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td><strong>marking</strong> $(F_4)$</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi</strong> $(F_5)$</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity</strong> $(F_6)$</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td><strong>assignment</strong> $(F_7)$</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

**ad F_1:** Not all nouns are obligatorily classified in Bora (cf. Seifart 2005: 113). Unclassified nouns are non-countables and may optionally be “unitised” by classifiers in order to be pluralised and countable (cf. [8]).

(8)  
A  úñbi, pìːka  
basket  manioc  
"basket(s), manioc"
B  *úñbi -ːne, *pìːka -ːne  
basket  PL  manioc  PL  
"baskets, manioc tubers/plants/seeds"
C  úñbi -ːbaj, pìːka -ːu  
basket  CL:CONTAINER  manioc  CL:3D.ROUND  
"(a) basket, manioc tuber"
D  úñbi -ːbaj -ːne, pìːka -ːu -ːne  
basket  CL:CONTAINER  PL  manioc  CL:3D.ROUND  PL  
"baskets, manioc tubers" (Bora; cf. Seifart 2005: 113f.)

**ad F_2:** It is possible to mark animate and inanimate nouns and their satellites with various class markers or with a mixed choice of classifiers (Seifart 2005: 169; cf. [4]).
ad F₃: Bora employs 72 class markers (cf. Seifart 2005: 86ff. and table 4 for a list of selected markers) and a range of 53 repeaters (Seifart 2005: 96ff.), which each represent an own referential class. Therefore, the class inventory of the Bora system of nominal categorisation is a large one.

Table 4: Selected polysyllabic specific classifiers of Bora (cf. Seifart 2005: 91ff.).

<table>
<thead>
<tr>
<th>classifier</th>
<th>glossing</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>-boh</td>
<td>POND</td>
<td>“stationary pond of liquid”</td>
</tr>
<tr>
<td>-dgi:hu</td>
<td>POWDER</td>
<td>“ash, powder”</td>
</tr>
<tr>
<td>-gya:jko</td>
<td>HOOK</td>
<td>“hook”</td>
</tr>
<tr>
<td>-gwa:j:te</td>
<td>CRUMBS</td>
<td>“crumbs”</td>
</tr>
<tr>
<td>-hpa:jko</td>
<td>LIQUID</td>
<td>“liquid”</td>
</tr>
<tr>
<td>-htoi</td>
<td>CURVE</td>
<td>“curve”</td>
</tr>
<tr>
<td>-hu:?:o</td>
<td>PALMLEAF</td>
<td>“palm leaf”</td>
</tr>
<tr>
<td>-kaha</td>
<td>CREEK</td>
<td>“little creek”</td>
</tr>
<tr>
<td>-nakah</td>
<td>SIDE</td>
<td>“one side”</td>
</tr>
<tr>
<td>-mi:?:o</td>
<td>HARD.SHELL</td>
<td>“hard shells”</td>
</tr>
<tr>
<td>-padgi</td>
<td>GARDEN</td>
<td>“garden (archaic)”</td>
</tr>
<tr>
<td>-pohi:si</td>
<td>RING</td>
<td>“roll, ring”</td>
</tr>
<tr>
<td>-ra:ra</td>
<td>BROKEN</td>
<td>“broken, rotten”</td>
</tr>
<tr>
<td>-ra:ra:aj</td>
<td>GRAINS</td>
<td>“grains”</td>
</tr>
<tr>
<td>-fiu:ru:</td>
<td>CHUNK</td>
<td>“thick and short chunk”</td>
</tr>
<tr>
<td>-ra:tsi</td>
<td>CLEARING</td>
<td>“clearing”</td>
</tr>
<tr>
<td>-pa:mi</td>
<td>LEAF</td>
<td>“thin and flexible”</td>
</tr>
<tr>
<td>?babaj</td>
<td>BAG</td>
<td>“bundle, bag”</td>
</tr>
<tr>
<td>-be:guwa</td>
<td>STACK</td>
<td>“stack of thin objects”</td>
</tr>
<tr>
<td>-da:ti</td>
<td>PIECE</td>
<td>“small piece”</td>
</tr>
<tr>
<td>-he:hu</td>
<td>HOLE</td>
<td>“hole”</td>
</tr>
<tr>
<td>-?a:ba</td>
<td>SMALL.PALMTREE</td>
<td>“small palm tree”</td>
</tr>
<tr>
<td>-?a:ba</td>
<td>GABLE</td>
<td>“gable of a roof”</td>
</tr>
<tr>
<td>-?a:guwa</td>
<td>DOORWAY</td>
<td>“opening”</td>
</tr>
</tbody>
</table>

ad F₄: A number of properties suggests that Bora class markers do not merely copy information from the head noun, but contribute semantically to the noun phrase. First, there are cases of variable class ‘disagreement’ (cf. 4), which is not present in concordial class systems. Second, the class markers are used as referential elements themselves (cf. 5), which is not possible for agreement markers. In addition, there is no over linguistic element that the markers could copy information from. Third, classifier ‘stacking’ is possible in Bora (cf. 1 D and 6 D); in these cases, both markers contribute semantically to their host noun phrase. This is again not possible in concordial class systems.

ad F₅: Since Bora class markers may occur on nouns (cf. 6), nominal modifiers (cf. 4), and main clause predicates (cf. 9), class membership is marked NP-internal and -external (cf. 5).
ad F₆: Bora class markers may contribute to their noun phrase’s semantics; if this is the case, the noun phrase’s semantics is composed out of the combination of class marker and NP-head (cf. Seifart 2005: 200 and 10). There are non-compositional classified nouns as well, the semantics of which are not recognisably influenced by the class marker in use (ib.; cf. 11).

(10) A tākī-ʔa
banana-CL:3D.LONG.OBJ
“banana fruit”
B лэ-хэ-ʔa
till.NOMINALISED-CL:ENCLOSURE
“bed for horticulture”
C мэбко-ʔa
corral-CL:ENCLOSURE
“corral/room” (Bora; cf. Seifart 2005: 195)
D тэбэ-ʔa
shoot.NOMINALISED-CL:LITTLE.STICK
“arrow” (Bora; cf. Seifart 2005: 203)

ad F₇: The class assignment of the Bora nouns to either general or specific classes is almost fully semantically motivated (Seifart 2005: 220). Exceptions are animal names, the assignment of which is always opaque, and “a minority of inanimate classified nouns that are built with seven class markers from the core set of specific class markers” (ib.). Cf. also 11 for a number of opaque class assignment.
1.3.3 Mapping

Figure 3: Bora.

Table 5: The formalisation and transparency value of Bora.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obrigatoriness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td>x-value</td>
<td>0.542</td>
<td>y-value</td>
<td>0.726</td>
</tr>
</tbody>
</table>

**ad X₁:** The Bora system possesses 125 classes (Seifart 2005: 86ff. and 96 ff.), which makes it a large system.

**ad X₂:** Bora knows a large number of targets of classificatory marking, including nouns, adjectives, pronouns (demonstratives, interrogatives possessives, relatives), numerals, and the verb (Seifart 2005).

**ad X₃:** The class markers occur in constructions that host the classified noun (cf. [7] A), in noun phrases that lack a classified noun (cf. [7] B), and on NP-external elements such as the predicate (cf. [9]).
ad X₄: The classification is non-obligatory in that not all nouns in all contexts are classified (cf. 12 where classificatory marking is omitted altogether). However, a restricted set of hosts is obligatorily marked for a noun’s class in discourse, for instance demonstratives (cf. Seifart 2005: 3).

(12) táj -nādhe -tābū =pe
POSS.1SG brother DECEASED PST
“my late brother” (Bora; cf. Seifart 2005: 57)

ad X₅: The markers take the form of affixes (cf. table 4).

ad X₆: This study argues that seeming cases of class agreement such as in 4 represent adjacent noun phrases, which each feature a single classifier. This analysis is confirmed by Seifart (2005: 5): “[T]wo coreferential nominal expressions in a clause are typically in a relation of loose apposition, rather than forming a tightly integrated constituent[...]”.

ad X₇: Not all nouns can be classified: Non-countables for instance generally remain unclassified (Seifart 2005: 113).

ad Y₁: The Bora classes are semantically bijective (cf. table 4).

ad Y₂: Class assignment in Bora is prevailingly semantic; only a smaller number of nouns (e.g. animal names) are reported to show an opaque class assignment (cf. 11).

ad Y₃: The high number of classes results in a high number of assignment rules.

ad Y₄: The Bora class markers express class information only (cf. table 4).

ad Y₅: Each Bora class has its own designated marker (cf. table 4).

ad Y₆: Bora class marking can be informative (cf. 6 where the class marker alters the noun phrase’s semantics and markers can even be stacked to further specify the nominal referent) or redundant (cf. 7 where a class reassignment does not alter the semantics).

ad Y₇: Nouns can take different class markers (cf. 7).
1.4 Grammaticalisation

A comparison of the Witotoan family and neighbouring languages shows that the Bora system of nominal categorisation can be traced back to earlier stages (at least to Proto-Bora-Muinane; cf. Seifart 2007: 427ff. and 434ff.) in its core. The Proto-Bora-Muinane system of nominal categorisation and the one of its daughters display a higher degree of formalism as compared to its sister Proto-Witoto-Ocaina (cf. e.g. Seifart 2007: 434ff.). The varying specificity of the individual Bora class markers indicates that the class markers are of different age: The more specific a marker is in its meaning and the more transparent its relation to a fully nominal counterpart is, the more likely it is that it entered the class inventory more recently than other classifiers.

Bora shows a medium degree of formalisation: On the one hand, it possesses a large inventory, and the system is not able to classify the full nominal lexicon. Together with the single marking, these properties are a sign of a low degree of formalisation. On the other hand, the bound markers attach to a broad range of nominal satellites, and the classification is obligatory in a number of morphosyntactic contexts, which is a sign of a higher degree of formalisation. The system is fairly transparent, despite the high number of assignment rules: Its classes are semantically specific and straightforward, the majority of nouns is flexibly assigned on a semantic basis, the markers are categorially independent and each denote a single class. The semantic information conveyed by the markers is productive and informative in many cases, but it can also be redundant.

2 Boumaa Fijian

2.1 Background

Boumaa Fijian is “a dialect of the eastern [Fijian; MP] language that is mutually intelligible with Standard Fijian” (Dixon 1988: 1). In this dossier, reference to Boumaa Fijian is made by the label “Fijian”; if a reference to another dialect or Standard Fijian is made, this is indicated in situ. This dossier is mainly based upon four works on Fijian: Dixon 1988, Lichtenberk 1983 and 2009, and Schütz 1985. The Fijian possessive classifier system (which is typically referred to as a relational classifier system) behaves in most aspects according to the expectations of a classifier system; only the opaque origin of its markers
and the small class inventory are properties that do not correspond to the definition of a prototypical classifier system.

nominal classification devices

CLBs

CCBs

noun
numeral
possessive
multiple

Figure 4: Classification of the Boumaa Fijian system of nominal classification.

2.2 Classification Typology

2.2.1 Noun Classifiers

The Boumaa Fijian classifiers do not attach to the bare noun (cf. [13] where no classifier occurs), but occur in possessive contexts. Therefore, Boumaa Fijian does not employ a numeral classifier system.

(13) A a ulu.-ni.-vanua rewa dua.dua
    ART mountain high uniquely
    “the highest mountain” (Boumaa Fijian; cf. Dixon 1988: 118)

B na 'oro tuaraga mai Ositerela, o Canberra
    ART village chief at Australia ART Canberra
    “Canberra, the capital city of Australia” (Boumaa Fijian; cf. Dixon 1988: 184)

C sa maruma sava.savaa sara gaa o Aneta
    ASPECT.MARKER lady clean INTENSIFIER ART Aneta
    “Aneta is a very clean lady.” (Boumaa Fijian; cf. Dixon 1988: 66)

2.2.2 Numeral Classifiers

The Boumaa Fijian classifiers do not occur in quantificational contexts (cf. [21] where no classifier occurs), but in possessive contexts. Therefore, Boumaa Fijian does not employ a numeral classifier system.

(14) A e tolu a gone
    3SG three ART child
    “three children”

B a gone e tolu
    ART child 3SG three
    “three children” (Boumaa Fijian; cf. Dixon 1988: 144)
2.2.3 Possessive Classifiers

There are five morphosyntactically possible possessive constructions in Fijian (cf. Dixon 1988: 122), two of which contain classifiers:

1. constructions employing the possessive suffix -i (cf. [15])

   (15) a liga -i Jone
       ART hand POSS.MARKER John
   “John’s hand” (Boumaa Fijian; cf. Dixon 1988: 120)

2. constructions employing a classifier plus -i (cf. [16])

   (16) a waqona me- i Jone
       ART kava CL:DRINKABLE POSS.MARKER John
   “John’s kava (something drunk)” (Boumaa Fijian; cf. Dixon 1988: 120)

3. constructions employing a pronominal suffix (cf. [17])

   (17) na tama -na
       ART father 3SG.POSS
   “her father” (Standard Fijian; cf. Schütz 1985: 445)

4. constructions employing a classifier plus a possessor pronoun (cf. [18])

   (18) na me- na moli
       ART CL:DRINKABLE 3SG.POSS citrus
   “her citrus” (Standard Fijian; cf. Schütz 1985: 445)

5. constructions employing the ‘relator’ ni “of” (cf. [19])

   (19) a cauravou ni Boumaa
       ART youth of Boumaa
   “the youth(s) of Boumaa” (Boumaa Fijian; cf. Dixon 1988: 125)

Which possessive construction occurs depends on the kind of noun (personal/place name, pronoun, common noun with reference to a human/nonhuman/inanimate reference and bound or free noun; cf. Dixon 1988: 119 and table [7]). In case the possessor is referred to by a human noun, the constructions (3) in the case of a bound noun and (4) in the case of a free noun are expanded by a post-head possessor noun phrase, which is indicated in table [7] by an asterisk (cf. Dixon 1988: 120). If there is more than one construction possible, the most frequently used one is marked by an exclamation mark.

1 As Dixon (1988:124) points out, a glossing of ni as “associated with” corresponds closer with the Fijian semantics of this marker. Since the glossing of this marker does not affect the interests or outcomes of this dossier, the glossing “of” is maintained here.
Table 6: Distribution of possessive constructions in Boumaa Fijian (cf. Dixon 1988: 120).

<table>
<thead>
<tr>
<th>possessor</th>
<th>possessed: bound noun</th>
<th>possessed: free noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>personal/place name</td>
<td>(1)</td>
<td>(1) or (2!)</td>
</tr>
<tr>
<td>pronoun</td>
<td>(1) or (3)</td>
<td>(4)</td>
</tr>
<tr>
<td>human noun</td>
<td>(1), (3)!*, or (5)</td>
<td>(4)*</td>
</tr>
<tr>
<td>animate noun</td>
<td>(3) or (5!)</td>
<td>(4) or (5!)</td>
</tr>
<tr>
<td>inanimate noun</td>
<td>(5)</td>
<td>(5)</td>
</tr>
</tbody>
</table>

This system is described in section 2.3.

2.2.4 Multiple Classifier Systems

The Boumaa Fijian classifiers occur only in possessive contexts. Therefore, Boumaa Fijian does not employ a multiple classifier system.

2.2.5 Concordial Class Systems

The nominal classification system of Boumaa Fijian is neither obligatory across contexts nor marks a variety of nominal satellites for the noun’s class membership. In addition, nouns are assigned to a particular referential class by the speaker according to his or her discoursal demands (Dixon 1988: 135), which does not correspond to this paper’s definition of a concordial class system. Therefore, Boumaa Fijian does not employ a concordial class system.

2.3 Analysis

2.3.1 Functional Setup

Table 7: The semantic and discourse properties of the Boumaa Fijian system of nominal classification.

<table>
<thead>
<tr>
<th>differentiating referents (C₁)</th>
<th>productive derivation, variable classification</th>
<th>✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability</td>
<td>✗</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
<td>✗</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
<td>✗</td>
</tr>
</tbody>
</table>

ad C₁: The classifiers occur on the possessive, not the noun. Therefore, they do not take part in productive nominal derivation. However, the classifiers are semantically contributing to their host NP and thus serve the differentiation of referents by means of variable classification (cf. 20 from Standard Fijian).
Appendix

(20) A na me- ART DRINKABLE 1SG.POSS kava
“my kava (to be drunk)”
B na na- ART GENERAL 1SG.POSS kava
“my kava (that I grew or sell)” (Standard Fijian; cf. Lichtenberk 1983: 157f.)

ad C₂: The classifiers do not obligatorily occur in quantificational constructions (cf. again 21); this is only the case when possession is quantified (cf. 21). Since the classifier does not semantically take an active part in quantification, the Boumaa Fijian system does not serve individuation.

(21) e tini ‘a lima a levu ni we- irau ‘oli 3SG ten and five and number of GENERAL POSS.1DL dog
“We have fifteen dogs.” (literally: “the number of our dogs is fifteen”; Boumaa Fijian; cf. Dixon 1988: 143)

ad C₃: The Boumaa Fijian classifier inventory is semantically straightforward in entertaining the four categories [+edible], [+drinkable], [+general], and [+inalienable possession] (Dixon 1988: 135f.). Therefore, the system’s semantics does not allow the speaker to freely attribute values to an entity by classificatory means.

ad C₄: The classifiers occur in possessive constructions together with the possessed noun (or other parts of speech whose referent is possessed; cf. 22, where the possessed is referred to by an adjective) that is being classified. Therefore, the Boumaa Fijian classifier system does not serve reference tracking by means of nominal classification.

(22) o -na we-vuni.vuni CL:GENERAL 3SG.POSS soul.hide.hide=modest
“his/her modesty” (Boumaa Fijian; cf. Dixon 1988: 138)

2.3.2 Formal Setup

Table 8: The formal properties of the nominal classification system of Boumaa Fijian.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>obligatory assignment of each noun</td>
<td>x</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>single class membership</td>
<td>x</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>limited class number</td>
<td>√</td>
</tr>
</tbody>
</table>
Appendix

marking (F₁)  CLS-like: marked once ✓
               CCS-like: multiply marked ×

locus operandi (F₅)  CLS-like: locus operandi: NP-internal ✓
                    CCS-like: locus operandi: NP-internal and -external ×

semanticity (F₆)  CLS-like: semantic contribution ✓
                CCS-like: semantically vain ×

assignment (F₇)  CLS-like: chosen freely by the speaker ×
                CCS-like: mechanic assignment (semantic) ✓

ad F₁: Nominal classification is obligatory in possessive contexts only (cf. again 13).

ad F₂: Certain nouns can take different classifiers, depending on the context and its semantics (cf. again 20).

ad F₃: The classifier inventory is a smallish one with only four classes (cf. table 9).


<table>
<thead>
<tr>
<th>Not Consumed</th>
<th>Marker</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'e-/ke-</td>
<td>general possession (&quot;owned by the possessor&quot;)</td>
</tr>
<tr>
<td></td>
<td>'e-/ke-</td>
<td>inalienable (&quot;related to the possessor&quot;)</td>
</tr>
<tr>
<td>Consumed</td>
<td>'e-/ke-</td>
<td>edible/chewable/smokable</td>
</tr>
<tr>
<td></td>
<td>'e-/ke-</td>
<td>drinkable/suckable/llickable</td>
</tr>
</tbody>
</table>

ad F₄: The classifiers occur once in the context of classification, not multiply (cf. again 20).

Note that the Boumaa Fijian inalienable and edible category are marked by the same classifier. Dixon (1988: 137) notes that "there appears to be no connection between them; the two 'e- may conceivably have different historical origins". Even if they are etymologically no distinct forms and we are dealing with an actual tripartite system the " edible" category of which has semantically expanded in a way that we cannot immediately comprehend, this does not pose a direct problem for its analysis. A possible classificatory parallel can be found in Mota (Vanuatu; ISO 639-3: mtt), where the possessive classifier "ga denoted close belongings and edible possessions" (Franjieh in prep.: 2). There also has been much discussion about the semantic analysis of especially 'exotic' languages (from a Western point of view). Culturally very specific conceptions are not unusual: For instance, traditional articles for clothing in Manam are "possessed inalienably when being actually worn, but are otherwise possessed generally" Lichtenberk (1983: 159), and the same is found in Kilivila (Senft 1986: 54), which clearly is a culture-specific view on clothing. Lichtenberk (1983: 174) further argues that a researcher’s analysis is necessarily an outsider’s interpretation that “need[s] not fully coincide with that of the speakers of the language, and that some items that to him are exceptional would no longer appear as such if he were able to view the world through the eyes of the members of that culture".
ad F5: The classifiers occur within the NP where the classification takes place (cf. 23).

(23) au aa rai -ca [a o -na rubi -ci Mere
1SG PAST see TR ART CL:GENERAL 3SG.POSS hit TR Mary
a cauravou yai]
ART youth this

“I saw this youth’s hitting Mary (i.e. the hiding he gave Mary)” (Boumaa Fijian; cf. Dixon 1988: 139)

ad F6: The classifiers carry a semantic potential which they add to the classifiers (cf. again 20). The lexical groups of “body parts, kin terms, and parts of a whole [ … ] can be said to participate in inalienable relationships” (Schütz 1985: 448). The semantic basis for the differentiation of the manner of consumption is most likely a “change of state as it is being consumed; e.g. food is bitten and chewed before being swallowed, tobacco is burnt and only the smoke is inhaled” (Dixon 1988: 136) and therefore marked by me-, whereas liquids and several other goods of consumption (e.g. pills; cf. ib.) that are not changing their physical state during consumption are marked by ‘e-.

ad F7: The semantic potential of the classifiers prevents a free assignment of classifiers, but allows a speaker-sided assignment. For instance, artifacts are by nature not edible or drinkable and thus are not a natural member of the category [+edible] or [+drinkable]. Similarly, body parts, kin terms, and parts of a whole are considered inalienable (Schütz 1985: 448) and thus are not assigned to another category.
2.3.3 Mapping

Figure 5: Boumaa Fijian.

Table 10: The formalisation and transparency value of Boumaa Fijian.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size X₁</td>
<td>1</td>
<td>semantic structure Y₁</td>
<td>1</td>
</tr>
<tr>
<td>host number X₂</td>
<td>0</td>
<td>rule types Y₂</td>
<td>0.666</td>
</tr>
<tr>
<td>locus operandi X₃</td>
<td>0</td>
<td>rule number Y₃</td>
<td>1</td>
</tr>
<tr>
<td>obligatoryness X₄</td>
<td>0.333</td>
<td>categorial independence Y₄</td>
<td>1</td>
</tr>
<tr>
<td>boundedness X₅</td>
<td>1</td>
<td>discreteness of markers Y₅</td>
<td>1</td>
</tr>
<tr>
<td>marking X₆</td>
<td>0</td>
<td>redundancy Y₆</td>
<td>1</td>
</tr>
<tr>
<td>exhaustivity X₇</td>
<td>1</td>
<td>flexibility Y₇</td>
<td>1</td>
</tr>
<tr>
<td><strong>x-value</strong></td>
<td>0.476</td>
<td><strong>y-value</strong></td>
<td>0.952</td>
</tr>
</tbody>
</table>

ad **X₁**: Boumaa Fijian employs a small class inventory of three or four classes, depending on the analysis of the marker 'e-/ke-' (cf. table 9).

ad **X₂**: The classifiers attach only to the possessive marker (cf. Dixon 1988: 125 and [19]).

ad **X₃**: Due to their occurrence in possessive constructions, the Boumaa Fijian classifiers operate within the borders of the noun phrase. The consulted sources and available data do not suggest an occurrence outside of the classificatory context.
ad X₄: The classifiers occur only in possessive constructions (cf. Dixon 1988: 125).

ad X₅: The classifiers take the form of affixes (cf. table 9).

ad X₆: The classifiers occur once per context, as no other part of speech takes classificatory marking than the possessive.

ad X₇: Boumaa Fijian inalienable constrictions typically do not host classifiers (cf. 19). Nevertheless, the system allows a classification of the full nominal lexicon by employing a generic category for alienable items and a category for things that are related to the possessor (cf. 23 and 24).

(24) A a o- CL:OWNED na i- POSS.MARKER talanoa
“his story (he tells it)”

B a 'e- CL:RELATED na i- POSS.MARKER talanoa
“his story (it’s about him)” (Boumaa Fijian; cf. Dixon 1988: 137)

ad Y₁: The Boumaa Fijian classes are semantically bijective (cf. table 9). This includes the general class, which applies to all entities that are not assignable to the other categories.

ad Y₂: While the assignment is generally carried out on semantic grounds, it is partly delimited by conventionalisation: Nominalised verbs are classified by we-/o-only (Dixon 1988: 139f.).

ad Y₃: The class assignment requires only a single rule per class (cf. table 9), which comes down to a number of four assignment rules.

ad Y₄: The markers convey class information only (cf. table 9) and thus are categorically independent.

ad Y₅: The evaluation of the discreteness of the Boumaa Fijian markers depends on the analysis of the class inventory: Typically, the inalienable class and the edible class are distinguished into two discrete categories. In this case, Boumaa Fijian would mark two categories with the same marker ’e-/ke-. Other Oceanic systems of possessive classification such as the ones of Mota (Franjieh in prep.: 2) and Kilivila mark an “intimate, intrinsic connection” (Senft 1986: 51) of the classified to the possessor by means of an edible category among others, which suggests a reanalysis of the system as possessing an inventory of three classes. In this case, the Boumaa Fijian system possesses a bijective and discrete class marking.
ad Y₆: The classificatory marker adds semantic content to its host NP and is therefore semantically informative (cf. 20 and 25).

\[(25)\]  
\[
\begin{array}{ll}
A & a \quad o-mu \quad da’ai \\
\text{ART} & \text{CL:OWNED-2SG.POSS} & \text{gun} \\
\end{array}
\]

“your gun (which belongs to you)”

\[
\begin{array}{ll}
B & a \quad ‘e-mu \quad da’ai \\
\text{ART} & \text{CL:RELATED-2SG.POSS} & \text{gun} \\
\end{array}
\]

“your gun (which will be used to shoot you)” (Boumaa Fijian; Dixon 1988: 137)

ad Y₇: The same noun can take different classifiers (cf. 20). This makes the Boumaa Fijian classification a flexible one.

2.4 Grammaticalisation

The Boumaa Fijian possessive classifier system displays a few formal properties that correspond to a general grammaticalisation cline: For instance, it has a small inventory of three or four classes, employs affixes as its markers, and is able to classify the whole nominal lexicon. Other formal properties are corresponding to a lower degree of grammaticalisation, such as the single occurrence of the markers in classificational contexts only and the possessive being the only morphosyntactic host for classifiers. This results in a medium value of formalisation (0.476). The system is highly transparent (value: 0.952) in that the classes are semantically bijective and framed by a small number of prevailing semantic assignment rules. The markers are categorically independent and discrete per class. They convey semantically informative information, and may be assigned in a flexible way. The Boumaa Fijian system is one representative of a fairly widespread Oceanic possessive classification type, which includes one or two alimentary classes. Data from comparative and historical studies suggests a common ancestral Proto-Oceanic system (cf. e.g. Franjieh 2012: 206f. for a discussion). This proto-system does not possess formal or other properties that differ to a greater degree from the one of Boumaa Fijian. This suggests that the system did not grammaticalise notably in both its formal and transparent dimension.

3 Burushaski

3.1 Background

< language isolate (ISO 639-3: bsk; cf. Lewis et al. 2013: s.v.)
< language isolate (cf. Ruhlen 1987)

Burushaski is a language isolate that is spoken in northern Pakistan and India (“North Gilgit district areas, Hunza-Nagar and Yasin areas; scattered in Gilgit, Kashmir, and various cities”; Lewis et al. 2013: s.v.). This dossier
Appendix

is mainly based on Berger 1998, Munshi 2006, and Tiffou/Pesot 1989. The Burushaski concordial class system displays ‘the best of both worlds’ of nominal categorisation: It employs agreement and serves a discourse’s reference tracking and patterns the mental lexicon formally, and it also productively operates in the semantic domain.

![Nominal classification devices]

Figure 6: Classification of the Burushaski system of nominal classification.

3.2 Classification Typology

3.2.1 Noun Classifiers

Burushaski does not employ a noun classifier system (cf. 26, where no classifier occurs).

(26) we gušiŋa kančéni(-tiŋ) ban.
    DEM:3PL.ANIMATE woman:PL.ANIMATE volatile(-PL.ANIMATE)
    be:PRS3PL.ANIMATE
    “These women are volatile.” (Burushaski; cf. Tiffou/Pesot 1989: 54)

3.2.2 Numeral Classifiers

Burushaski does not employ a numeral classifier system (cf. 27, where no classifier occurs).

(27) A iské-kuc, thalé-kuc
    three-day seven-day
    “three days, seven days”
B æltal, wálkuc
    two.day four.day
    “two days, four days” (Burushaski; cf. Tiffou/Pesot 1989: 27)
C han kuc, cendó kuc
    one day five day
    “one day, five days” (Burushaski; compiled after Tiffou/Pesot 1989)
3.2.3 Possessive Classifiers

Burushaski does not employ a possessed classifier system (cf. 28 where no classifier occurs).

(28)  
A  
\[múimo \quad háalar \quad i-\dot{c}h\dot{a}-m-o.\]  
POSS3SG.F  house-GEN.RESIDUE  3SG.ANIMATE-brought-IPF-3SG.F  
“She brought him to her (own) house.”

B  
iné-mo  
DEM3SG.HUMAN-POSS  háalar  
i-\dot{c}h\dot{a}-m-o.  
3SG.ANIMATE-brought-IPF-3SG.F  
“She brought him to her [i.e. another woman’s; MP] house.” (Burushaski; cf. Berger 1998: 80)

3.2.4 Multiple Classifier Systems

Burushaski does not employ a multiple classifier system (cf. 26, 27, and 28 where no classifier occurs).

3.2.5 Concordial Class Systems

Burushaski possesses a concordial class system, which operates by means of class agreement (cf. 26 where all elements are marked for the animate plural class). This system is analysed in section 3.3.

3.3 Analysis

3.3.1 Functional Setup

Table 11: The semantic and discourse properties of the Burushaski system of nominal classification.

| **differentiating referents** (C1) | productive derivation, variable classification | ✓ |
| **individuation** (C2) | unitisation for countability | ✓ |
| **attribution of properties** (C3) | speaker-sided values | ✓ |
| **reference identification** (C4) | reference tracking via AGR, anaphora | ✓ |

**ad C1:** The flexibility of class membership in Burushaski is quite productive and mainly based on a count-mass distinction. This process allows that “[a]morphous solids may fall under [+concrete] or [-concrete] class [sic!] depending on speakers’ perception in terms of the semantic feature “formlessness/shapelessness”” (Munshi 2006: 163). This and the differentiation of epicene nouns such as -yi “offspring”, which can be distinguished by male and female gender (cf. the examples in Munshi 2006: 154) comes down to a free and variable classification rather for the purpose of referent differentiation.
ad C₂: Concordial class membership in Burushaski can serve as a means for the differentiation count and mass referents; e.g., trees are always members of the “residue” class, their fruits are members of the “concrete” class; crops are members of the “residue” class, the individual grain is a member of the “concrete” class (Berger 1998: 36).

ad C₃: A nominal referent can be attributed properties by concordial class assignment; e.g., reference to humans can be made by a convergence of the noun to the “concrete” class in order to achieve a pejorative reference (cf. Berger 1998: 34f.).

ad C₄: The Burushaski concordial class system serves the purpose of reference identification in discourse (cf. 29).

(29) mu-del-i-m-i
3SG.F-hit-ÉPENTHETIC-IPF-3SG.M
“he hit her” (Burushaski; cf. Berger 1998: 112)

3.3.2 Formal Setup

Table 12: The formal properties of the nominal classification system of Burushaski.

<table>
<thead>
<tr>
<th>Feature</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanically assigned</td>
</tr>
</tbody>
</table>

ad F₁: Every noun in Burushaski belongs to a concordial class (Berger 1998: 33).
**ad F₂**: Every noun in Burushaski typically belongs to one concordial class (Berger 1998: 33). Though this is the canonical case, there is a certain amount of flexibility involved in the concordial class assignment of the Burushaski nominals: Depending on the real world referent’s consistency (continuous vs. individuated), object nouns generally may take either the “concrete” or the “residue” class marker (cf. Berger 1998: 36ff. and Munshi 2006: 163). Furthermore, epicene nouns distinguish their referents’ biological sex by means of class marking (Munshi 2006: 154).

**ad F₃**: Burushaski discriminates four concordial classes, which respectively comprise human males (“hm”), human females (“hf”), animals and some inanimates (“x”), and abstracts, liquids, and the remaining inanimates (“y”). For a better understanding and comparability, these are referred to by “masculine”, “feminine”, “concrete”, and “residue” within this dossier. Due to formal syncretisms, the classes partially are subsumed under “animate” (masculine and feminine) or “concrete/animate” (masculine, feminine, and concrete; cf. Berger 1998: 33); the inanimate class comprises the “concrete” and “residue” class. The only forms that formally discriminate all four classes are the third person singular forms of the auxiliary verb (cf. Berger 1998: 34); they are given in table 13.

Table 13: The forms of the third person singular of the Burushaski auxiliary *b(á)-* (Berger 1998: 144).

<table>
<thead>
<tr>
<th>class</th>
<th>singular form</th>
<th>plural form</th>
</tr>
</thead>
<tbody>
<tr>
<td>masculine</td>
<td>báí</td>
<td>báan</td>
</tr>
<tr>
<td>feminine</td>
<td>bo</td>
<td>báan</td>
</tr>
<tr>
<td>concrete</td>
<td>bi</td>
<td>bié</td>
</tr>
<tr>
<td>residue</td>
<td>bilá</td>
<td>bićá</td>
</tr>
</tbody>
</table>

**ad F₄**: The Burushaski system of nominal categorisation employs agreement on a range of modifiers such as pronominals, demonstratives, adjectives, several numerals, etc. (cf. 26 and Munshi 2006: 163).

**ad F₅**: Among others, the Burushaski pronominal verb affixes agree with a noun in its concordial class. Therefore, class agreement is marked even outside the NP-borders (cf. 28 A).

**ad F₆**: A noun’s membership of a concordial class mirrors its humanness and biological sex in the case of the “masculine” and “feminine” class and its degree of individuation, concreteness, or countability in the case of the “concrete” and “residue” class (e.g. čhumár c. “(metal) pan”/r. “iron”, bayú c. “leavened

**ad F7:** The Burushaski system of nominal categorisation is very transparent with respect to its semantics. In general, a speaker is able to assign more or less all nouns spontaneously to the individual classes (cf. Munshi 2006: 162f.). Nevertheless, we find “some exceptions which either fall under the class of [+concrete] or [-concrete] [i.e. the residue class; MP] nouns, [sic] viz., ‘tree’, ‘book’, ‘clothes’, ‘house’, ‘field’, etc. Thus, there may also be inter-speaker variations with respect to this group” (Munshi 2006: 166).

3.3.3 Mapping

![Figure 7: Burushaski.](image)

Table 14: The formalisation and transparency value of Burushaski.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size X₁</td>
<td>1</td>
<td>semantic structure Y₁</td>
<td>1</td>
</tr>
<tr>
<td>host number X₂</td>
<td>1</td>
<td>rule types Y₂</td>
<td>0.666</td>
</tr>
<tr>
<td>locus operandi X₃</td>
<td>1</td>
<td>rule number Y₃</td>
<td>1</td>
</tr>
<tr>
<td>obligatoriness X₄</td>
<td>1</td>
<td>categorial independence Y₄</td>
<td>0.666</td>
</tr>
<tr>
<td>boundedness X₅</td>
<td>1</td>
<td>discreteness of markers Y₅</td>
<td>0.5</td>
</tr>
<tr>
<td>marking X₆</td>
<td>1</td>
<td>redundancy Y₆</td>
<td>1</td>
</tr>
<tr>
<td>exhaustivity X₇</td>
<td>1</td>
<td>flexibility Y₇</td>
<td>0.666</td>
</tr>
<tr>
<td>x-value</td>
<td>1</td>
<td>y-value</td>
<td>0.786</td>
</tr>
</tbody>
</table>

378 Appendix
Appendix

**ad X**: The Burushaski class inventory comprises four categories (cf. table 13), which renders it a small one.

**ad X**: Class distinctions occur on verbal affixes and auxiliary forms, the nominal plural endings, numerals, adjectives, and a range of pronouns (demonstratives, interrogatives, and possessives; cf. Munshi 2006).

**ad X**: The class markers occur NP-internally, either together with the classified noun (cf. 26) or without it (cf. 28). They also occur on the verb and thus NP-externally (cf. 29).

**ad X**: Every Burushaski noun belongs to a class (Berger 1998: 33), which makes the classification an obligatory one. The agreement markers occur obligatorily across contexts (cf. 29).

**ad X**: The markers occur in bound form (cf. e.g. 29).

**ad X**: The markers may occur multiply in the same context (cf. 26).

**ad X**: The system assigns all nouns to a concordial class (Berger 1998: 33).

**ad Y**: The semantics underlying the individual classes are bijective: The “masculine” class comprises all nouns for human males, the “feminine” class does so for all human and supernatural females (Berger 1998: 34). The “concrete” class comprises all non-human (besides supernatural females), countable objects and animates, whereas the “residue” class mainly contains masses and collectives, but also objects that are categorised by their consistency (of a certain mass that belongs to the “residue” class), but also abstracts (cf. Berger 1998: 33 and 35).

**ad Y**: The speaker of Burushaski is able to spontaneously assign nouns to the individual classes (Munshi 2006: 162f.). The assignment is carried out on semantic grounds, with only a few reported exceptions, which display an opaque assignment (Munshi 2006: 166).

**ad Y**: Due to the small class inventory and the bijective class semantics, the number of assignment rules is low.

**ad Y**: The class markers are fused with the category number (cf. 26).
ad Y₅: The class inventory shows large syncretisms in different domains: Only the third person singular form of the auxiliary verb distinguishes all four classes (cf. Berger 1998: 34). In other contexts, the classes show different syncretic groupings such as “animate”, “animate/concrete”, etc. (cf. the discussion under F₃).

ad Y₆: The class markers convey informative and relevant semantics, as the productive reassignment of Burushaski nouns proves: In a pejorative use, reference to humans can be made by a convergence of the noun to the “concrete” class (cf. Berger 1998: 34f.). Big trees are always members of the “residue” class, their fruits are members of the “concrete” class (just as fruits and fruit-like plants such as turnips), such as bāalt c. “apple” vs. bāalt r. “apple tree” (cf. Berger 1998: 36). Smaller plants such as shrubberries, flowers, or herbs are “residue” members; crops are usually also members of the “residue” class, but the individual grain is treated as a member of the “concrete” class. The word for “seed”, ġunó, is also a member of the “concrete” class, except for its variant that refers to “semen” (Berger 1998: 36). Wood, textiles (but not cloth as such), and objects that are made of either, the head, bones, and other body parts that are characterised as “bony”, containers and pipes are usually members of the “concrete” class. Metals (and metal objects), cloth, hand/fist, arm/leg, claw/paw, wings, skin/fur/hair, leather (and leather objects), fleshed body parts (e.g. shank and breast), holes and cavations, places and plains, ways, rooms, buildings, toponyms, and abstracts are all members of the “residue” class (cf. Berger 1998: 36ff.). The main characteristic of the semanticity of the Burushaski system of nominal categorisation seems to be the highly productive mass-count distinction. Masses and collectives may share the same noun stem with individual items that are related to or derived from them. The overall semantics behind this system is thus [(human) animate (m. vs. f.) - inanimate (c. vs. r.)].

ad Y₇: The classification of Burushaski nouns is largely flexible (cf. the examples under Y₆).

3.4 Grammaticalisation

The Burushaski system of nominal categorisation formally displays a high degree of formalism by employing a sophisticated agreement system with different syncretic agreement patterns of the four concordial classes (value: 1). It also possesses a high degree of productive and transparent semanticity (value: 0.786), and allows a largely flexible classification for the purpose of individuation and sex distinction of nominal referents. The formal syncretisms of class indicate a highly aged system, which nevertheless maintains a high degree of transparency and semantic productivity.
4 Chichewa

4.1 Background

Chichewa (or Chewa, also known as (Chi-)Nyanja) is a language of East, Central, and Southern Africa, which is spoken in Malawi, Botswana, Mozambique, Swaziland, Zambia, and Zimbabwe (cf. Lewis et al. 2013: s.v. and Mchombo 2004: 1). This dossier is mainly based on Mchombo 2004, Stevick and Hollander 1965 (and its subsequent Peace Corps edition from 1995), and Watkins 1937. The Chichewa system meets the definition of a concordial class system.

![nominal classification devices](image)

Figure 8: Classification of the Chichewa system of nominal classification.

4.2 Classification Typology

4.2.1 Noun Classifiers

Chichewa does not employ a noun classifier system (cf. (30) where class agreement occurs instead of a classifier).

(30) A- sodzi a- dzó- bá mi- kángó
y- anú i- tátu yíi
SUBJ.MARKER:4:PL your SUBJ.MARKER:4:PL three 4:DEM
i- méné i- ku- sáuls
SUBJ.MARKER:4:PL REL SUBJ.MARKER:4:PL PRS bother
-á alenje.
FINAL.VOWEL 2:PL: hunter

“The fishermen will steal these three lions of yours which are bothering the hunters.”

(Chichewa; cf. Mchombo 2004: 24f.)
4.2.2 Numeral Classifiers
Chichewa does not employ a numeral classifier system (cf. 30, where class agreement occurs instead of a classifier).

4.2.3 Possessive Classifiers
Chichewa does not employ a possessive classifier system (cf. 30, where class agreement occurs instead of a classifier).

4.2.4 Multiple Classifier Systems
Chichewa does not employ a multiple classifier system (cf. 30, where class agreement occurs instead of classifiers).

4.2.5 Concordial Class Systems
Chichewa employs a concordial class system, which manifests itself in form of class agreement (cf. 30). This system is analysed in section 4.3.

4.3 Analysis
4.3.1 Functional Setup

<table>
<thead>
<tr>
<th>differentiating referents (C₁)</th>
<th>productive derivation, variable classification ✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability x</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values ✓</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora ✓</td>
</tr>
</tbody>
</table>

ad C₁: The Chichewa system is not semantically motivated (Mchombo 2004: 4, Watkins 1937: 23), except for the classes 15 to 18. The lexicon can be expanded by derivational means of nominal classification only in these cases. Since the class change manifests itself in a change of agreement marking, this change is derivational rather than a case of variable classification.

ad C₂: The consulted sources provide no evidence for individuation to play a role in the Chichewa concordial class system.
ad C₃: By an assignment to the classes 15 to 18, the speaker can assign e.g. diminutive values to the referent (cf. [31] on page [383]).

(31)  

A m-
 lenje w-
 ánú u-
 ja
 TSG hunter SUBJ.MARKER:1:SG your SUBJ.MARKER:1:SG that
 w-
 á
 nthábula
 TSG ASSOCIATIVE.MARKER 10:PL.humor

“that humorous hunter of yours” (Chichewa; cf. Mchombo 2004: 4)

B ka-
 m-
 lenje k-
 ánú
 T2:SG 1:SG hunter SUBJ.MARKER:12:SG your
 ka-
 ja k a
 ASSOCIATIVE nthábula
 SUBJ.MARKER:12:SG that T2:SG ASSOCIATIVE 10:PL.humor

“that small humorous hunter of yours” (Chichewa; cf. Mchombo 2004: 5)

ad C₄: The Chichewa class markers serve as a means of reference tracking (cf. [32], where the class marking on pronouns allows an unambiguous reference identification).

(32)  

Izi
 ndi
 mbúzi
 zi-
 méné a-
 nyaní
 á-
 ná-
 zi-
 phík -á
 SUBJ.MARKER:2:PL PAST OBJ.MARKER:10:PL cook APPLICATIVE -a
 chí-
 túnábwa.
 FINAL.VOWEL 7:SG pancake

“These are the goats that the baboons cooked the pancake for.” (Chichewa; cf. Mchombo 2004: 86)

4.3.2 Formal Setup

Table 16: The formal properties of the nominal classification system of Chichewa.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>√</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>√</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>×</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>×</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>×</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>×</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker</td>
<td>×</td>
</tr>
</tbody>
</table>
**ad F1:** The Chichewa noun has “at the minimum [...] a bimorphemic structure” (Mchombo 2004: 3), which consists of the stem and a concordial class prefix. Conversely argued, each noun is assigned to a concordial class.

**ad F2:** The consulted sources did not reveal information on the flexibility of a noun’s class membership in Chichewa. However, the literature and the examples therein displayed a consistent class membership of the nouns listed. The Chichewa noun thus appears to be inflexible with respect to its concordial class membership. A systematic exception to this rule are the classes 15 to 18, which may attach to a noun’s original class in order to express diminutive and locative meaning (cf. Mchombo 2004: 5).

**ad F3:** The Chichewa system partly interacts with the category of number (cf. Mchombo 2004: 3). The consulted sources differ from each other with respect to the internal structure of the class inventory, to the class markers’ form, and to the numeric labels of the single classes. As this study aims at a separation of number and nominal classification, Watkin’s (1937) account, where number is ignored for the labelling of the individual classes, is taken up in this dossier. Note however that glossed examples follow the glossing of their source and thus may display a different numbering from e.g. Mchombo (2004). Watkins’ (1937) classes 1 to 7 each distinguish singular and plural, the classes 8 to 10 do not (cf. Watkins 1937: 24). A list of the markers is given in table 17; the qualitative marking corresponds to relative (or possessive; cf. Watkins 1937: 124) and adjectival marking (cf. Watkins 1937: 25).

Table 17: The concordial class markers of Chichewa according to Watkins (1937: 24).

<table>
<thead>
<tr>
<th>class</th>
<th>noun</th>
<th>V&lt;sub&gt;SUBJ&lt;/sub&gt;</th>
<th>V&lt;sub&gt;OBJ&lt;/sub&gt;</th>
<th>qualitative 1</th>
<th>qualitative 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:SG</td>
<td>mu-</td>
<td>a-</td>
<td>-mu-</td>
<td>wamu-</td>
<td>waamu-</td>
</tr>
<tr>
<td>1:PL</td>
<td>ma-</td>
<td>pa-</td>
<td>-pa-</td>
<td>pa-</td>
<td>pa-</td>
</tr>
<tr>
<td>2:SG</td>
<td>mu-</td>
<td>u-</td>
<td>-(w)u-</td>
<td>wa(w)u-</td>
<td>wa(w)u-</td>
</tr>
<tr>
<td>2:PL</td>
<td>mi-</td>
<td>i-</td>
<td>-(y)i-</td>
<td>ya(y)i-</td>
<td>ya(y)i-</td>
</tr>
<tr>
<td>3:SG</td>
<td>φ-</td>
<td>li-</td>
<td>-li-</td>
<td>la-</td>
<td>la-</td>
</tr>
<tr>
<td>3:PL</td>
<td>ma-</td>
<td>pa-</td>
<td>-pa-</td>
<td>papa-</td>
<td>papa-</td>
</tr>
<tr>
<td>4:SG</td>
<td>φ-</td>
<td>(b)u-</td>
<td>-(b)u-</td>
<td>(b)wa-</td>
<td>(b)wa-</td>
</tr>
<tr>
<td>4:PL</td>
<td>ma-</td>
<td>pa-</td>
<td>-pa-</td>
<td>pa-</td>
<td>pa-</td>
</tr>
<tr>
<td>5:SG</td>
<td>φ-</td>
<td>i-</td>
<td>-(y)i-</td>
<td>ya(y)i-</td>
<td>ya(y)i-</td>
</tr>
<tr>
<td>5:PL</td>
<td>φ-</td>
<td>zi-</td>
<td>-zi-</td>
<td>za-</td>
<td>za-</td>
</tr>
<tr>
<td>6:SG</td>
<td>tci-</td>
<td>tci-</td>
<td>tci-</td>
<td>tcalc-</td>
<td>tcalc-</td>
</tr>
<tr>
<td>6:PL</td>
<td>vi-</td>
<td>vi-</td>
<td>vi-</td>
<td>va-</td>
<td>va-</td>
</tr>
<tr>
<td>7:SG</td>
<td>ka-</td>
<td>ka-</td>
<td>ka-</td>
<td>ka-</td>
<td>ka-</td>
</tr>
<tr>
<td>7:PL</td>
<td>tu-</td>
<td>tu-</td>
<td>tu-</td>
<td>twa-</td>
<td>twa-</td>
</tr>
<tr>
<td>8</td>
<td>pa-</td>
<td>pa-</td>
<td>pa-</td>
<td>pa-</td>
<td>pa-</td>
</tr>
<tr>
<td>9</td>
<td>mu-</td>
<td>mu-</td>
<td>mu-</td>
<td>mu-</td>
<td>mu-</td>
</tr>
<tr>
<td>10</td>
<td>(k)u-</td>
<td>ku-</td>
<td>ku-</td>
<td>kwa-</td>
<td>kwa-</td>
</tr>
</tbody>
</table>
ad F₄: Class markers may occur multiply in Chichewa (cf. [31]).

ad F₅: Class marking may occur within the noun phrase in the presence or absence of the classified noun (cf. [31] and [33] C). It may also occur on the predicate and thus NP-externally (cf. [33] B).

(33)  
A  Fisi anadyá mkângo.  
hyena:3:SG ate lion:3:SG  
“The hyena ate the lion.”

B  Ā-tá-dyá anapítá ku San Francisco.  
he-SER-eat it:3:SG-eat he-went to San Francisco  
“Having eaten it, he went to San Francisco.”

C  Ā-tá-dyá įwo anapítá ku San Francisco.  
he-SER-eat it:3:SG he-went to San Francisco  
“Having eaten it (something other than the lion), he went to San Francisco.”  
(Chichewa; cf. Bresnan and Mchombo 1987: 749)

ad F₆: A change in concordial class may alter the semantics of the noun phrase (cf. [31]). Nevertheless, nouns seem not to be able to change their concordial class except for the (additional) attachment of the classes 15 to 18 (for diminutives and locative constructions).

ad F₇: Since the system is not semantically motivated (cf. Mchombo 2004: 4 and Watkins 1937: 23) and since the consulted sources did not list transparent formal strategies of concordial class assignment (which is a strong indication that a noun’s concordial class has to be learned), the assignment appears to be performed by the system and not by the speaker.
4.3.3 Mapping

Table 18: The formalisation and transparency value of Chichewa.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size $X_1$</td>
<td>0.875</td>
<td>semantic structure $Y_1$</td>
<td>0.5</td>
</tr>
<tr>
<td>host number $X_2$</td>
<td>1</td>
<td>rule types $Y_2$</td>
<td>0.666</td>
</tr>
<tr>
<td>locus operandi $X_3$</td>
<td>1</td>
<td>rule number $Y_3$</td>
<td>0</td>
</tr>
<tr>
<td>obligatoriness $X_4$</td>
<td>1</td>
<td>categorial independence $Y_4$</td>
<td>0.666</td>
</tr>
<tr>
<td>boundedness $X_5$</td>
<td>1</td>
<td>discreteness of markers $Y_5$</td>
<td>0</td>
</tr>
<tr>
<td>marking $X_6$</td>
<td>1</td>
<td>redundancy $Y_6$</td>
<td>0.75</td>
</tr>
<tr>
<td>exhaustivity $X_7$</td>
<td>1</td>
<td>flexibility $Y_7$</td>
<td>0.666</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x-value</th>
<th>y-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.982</td>
<td>0.512</td>
</tr>
</tbody>
</table>

ad $X_1$: Chichewa employs ten concordial classes (cf. table 17), which is a small inventory.

ad $X_2$: The verb, adjective, numeral, demonstrative and relative pronoun as well as other pronouns agree with their referent noun in concordial class and number (cf. 30).

ad $X_3$: The predicate takes class markers, which thus transcend the borders of the noun phrase (cf. 33 B). NP-internal class marking occurs regardless of the presence of the classified noun (cf. 33 C).
A range of nominal modifiers (demonstratives, numerals, relative constructions, possessives, and adjectives) obligatorily agrees with their referent noun’s concordial class (cf. Mchombo 2004: 24).

Class marking occurs in form of affixes (cf. table [17]).

Class marking may occur multiply in the same context (cf. 31).

The Chichewa noun has “at the minimum […] a bimorphemic structure” (Mchombo 2004: 3), which consists of the stem and a concordial class prefix. Conversely argued, each noun is assigned to a concordial class.

The concordial class system of Chichewa is inherited from at least the Proto-Bantu stage (cf. Watkins 1937: 23), and the underlying semantic principles have eroded considerably. At its contemporary state, the system is not semantically motivated (Mchombo 2004: 4 and Watkins 1937: 23).

As most nouns seem to be assigned opaquely, only the diminutive and locative classes seem to employ a semantic assignment. The consulted sources do not reveal information on morphological or phonological assignment.

The number of assignment rules is high due to the large number of nouns with a semantically unmotivated class membership.

The individual class markers convey information on both classification and number (cf. table [17]).

The same class marker may index more than one class (cf. table [17]).

Class marking in Chichewa may be either informative (cf. 31) or vain (Mchombo 2004: 4 and Watkins 1937: 23). As especially the semantically motivatable classes show a productive profile, the informativeness is assessed in this dossier as more influential for the system than its redundancy, which results in an intermediate value of 0.75.

The locative and diminutive classes allow a largely flexible classification (cf. Mchombo 2004: 5).

4.4 Grammaticalisation

Chichewa displays a high degree of formalisation (value: 0.982), as is common for African concordial class systems. Its classes are partly semantically productive, which results in an intermediate degree of transparency (value: 0.512).
5 Chukchi

5.1 Background

Chukchi (also known as Chukot or Luoravetlan) is a language spoken in Northeast Siberia, on the Chukchi peninsula, the Chukotka Autonomous Region, and the Kamchatka District (cf. Lewis et al. 2013: s.v.). Chukchi is a special case with respect to its nominal classification: It is typically considered not to employ a classifier system or a concordial class system. As Aikhenvald (2000: 438f.) notes, Chukchi nouns “are divided into declensional paradigms according to the animacy of the referent”. Therefore, its system’s operating mechanisms and functions are investigated and compared to those of acknowledged systems of nominal categorisation. Though it is not commonly acknowledged as a classifier system or concordial class system, it will be treated as a concordial class system within this dossier, because it interacts with the declensional paradigms of the Chukchi nominal inflection (which is a typical feature of concordial class systems) and displays some involved semantics. The study of Chukchi shows that the system does not fall under the scope of nominal classification, though it might have been a productive nominal classification device at earlier stages. This dossier is mainly based on Bogoras 1922, Comrie 1981, Dunn 1999 and Weinstein 2010.

![Figure 10: Classification of the Chukchi system of nominal classification.](image)

5.2 Classification Typology

5.2.1 Noun Classifiers

Chukchi does not employ a noun classifier system (cf. Aikhenvald 2000: 438f.).
5.2.2 Numeral Classifiers
Chukchi does not employ a numeral classifier system (cf. Aikhenvald 2000: 438f. and Weinstein 2010: 837ff.).

5.2.3 Possessive Classifiers
Chukchi does not employ a possessive classifier system (cf. Aikhenvald 2000: 438f.).

5.2.4 Multiple Classifier Systems
Chukchi does not employ a multiple classifier system (cf. Aikhenvald 2000: 438f.).

5.2.5 Concordial Class Systems
Chukchi distinguishes two declensional classes on the basis of the animacy of the nominal referent (cf. Aikhenvald 2000: 438 and Comrie 1981: 246). It is common sense that Chukchi does not employ a concordial class system (cf. Aikhenvald 2000: 438f. and Weinstein 2010: 22 and 38). Since the formal classification of its nominal lexicon is comparable to acknowledged concordial class systems and distinguishes two classes, its proximity to nominal classification devices is analysed in section 5.3.

5.3 Analysis
5.3.1 Functional Setup

Table 19: The semantic and discourse properties of the Chukchi system of nominal classification.

<table>
<thead>
<tr>
<th>differentiating referents (C_1)</th>
<th>productive derivation, variable classification</th>
<th>✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C_2)</td>
<td>unitisation for countability</td>
<td>✗</td>
</tr>
<tr>
<td>attribution of properties (C_3)</td>
<td>speaker-sided values</td>
<td>✗</td>
</tr>
<tr>
<td>reference identification (C_4)</td>
<td>reference tracking via AGR, anaphora</td>
<td>✓</td>
</tr>
</tbody>
</table>

**ad C_1:** Since common nouns with human referents may optionally be assigned to either the “higher animates” class or to the residue class (with the referent remaining the same), there is no expansion of the nominal lexicon by means of derivation present in Chukchi, but a variable classification.
ad C₂: The consulted sources do not reveal information about individuation of referents playing a role in the Chukchi system. However the semantics of the system (higher animates class vs. residue class) do not provide an opportunity of the system to be involved in individuation processes.

ad C₃: Besides for nouns denoting human beings, which may be assigned to either the “high animate” class (optionally) or the residue class, a noun’s concordial class is not variable in Chichewa (Comrie 1981: 246, Dunn 1999: 65). This prevents the speaker’s possibility to attribute values to the referent by classificatory means.

ad C₄: Since the pronoun is marked for concordial class in Chukchi, referent identification is by means of class marking is possible.

5.3.2 Formal Setup

Table 20: The formal properties of the nominal classification system of Chukchi.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>×</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>✓</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>×</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>✓</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>✓</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>×</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker</td>
<td>✓</td>
</tr>
</tbody>
</table>

ad F₁: Each noun is assigned to one of the two referential classes according to their referent’s degree of animacy (cf. e.g. Comrie 1981: 246).

ad F₂: While the class membership is fixed for the bulk of the nominal lexicon, common human nouns may be assigned to either class (cf. Comrie 1981: 246 and Dunn 1999: 65).
ad F₃: Chukchi possesses two referential classes (cf. e.g. Aikhenvald 2000: 438f.), which is a clearly delimited set.

ad F₄: The declensional affixes occur only on the noun or pronoun and thus is only marked once per context (cf. the respective sections on the nouns, adjectives, determiners, numerals, etc. in Weinstein 2010).

ad F₅: Since the class membership of a noun is only marked on the noun or pronoun (cf. the respective sections on the nouns, adjectives, determiners, numerals, etc. in Weinstein 2010), their occurrence is limited to the noun phrase.

ad F₆: While the assignment is based on semantics and semantics only, the class membership does not seem to shape the meaning of the affected noun phrase, since their referents’ inherent degree of animacy determines their referential class. Common nouns are able to freely shift in their class membership (cf. Comrie 1981: 246 and Dunn 1999: 65), which suggests that the class membership does not exert influence on the noun’s semantics.

ad F₇: The assignment criteria of the Chukchi referential classes are purely semantic and so simple, that a noun’s class membership does not have to be learned individually: Proper nouns and certain kin terms are assigned to class I, humans to class I or II according to the speaker’s will, and all other nouns are assigned to class II (cf. Comrie 1981: 246).

5.3.3 Mapping

![Figure 11: Chukchi.](image)
Table 21: The formalisation and transparency value of Chukchi.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>1</td>
<td>semantic structure</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>0.25</td>
<td>rule types</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>0.5</td>
<td>rule number</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>1</td>
<td>categorial independence</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>1</td>
<td>discreteness of markers</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>0</td>
<td>redundancy</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>1</td>
<td>flexibility</td>
</tr>
<tr>
<td>x-value</td>
<td>0.679</td>
<td>y-value</td>
<td>0.702</td>
</tr>
</tbody>
</table>

**ad X_1**: The two referential classes of Chukchi form a small set in this study’s definition (Comrie 1981: 246).

**ad X_2**: The marking occurs only on two hosts, the noun and the pronoun (cf. the respective sections on the nouns, adjectives, determiners, numerals, etc. in Weinstein 2010).

**ad X_3**: As the marking occurs on nouns and pronouns, it operates NP-externally and independent of the presence of the classified noun.

**ad X_4**: As the class distinction derives from different obligatory case and number markers (Comrie 1981: 246), it is obligatory.

**ad X_5**: The classes are distinguished by portmanteau affixes of the nominal declensional paradigm of Chukchi (Comrie 1981: 246f.).

**ad X_6**: As the marking occurs only on nouns and pronouns, multiple marking in the same context is not possible.

**ad X_7**: Each Chukchi noun is assigned to one of the two referential classes (cf. e.g. Comrie 1981: 246).

**ad Y_1**: Class I hosts proper names of humans and named animals (‘high animate’ class; cf. Dunn 1999: 64f.), certain kin terms, and optionally other human nouns. Class II hosts human common nouns and all nonhuman nouns (cf. Comrie 1981: 246). The class structure is therefore bijective.

**ad Y_2**: There are only semantic assignment rules necessary to master the assignment of all Chukchi nouns.
ad Y<sub>3</sub>: The straightforward class structure results in a small number of assignment rules: ‘Higher animates’ are assigned to class I, all other nouns are assigned to class II, and nouns denoting humans may optionally be assigned to class I (cf. Comrie 1981: 246 and Dunn 1999: 65).

ad Y<sub>4</sub>: Only a number of cases display a class distinction, e.g. the locative, which is marked by -ne/-na for class I and by -k for class II (cf. Weinstein 2010: 41). Ergativity is expressed differently for both classes: Class I takes the locative suffix -ne/-na, class II takes the instrumental suffix -e/-a (cf. Weinstein 2010: 40), and personal pronouns of all persons taking an independent ergative suffix to mark it (cf. 34 and Comrie 1981: 247).

(34) <em>gum-na’n</em>, <em>gr-na’n</em>, <em>kn’a’n</em>, <em>morg-na’n</em>, <em>torg-na’n</em>, <em>ergr-na’n</em>
1SG-ERG, 2SG-ERG, 3SG-ERG, 1PL-ERG, 2PL-ERG, 3PL-ERG

“<em>I, thou, he/she/it, we, ye, they</em>” (Chukchi; cf. Bogoras 1922: 719f.)

According to Dunn (1999: 65), Comrie (1981: 246) and Bogoras (1922: 694), class I furthermore distinguishes singular from plural in all cases, whereas class II distinguishes singular and plural only in the absolutive. Weinstein’s (2010: 34f.) conclusions slightly differ from this description: There are plural forms of both class II common nouns (ib.: 34), and class I proper nouns and close kinship terms (by insertion of -p(r); ib.: 39), but they are altogether facultative (ib.: 35, 37, and 39). The portmanteau markers thus encode three categories: Case, number, and class.

ad Y<sub>5</sub>: The ‘high animate’ class is marked by -ne/-na in a number of cases, which distinguishes it from class II (cf. Weinstein 2010: 40f.).

ad Y<sub>6</sub>: The semantics conveyed by the class markers are either redundant, as the information expressed is already present in the classified noun (proper names and kinship terms typically refer to higher animates), or vain (class II nouns, common human nouns).

ad Y<sub>7</sub>: The bulk of nouns is not flexible with respect to its class membership; only common nouns denoting humans can be assigned to either class (cf. Comrie 1981: 246 and Dunn 1999: 65).

5.4 Grammaticalisation

It is hard to pigeonhole the Chukchi system of nominal categorisation, since it displays equally mixed properties of both classifier systems and concordial class
systems. At the same time, it lacks the core properties of either type of system: There is neither a contribution to the noun phrase’s semantics nor agreement. Since the involved semantics only affects the assignment of the Chukchi nouns to a referential class, the degree of formal character appears to prevail. Formal criteria like the marking by affixes and the opacity of the markers also point to a fairly grammaticalised system. The phenomenon resembling nominal classification in Chukchi is not very well outlined in the literature on Chukchi, yet the available data and sources suggest that Chukchi does not employ a nominal classification device sensu stricto. Instead, it represents either a case of a decayed system, or a type of discoursal ‘nominal case marking register’, which emphasises the animacy of kin and individuals.

6 Dongo-Ko

6.1 Background

Dongo-Ko (or Donga) is a Mba-language spoken in Central Africa. According to Lewis et al. (2013: s.v.), it is located in the Orientale Province and in the east of Watsa. According to Pasch (1986: 17f.), the main source of this dossier, its location is the North of Zaire (north of the Zaire river). Dongo-Ko employs two systems of nominal categorisation, a concordial class system and a possessive classifier system. This is an instance of multiple systems of nominal classification within the same language, not a mixed system.

---

![Diagram of nominal classification devices]

Figure 12: Classification of the Dongo-Ko system of nominal classification.
6.2 Classification Typology

6.2.1 Noun Classifiers
Non-possessive noun phrases in Dongo-Ko do not host classifiers (cf. 35, where no classifier occurs).

(35) nye-nyó kókókó-ke-nyó
woman-2a:PL bad-ATTRIBUTIVE-2a:PL
“bad women” (Dongo-Ko; cf. Pasch 1986: 232)

6.2.2 Numeral Classifiers
Dongo-Ko does not employ a numeral classifier system. Numerals, which are not inflected, follow the noun they modify “without any addition” (Pasch 1986: 255; cf. 36, where no classifier occurs).

(36) (gbu) fu-so bbo
big island-4:PL two
“two (big) islands” (Dongo-Ko; cf. Pasch 1986: 255)

6.2.3 Possessive Classifiers
Dongo-Ko employs a possessive classifier system (cf. 37). This system is analysed in section 6.3.

(37) A ggy-kó ky-á ná tátámbù-gù
ear-5:SG 5:SG-POSS CL:ANIMATE lion-7:SG
“the lion’s ear” (Dongo-Ko; cf. Pasch 1986: 249)
B bibú-yí y-á buá rá
bowel:2:PL y-á 2:PL-POSS CL:BELLY 1SG
“my bowels” (Dongo-Ko; cf. Pasch 1986: 251)

6.2.4 Multiple Classifier Systems
Dongo-Ko classifiers occur only in possessive constructions (cf. 35 and 36, where no classifiers occur on other nominal modifiers).

6.2.5 Concordial Class Systems
Dongo-Ko employs a concordial class system (cf. 38). This system is analysed in section 6.4.

(38) A kpéli-kó ky-ì ngári-gó
arm-5:SG 5:SG-POSS child-7:SG
“the child’s arm” (Dongo-Ko; cf. Pasch 1986: 244)
B kába-à kókókó-ke-à
tobacco-1a:SG small-1a:SG
“hashish” (literally “small tobacco”; Dongo-Ko; cf. Pasch 1986: 231)
6.3 Analysis: Possessive Classifier System

Dongo-Ko employs a possessive classifier system, which applies not to all kinds of possessive constructions. The two basic strategies of forming possessive constructions comprehend constructions with classifiers (for inalienable and alienable nouns) and constructions lacking a classifier (for specific kin terms and body part terms; cf. Pasch 1986: 239ff.). The focus of this dossier lies on the classifier constructions, which are explored after a short sketch of the full system of possessive constructions in Dongo-Ko (cf. Pasch 1986: 240-255). Possessive constructions lacking a classifier may be:

- absolute constructions (for certain kin terms of the first and second person singular (only members of the class pair 1a/6); cf. 39 A)
- juxtapositions (for kinship terms; cf. 39 B)
- constructions employing the possessive marker í (for kin and body part terms; cf. 39 C and D)
- constructions employing the possessive marker tí (for kin and body part terms; cf. 39 E and F); this type has the structure [possessed entity — possessive marker tí — possessor-class marker] (cf. Pasch 1986: 244f.).

(39) A  ngá(-zó) your.mother
   “your mother” (Dongo-Ko; cf. Pasch 1986: 240)
B  ngé=-rú-ngó wife=chief-2a:PL
   “the wives of the chief” (Dongo-Ko; cf. Pasch 1986: 241)
C  b÷-ng-í àndjú son-9:SG-POSS brother.of.mother
   “my cousin” (Dongo-Ko; cf. Pasch 1986: 243)
D  kpéll-í-kó ky-í  9gáár-í child-7:SG
   “the arm of the child” (Dongo-Ko; cf. Pasch 1986: 244)
E  té-za tí wáší=řé
   mother=3SG POS woman=1:SG
   “my mother-in-law”
F  má-ló tí dů=Šngó
   head-3:SG POSS young.man-9:SG
   “the young man’s head” (Dongo-Ko; cf. Pasch 1986: 245)

5 Kinship terms and body part terms of this constructional type differ in their structure. Kinship terms display the structure [possessed stem-class suffix-possessive marker í — possessor-class marker], body part terms [possessed stem-class suffix — concordial morpheme — possessive marker tí — possessor-class marker] (cf. Pasch 1986: 243f.).
As for the constructions employing classifiers, only two of the ten classifiers serve the categorisation of alienable nominal referents. Both of them are also (though occasionally and arguably) used to categorise inalienable real world referents (cf. [40]), which is only recognised for the classifier \( (m)b \) by Pasch (1986: 248).

(40) A \( wùngá-r̥-b \) l-a \( mb\̣-r̥ \)
    baby-3:SG 3:SG-POSS CL:PROPERTY-1SG
  “my baby” (Dongo-Ko; cf. Pasch 1986: 247)
B \( t̥r̥m̥-m̥ \) m-a \( ḅ \) r̥
    forge-11:SG 11:SG-POSS CL:PROPERTY 1SG
  “my smithery”
C \( r̥-g̣ \) \( ð̣-a \) \( d^a \) r̥
    chief-7:SG 1a:SG-POSS CL:HOUSE 1SG
  “my chief” (Dongo-Ko; cf. Pasch 1986: 248)

6.3.1 Functional Setup

Table 22: The semantic and discourse properties of the Dongo-Ko possessive classifier system.

<table>
<thead>
<tr>
<th>differentiating referents (C_1)</th>
<th>productive derivation, variable classification</th>
<th>✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C_2)</td>
<td>unitisation for countability</td>
<td>×</td>
</tr>
<tr>
<td>attribution of properties (C_3)</td>
<td>speaker-sided values</td>
<td>×</td>
</tr>
<tr>
<td>reference identification (C_4)</td>
<td>reference tracking via AGR, anaphora</td>
<td>×</td>
</tr>
</tbody>
</table>

**ad C_1:** The possessed classifiers occur as independent elements and thus do not serve derivational means. It is however possible to use different classifiers for the same referent (cf. again [11] on page [397]). Due to this variable classification, an ad hoc differentiation of referents seems possible.

(41) A \( ṇ̃-ḳo \) ky-a \( ḅ \) r̥
    belly-5:SG 5:SG-POSS CL:PROPERTY 1SG
  “my belly”
B \( ṇ̃-ḳ ṇ̃ \) ky-a \( ḅ̣̣ \) r̥
    belly-5:SG 5:SG-POSS CL:BELLY 1SG
  “my belly” (Dongo-Ko; cf. Pasch 1986: 248)

**ad C_2:** There is no information available that suggests that the possessed classifiers are involved in an individualising function.

**ad C_3:** The examples in Pasch (1986: 247-252) did not reveal a semantic contribution to the NP where the possessed classifiers occur (cf. [11]); an attribution of properties thus is unlikely.
ad C₄: The classifiers occur once and within the NP where they categorise their nominal referent. Since the consulted sources did not provide information on a potential anaphoric use of the classifiers or a comparable function, reference identification seems not to play a role for the Dongo-Ko possessed classifier system.

6.3.2 Formal Setup

Table 23: The formal properties of the possessive classifier system of Dongo-Ko.

<table>
<thead>
<tr>
<th>Feature (Fᵢ)</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong> (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility</strong> (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td><strong>inventory</strong> (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td><strong>marking</strong> (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi</strong> (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity</strong> (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td><strong>assignment</strong> (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad F₁: The possessed classifier system does not affect all nouns of Dongo-Ko (cf. [42] where no classifier occurs despite the possessive context). Generally speaking, the classifiers categorise mostly body parts with respect to their inherent properties (cf. Pasch 1986: 245ff.).

(42) ɪẹ=nyá
son=your.mother

"your brother" (literally "son your mother"; Dongo-Ko; cf. Pasch 1986: 241)

ad F₂: The general classifiers may facultatively replace a semantically more specific classifier (cf. [41]).

ad F₃: Dongo-Ko employs ten possessed classifiers, which are listed in table [24]. Even if one takes into account the rare classifier ŋgì, the inventory of the
Appendix

Dongo-Ko set of class markers counts only eleven members (cf. Pasch 1986: 246-252), which is a limited and smallish number according to our definition.


<table>
<thead>
<tr>
<th>classifier</th>
<th>semantics</th>
<th>lexical source</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m)bà</td>
<td>“property”</td>
<td>mbàmè-m`ó</td>
</tr>
<tr>
<td>dà</td>
<td>“house”</td>
<td>dà-wò</td>
</tr>
<tr>
<td>nì</td>
<td>“animate”</td>
<td>nìggà-gà/nì-g`ì</td>
</tr>
<tr>
<td>kò</td>
<td>“body”</td>
<td>kò-(w`ì)</td>
</tr>
<tr>
<td>mà</td>
<td>“head”</td>
<td>mà-l`ò</td>
</tr>
<tr>
<td>lò</td>
<td>“eye”</td>
<td>lò-lò</td>
</tr>
<tr>
<td>nò</td>
<td>“mouth”</td>
<td>nò-wò</td>
</tr>
<tr>
<td>rò</td>
<td>“neck”</td>
<td>rò-uò</td>
</tr>
<tr>
<td>bwà</td>
<td>“belly”</td>
<td>bwà-mò-kì</td>
</tr>
<tr>
<td>mòggà</td>
<td>“back”</td>
<td>mòggà</td>
</tr>
</tbody>
</table>

ad F₄: The Dongo-Ko possessed classifiers occur once (cf. [43]).

(43) fù-lò  l-á  bà  rè  
island-3:SG  3:SG-POSS  CL:PROPERTY  1SG  
“my island” (Dongo-Ko; cf. Pasch 1986: 247)

ad F₅: The classifiers occur only in possessive constructions and thus NP-internally (cf. [44]).

(44) nzó-mò  m-á  bà  rè  
flour-11:SG  11:SG-POSS  CL:PROPERTY  1SG  
“my flour” (Dongo-Ko; cf. Pasch 1986: 247)

ad F₆: There is no information available on the semantic potency of the classifiers to possessive constructions. The examples in Pasch (1986: 247-252) do not reveal a semantic contribution to the noun phrase. Instead, the classifier is chosen due to the properties of the nominal referent. It is e.g. imaginable that a body part such as “bowels” can be referred to by the respective body part classifier bwà in case of a reference to one’s own bowels, but with a general classifier such as dà in case of a reference to a butchered animal. Nevertheless, there was no indication of a suchlike classification in the source at use.

ad F₇: The classifier is chosen due to the inherent properties of the entity that is referred to. Since the semantic assignment criteria are transparent (cf. Pasch 1986: 247-252), the class membership of a noun does not have to be learned.
6.3.3 Mapping

Figure 13: Possessive classification in Dongo-Ko.

Table 25: The formalisation and transparency value of possessive classification in Dongo-Ko.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>semantic structure</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>rule types</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>categorical</td>
<td>$Y_4$</td>
</tr>
<tr>
<td>obligatoryness</td>
<td>$X_4$</td>
<td>discreteness of</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>markers</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>redundancy</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>flexibility</td>
<td>$Y_7$</td>
</tr>
<tr>
<td><strong>x-value</strong></td>
<td>0.208</td>
<td><strong>y-value</strong></td>
<td>0.929</td>
</tr>
</tbody>
</table>

ad $X_1$: With a maximum of eleven classes (cf. table 24), the Dongo-Ko possessive classifier system is a small one.

ad $X_2$: The classifiers occur together with the possessive as their morphosyntactic host (cf. 44).

ad $X_3$: Due to their occurrence in possessive constructions only, the classifiers are confined to the borders of the noun phrase. The consulted sources do not comment on a potential anaphoric function of the classifiers, and the
Appendix

available data from Dongo-Ko features the classifiers together with the classified noun in all instances. Therefore, the presence of the classifiers appears to depend on the presence of the classified noun.

ad X₄: Certain possessive constructions require a classifier insertion (cf. Pasch 1986: 245f.).

ad X₅: The classifiers occur as bound (cf. 40 A) and unbound elements (cf. 40 B and C).

ad X₆: The classifiers occur once per context (cf. e. g. 10).

ad X₇: The classifiers mainly categorise body parts (cf. Pasch 1986: 245ff.), while e. g. kinship terms are encoded by possessive constructions lacking a classifier (cf. 12 and Pasch 1986: 252).

ad Y₁: The semantic structure of the individual classes is straightforward (cf. table 24 and Pasch 1986: 247ff.).

ad Y₂: The assignment is carried out on semantic grounds by transparent criteria (cf. Pasch 1986L: 247ff.), which enables the speaker to assign the nouns him- or herself.

ad Y₃: Due to the low number of classes and their bijective semantics, the number of assignment rules is low.

ad Y₄: The class markers encode classificatory semantics only (cf. table 24).

ad Y₅: Each class has its own distinct marker (cf. table 24).

ad Y₆: The classifiers are assigned on basis of properties of the real world referent. In case of variable classification (cf. 41), there seems no change in semantics to be involved. Therefore, the semantics that are present in the classifiers, appear to be redundant and already present in the categorised noun.

ad Y₇: A more general classifier can vary with a more specific classifier for the same noun (cf. 41).
6.4 Analysis: Concordial Class System

6.4.1 Functional Setup

Table 26: The semantic and discourse properties of the Dongo-Ko concordial class system.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification</td>
</tr>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
</tr>
</tbody>
</table>

ad C₁: The class markers serve the expansion of the lexicon’s referential power, e.g. in case of nominalisation of a verb stem by attachment of a classifier marker (cf. 45).

(45) A ̀zilè-̀, ̀zilè-̀dò, ̀zilè-̀ño, ̀zilè-̀ygò, ̀zilè-̀yògò
steal steal-3:SG steal-4:PL thief-7:SG thief-2a:PL
“to steal, theft, thefts, thief, thieves”
B ̀gò̀lè-̀ndàn, ̀gù̀lè-̀gù̀lè
heat-3:SG
“hot, heat” (Dongo-Ko; cf. Pasch 1986: 194f.)

ad C₂: The consulted sources do not report individuation to play a role in the Dongo-Ko concordial class system. The class marking occurs in all instances stable.

ad C₃: Since the concordial classes “are not semantically determined” (Pasch 1986: 194), there is no possibility for the speaker to attribute properties by class assignment.

ad C₄: Despite Pasch (1986: 194) reporting a lack of semantic motivation of the individual classes, the Dongo-Ko system displays an underlying semantics of animateness. Reference identification is possible by e.g. affixing animacy markers to the verb (cf. 46).

(46) A ̀z̀-̀rà-̀wè
3SG:ANIMATE-DEF-come-3SG
“He/she [+animate] comes.”
B ̀b̀-̀rà-̀wè
3INANIMATE-DEF-come-3SG
“It [-animate] comes.”
C ̀zw-̀rà-̀nymè
3PL:ANIMATE-DEF-come-3PL
“They [+animate] come.”
Appendix

D bá-rá-nyé
3:INANIMATE-DEF-come-3PL

6.4.2 Formal Setup

Table 27: The formal properties of the concordial class system of Dongo-Ko.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised X</td>
<td>obligatory assignment of each noun √</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership X</td>
<td>single class membership √</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number X</td>
<td>limited class number √</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once X</td>
<td>multiply marked X</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal X</td>
<td>locus operandi: NP-internal and -external √</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution X</td>
<td>semantically vain X</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker X</td>
<td>mechanic assignment √</td>
</tr>
</tbody>
</table>

ad F₁: The nouns of Dongo-Ko are all members of a certain concordial class, and are typically marked by a class suffix (cf. Pasch 1986: 181).

ad F₂: The consulted sources and available data suggest a rigid class membership. Note however that number of nouns knows two singular or two plural forms: The classes 1a and 7 host a number of these nouns, which are pluralised in the class 2(a) or 6. Class 5 hosts some nouns that can be pluralised in both class 6 or 2(a) (cf. Pasch 1986: 197f. and [47]). These cases are analysed as exceptions here, and not as a truly flexible classification.

(47)  
A bá, bá-ngó, bá-yì  
night:1:SG night-7:SG night-2:PL  
“night, night, nights” (Dongo-Ko; cf. Pasch 1986: 197)  
B bwá-kò, bwá-nzó, bwá-i  
belly-5:SG belly-6:PL belly-2:PL  
“belly, bellies, bellies” (Dongo-Ko; cf. Pasch 1986: 198)
ad F₃: The concordial class system of Dongo-Ko employs eleven classes. Their markers are listed in table 28; their correspondences in singular and plural in table 29.

Table 28: The inventory of Dongo-Ko class markers (cf. Pasch 1986: 193 and 229f.)

<table>
<thead>
<tr>
<th>class suffix</th>
<th>class prefix</th>
<th>class</th>
</tr>
</thead>
<tbody>
<tr>
<td>-((w)o)</td>
<td>w- (inanimate)</td>
<td>1 (sg.)</td>
</tr>
<tr>
<td>-∅</td>
<td>(w-) (animate)</td>
<td>1a (sg.)</td>
</tr>
<tr>
<td>-yo</td>
<td>y- (inanimate)</td>
<td>2 (pl.)</td>
</tr>
<tr>
<td>-nyo</td>
<td>ny- (animate)</td>
<td>2a (pl.)</td>
</tr>
<tr>
<td>-lo</td>
<td>l-</td>
<td>3 (sg.)</td>
</tr>
<tr>
<td>-xo</td>
<td>s-</td>
<td>4 (pl.)</td>
</tr>
<tr>
<td>-ko</td>
<td>k-</td>
<td>5 (sg.)</td>
</tr>
<tr>
<td>-(n)zo</td>
<td>z- (inanimate)</td>
<td>6 (pl.)</td>
</tr>
<tr>
<td>-go</td>
<td>(g)y- (inanimate)</td>
<td>7 (sg.)</td>
</tr>
<tr>
<td>-ŋgo</td>
<td>g(y)-, ŋg- (inanimate)</td>
<td>9 (sg.)</td>
</tr>
<tr>
<td>-mo</td>
<td>m-</td>
<td>11 (sg.)</td>
</tr>
</tbody>
</table>

Table 29: Correspondences between singular and plural classes in Dongo-Ko. (cf. Pasch 1986: 193)

<table>
<thead>
<tr>
<th>class</th>
<th>concording class(es)</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1a</td>
<td>2a, 6</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>2, 2a, 6</td>
</tr>
<tr>
<td>9</td>
<td>2, 2a</td>
</tr>
<tr>
<td>11</td>
<td>2, 6</td>
</tr>
<tr>
<td>plural</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1, 7, 9, 11</td>
</tr>
<tr>
<td>2a</td>
<td>1a, 7, 9</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

ad F₄: The class membership of a Dongo-Ko noun may be multiply marked (on adjectives, demonstratives, and even verbs; cf. 48).
(48) zu₃PL.ANIMATE loe₃SG/story₅SG ky-e₅SG-REL.PARTICLE lŋgo₃PL.ANIMATE

“They told the story of what had happened between them.” (Dongo-Ko; cf. Pasch 1986: 239)

**ad F₅:** Since a noun’s class membership may be marked on the verb (cf. [46]), the borders of the noun phrase do not play a role in the class marking of Dongo-Ko.

**ad F₆:** Due to the mixed character of the individual classes, they “are not distinctly semantically determined” (Pasch 1986: 194) and do not contribute to the noun phrase’s semantics.

**ad F₇:** Since the classes are not semantically determined (cf. Pasch 1986: 194), the nouns are assigned mechanically to their respective concordial class.

### 6.4.3 Mapping

![Figure 14: Concordial classification in Dongo-Ko.](image)

Figure 14: Concordial classification in Dongo-Ko.
Table 30: The formalisation and transparency value of concordial classification in Dongo-Ko.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>X₄</td>
<td>categorical independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td>x-value</td>
<td>0.982</td>
<td>y-value</td>
<td>0.345</td>
</tr>
</tbody>
</table>

ad X₁: With eleven classes (cf. table 28), the system is a small one.

ad X₂: The markers attach to the noun itself (cf. 48), adjectives (cf. 38 B), the proximal and distal form of the demonstrative (cf. 49 A and B), the possessive (cf. 44), and the verb (cf. 49 C and D); furthermore, këlə “other”, the interrogative “which”, and the relative marker agree with their head noun’s class (Pasch 1986: 229).

(49) A limê-mê  m-ê
tear-11:SG T़:SG-DEM.CLOSE
“this tear (here)”

B limê-sê  s-ê
tear-4:PL T़:PL-DEM.DISTANT
“that tears (there)” (Dongo-Ko; cf. Pasch 1986: 234)

C mâti-yê  ê-ô-ti
bow-2:PL T़:ANIMATE-DEF-fall
“The bows fall.”

D bêlê-nyê  z-ô-ti
dog-2a:PL T़:ANIMATE-DEF-fall

ad X₃: The markers occur among others on the verb (cf. 49 C and D). Pasch (1986) does not report on marking in absence of the classified noun, yet 46 suggests a general availability of class marking in that case.

ad X₄: As case marking is typically obligatory and as the Dongo-Ko class markers are fused with the case markers, concordial classification in Dongo-Ko is obligatory.

ad X₅: The classes are marked by prefixes (cf. table 28).
Appendix

ad X₆: The concordial class system of Dongo-Ko allows multiple marking in the same context (cf. 50).

(50) bì-gó kɔdɔkɔdɔ-bì-gó
leopard-7:SG bad-ATTRIBUTIVE-7:SG
"a bad leopard" (Dongo-Ko; cf. Pasch 1986: 232)

ad X₇: All nouns are assigned to a concordial class (Pasch 1986: 181).

ad Y₁: The system in general lacks a semantic underpinning (Pasch 1986: 194). Nevertheless, a number of tendencies of earlier semantically motivated class membership is observable: The class pair 1a/2a hosts mostly animates, the class pair 3/4 hosts body part and fruit terms as well as derived animal names (which can be further designated by the suffixes i- (f.) and ki- (m.); cf. Pasch 1986: 210). The class pairs 5/6 also hosts body part terms, language names (cf. Pasch 1986: 217), and measure terms (Mengenbegriffe). The class pair 11/2 also hosts measure terms and abstracts (cf. Pasch 1986: 194). The class pair 7/2a hosts animates, which may be derived or simple nouns; it comprehends agents (e.g. tɔ=nà-gó “hunter”, tɔ-è “hunt”), persons with abstract properties (e.g. bá=màndú-ò “poor person”, màndú-ò “poverty”), ethnic denotations of individuals (e.g. màmbùlè-gó “pygmy”), and denotations of individual animals (esp. insects; e.g. mbòmbò=lá-gó “bee”; cf. Pasch 1986: 218f.).

ad Y₂: There are no transparent general semantic assignment rules, since the individual classes “partly host nouns from very differing areas” (Pasch 1986: 194). The consulted sources do not report on other productive assignment rules than semantic ones. Therefore, the assignment appears to be opaque in general.

ad Y₃: The lack of general assignment rules results in a high number of opaque assignment rules.

ad Y₄: The class markers are portmanteau morphemes, which encode information on both class and number (cf. e.g. 49).

ad Y₅: The individual ‘singular classes’ and ‘plural classes’ have distinct class markers (cf. table 28). As the same ‘plural class’ may serve more than one ‘singular class’ (cf. table 29), individual markers may denote more than one class.

ad Y₆: According to Pasch (1986: 194), the classes lack a semantic underpinning.

ad Y₇: Despite a few nouns that exceptionally display either a flexible singular or plural marking (cf. 47), the Dongo-Ko concordial class system does not all a flexible realignment of its nouns.
6.5 Grammaticalisation

Despite the lack of diachronic data, Pasch (1986) attempts a description of the development of the Dongo-Ko systems of nominal categorisation on the basis of a comparison of other Mba languages: The possessive markers í (for close kinship and certain body parts that are not categorised by a classifier) and á (employed in classifier constructions) indicate a possible earlier distinction of inherent relations (marked by í) as opposed to inalienable possession (marked by á; cf. Pasch 1986: 253). Besides these markers, the actual classifiers give Pasch (1986: 252ff.) reason to make suggestions about their differing ages. Since this description is a tentative one and not based on diachronic records or other data, it will be sketched out briefly here: The initial stage of possessive classification probably differentiated externalisable property (e.g. mbá) from other property. By the course of time, the internal classification of body part possession developed next to the externalisable markers and grew more and more specific (Pasch (1986: 254) notes that nj might be the oldest marker of body part possession). The distinction of animacy within the concordial class system is probably younger than the class suffixes and well-integrated into (if not even developed from) the class suffixes, since the marking of animacy and concordial class cooperates perfectly (except for minor differences in the classes 1a and 2a; cf. Pasch 1986: 257). The sketched development of the whole concordial class system (Pasch 1986: 372ff.) is somewhat contradictory to this assumption: Pasch assumes here an initial twofold distinction of [±animate] in Proto-Mba, which is extended to “at least ten singular and five plural classes” (ib.: 372), which is given up in its later progress, except for the Dongo-Ko branch.

The two systems at their contemporary state contrast with respect to their formalisation and transparency: The possessive classifier system is lowly formalised (value: 0.208) and highly transparent (value: 0.929), while the concordial class system is highly formalised (value: 0.982), but shows a low degree of transparency (value: 0.345).

7 Dyirbal

7.1 Background

Dyirbal (or Djirbal) is a Pama-Nyungan language spoken by less than 30 people in Northeast Queensland (“Herberton south to Herbert river headwaters to Cashmere, at Ravenshoe, Millaa Millaa, and Woodleigh, east to Tully falls; Lewis et al. 2013: s.v.). This dossier presents Traditional Dyirbal, which was documented by Dixon in 1963, and not Younger Dyirbal, which was documented by Schmidt in 1982. It is based primarily on Dixon (1968, 1972, and 1982) and
Lakoff (1990). The Dyirbal system is considered a textbook example of an Australian concordial class system, yet it displays an intermediate systemic setup: It does not employ agreement ‘proper’ and semantics play a dominant role in the distribution of its nouns to its concordial classes. This study reanalysed the Dyirbal system as a highly formalised noun classifier system.

![nominal classification devices](image)

Figure 15: Classification of the Dyirbal system of nominal classification.

7.2 Classification Typology

7.2.1 Noun Classifiers

The Dyirbal classification system features its markers adjacent to the noun, on the so-called ‘noun marker’ (cf. 51). This study analyses the Dyirbal system as a highly formalised noun classifier system. It is analysed in section 7.3.

(51) balan ñugumbil baggal yaranggu bangu
  there:NOM.II  woman:NOM  there:ERG.I  man:ERG  there:INSTR.IV
  yaranggu balgan
  stick:INST  hit:PRS/PST

“man is hitting woman with stick [sic!]” (Dyirbal; cf. Dixon 1972: 93)

7.2.2 Numeral Classifiers

None of the consulted works on Dyirbal revealed the existence of number as a part of speech in Dyirbal. Since “[m]ost Australian languages lack a separate class of numbers” (Dixon 2002: 67) and this situation is found in Dyirbal, a numeral classifier system is not existing there.

7.2.3 Possessive Classifiers

Dyirbal does not employ a possessed classifier system (cf. 52 where no classifier occurs with the possessive).

(52) gipabawal wanyuru ñugi ñavanu
  this:NOM.IV.LONGWAY  who:POSS  big:NOM  stand:PRS/PST

“Whose big thing is standing over there?” (Dyirbal; cf. Dixon 1972: 107)
7.2.4 Multiple Classifier Systems

Dyirbal employs a classification system, which attaches its markers to the so-called ‘noun marker’ in its plain, interrogative, and demonstrative form (Dixon 1968: 107). The system is thus not a multiple classifier system.

7.2.5 Concordial Class Systems

Dyirbal is commonly attributed the status of a concordial class system. This study however analyses Dyirbal as a highly formalised noun classifier system, as it lacks agreement ‘proper’.

7.3 Analysis

7.3.1 Functional Setup

Table 31: The semantic and discourse properties of the Dyirbal system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification ✓</td>
</tr>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability ×</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values ✓</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora ✓</td>
</tr>
</tbody>
</table>

ad C₁: Dixon (1968) describes a number of exceptions from the semantic assignment rules in Dyirbal. These are not used spontaneously and variable, but assign the noun permanently to a distinct class: For instance, “if a subset of a certain set of nouns has a particularly important property then it may be assigned a different noun class from the main set, to emphasise this property” (Dixon 2002: 467). Since this is not performed spontaneously (which would correspond to variable classification), this phenomenon corresponds to the definition of the expansion of the lexicon’s referential power. Dyirbal nouns are assigned to one class and one class only, the only exception being the reference to the sex of an animal, which allows the assignment of either class I or class II to respectively male or female referents (cf. Dixon 1968: 110f. and 123). Since there is no ‘sex-neutral’ class in Dyirbal, this may be accounted for by productive derivation and not variable classification.

ad C₂: The sources did not reveal information on a potential individualising function of the classificatory system of Dyirbal. Since many Australian languages, among them Dyirbal, lack numbers as a part of speech (Dixon 2002: 67), an individualising function for counting purposes is not likely to exist in Dyirbal.
ad C₃: The speaker of Dyirbal may “subtly […] ‘manipulate’ noun class to describe unusual situational events” (Dixon 1968: 111). For instance, joking reference to a hermaphrodite can be made by assigning the noun for “man” to the (feminine) class II: *balan yara* “here.NOM.F man”.

ad C₄: The Dyirbal NP typically contains a noun and a ‘noun marker’, yet “it is perfectly normal for NPs in Dyirbal to contain only a marker, or only a noun, or only an adjective” or any combination thereof (Dixon 1972: 60). Therefore it can be concluded that the ‘noun markers’, which indicate a noun’s concordial class, are used as a device for reference identification in discourse.

### 7.3.2 Formal Setup

Table 32: The formal properties of the nominal classification system of Dyirbal.

<table>
<thead>
<tr>
<th>Feature</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong> (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility</strong> (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td><strong>inventory</strong> (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td><strong>marking</strong> (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi</strong> (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity</strong> (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td><strong>assignment</strong> (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad F₁: Each noun is assigned to a concordial class with respect to the individual class features [± masculine], [± animate], [± feminine], [± fire-related], [± water-related], [± harmful], [± (bearer of) non-flesh food]. If these features do not apply, a noun is assigned to the residual class IV (cf. Dixon 1968: 124).

ad F₂: Dyirbal nouns are assigned to one class and one class only, the exception being the reference to the sex of an animal, which allows the assignment of either class I or class II to respectively male or female referents (cf. Dixon 1968: 110f. and 123).
ad F₃: Dyirbal has four concordial classes (cf. e.g. Dixon 1982: 178), which is a clearly delimited set.

ad F₄: Nouns are marked for their class membership on the ‘noun marker’ only (cf. e.g. Dixon 1968: 106), which prevents multiple marking.

ad F₅: Dyirbal nouns themselves are not marked for their concordial class; neither are predicates (cf. Dixon 1968: 106). The only source displaying a concordial class is the ‘noun marker’, which by definition occurs NP-internally.

ad F₆: Due to the dominant role of semantics in the Dyirbal concordial class assignment, the different classes are closely related to their inherent semantics. Nevertheless, a noun may change its noun class (and thus its semantic facets) only in clearly delimited semantic areas. This may again be illustrated by the (joking) reference to a hermaphrodite by the noun for “man” plus the (feminine) marker of class II: *balan yara “here.NOM.F man”* (cf. Dixon 1968: 111).

ad F₇: Before its reanalysis after the 1960s (cf. Lakoff 1990: 97 and Schmidt 1985), the Dyirbal concordial class assignment in its full extent was transparent to its speakers, which e.g. were able to uniformly and spontaneously assign new loan words to the same class (cf. Dixon 1968: 119f.).

7.3.3 Mapping

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{transparency value} & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1 \\
\hline
\text{formalisation value} & 0 & 0.2 & 0.4 & 0.6 & 0.8 & 1 \\
\hline
\end{array}
\]

Figure 16: Dyirbal.
Table 33: The formalisation and transparency value of Dyirbal.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>1</td>
<td>semantic structure</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>0.5</td>
<td>rule types</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>0.5</td>
<td>rule number</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>X₄</td>
<td>0.666</td>
<td>categorial independence</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>1</td>
<td>discreetness of markers</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>0</td>
<td>discreteness of markers</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>1</td>
<td>flexibility</td>
</tr>
<tr>
<td>x-value</td>
<td>0.667</td>
<td>y-value</td>
<td>0.75</td>
</tr>
</tbody>
</table>

ad X₁: Dyirbal’s inventory comprises four classes (cf. table 34), which renders it a small system.


<table>
<thead>
<tr>
<th>class</th>
<th>class marker</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-l</td>
<td>“masculine; animate”</td>
</tr>
<tr>
<td>II</td>
<td>-n</td>
<td>“feminine; water; fire; fighting”</td>
</tr>
<tr>
<td>III</td>
<td>-m</td>
<td>“non-flesh food”</td>
</tr>
<tr>
<td>IV</td>
<td>-</td>
<td>residue class</td>
</tr>
</tbody>
</table>

ad X₂: The class marking occurs on the ‘noun marker’ in its plain, interrogative, and demonstrative form (Dixon 1968: 107).

ad X₃: The only source marked for class is the NP-internal ‘noun marker’, which precedes the noun and combines three morphemes respectively for the noun’s case, its concordial class (besides the ablative, allative, or locative case; cf. Dixon 1982: 163), and its visibility (cf. ib. and 53).

(53) A bala-∅∅ diban-∅ ∅-ŋgu-∅∅ yibi-ŋgu ∅ yibi-ŋgu ∅ buran there-ABS-IV stone-ABS here-ERG-II woman-ERG look.at

“The woman here is looking at the stone there.” (Dyirbal; cf. Dixon 1982: 161)

B yala-∅∅ diban-∅ ∅-ŋgu-∅∅ yibi-ŋgu ∅ ba-ŋgu l buran there-ABS-IV stone-ABS there-ERG-IV man-ERG look.at

“The man there is looking at the stone here.” (Dyirbal; cf. Dixon 1968: 107)

Since “it is perfectly normal for NPs in Dyirbal to contain only a marker, or only a noun, or only an adjective” or any combination thereof (Dixon 1972: 60), and since the interrogative and demonstrative form of the ‘noun marker’ also display class marking, the Dyirbal system’s locus operandi is NP-internal, yet independent of the presence of the classified noun.
Appendix

**ad X4:** The classification is obligatory in the sense that the plain, interrogative, and demonstrative form of the ‘noun marker’ obligatorily carry a class marker. These elements are however frequently omitted, which makes the Dyirbal system ‘quasi-obligatory’.

**ad X5:** The class markers are affixed elements (cf. table 34).

**ad X6:** The marking occurs on the plain, interrogative, and demonstrative form of the ‘noun marker’, which prevents multiple marking, as these elements are not likely to cooccur in discourse.

**ad X7:** All Dyirbal nouns are assigned to one of the system’s classes (Dixon 1968: 124).

**ad Y1:** Next to its fairly specified classes I and III, the system’s classes II and IV are of a rather mixed nature (cf. table 34). Furthermore, there is a high number of mythological associations with real world referents, which leads to a reassignment of the nouns denoting those referents. This blurs the semantic structure of the classes.

**ad Y2:** Most assignment is carried out on semantic grounds. As Lakoff (1990: 95) notes, “Dixon does not [...] claim that all Dyirbal classification works by his principles”. Nevertheless, the number of nouns that displays an unexpected class membership is small and does not falsify the clear semantic distribution of the bulk of the Dyirbal nominal lexicon. Examples for exceptions from the rules may be due to opaque mythological connotations and comprise e.g. the nouns for dog, bandicoot, platypus, echidna (all class II instead of class I; cf. Lakoff 1990: 95 and Dixon 1968: 122)), and money (a loanword, which is assigned to class I for unknown reasons; cf. Lakoff 1990: 95 and Dixon 1968: 123).

**ad Y3:** Due to the small class inventory and semantically transparent assignment, the number of assignment rules is small.

**ad Y4:** The class markers mark class membership only (cf. table 34).

**ad Y5:** Each class has a distinct marker (cf. table 34).

**ad Y6:** Class membership can be either informative (e.g. *balan yara* “here (NOM.F) man” ∼ “this effeminate man” ∼ “this hermaphrodite”; cf. Dixon 1968: 111), redundant (e.g. if *yara* “man” is marked for the masculine class), or vain (e.g. for members of the residue class). Since semantics play a highly productive role in class (re-)assignment in Dyirbal, the informativeness of the markers is set at the medium point between redundant and informative marking.
7.4 Grammaticalisation
The Dyirbal system of nominal categorisation shows mixed properties of both classifier systems and concordial class systems. The formal appearance of the four class markers and their opaque origin indicates concordial class-like properties, just as the obligatory assignment of each noun to a concordial class and the inflexibility of the concordial class membership does. The strong semantic component of the system (which has been reanalysed by subsequent generations in order to keep the system transparent), the lack of agreement and the restriction to a single NP-internal occurrence of concordial class markers indicates classifier-like properties. The Dyirbal system may be preliminarily interpreted as an erstwhile classifier system, that experienced a process of formalisation. Before this process could turn the whole system into a concordial class system, it has stopped and left an intermediate system with strong formalised traits behind (value: 0.667). The system has been transparent to its speakers, as long as the traditional Dyirbal culture and mythology was a part of their speaker’s daily life. According to Schmidt (1985), this was not the case during her data collection in 1982. Consequently, the then opaque semanticity of the system had been reanalysed into a system the semantic basis of which closely resembles the one of English.

8 German
8.1 Background

German (Deutsch, Tedesco, Allemand) is a West Germanic language spoken in Central Europe, mainly in the countries of Austria, Germany, and Switzerland. It employs a tripartite concordial class system, which it inherited from Proto-Indo-European.
8.2 Classification Typology

8.2.1 Noun Classifiers

German does not employ a noun classifier system (cf. 54, where no classifier occurs).

(54) Der Mensch bewundert auftichtig nur das Unverdiente: Talent, Abstammung, Schönheit.

“The human being sincerely admires only the undeserved: talent, ancestry, beauty.”

(54) (German)

8.2.2 Numeral Classifiers

German does not employ a numeral classifier system (cf. 55, where no classifier occurs).

(55) Drei Chinesen mit 'nem Kontrabaß saßen auf der Straße...]

“Three Chinese with a doublebass sat on the street...”

(55) (German)

8.2.3 Possessive Classifiers

German does not employ a possessed classifier system (cf. 56, where no classifier occurs).

(56) Der Optimist hofft, daß seine Träume wahr werden, der Pessimist fürchtet sich davor.

(56) (German)
8.2.4 Multiple Classifier Systems

German does not employ a multiple classifier system (cf. 54, 55, and 56 where no classifier occurs).

8.2.5 Concordial Class Systems

German employs a concordial class system, which is marked by means of class agreement on a range of nominal modifiers (cf. 57 where all nominal satellites are marked for the head noun’s class). This system is analysed in section 8.3.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th></th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ein</td>
<td>braun-er</td>
<td>Hut,</td>
</tr>
<tr>
<td>a:NOM.M.SG</td>
<td>brown-NOM.M.SG</td>
<td>hat:NOM.M.SG</td>
<td></td>
</tr>
<tr>
<td>braun-e</td>
<td>Tusche,</td>
<td>ein</td>
<td></td>
</tr>
<tr>
<td>brown-NOM.F.SG</td>
<td>bag:NOM.F.SG</td>
<td>a:NOM.N.SG</td>
<td></td>
</tr>
<tr>
<td>Buch</td>
<td>book:NOM.N.SG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a brown hat, a brown bag, a brown book</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(57) ein braun-e Hut, ein-e braun-es Buch

8.3 Analysis

8.3.1 Functional Setup

Table 35: The semantic and discourse properties of the German system of nominal classification.

| differentiating referents (C1) | productive derivation, variable classification | ✓ |
| individuation (C2) | unitisation for countability | ✗ |
| attribution of properties (C3) | speaker-sided values | ✗ |
| reference identification (C4) | reference tracking via AGR, anaphora | ✓ |

This table lists the properties and their status: ✓ for present, ✗ for absent.

ad C1: German does not allow a free, variable classification of its nominal lexicon. Nevertheless, it employs a number of semantic niches that allow the speaker to lexically distinguish between referents such as cars (default gender: masculine; e.g. der BMW m.) and motorcycles (default gender: feminine; e.g. die BMW f.). Another niche are ship names, which are always feminine (cf. the
male person Bismarck m. “Bismarck” as compared to the vessel die Bismarck f.). Furthermore, different concordial class membership can help distinguishing nominal referents (e.g. der Kunde m. “the customer” and die Kunde f. “lore, news, knowledge”, der Schild m. “(battle-)shield”, das Schild n. “sign”), though this is an instance of lexicalisation and not productive. German also differentiates the sex of humans and other animates by concordial class membership (e.g. der Verantwortliche m. “the (male) responsible person” vs. die Verantwortliche f. “the (female) responsible person”).

ad C2: Unlike its sister language Dutch, the German system is not involved in individuation.

ad C3: In general, German does not allow for a speaker-sided attribution of properties to nominal referents by means of its concordial class system. Cases of systemically conflicting gender assignment (e.g. Mädchen n. “girl”, which conflicts with the semantically demanded feminine and thus is regularly referred back to by a feminine pronoun in spoken language) are not a solid fundament for assuming an exception. Yet, especially violating the ‘correct’ gender that denotes the sex of (human) animates carries a pejorative connotation, which is comparable to English constructions such as If you ladies leave my island, if you survive recruit training, you will be a weapon, where the gender mismatch is indicated by ladies, whereas it is men that are addressed. For this purpose, it is always a lexical item that varies, and not a single item’s concordial class.

ad C4: By employing agreement on a range of nominal satellites, the German concordial class system is involved in reference identification (cf. 58).

(58) Ein Auto stieß mit einem Zug zusammen. Er trug erheblichen Schaden davon. “A car collided with a train. It [the train] took considerable damage.” (German)

8.3.2 Formal Setup

Table 36: The formal properties of the nominal classification system of German.

<table>
<thead>
<tr>
<th>Obligatoriness (F₁)</th>
<th>CLS-like: nouns may remain uncategorised</th>
<th>☒</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CCS-like: obligatory assignment of each noun</td>
<td>☒</td>
</tr>
</tbody>
</table>

| Flexibility (F₂) | CLS-like: flexible class membership | ☒ |
|------------------| CCS-like: single class membership | ✓ |
Appendix

inventory (F₃)  CLS-like: largish class number ✗
CCS-like: limited class number ✓
marking (F₄)  CLS-like: marked once ✗
CCS-like: multiply marked ✓
locus operandi (F₅)  CLS-like: locus operandi: NP-internal ✗
CCS-like: locus operandi: NP-internal and -external ✓
semanticity (F₆)  CLS-like: semantic contribution ✗
CCS-like: semantically vain ✓
assignment (F₇)  CLS-like: chosen freely by the speaker ✗
CCS-like: mechanic assignment ✓

ad F₁: All German nouns belong to a concordial class. In the case of pluralia tanta (e.g. die Eltern “the parents”), this concordial class may not be determinable, while this is not necessary, since they never occur in the singular.

ad F₂: A noun typically belongs to one concordial class and one concordial class only. No noun can shift from concordial class to concordial class arbitrarily.

ad F₃: German employs three concordial classes, which is a delimited and small set by our definition.

ad F₄: A noun’s concordial class membership is expressed on a range of nominal satellites and may be multiply marked (cf. 59).

(59)  Die,i  ist  die,k  ist  die,k  ist
the:NOM.F.SG  be:PRS3SG  die:GEN.M.SG  be:PRS3SG  die:GEN.M.SG
Freiheit,i  des,l  Buchdrucks,j  Forderung,n  der,l  entstehenden,l  der,l  reifen,l
erste,m  das,m  erste,m  das,m  erste,m
first:NOM.N.SG  the:NOM.N.SG  first:NOM.N.SG  the:NOM.N.SG  first:NOM.N.SG
Opfer,m  Demokratie,l  OPfer,m  Demokratie,l  Opfer,m  Demokratie,l

“The freedom of letterpress printing is the first demand of an emerging and the first victim of a mature democracy.” (German)

ad F₅: The marking of concordial class is not limited to the boundaries of the noun phrase where the classified noun occurs (cf. 60).

(60)  A  ist  ist
Meine,Rüstung,i  ist  be:PRS1SG
my armour:NOM.F.SG  be:PRS1SG
eine,ganz,aus,Fehlern,gewebe,t,i
one fully woven from mistakes: NOM.F.SG
“My armour is one [that is] woven fully from mistakes.”

Sylvia Maria Kristel: war eine niederländische Schauspielerin.

Sylvia Maria Kristel: NOM.F SG was a Dutch actress: NOM.F SG

Sie spielte in mehr als fünfzig internationalen Spielfilmen.

She: NOM.F SG played in more than fifty international movies.

“My armour is one [that is] woven fully from mistakes.” (German)

ad F6: A noun’s class membership does not influence its semantics, the only exception being cases of lexicalised ‘multiple gender’ such as der Schild m. “the (battle-)shield” vs. das Schild n. “the sign” and the distinction of sex in epicene nouns such as der/die Gläubige “the believer (m./f.)”.

ad F7: A noun’s concordial class has to be learned, as there are no productive and transparent assignment mechanisms present in contemporary German (cf. also Weinrich 1993: 325)

8.3.3 Mapping

![Figure 18: German.](image-url)
Table 37: The formalisation and transparency value of German.

<table>
<thead>
<tr>
<th>x-criteria value</th>
<th>y-criteria value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>semantic structure</td>
</tr>
<tr>
<td>host number</td>
<td>rule types</td>
</tr>
<tr>
<td>locus operandi</td>
<td>rule number</td>
</tr>
<tr>
<td>obligatoryness</td>
<td>categorial independence</td>
</tr>
<tr>
<td>boundedness</td>
<td>discreteness of markers</td>
</tr>
<tr>
<td>marking</td>
<td>redundancy</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>flexibility</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x-value</th>
<th>y-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.929</td>
<td>0.202</td>
</tr>
</tbody>
</table>

**ad X₁:** The inventory of three classes renders the German system a small one.

**ad X₂:** Class information is expressed on nouns, adjectives, numerals, articles, and different pronouns (cf. [60]).

**ad X₃:** Since there is no marking on the verb, adverb, or other NP-external constituents, the marking occurs only NP-internally, but is not bound to the noun phrase where the classified noun occurs (cf. [61], where class marking occurs transclausally, yet NP-internally).

(61)

A Hast have:PRS2SG Du 2SG die ART_{def:ACC.F.SG} Frau:ACC.F.SG mit with der Gesichtstäubierung seen been

“A have you seen the woman with the facial tattoo?”

B Welche? INTERROGATIVE2SG:ACC.F.SG

“Which [one]?”

C Die ART_{def:ACC.F.SG} da there vorne ahead

“The [one] over there.”

**ad X₄:** As agreement morphology on nominal satellites is obligatory, the marking of the German noun’s inherent class value is obligatory, too.

**ad X₅:** A noun’s class membership is expressed by means of portmanteau affixes.

**ad X₆:** The class markers may occur multiply in the same context (cf. [57]).

**ad X₇:** Every German noun has a class value, i.e. all German nouns are inherently classified.
ad Y₁: The individual classes display neither a bijective nor an opaque semantic value, but comprise nouns from all areas (animates, inanimates, masses, abstracts, etc.).

ad Y₂: German employs semantic, morphological, phonological, and opaque assignment rules.

ad Y₃: The number of assignment rules in German is so high that each noun’s inherent class value has to be learned separately.

ad Y₄: A noun’s class membership is expressed by means of portmanteau morphemes, which additionally encode information on a noun’s case and number.

ad Y₅: The class markers are employed across the German classes (cf. table 38 for a representative example of the German articles).

Table 38: Definite and indefinite articles in German

<table>
<thead>
<tr>
<th>case</th>
<th>Mdef</th>
<th>Fdef</th>
<th>Ndef</th>
<th>PLdef</th>
<th>M indef</th>
<th>F indef</th>
<th>N indef</th>
<th>PL indef</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>der¹</td>
<td>die⁵</td>
<td>das</td>
<td>die⁵</td>
<td>ein⁶</td>
<td>eine</td>
<td>ein⁶</td>
<td>∅</td>
</tr>
<tr>
<td>GEN</td>
<td>des²</td>
<td>der¹</td>
<td>des²</td>
<td>der⁴</td>
<td>eines⁷</td>
<td>eimer</td>
<td>eines⁷</td>
<td>∅</td>
</tr>
<tr>
<td>DAT</td>
<td>dem³</td>
<td>dem¹</td>
<td>dem³</td>
<td>dem⁴</td>
<td>einem⁸</td>
<td>eimer</td>
<td>einem⁸</td>
<td>∅</td>
</tr>
<tr>
<td>ACC</td>
<td>den⁴</td>
<td>die⁵</td>
<td>das</td>
<td>die⁵</td>
<td>einen</td>
<td>eine</td>
<td>ein⁶</td>
<td>∅</td>
</tr>
</tbody>
</table>

ad Y₆: Class membership can be semantically vain (e.g. der Löffel m. “the spoon”, die Gabel f. “the fork”, das Messer n. “the knife”), redundant (e.g. der Hengst m. “the stallion”, die Stute f. “the mare”), or informative (e.g. der/die Gläubige “the believer (m./f.)”).

ad Y₇: A German noun cannot change its class at the speaker’s free will and is in general rigid. Only a few lexical groups such epicene nouns can change their class in order to e.g. mark the referent’s biological sex (e.g. der/die Gläubige). Other ‘multiple gender’ nouns are either nouns that show a different class membership in different varieties of German (e.g. New High German der Teller m. “the plate” vs. Austrian German das Teller n. “the plate”), or they are lexicalised homonyms that are distinguished by their class (e.g. der Schild m. “the (battle-)shield” vs. das Schild n. “the sign”).

8.4 Grammaticalisation

Most of the Indo-European languages represent prototypical concordial class systems. This may be due to the fact that the Indo-European languages are
very well researched from the beginnings of scientific linguistics as a research tradition on. Their study has considerably shaped our perception of nominal classification, and most subsequent studies on systems of non-Indo-European languages were performed from a Eurocentric perspective (cf. Grinevald 2000: 57). Unsurprisingly, the German concordial class system corresponds to the notion of a prototypical one, which only lacks an NP-external class marking. It is highly intransparent due to the various types of assignment rules that may apply, and which are blurring the assignment to such a degree, that a second language learner has to learn each noun’s inherent class value individually. Furthermore, the markers are not specific to an individual class, but typically occur in the inflectional paradigms of more than one class.

9 Great Andamanese

9.1 Background

Great Andamanese is a koiné from four northern varieties (Jeru, Sare, Khora, and Bo) of Andamanese, a genuine language of the Andaman islands (Abbi 2013: 9f.). This dossier is mainly based on Abbi 2011 and 2013. The Great Andamanese system of nominal classification displays a highly productive expansion to other grammatical functions such as the classification of action manners or deixis, which apply to the verbal semantics and thus are not original to nominal classification. Abbi (2011 and 2013) assumes possessive classification to be the original function of the system. In this study, the system is assumed to have undergone a functional split, which results in an employment of the same classificatory markers as nominal and verbal classifiers, depending on their host (nouns, possessives, and adjectives ∼ nominal classification, verbs and adverbs ∼ verbal classification). This dossier thus analyses the ‘nominal classification part’ of the system and only features a description of the ‘verbal classification part’.
9.2 Classification Typology

9.2.1 Noun Classifiers

Great Andamanese does not employ a noun classifier system (cf. 62, where no classifier occurs).

(62) cao-nu ra-be pa-om
  dog-PL pig-ABS bark-NONPAST
  “Dogs are barking at the pigs.” (Great Andamanese; cf. Abbi 2013: 114)

9.2.2 Numeral Classifiers

Great Andamanese distinguishes only the numerals “one” and “two” and does not employ a numeral classifier system (cf. Abbi 2013: 114f. and 63, where only possessive classifiers occur).

(63) A \( t^\text{b}u(t) = t^\text{\textit{v}ir} \) p\( \text{h} \)o
  1SG.CL4.POSS=child one
  “I have one child.”

B \( t^\text{b}u(t) = t^\text{\textit{v}irin} \) n\( \text{\textit{r}tap}^b \)ul
  1SG.CL4.POSS=child-PL two
  “I have two children.”

C \( t^\text{b}u(t) = t^\text{\textit{v}ire} \) n\( \text{\textit{r}akamu} \)
  1SG.CL4.POSS=child many
  “I have many children.” (Great Andamanese; cf. Abbi 2013: 115)

9.2.3 Possessive Classifiers

Great Andamanese employs classifiers in possessive constructions (cf. 63). These classifiers occur also on adjectives and nouns in their function as nominal classifiers (cf. 64), which renders the system a multiple classifier system.
(64) A  a-lep<sup>h</sup><sub>ai</sub>  er=cak  nol  e=cai  
  ARGUMENT-Lephai  CL:2=face  good  CL:5=bad  
  un=tabol=a  
  REFLEXIVE-naughty-PST  
  “Lephai was goodlooking, but naughty.” (Great Andamanese; cf. Abbi 2011: 774)  

B  ot=cala,  er=cala,  an=cala,  e=tei,  at=tei,  
  CL:2=scar  CL:3=scar  CL:5=blood  CL:4=blood  
  an=tei  CL:3=blood  
  “scar left by arrow-head, scar on the head, scar on the limbs, blood inside the body, blood outside the body, blood on/from finger” (Great Andamanese; cf. Abbi 2011: 764)  

9.2.4 Multiple Classifier Systems  
The Great Andamanese classifiers occur as nominal classifiers in possessive constructions and with adjectives and nouns (cf. 63 and 64). This renders it a multiple classifier system. It is analysed in section 9.3.  

9.2.5 Concordial Class Systems  
Great Andamanese does not employ a concordial class system (cf. Abbi 2013: 115 and 65, where no class agreement occurs e.g. on the adjective).  

(65) pʰr  lɔbiŋ-bi  kʰulɔl  jio  
  bamboo  long-ABS  there  exist  
  “There is a long bamboo.” (Great Andamanese; cf. Abbi 2013: 132)  

9.3 Analysis  

9.3.1 Functional Setup  

Table 39: The semantic and discourse properties of the Great Andamanese system of nominal classification.  

| differentiating referents (C<sub>1</sub>) | productive derivation, variable classification ✓ |  
| individuation (C<sub>2</sub>) | unitisation for countability ✗ |  
| attribution of properties (C<sub>3</sub>) | speaker-sided values ✗ |  
| reference identification (C<sub>4</sub>) | reference tracking via AGR, anaphora ✓ |  

ad C<sub>1</sub>: The classifiers occur as bound elements on the noun (cf. e.g. 63). Given that a part of the Great Andamanese nominal lexicon are ‘bound nouns’ which demand an obligatory classifier (Abbi 2011: 763f.), this phenomenon serves the expansion of the lexicon’s referential power. Since classifier-noun
Appendix

combinations may vary and since this variation affects the noun phrase’s semantics, a differentiation of referents by the means of classifier variation is possible in Great Andamanese. In the case of nouns that are not bound, this appears to be variable classification due to the clitic nature of the markers.

**ad C₂**: The consulted sources do not reveal information about the Great Andamanese classifier system to be involved in the individuation of nominal referents.

**ad C₃**: There is no information available which suggested that Great Andamanese allows its speaker to attribute properties to referents by classificatory means.

**ad C₄**: The Great Andamanese classifiers appear to contribute to a noun phrase’s semantics rather than to serve the purpose of reference identification.

### 9.3.2 Formal Setup

Table 40: The formal properties of the nominal classification system of Great Andamanese.

<table>
<thead>
<tr>
<th>Feature (Fₙ)</th>
<th>CLS-like:</th>
<th>CCS-like:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong> (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility</strong> (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td><strong>inventory</strong> (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td><strong>marking</strong> (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi</strong> (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity</strong> (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td><strong>assignment</strong> (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

**ad F₁**: Not all nouns are categorised under all circumstances (cf. 66 where no nominal classifier occurs).
Appendix

(66) tʰu bol-ta cokbi bi i=thudo∅
1SG bol-INSTR turtle-ABS CL:5=kill-PST
“I killed the turtle with the help of a Bol-rope.” (Great Andamanese; cf. Abbi 2013: 123)

ad F2: The same noun (referring to the same entity) may take various or no classifier according to the context (cf. 64 B and 67).

(67) A cokbi t-ot=tʰ omu
turtle INANIMATE-CL:4=flesh
“turtle meat (cut up for consumption)"
B cokbi ot=tʰ omu
turtle CL:4=flesh
“turtle meat (still on its body)"
C cokbi tʰ omu
turtle flesh
“turtle meat (general, non-specific)” (Great Andamanese; cf. Abbi 2011: 762)

ad F3: Great Andamanese employs seven classifiers (cf. table 41), which is a small and limited set in our definition.

Table 41: The inventory of the Great Andamanese possessive classifiers (cf. Abbi 2011: 745).

<table>
<thead>
<tr>
<th>class</th>
<th>marker</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a-</td>
<td>“mouth and its semantic extensions” (9%)</td>
</tr>
<tr>
<td>2</td>
<td>e-r-</td>
<td>“major external body parts” (36%)</td>
</tr>
<tr>
<td>3</td>
<td>oŋ-</td>
<td>“extreme ends of the body (e.g. toes, fingernails)” (6%)</td>
</tr>
<tr>
<td>4</td>
<td>ut-</td>
<td>“bodily products, part-whole relationships” (13%)</td>
</tr>
<tr>
<td>5</td>
<td>e-</td>
<td>“inner organs” (17%)</td>
</tr>
<tr>
<td>6</td>
<td>ara-</td>
<td>“parts designating round shape and sexual organs” (12%)</td>
</tr>
<tr>
<td>7</td>
<td>o- o-</td>
<td>“parts for legs and related terms” (7%)</td>
</tr>
</tbody>
</table>

ad F4: Though examples of multiple occurrence of classifiers in Great Andamanese exist, this is not an instance of multiple marking. The classifiers’ scope is only over its morphosyntactic host and depends on its semantics, not the semantics (or form) of the head noun (cf. 68).

(68) aka=pʰ up at=tʰ ey e=cey
his=sputum CL:4=odour CL:5=bad
“He’s bad smelly sputum.” (Great Andamanese; cf. Abbi 2013: 244)

ad F5: The Great Andamanese classifiers occur in possessive constructions (cf. 63) on adjectives, and nouns themselves (cf. 68). They also occur on verbs and adverbs, where they categorise the action, not a nominal referent (cf. table
In their function as nominal classifiers, they thus occur only NP-internally.


<table>
<thead>
<tr>
<th>class</th>
<th>classifier</th>
<th>verbal semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a-</td>
<td>“mouth-related activity, origin”</td>
</tr>
<tr>
<td></td>
<td>e.g. a=jire “abuse”, a=kap “sprout”</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>er-</td>
<td>“action involving the front part of the body”</td>
</tr>
<tr>
<td></td>
<td>e.g. era=luk “weigh”</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>aq-</td>
<td>“hand-related activity”</td>
</tr>
<tr>
<td></td>
<td>e.g. aq=c “stitch”, an=tuwe “trembling of hands”</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ut-</td>
<td>“directional, experiential”</td>
</tr>
<tr>
<td></td>
<td>e.g. ot=c “leave”, t\textsuperscript{b}e=bom “be hungry”</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>e-</td>
<td>“action involving the interior of an object”</td>
</tr>
<tr>
<td></td>
<td>e.g. e=leko “suck”, e=mrno “tear”</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ara-</td>
<td>“action involving the middle portion of the body”</td>
</tr>
<tr>
<td></td>
<td>e.g. ara-delo “be pregnant”</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>o-</td>
<td>“resultative state”</td>
</tr>
<tr>
<td></td>
<td>e.g. o=mrno “make nest”, o=beo “sting”</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>adverbial semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

(69)

A \textsuperscript{a} am=p\textsuperscript{b-o}, er=p\textsuperscript{b-o}, ut=p\textsuperscript{b-o}, er=bate, ek=bate, ut=bate
CL:6=cut CL:2=cut CL:4=cut CL:2=slap OBJ=slap CL:4=slap
“cut down/fell, hit with a stick (in the front), separate from the source, slap on the face, slap suddenly, slap hard” (Great Andamanese; cf. Abbi 2011: 770)

B e-jone, ut=jone, e=jet\textsuperscript{b}, e=colol, e=biye
CL:5=scare CL:4=scare CL:1=stop CL:5=roll CL:5=think
“be afraid, get startled, feel nauseated/uneasy, roll down, think” (Great Andamanese; cf. Abbi 2011: 772)

C ut=t\textsuperscript{b}-e, er=t\textsuperscript{b}-e
CL:4=break-IMP CL:2=break-IMP
“Tattoo the back of the body. Tattoo the forehead.” (Great Andamanese; cf. Abbi 2011: 771)
The Great Andamanese classifiers may contribute to the semantics of the noun phrase (cf. 70).

(70) A: $ot=ca\ddot{a}$, $er=ca\ddot{a}$, $\alpha\gamma=ca\ddot{a}$, $e=tei$, $ot=tei$, 
\[ CL:4=\text{scar} \quad CL:2=\text{scar} \quad CL:3=\text{scar} \quad CL:5=\text{blood} \quad CL:4=\text{blood} \]
\[ \alpha\gamma=tei \]
\[ CL:3=\text{blood} \]
“scar left by arrow-head, scar on the head, scar on the limbs, blood inside the body, blood outside the body, blood on/from finger”

B: $ara=m\ddot{i}k^b\text{-}tei$, $ot=tei$, $e=\text{teye}$
\[ CL:6=\text{middle-blood} \quad CL:4=\text{blood} \quad CL:2=CL:5=\text{measles} \]
“stomach ache, splitting headache, measles”

C: $f\ddot{a}ro$ $\text{ter}=\text{li}k^b\text{-}u(i)$, $f\ddot{a}ro$ $\text{tara}=#\ddot{a}x=rel$
\[ \text{sea} \quad CL:2=\text{lap} \quad \text{sea} \quad CL:6=\text{green/blue} \]
“deep sea, open sea” (Great Andamanese; cf. Abbi 2011: 764)

D: $\text{buruj}$ $\text{ter}=p^b\text{-}el$, $\text{foc}$ $\text{ta}=p^b\text{-}eq$
\[ \text{mountain} \quad CL:2=\text{back} \quad \text{vessel} \quad CL:1=\text{cavity} \]
“the back of the mountain, the mouth of the vessel” (Great Andamanese; cf. Abbi 2013: 765)

Both lexicalised class membership and variable class membership (with inter-speaker variation; cf. Abbi 2011: 761) exist in Great Andamanese. Depending on the constructional context and the degree of grammaticalisation of the individual classifier, the assignment may be more or less flexible.

9.3.3 Mapping

![Figure 20: Great Andamanese.](image-url)
Table 43: The formalisation and transparency value of Great Andamanese.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td>x-value</td>
<td>0.649</td>
<td>y-value</td>
<td>0.714</td>
</tr>
</tbody>
</table>

ad X₁: The system’s seven categories render it a small system (cf. table 41).

ad X₂: Next to verbs and adverbs, which constitute the morphosyntactic hosts for the ‘verbal classification part’ of the Great Andamanese system (cf. table 42), classifiers may occur on adjectives, nouns, and in possessive constructions as nominal classifiers (cf. 63 and 64).

ad X₃: Due to the functional expansion as verbal classifiers (cf. table 42), the markers occur outside of the noun phrase.

ad X₄: Bound nouns and certain possessive contexts obligatorily require the insertion of a classifier (Abbi 2011: 754 and 762f.).

ad X₅: The markers occur in the form of bound clitics (cf. 70).

ad X₆: The markers have a scope over their morphosyntactic host (cf. 68). Seeming occurrences of multiple marking such as in 71 are thus actually instances of multiple classification.

(71) k₃t e=mèle e=nō thit=cəyə o
soil CL:5=fine CL:5=good earth.OBJ=find-PST
“He found good and fine soil on the ground.” (Great Andamanese; cf. Abbi 2013: 244)

ad X₇: Each class originally denotes a distinct division or area of the body. The classes are expanded through by ‘ethno-semantic’ associations (Abbi 2011: 746). In general, only nouns denoting entities that are inherent to another entity can be categorised (Abbi 2011: 768).
The individual classes are semantically straightforward per host, i.e. the same classifier varies its semantic connotation with respect to the part of speech that it attaches to (cf. tables 44 and 45). This merging of semantic domains across morphosyntactic contexts results in a mixed structure of the individual classes.


<table>
<thead>
<tr>
<th>class</th>
<th>bodily reference</th>
<th>possessive axioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a-</td>
<td>“mouth and semantic extensions”</td>
<td>“(extension of) mouth, source”</td>
</tr>
<tr>
<td>2 er-</td>
<td>“major external body parts”</td>
<td>“external organs, deictic”</td>
</tr>
<tr>
<td>3 or-</td>
<td>“extreme ends of the body”</td>
<td>“(extension of) hand, extremities”</td>
</tr>
<tr>
<td>4 ut-</td>
<td>“bodily products, part-whole”</td>
<td>“(extension/products of) self”</td>
</tr>
<tr>
<td>5 e-</td>
<td>“inner organs”</td>
<td>“internal organs”</td>
</tr>
<tr>
<td>6 ara-</td>
<td>“round shape, sexual organs”</td>
<td>“relational, circular, curved”</td>
</tr>
<tr>
<td>7 o-~o-</td>
<td>“parts for legs and related terms”</td>
<td>“lower body part”</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>class</th>
<th>spatial reference</th>
<th>reference point</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a-</td>
<td>“surface”</td>
<td>“front”</td>
</tr>
<tr>
<td>2 er-</td>
<td>“anterior, exterior”</td>
<td>“front, out”</td>
</tr>
<tr>
<td>3 or-</td>
<td>“posterior, superior”</td>
<td>“up”</td>
</tr>
<tr>
<td>4 ut-</td>
<td>“interior, centre”</td>
<td>“in”</td>
</tr>
<tr>
<td>5 e-</td>
<td>“periphery”</td>
<td>“edge”</td>
</tr>
<tr>
<td>6 ara-</td>
<td>“inferior”</td>
<td>“down”</td>
</tr>
</tbody>
</table>

The transparent class semantics allow a semantic assignment in many cases. There are however opaque cases of classification (cf. 70 C)

The class semantics are transparent, but differ with respect to their morphosyntactic host. The individual semantic paradigms are homogeneous (cf. table 41). The inventory is a small one, yet there are opaque cases of class assignment (cf. 70 C). All these factors in combination result in a medium number of assignment rules in Great Andamanese.

The markers convey classificatory semantics only (cf. table 41).

Each class has its own, distinct marker (cf. table 41).
ad Y₆: Considering the available data, the bulk of classifier insertions involves a semantic manipulation of its host noun phrase (cf. 70).

ad Y₇: Depending on the individual noun and the discoursal context, the classification is flexible or even absent (cf. 67 and 70).

9.4 Grammaticalisation

All in all, the Great Andamanese system of nominal classification displays an intermediate formalisation value (0.649) and a value just above the ‘transparency threshold’ (0.714). There are a few properties of the Great Andamanese classifier system that are not prototypical for a classifier system. These are the intransparent origin of its classifiers, the small number of referential classes, and the assignment of nominal referents to a respective class, which is fixed in some cases, whereas it is depending on the speaker’s discourse demands in others. The expansion to morphological hosts such as bound nouns, verbs, adverbs, and adjectives, and the development of a specific semantic structure for each morphosyntactic host indicates a functional split in the system’s history (cf. figure 21).

The individual classifiers display a differing degree of grammaticalisation (e.g. the class 6 and 7 markers are more grammaticalised than the class 5 marker; cf. Abbi 2011: 777) and so do the individual functional sets of classifiers (e.g. the classifiers cooccurring with adverbs and verbs are more grammaticalised than those cooccurring with nouns and adjectives; cf. ib.). The number of categorised nouns (and adjectives, verbs, and adverbs; cf. Abbi 2011: 777) outweighs the number of these elements that are not subject to categorisation. If a further shift into the more formal area of the functional continuum took
Appendix

place, a loss of optional categorisation would be expected (i.e., all nouns are categorised by the classifiers from a certain moment on), and agreement would be expected to be introduced rather than the occurrence of e.g. classifier-adjective combinations in which the classifier categorises the adjective.

10 Hadza

10.1 Background

< Hatsa < Hadza (ISO 639-3: hts; cf. Lewis et al. 2013: s.v.)
< Khoisan (cf. Ruhlen 1987)

Hadza is a language spoken by one of the last remaining hunter-gatherer-tribes in north central Tanzania (southeast of Lake Eyasi; cf. Blench 2008: 1). It is perceived as a language isolate and not belonging to the Khoesan languages by the Hadza specialists (cf. ib.). This dossier is based on Blench 2008, Edenmyr 2004, and Sands 2013. The Hadza concordial class system meets all requirements of a prototypical concordial class system.

10.2 Classification Typology

10.2.1 Noun Classifiers

Hadza does not employ a noun classifier system (cf. [72] where no classifier occurs).

(72) bami-ya, pa’anakwete, ‘ashapo sa mana-ko
3SG.M-3SG.M.COP old.man:M.SG expert:M.SG meat-3SG.F

“he is an old man, a real meat expert” (Hadza; cf. Bala 1998: 26, cited in Sands 2013: 268)
10.2.2 Numeral Classifiers

Hadza does not employ a numeral classifier system. Cf. \[73\] where no classifier occurs (the consulted sources do not provide an example with a numeral; the quantifier *wa’ina-* “all” however behaves in the same way as numerals do; cf. Sands 2013: 111).

\[73\]  

\begin{verbatim}
Tuma-ko mana-ko habi ts’if \( \text{wa’ina-echa.} \)
\end{verbatim}

“Where is the meat all these nights.” (Hadza; Bala 1998: 26, cited in Sands 2013: 111)

10.2.3 Possessive Classifiers

Hadza does not employ a possessed classifier system (cf. \[74\] where no classifier occurs).

\[74\]  

\begin{verbatim}
A A hadisi-ko kota bahe Iyeye- mika, and story-3SG.F 3SG.F.PST end Iyeye 3SG-M:GEN-3SG.F
\end{verbatim}

“And the story of Iyeye is finished.” (Hadza; cf. Bala 1998, cited in Sands 2013: 113)

\begin{verbatim}
B hits’a bek’ahu- hicke
\end{verbatim}

“the fat of the elephant” (Hadza; cf. Berger 1943: 115, cited in Sands 2013: 113)

10.2.4 Multiple Classifier Systems

Hadza does not employ a multiple classifier system (cf. \[72\] \[73\] and \[74\] where no classifier occurs).

10.2.5 Concordial Class Systems

Hadza employs a concordial class system, which manifests itself in discourse by means of agreement on a range of nominal satellites (cf. \[72\] \[73\] and \[74\]).

10.3 Analysis

10.3.1 Functional Setup

| differentiating referents (C₁) | productive derivation, variable classification | ✓ |
| individuation (C₂) | unitisation for countability | ✗ |
| attribution of properties (C₃) | speaker-sided values | ✗ |
| reference identification (C₄) | reference tracking via AGR, anaphora | ✓ |
The Hadza system knows a certain fluidity of nouns and classes: The speaker may assign a noun referring to a human or higher animal to the class denoting natural sex (cf. Edenmyr 2004: 5 and 9). The system also enables the speaker to expand the lexicon’s referential power in a group of nouns that are mostly semantically close to each other, but discriminate e.g. their referent’s size by classificatory means (Edenmyr 2004: 15).

The consulted sources did not reveal the Hadza system to be involved in the individuation of referents.

The consulted sources did not reveal information on the speaker’s ability to attribute properties to referents by classificatory means.

Hadza employs agreement on a range of nominal satellites including pronouns and verbs. Therefore, the Hadza class marking serves the purpose of reference identification (cf. 75).

(75)  
A  /ˈa-nə’a  
kiss-3SG.M.OBJECT\_direct\_1SG  
“I kissed him.”
B  /ˈi-na-nə’a  
kiss-3SG.F.OBJECT\_direct\_1SG  
“I kissed her.”
C  /ˈtə-nə’a  
kiss-3PL.M.OBJECT\_direct\_1SG  
“I kissed them.” (male reference)
D  /ˈi-təh-nə’a  
kiss-3PL.F.OBJECT\_direct\_1SG  
“I kissed them.” (female or mixed reference; Hadza, compiled after Sands 2013: 120f.)

10.3.2 Formal Setup

Table 47: The formal properties of the nominal classification system of Hadza.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like:</th>
<th>CCS-like:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong></td>
<td>nouns may remain uncategorised</td>
<td>mark once</td>
</tr>
<tr>
<td>( F_1 )</td>
<td>obligatory assignment of each noun</td>
<td></td>
</tr>
<tr>
<td><strong>flexibility</strong></td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>( F_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>inventory</strong></td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>( F_3 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>marking</strong></td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>( F_4 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Appendix

<table>
<thead>
<tr>
<th>locus operandi (F₅)</th>
<th>CLS-like: locus operandi: NP-internal</th>
<th>×</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS-like: locus operandi: NP-internal and -external</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>semanticity (F₆)</th>
<th>CLS-like: semantic contribution</th>
<th>×</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS-like: semantically vain</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>assignment (F₇)</th>
<th>CLS-like: chosen freely by the speaker</th>
<th>×</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS-like: mechanic assignment</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

**ad F₁:** Every Hadza noun is assigned to the masculine or feminine class (cf. Sands 2013: 108).

**ad F₂:** In general, the Hadza noun belongs to one single class (as can be deduced from Edennyryr 2004: 5). There are two niches that allow a certain degree of flexibility with respect to the concordial class membership: Nouns referring to humans (and animals) may take the gender of their real life referent (cf. Edennyryr 2004: 5 and 9) and a certain group of ‘noun pairings’. Among the about 1400 nouns that were investigated by Edennyryr (2004), there are about 40 noun pairs with differing gender and -though closely related- in meaning, too. While the greater part of these nouns seems to reflect a discrimination of size (the feminine gender marking the bigger of the two parts), “[s]ome of these do not conform to any obvious pattern of semantic differences” and “do not always point in the same direction” (both Edennyryr 2004: 15). Therefore, the speaker cannot be assumed to to be able to use this pattern productively.

**ad F₃:** Hadza employs two concordial classes (cf. e.g. Edennyryr 2004: 5 and table 48), which is a smallish and delimited class set.

Table 48: The inventory of the Hadza concordial class markers (cf. Sands 2013: 109)

<table>
<thead>
<tr>
<th>class</th>
<th>marker</th>
<th>copula markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>F.SG &amp; collective</td>
<td>-ko</td>
<td>-ako</td>
</tr>
<tr>
<td>F.PL &amp; paucal</td>
<td>-be’e</td>
<td>-pe’e</td>
</tr>
<tr>
<td>M.SG &amp; collective</td>
<td>ø</td>
<td>-a</td>
</tr>
<tr>
<td>M.PL &amp; paucal</td>
<td>-bi’i</td>
<td>-pi’i</td>
</tr>
</tbody>
</table>

**ad F₄:** Hadza employs agreement and thus allows for multiple marking in the same context (cf. 76).

(76) | uwi-yako | k’uni-ko | n’apa-ko | pakapa’a-ako |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mosquito-3SG.F.COP</td>
<td>small-3SG.F</td>
<td>Ndorobo-3SG.F</td>
<td>big-3SG.F.COP</td>
<td></td>
</tr>
</tbody>
</table>

“[A] mosquito is small, Ndorobo fly is big.” (Hadza; cf. Sands 2013: 110)

**ad F₅:** Among others, the class markers occur on the verb (cf. 75). Class marking is thus not restricted to the noun phrase.
ad F₆: Leaving aside animates, where the discrimination of grammatical gender may mean a discrimination of the referent’s natural sex, there are only perceivable tendencies of an underlying semantics of the concordial classes. Edenmyr (2004: 15f.) provides (seemingly unproductive) noun pairings: The same nominal stem shows a difference in meaning depending on its assigned class. However, as there are no discrete and systematic semantics derivable, one must conclude that these cases display the semantic relics of an earlier stage. At the current point of description, there is no identifiable semantic contribution to the noun phrase that is performed by the means of the concordial class system.

ad F₇: The speaker has no influence on a noun’s class membership (besides the niche of human and higher animate referent nouns, where he or she can assign the concordial class that corresponds to the referent’s natural sex; cf. again Edenmyr 2004: 5 and 9).

10.3.3 Mapping

![Diagram](image.png)

Figure 23: Hadza.
Table 49: The formalisation and transparency value of Hadza.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obligatoryness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
</tbody>
</table>

x-value 1

y-value 0.583

**ad X₁:** The Hadza system possesses two concordial classes (cf. table 48), which renders it a small system.

**ad X₂:** Hadza marks a noun’s class membership on the noun itself and several nominal modifiers such as demonstratives, pronouns, possessives, adjectives, verbs, and numerals (cf. Sands 2013: 108ff.).

**ad X₃:** Class marking occurs among others on the verb (cf. 75), which is an NP-external element. Personal pronouns are also marked for class (cf. table 50). As personal pronouns typically replace lexical nouns in discourse, the NP-internal presence of class marking does not depend on the presence of the classified noun.

Table 50: The Hadza personal pronouns (cf. Sands 2013: 112)

<table>
<thead>
<tr>
<th>class</th>
<th>independent form</th>
<th>copula form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG.M</td>
<td>'ono</td>
<td>'one-ne’e</td>
</tr>
<tr>
<td>1SG.F</td>
<td>'onoko</td>
<td>'one-neko</td>
</tr>
<tr>
<td>2SG.M</td>
<td>te</td>
<td>te-te’e</td>
</tr>
<tr>
<td>2SG.F</td>
<td>teko</td>
<td>te-teko</td>
</tr>
<tr>
<td>3SG.M</td>
<td>bami</td>
<td>bami-ya</td>
</tr>
<tr>
<td>3SG.F</td>
<td>boko</td>
<td>boko-wako</td>
</tr>
<tr>
<td>1PL.Mincl</td>
<td>'unibi’i</td>
<td>‘unibi’i</td>
</tr>
<tr>
<td>1PL.Mexcl</td>
<td>‘ubi’i</td>
<td>‘ubi-upi’i</td>
</tr>
<tr>
<td>1PL.Fincl</td>
<td>‘onebe’e</td>
<td>‘onebe’e</td>
</tr>
<tr>
<td>1PL.Fexcl</td>
<td>‘obe’e</td>
<td>‘obe-ape’e</td>
</tr>
<tr>
<td>2PL.M</td>
<td>'iti’bi</td>
<td>'iti’bi</td>
</tr>
<tr>
<td>2PL.F</td>
<td>‘etebe’</td>
<td>‘etebe’</td>
</tr>
<tr>
<td>3PL.M</td>
<td>‘i’</td>
<td>‘i’</td>
</tr>
<tr>
<td>3PL.F</td>
<td>be’e</td>
<td>be’e</td>
</tr>
</tbody>
</table>

**ad X₄:** As the markers are fused with the categories of number and case (cf. table 48), class marking on nominal modifiers in discourse is obligatory.
ad X5: The markers occur either as affixes (cf. 76) or in fused form (cf. table 50).

ad X6: The Hadza system allows multiple marking in the same context (cf. 76).

ad X7: All Hadza nouns are assigned to one of the concordial classes (Sands 2013: 108).

ad Y1: The Hadza classes display strong semantic tendencies to host certain related groups of nouns. The class structure is a mixed one, as several semantic criteria apply. All of these criteria represent strong tendencies that know a few exceptions; the following description is based on Edenmyr (2004: 9-16). Nouns referring to humans are either masculine or feminine (e.g. kinship terms) or may flexibly take the concordial class that corresponds to the noun’s real life referent. If reference to a group is made, the Hadza speaker either combines a male and a female plural noun or makes use of the generic feminine. If a noun has a distinct semantics with respect to its referent’s sex, the class marker occasionally is omitted (e.g. lu'ja(-ko) “woman”). Names for animals may be distinguished by their referent’s natural sex, but do have an unmarked gender.

Trees and plants, tubers, and edible roots are typically feminine, whereas most parts of plants, berries and fruits, as well as the closely related category of food is masculine. Body part terms often occur in both masculine and feminine in a similar meaning, but hair and body cover, genitalia, blood vessels, intestines, bones, and diseases and ailments are masculine, whereas muscles are feminine. The distinction in terms of shape seems to play a role within this group as well. Substances such as liquids and food are masculine, the only systematic exception being honey, the terms of which are “rather evenly distributed” (Edenmyr 2004: 13). Weather phenomena are masculine, whereas geographical terms that are related to water (as well as holes and caves) are feminine. Other geographical landmarks are distributed over both classes. The words ti’ja-ko “sun” and ts’i’ja-ko “day” are feminine, their counterparts seta “moon” and ts’i’(i) “night” are masculine. Other times and seasons are masculine. Nouns indicating directions are feminine.

Besides and within the semantic groupings of the Hadza nouns, two tendencies for the distribution of the Hadza nouns over the concordial classes play a role: long and thin objects tend to fall in the masculine class, whereas short and stout objects are assigned to the feminine class. Furthermore larger objects prefer the feminine class, and smallish objects the masculine class.

6This is the masculine for carnivorous mammals, bush- and reedbucks, doves and pigeons, herons and storks, mongooses, and bees and wasps. Large and domestic grass-eaters, small duikers and dik-diks, and spiders and termites are feminine. Lower animals appear to be categorised by their shape, where long and thin animals are assigned to the masculine gender (e.g. snakes and long, thin insects).
ad Y₂: Semantic assignment is clearly involved in the Hadza classification system. As all semantic principles represent tendencies with unmotivated exceptions to them, opaque assignment is also present. The consulted sources do not report on other assignment mechanisms.

ad Y₃: The strong semantic underpinning of the Hadza classes is never fully met; each semantic principle knows exceptions (cf. Edenmyr 2004: 9-16). This results in a medium number of assignment rules.

ad Y₄: The markers are fused with the categories case and number (cf. table 48).

ad Y₅: Each class has its own, distinct markers (cf. the tables 48 and 50).

ad Y₆: The Hadza classes possess a strong semantic expressive force, which however cannot be productively applied by the speaker at large (cf. Y₇). A noun’s class membership thus conveys in many cases semantics, which can be either redundant or informative.

ad Y₇: In general, the Hadza noun belongs to a single class (as can be deduced from Edenmyr 2004: 5). There are two niches that allow a certain degree of flexibility with respect to the concordial class membership: Nouns referring to humans (and animals) may take the gender of their real life referent (cf. Edenmyr 2004: 5 and 9) and a certain group of ‘noun pairings’. Among the about 1400 nouns that were analysed in Edenmyr (2004), there were about 40 noun pairs with differing gender and -though closely related- in meaning, too. While the greater part of these nouns seems to reflect a discrimination of size (the feminine gender marking the bigger of the two parts), “some of these do not conform to any obvious pattern of semantic differences” and “do not always point in the same direction” (both Edenmyr 2004: 15). Therefore, one cannot assume the speaker to be able to use this pattern productively.

10.4 Grammaticalisation
The Hadza concordial class system appears to be a formerly semantically productive system of nominal categorisation, as a number of (non-general) semantic criteria (e.g. the class groupings of words for carnivorous animals, substances, weather phenomena etc.; cf. Edenmyr 2004: 9ff.) shows. These recognisable semantic areas are not exclusive, i.e. there are exceptional nouns to every area, that do not belong to the respective predictable concordial class. Other criteria, such as the assignment of concordial class due to an item’s shape (long, thin, or smallish items tend to correspond to the masculine gender, short, stout, and largish items to the feminine gender) seem to have played an important role in the assignment at earlier stages, but cannot be productively used to
Appendix

evoke a semantic discrimination (cf. Edenmyr 2004: 15f.). The extensive agreement marking (in and beyond the NP-borders), the obligatory assignment of each noun to a concordial class, the smallish class number, and the inflexibility of the vast part of the Hadza nouns mirror its high degree of formalisation.

11 Hausa

11.1 Background

Hausa is a major African language that is spoken in Western Central Africa, more precisely in “Sokoto, Kaduna, Katsina, Kano, Bauchi, Jigawa, Zamfara, Kebbi, and Gombe states. Also in Benin, Burkina Faso, Cameroon, Central African Republic, Chad, Congo, Eritrea, Germany, Ghana, Niger, Sudan, Togo” (Lewis et al. 2013: s.v.). This dossier is mainly based upon Abdoulaye 1992, Jaggar 2001, and Newman 2000. The Hausa system of nominal classification is another example for a prototypical concordial class system.

![Diagram of nominal classification devices](http://example.com/diagram.png)

Figure 24: Classification of the Hausa system of nominal classification.

11.2 Classification Typology

11.2.1 Noun Classifiers

Hausa does not employ noun classifiers (cf. where no classifiers occur).

(77) Zaki, bari, demisa, bari, jafanku na manga, wa shike shiga, sai wawa, sai mahaukaci?

“Lion, stop! Leopard, stop! Your quarrelling is the quarrelling of the mighty ones. Who is going to mix himself up in it except a fool, except a madman?” (Hausa; cf. Jaggar 2001: 709)
11.2.2 Numeral Classifiers

In Hausa, “it is normal [with many nouns] to use the singular form with numerals and other quantifiers even when the noun has a corresponding plural form” (Newman 2000: 382). According to Newman (2000: 382), this trend seems to decrease in favour of the use of plural nouns in suchlike constructions (cf. 78).

(78)  
A  kàtīf haufa  
mattress:F.SG four  
“four mattresses” (Hausa; cf. Newman 2000: 382)  
B  kàtīf haufa  
mattress:PL four  
“four mattresses” (Hausa; compiled after Newman 2000)

As the above example illustrates, no classifier-like elements occur in quantified constructions.

11.2.3 Possessive Classifiers

Hausa does not employ a possessive classifier system (cf. 79, where no classifier occurs, and Jaggar 2001: 332ff.).

(79)  
A  zant na Hâdziâ  
cloth:M.SG of:M.SG Hadiza  
“a cloth of Hadiza’s”  
B  mûtâ ta huyu  
car:F.SG of:F.SG two  
“the second car”  
C  wàsànni na yârâ  
game:PL of:PL children  

11.2.4 Multiple Classifier Systems

Hausa does not employ a multiple classifier system (cf. 77, 78 and 79 where no classifier occurs, and Jaggar 2001: 332ff.).

Note that the combination of numerals and singular nouns is the typical environment of numeral classifiers. The occurring change might be explained by Rijkhoff’s (e.g. 2002a and 2002b) theory of a nominal aspect, which would currently be reanalysed in the case of Hausa. Rijkhoff (2002b: 214) defines nouns by two features of the referent they denote, [+shape] and [+homogenity]. If a language’s nouns are perceived as wearing the feature [+shape], they are perceived as being definitely outlined in the spatial dimension and can directly be pluralised, whereas the feature [−shape] cannot be directly connected with a numeral, but needs a classifier in order to ‘unitise’ the nominal concept referred to (ib.).
11.2.5 Concordial Class Systems

Hausa employs a concordial class system (cf. [80], where class membership of the noun is marked by an agreeing auxiliary). This system is analysed in section [11.3].

(80) A ṛagō nē. STABILISER:M.SG đākā nē. STABILISER:M.SG
      "It’s a ram. It’s a room."

B tunīyā tāgā cē. STABILISER:F.SG window:F.SG
      "It’s a sheep. It’s a window." (Hausa; cf. Newman 2000: 216)

11.3 Analysis

11.3.1 Functional Setup

Table 51: The semantic and discourse properties of the Hausa system of nominal classification.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents</td>
<td>productive derivation, variable classification</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>indiduation</td>
<td>unitisation for countability</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>attribution of properties</td>
<td>speaker-sided values</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>reference identification</td>
<td>reference tracking via AGR, anaphora</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

ad C1: The Hausa nominal lexicon employs some 75 nouns of epicene class (cf. [86], i.e. the noun’s class is contextually assigned with respect to the nominal referent’s natural sex.

(81) wādā, wūwā, sābiyā, kāmānā
      “dwarf, fool, albino, feeble person” (Hausa; cf. Newman 2000: 201)

ad C2: The consulted sources do not suggest that the Hausa system is involved in the individuation of nominal referents.

ad C3: Concordial class assignment in Hausa is carried out by the system, not the speaker (Jaggar 2001: 48), who thus has no opportunity to attribute properties to a nominal referent by classificatory means.

ad C4: Among others, the Hausa system marks a noun’s concordial class membership on pronouns. Therefore, it may be assumed to serve the purpose of reference identification (cf. [82] where the class marking allows the identification of the intended referent).
(82) \textit{gayà mətə bə shi də âmfiə} \\

Telling her[, that (m.)] is of no use.” (Hausa; cf. Jaggar 2001: 289)

11.3.2 Formal Setup

Table 52: The formal properties of the nominal classification system of Hausa.

<table>
<thead>
<tr>
<th>Feature</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>obligatory assignment of each noun</td>
<td>✓</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>single class membership</td>
<td>✓</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>limited class number</td>
<td>✓</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>multiply marked</td>
<td>✓</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>locus operandi: NP-internal and -external</td>
<td>✓</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>semantically vain</td>
<td>✓</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>mechanic assignment</td>
<td>✓</td>
</tr>
</tbody>
</table>

\textbf{ad F₁:} In Hausa, “[a]ll singular NPs have gender” (Newman 2000: 200).

\textbf{ad F₂:} A Hausa noun’s membership of a concordial class is normally inflexible: “Out of a list of just over 5,000 common nouns, some 3,000 are masculine, 2,000 are feminine, and approximately 75 are epicene” (Newman 2000: 200).

\textbf{ad F₃:} Hausa employs two concordial classes (cf. e.g. Newman 2000: 200 and table 53), which is a clearly delimited set.

Appendix

ad F₄: A noun’s concordial class is marked on the Hausa ‘stabilisers’ (cf. 80), genitive linkers and independent possessives (cf. 79), the definite article (cf. 83A), the personal, demonstrative and nonpersonal pronouns (cf. 83 B and table 53), the adjective (cf. 83 C), and the diminutive marker (cf. 83 D and Newman 2000: 216ff.).

B wọ-n-can, DEICT-LINKER:M.SG-there.DISTANT wọ-c-can, DEICT-LINKER:F.SG-there.DISTANT wọ-dan-can DEICT-LINKER:PL-there.DISTANT “that [distant male], that [distant female], those [distant]” (Hausa; cf. Jaggar 2001: 324)

ad F₅: Since class is marked on a subject-agreement pronoun that partly represents the inflecting constituent of the verbal complex (cf. Jaggar 2001: 148), the marking occurs NP-internally and -externally (cf. 84).


ad F₆: Besides its reference to the biological sex of human and higher animate referents, the Hausa concordial class system does not productively contribute to the noun phrase (cf. Newman 2000: 201).

ad F₇: While there are several lexical groups that show a consistent semantic gender assignment (e.g. names of countries and cities; cf. Newman 2000: 203), this is an exception within the Hausa concordial class system. “[O]vert grammatical gender on simple nouns is lexically determined” (Jaggar 2001: 48) for the majority of nouns. If it comes to loanwords that cannot be analogically assigned to one of the semantic groups that are involved in semantic gender assignment, phonology is the basic assignment criterion (final -a(a) triggers assignment to the feminine class; cf. Jaggar 2001: 49).
11.3.3 Mapping

![Graph showing formalisation value and transparency value for Hausa.]

Figure 25: Hausa.

Table 54: The formalisation and transparency value of Hausa.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td>x-value</td>
<td>1</td>
<td>y-value</td>
<td></td>
</tr>
</tbody>
</table>

**ad X₁:** The two concordial classes of Hausa render it a small system (cf. 53).

**ad X₂:** Hausa marks a noun’s class membership on a broad range of nominal satellites (cf. F₄ and 83).

**ad X₃:** The marking occurs among others on the ‘stabiliser’, which is part of the VP (cf. 84).

**ad X₄:** The consulted sources and available do not suggest that the classificatory marking, which is virtually always fused with other categories, is not obligatory.
ad X₅: The marking occurs in affixed or fused form (cf. [83]).

ad X₆: Hausa allows multiple marking in the same context (cf. [85]).

(85) Audà nê yàfì kàntì.
Audu STABILISER.M.SG 3SG.M-PF go.to shop
“It’s Audu who’s gone to the shop.” (Hausa; cf. Jaggar 2001: 150)

ad X₇: All singular nouns have a class value (Newman 2000: 200).

ad Y₁: The class of the major part of the underived Hausa nominal lexicon is not predictable by their meaning (cf. Newman 2000: 201). The phonological form is the most reliable indicator of a noun’s class, despite a fair number of exceptions (cf. Newman 2000: 208f.). Nevertheless, there are several semantic areas that play a role in the class structure: The classes distinguish the biological sex of higher animates (cf. [86]).

(86) A bìjùnī, zàkarā, sarkī, ùbā
large.bull:M.SG cock:M.SG emir:M.SG father:M.SG
“large bull, cock, emir, father”
B kàza, kàharī-bà-kèjì, yà, uuâ
“hen, prostitute, elder sister, mother” (Hausa; cf. Newman 2000: 201)

Furthermore, several lexical groups of nouns have a default class: Geographical locations (names of countries, states, cities, towns, their quarters, and rivers), prayer times, temporal adverbials in nominal use, the cardinal directions, numerals, and the letters of the Roman alphabet are all feminine (cf. [87] and Newman 2000: 203ff.). Infinitive phrases and sentences, adverbial and prepositional phrases, as well as terms for fractions and the zero, months and weekdays besides saturday are masculine (cf. [88] and Newman 2000: 203ff.).

(87) A Afìrka, Báryn, Bága-daça, Dàmbàzau, Bàdunë
B àsùbâhì, ìzahàrì, lá’asàrì
first.prayer:F.SG second.prayer:F.SG third.prayer:F.SG
mùgàìrìhà, ìṣàhà
fourth.prayer:F.SG fifth.prayer:F.SG
“early morning prayer, early afternoon prayer, late afternoon prayer, sunset
prayer, evening prayer”
C yàlù, gòbé, jìyà, bàna,
today:F.SG tomorrow:F.SG yesterday:F.SG this.year:F.SG
bàìf
next.year:F.SG
“today, tomorrow, yesterday, this year, next year”
Appendix

D  kudù
   south:F.SG
   “south” (Hausa; cf. Newman 2000: 204)

E  biyu,  tålkin,  dārī
   two:F.SG  thirty:F.SG  hundred:F.SG
   “It’s two. It’s thirty. It’s on hundred.” (Hausa; cf. Jaggar 2001: 56)

F  bā,  tā,  jā,  hā,  kā

(88)  A  gayā  matā  bā  shī  dā  āmsfā
tell  INDIRECT.OBJECT.MARKER:3SG.F  NEG  3SG.M  with  use
   “Telling her[, that (m.]) is of no use.” (Hausa; cf. Jaggar 2001: 289)

B  à  zàune  nē  mukē
   in  sitting:M.SG  STABILISER:M.SG  we.were
   “Seated we were.” (Hausa; cf. Newman 2000: 205)

C  râbī,  ṭɑbū,  sîfîrī
   half:M.SG  quarter:M.SG  zero:M.SG
   “a half, a quarter, zero” (Hausa; cf. Newman 2000: 206)

D  Mārīs,  Yûlī,  Mūḥaḥrīm,  Lāḥâdi,
   Tālātā
   tuesday:M.SG
   “march, july, first Islamic month, sunday, tuesday” (Hausa; cf. Newman 2000: 204)

ad Y2: Semantics and phonology are both involved in the class assignment (cf. e.g. Newman 2000: 201ff.), class membership is mainly a lexical property and thus opaquely assigned (cf. Jaggar 2001: 48).

ad Y3: As class membership is mainly a lexical property (Jaggar 2001: 48), the number of assignment rules is high.

ad Y4: The class markers are fused with the category of number, and in their respective realisation virtually always an additional category (e.g. definiteness; cf. 83).

ad Y5: The singular class marking distinguishes class-specific markers in the second and third person (cf. e.g. table 53).

ad Y6: Class membership is informative in the few epicene nouns. Furthermore, it is redundant in the case of e.g. distinction of biological sex, as this information is already present in the lexical noun. The major part of the Hausa lexicon shows no semantic underpinning (cf. Newman 2000: 201).
ad Y₇: Except for a small number of epicene nouns, class membership is rigid (cf. Newman 2000: 200).

### 11.4 Grammaticalisation

Due to the marginal role of the semantics involved in the Hausa system of nominal categorisation, the reduced forms of its concordial class markers, its small class inventory, and the high degree of formal (i.e. phonological) criteria involved, the Hausa system displays an advanced degree of formalisation (value: 1). At the same time, it is fairly intransparent (value: 0.357).

### 12 Hmong Daw

#### 12.1 Background

Hmong Daw (White Hmong, Bai Miao, Hmoob Dawb) is spoken in the southern provinces of China and the northern areas of Laos, Thailand, and Vietnam (cf. Bisang 1993: 45). It is closely related to the dialect of Green Hmong (Hmoob Ntsuab, cf. ib.). Within this dossier, the language is referred to as Hmong. This dossier is mainly based on Bisang 1993 and 1999, Mottin 1978, and Aikhenvald 2000; Jaissé 1987 was not available for the author. Hmong Daw employs a comparatively small multiple classifier system, which discriminates seven categories.

![nominal classification devices diagram]

Figure 26: Classification of the Hmong Daw system of nominal classification.

---

8Hmong is usually classified as a Sino-Tibetan language, only rarely as an Austic language. Ruhlen (1987) lists it under its alternative name “Miao”.

12.2 Classification Typology

12.2.1 Noun Classifiers

The Hmong classifiers may occur adjacent to the bare noun, mainly for the purpose of marking of discoursal salience (cf. Bisang 1993: 30f. and 89). Since the same classifier set occurs in other morphosyntactic contexts as well (cf. 90 and 91), the system is a multiple classifier system, not a noun classifier system.

(89) **Mas** tus **tsen txawm ga ccev ccev los,** […]
then **CL:ANIMATE** tiger then **fly** swiftly **come**

“These tiger [+given] arrived swiftly, […]” (Hmong Daw; cf. Bisang 1993: 27)

12.2.2 Numeral Classifiers

The Hmong classifiers occur in quantificational constructions (cf. 90). Since the same classifier set occurs in other morphosyntactic contexts as well (cf. 89 and 91), the system is a multiple classifier system, not a noun classifier system.

(90) A **ib tug neeg**
one **CL:ANIMATE** man/person

“One man/person” (Hmong Daw; cf. Bisang 1999: 122)

B **ntau lab tsen**
many **CL:GENERAL** house

“many houses” (Hmong Daw; cf. Mottin 1978: 29)

12.2.3 Possessive Classifiers

The Hmong classifiers may occur in possessive constructions (cf. 91). Since the same classifier set occurs in other morphosyntactic contexts as well (cf. 89 and 90), the system is a multiple classifier system, not a noun classifier system.

(91) A **nus tus txiv ntzawm tus ntzais**
he **CL:ANIMATE/POSS** uncle **CL:ANIMATE/POSS** daughter

“the daughter of his uncle” (Hmong Daw; cf. Bisang 1988: 115, cited from Bisang 1993: 30)

B **nus rab riam ntaj**
he **CL:ARTEFACT/POSS** sword


12.2.4 Multiple Classifier Systems

The Hmong classifiers occur in a variety of morphosyntactic constructions (cf. 89, 90 and 91). This renders it a multiple classifier system. This system is analysed in section 12.3.

12.2.5 Concordial Class Systems

The consulted literature provides no evidence for a classification system alongside the multiple classifier system in Hmong Daw.
12.3 Analysis

12.3.1 Functional Setup

Table 55: The semantic and discourse properties of the Hmong Daw system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
<th>Involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification</td>
<td>✗</td>
</tr>
<tr>
<td>individualization (C₂)</td>
<td>unitisation for countability</td>
<td>✓</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
<td>✗</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
<td>✓</td>
</tr>
</tbody>
</table>

**ad C₁:** The Hmong Daw classifiers are independent elements and as such not involved in nominal derivation. The system allows for a certain variability with respect to a nominal referent’s classification (e.g. by variation of the general and a specific classifier). It is however not obvious that this variation serves the function of referent differentiation in a productive way. The classifier set is small, its semantics seemingly not very productive, and the classifiers interact with a referent’s discourse properties, which all are arguments at the least not in favour of an additional referent-differentiating function. Nevertheless, examples of seeming referent differentiation can be found in Mottin (1978; cf. 92) and (due to the lack of semantic potency) best related to the individualising function of the system. Note that the ‘classifier’ kauj “coil” is not part of the classificatory set listed in Bisang (1993).

(92)  
A  
\[ L̄ub \text{ kauj } x̄ov \]  
\[ \text{one } \text{ CL:COIL thread} \]  
“a cotton real”

B  
\[ L̄ub \text{ tx̄oj } z̄ov \]  
\[ \text{CL:1.DIMENSIONAL thread} \]  
“thread” (Hmong Daw; cf. Mottin 1978: 31 and 37)

**ad C₂:** The Hmong classifiers serve the function of individuation (cf. Bisang 1999: 117).

**ad C₃:** The consulted literature does not suggest the Hmong system to be involved in the speaker-sided attribution of properties by classificatory means.

**ad C₄:** The Hmong Daw classifiers are used as carriers of anaphora (cf. 93) and thus serve reference tracking.

(93)  
A  
\[ L̄ub \text{ tsev twg? } L̄ub \text{ no. } \]  
\[ \text{CL:GENERAL house which CL:GENERAL DEM} \]  
“Which house? That one.”
Appendix

12.3.2 Formal Setup

Table 56: The formal properties of the nominal classification system of Hmong Daw.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like:</th>
<th>CCS-like:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong></td>
<td>Nouns may remain uncategorised</td>
<td>Obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility</strong></td>
<td>Flexible class membership</td>
<td>Single class membership</td>
</tr>
<tr>
<td><strong>inventory</strong></td>
<td>Large class number</td>
<td>Limited class number</td>
</tr>
<tr>
<td><strong>marking</strong></td>
<td>Marked once</td>
<td>Multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi</strong></td>
<td>Locus operandi: NP-internal</td>
<td>Locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity</strong></td>
<td>Semantic contribution</td>
<td>Semantically vain</td>
</tr>
<tr>
<td><strong>assignment</strong></td>
<td>Chosen freely by the speaker</td>
<td>Mechanic assignment</td>
</tr>
</tbody>
</table>

**ad F1**: Depending on the context, nouns in Hmong Daw may remain uncategorised (cf. where no classifiers occur).

(94) *Sawv daws ris pob zeb mus wa txaj zeb rau pem ntav toj.*

Everybody carry stone go make pile stone at up middle mountain

“Everybody carried stones up to the mountain, in order to build piles of stones there.” (Hmong Daw; cf. Bisang 1988: 123, cited from Bisang 1993: 25)

**ad F2**: Within certain limits, it is possible to categorise the nouns of Hmong Daw by different classifiers (cf. 95).

(95) A  

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>neeg</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>ib</td>
<td>tag</td>
</tr>
<tr>
<td></td>
<td>one</td>
<td>CL: ANIMATE</td>
</tr>
<tr>
<td></td>
<td>“a man/person” (Hmong Daw; cf. Bisang 1999: 122)</td>
<td></td>
</tr>
</tbody>
</table>

B  

<table>
<thead>
<tr>
<th></th>
<th>neeg</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>leej</td>
</tr>
<tr>
<td></td>
<td>CL: HUMAN</td>
</tr>
<tr>
<td></td>
<td>“the man/person” (Hmong Daw; cf. Mottin 1978: 39)</td>
</tr>
</tbody>
</table>
Appendix

ad F₃: Hmong distinguishes seven categories (cf. table [57]).

Table 57: The inventory of Hmong Daw classifiers (cf. Bisang 1993: 31ff.).

<table>
<thead>
<tr>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>lub</td>
<td>general classifier</td>
</tr>
<tr>
<td>tus</td>
<td>“animate; long object”</td>
</tr>
<tr>
<td>kej</td>
<td>“human”</td>
</tr>
<tr>
<td>rab</td>
<td>“artefacts”</td>
</tr>
<tr>
<td>troj</td>
<td>“long, 1-dimensional object; abstract”</td>
</tr>
<tr>
<td>daim</td>
<td>“flat, 2-dimensional object”</td>
</tr>
<tr>
<td>txhais</td>
<td>“one of a natural pair”</td>
</tr>
</tbody>
</table>

Bisang (1993: 10) notes that there are elements that occur in the function of classifiers only with a few nominals. These are thooj (combining with siab “liver, seat of affections”), qhov (combining with chaw “land”), zaj (combining with kwv txhiaj “song” and npau suav “dream”), and tsab (nouns denoting messages). Bisang further notes that “we find some quantifiers which are almost identical to classifiers in combination with “underspecified nouns” such as ntawv ‘paper’ or las ‘speech’” (Bisang 1993: 10; cf. [96]), but does not assume them to be classifiers (cf. Bisang 1993: 24, referring to Ratliff 1989).

(96)  A ib phau QUANT ntawv paper
      “one book”

B ib lo QUANT las speech
      “one word” (Hmong; cf. Bisang 1993: 10)

ad F₄: The classifiers occur once per context (cf. e.g. [90]).

ad F₅: The Hmong Daw classifiers may occur as carriers of anaphora without their referent noun beyond the borders of the clause (cf. [93] Aikhenvald 2000: 329 and Mottin 1978: 29). The consulted sources and available data do not suggest the classifiers to occur outside of the noun phrase.

ad F₆: The classifiers do not contribute to the noun phrase’s general semantics (cf. [97] where tsav “tiger” occurs once with a classifier and once without a classifier without a change in meaning).

(97)  Muaj ib tug tsav caum los. […] Tsv ma ab noj noj
      there is one CL:ANIMATE tiger chase come tiger take eat eat
tag.
      all.of.it
      “There was a tiger who was running after it. When the tiger had eaten all of it, […]” (Hmong Daw; cf. Bisang 1988: 17, cited from Bisang 1999: 27)
The semantic function of the classifier is to individualise a transnumeral noun in order for it to be countable (cf. Bisang 1993: 3).

**ad F7:** The semantics of the Hmong Daw system of nominal categorisation are very transparent and even may vary in a few cases (cf. 95). Therefore, the speakers are able to assign nouns to referential classes on their own.

### 12.3.3 Mapping

![Figure 27: Hmong Daw.](image)

**Table 58: The formalisation and transparency value of Hmong Daw.**

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>0.75</td>
<td>semantic structure</td>
<td>0.75</td>
</tr>
<tr>
<td>host number</td>
<td>1</td>
<td>rule types</td>
<td>1</td>
</tr>
<tr>
<td>locus operandi</td>
<td>0.5</td>
<td>rule number</td>
<td>1</td>
</tr>
<tr>
<td>boundedness</td>
<td>0.666</td>
<td>categorial independence</td>
<td>1</td>
</tr>
<tr>
<td>marking</td>
<td>0</td>
<td>discreteness of markers</td>
<td>1</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>1</td>
<td>redundancy</td>
<td>0.5</td>
</tr>
</tbody>
</table>

| x-value          | 0.559  | y-value          | 0.798  |

**ad X1:** The seven classes of the Hmong system render it a small one (cf. table 57).
Appendix

ad X₂: The Hmong classifiers may occur not only with numerals (98 A), but also in possessive constructions (98 B), with demonstratives (98), quantifying adjectives (98 D), les mots interrogatifs impliquant une réponse définie ou numérale “the interrogative words that imply a definite or numeral answer” (Mottin 1978: 29; 98 E and F), and also (singular, definite) nouns themselves (98 G; cf. ib. and Aikhenvald 2000: 206f. and 215). Furthermore, classifiers may serve as an instance of anaphoric reference (cf. 98 E and F and Aikhenvald 2000: 329).

(98) A ob lub tsev
two CL:GENERAL house
“two houses”
B lawv lub tsev
3PL CL:GENERAL house
“their house”
C lub tsev no
CL:GENERAL house DEM
“that house there”
D ntav lub tsev
many CL:GENERAL house
“many houses”
E Lub tsev twg? Lub no.
CL:GENERAL house which CL:GENERAL DEM
“Which house? That one.”
F Pes tsaug lub? 3 lub.
how many CL:GENERAL three CL:GENERAL
“How many [houses]? Three.” (Hmong Daw; cf. Mottin 1978: 29)
G lub tsev
CL:GENERAL house
“the house (definite)”

ad X₃: The classifiers occur NP-internally and independent of the presence of the classified noun (cf. 98).

ad X₄: Nouns may remain uncategorised (cf. 94 where no classifier occurs).

ad X₅: The classifiers are morphosyntactically independent elements (cf. 98).

ad X₆: The classifiers occur once per context (cf. e.g. 98).

ad X₇: All Hmong nouns can be classified, as the system employs a general classifier (cf. table 57).
ad Y₁: Most of the classes are semantically straightforward. Individual classes (e.g. *tus* “animate; long object”; cf. table 57) show a mixed structure.

ad Y₂: Semantic assignment rules are sufficient to master the Hmong system (cf. table 57).

ad Y₃: The Hmong system employs a small number of assignment rules due to the small class inventory and its transparent semantics (cf. table 57).

ad Y₄: The classifiers convey classificatory semantics only (cf. table 57).

ad Y₅: Each classifier denotes a single class (cf. table 57).

ad Y₆: The system is semantically transparent, but does not semantically contribute to the noun phrase (cf. 97, where the same noun occurs with and without classifier without changing its meaning).

ad Y₇: There is little room for a variation of classifiers in Hmong (cf. table 57). The most obvious possibility for a flexible classification is thus the alternation of the general and specific classes.

12.4 Grammaticalisation

The Hmong Daw system of nominal classification has obviously experienced a process of grammaticalisation, as the lack of semanticity of the classifiers and the extension of their constructional hosts shows. Bisang (1993: 5f.) furthermore assumes a differing degree of grammaticalisation for different NP-constituents. These are listed in the following and order by an increasing degree of grammaticalisation: nouns > class nouns > quantifiers > intrinsic quantifiers > classifiers (Bisang 1993: 5f.). Class nouns are highly abstract nouns such as *ntoo* “tree”, *txiv* “fruit”, etc., and are important for word formation (ib.: 5). Quantifiers are elements that denote measures (e.g. *mev* “metre”), collectives (e.g. *pab* “group”), and kinds (e.g. *yam* or *hom*, both “(a) kind (of)”; ib.). Intrinsic quantifiers are measure nouns in our understanding, such as *lump* in a *lump* of sugar (e.g. *tsob* “bush”).

Hmong Daw is a multiple classifier language, since the classifiers may be hosted by a whole range of nominal satellites (demonstratives, possessives, numerals, etc.) and the noun itself. The extension of its morphological environment indicates the decrease or loss of the original categoriality of the Hmong Daw classifiers as a part of speech. An ongoing process of grammaticalisation could theoretically give rise to an agreement system in Hmong Daw. The system’s properties at their contemporary state indicate that we are dealing with a far progressed intermediate type.
13 Hungarian

13.1 Background

< Uralic (ISO 639-3: hun; cf. Lewis et al. 2013: s.v.)
< Ugric < Finno-Ugric < Uralic < Uralic-Yukaghir (cf. Ruhlen 1987)

Hungarian (Magyar) is the national language of Hungary and also spoken in parts of North America, other parts of Europe, and Asia (cf. Lewis et al. 2013: s.v.). This dossier is mainly based on Beckwith 1992, Csirmaz/Dékány 2010a and 2010b, Kenesei et al. 1998, and Rounds 2009. The Hungarian multiple classifier system in all contexts, but furthermore corresponds to the definition a classifier system.

![nominal classification devices](image)

Figure 28: Classification of the Hungarian system of nominal classification.

13.2 Classification Typology

13.2.1 Noun Classifiers

Hungarian does not employ a noun classifier system (cf. 99), where no classifier occurs).

(99) A nap el-olvaszt-ott-a a hav-at.
    ART sun PFX-melt-PAST-DEF3.SG ART snow-ACC
    “The sun melted the snow.” (Hungarian; cf. Kenesei et al. 1998: 195)

13.2.2 Numeral Classifiers

Hungarian employs a classifier system that allows the insertion of classifiers in quantificational contexts, but also features them together with demonstratives (cf. 105). This renders it a multiple classifier system in the definition of this dossier.

(100) A három fej saláta
      three CL:HEAD lettuce
    “three heads of lettuce” (Hungarian; cf. Csirmaz/Dékány 2010a: 5)
13.2.3 Possessive Classifiers

Hungarian does not employ a possessed classifier system (cf. 101), where no classifier occurs).

(101) Az a kávé az enyém.
that ART coffee the mine
“That coffee is mine.” (Hungarian; cf. Kenesei et al. 1998: 273)

13.2.4 Multiple Classifier Systems

Hungarian employs a multiple classifier system (cf. 105). This system is analysed in section 13.3.

13.2.5 Concordial Class Systems

Hungarian does not employ a noun class system (cf. 102, where no concordial class agreement occurs).

(102) A szobá-ban Anna sok könyvet olvasott.
ART room-INESSIVE Anna many book.ACC read
“In the room, Anna read many books.” (Magyar; cf. Kenesei et al. 1998: 176).

13.3 Analysis

13.3.1 Functional Setup

Table 59: The semantic and discourse properties of the Hungarian system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Classifier (C)</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents</td>
<td>(C1)</td>
<td>×</td>
</tr>
<tr>
<td>individuation</td>
<td>(C2)</td>
<td>✓</td>
</tr>
<tr>
<td>attribution of properties</td>
<td>(C3)</td>
<td>×</td>
</tr>
<tr>
<td>reference identification</td>
<td>(C4)</td>
<td>✓</td>
</tr>
</tbody>
</table>

ad C1: The Hungarian classifiers are independent elements and as such do not take part in word formation. Despite a selectional restriction existing between the lexical entry and the classifier to be inserted, there is a certain degree
of speaker-sided flexibility in classifier assignment, which serves the purpose of individuation (cf. 103) and not variable classification or re-presentation.

(103) A két szem kukorica
    two CL:EYE sweetcorn
    “two grains of sweetcorn”

B két cső kukorica
    two CL:TUBE sweetcorn
    “two ears of sweetcorn” (Hungarian; cf. Csirmaz and Dékány 2010b: 8)

ad C₂: Though optional, the classifiers serve the means of individuation, since they mainly occur in quantificational contexts (cf. 459).

ad C₃: The consulted sources and available data do not suggest the Hungarian classifiers to be involved in the speaker-sided attribution of properties by classificatory means.

ad C₄: The Hungarian classifiers may serve as carriers of anaphora (cf. 104 on page 459 B).

(104) A az a szem gyöngy
    that ART CL:EYE pearl
    “that pearl”

B az a szem
    that ART CL:EYE
    “that one” (Magyar; cf. Csirmaz and Dékány 2010a: 11)

13.3.2 Formal Setup

Table 60: The formal properties of the nominal classification system of Hungarian.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>obligatory assignment of each noun</td>
<td>✗</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>single class membership</td>
<td>✗</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td>limited class number</td>
<td>✓</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>multiply marked</td>
<td>✗</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>locus operandi: NP-internal and -external</td>
<td>✗</td>
</tr>
</tbody>
</table>
Appendix

**semanticity** (F₆)  
CLS-like: semantic contribution ✓  
CCS-like: semantically vain ✗

**assignment** (F₇)  
CLS-like: chosen freely by the speaker ✓  
CCS-like: mechanic assignment ✗

**ad F₁:** The Hungarian noun is optionally categorised in all contexts (cf. 105).

(105)  
A hét (fej) saláta  
seven CL:HEAD lettuce  
“seven lettuces”

B hét (szem) cukor  
seven CL:EYE candy  
“seven pieces of candy”

C hét (szál) gyertya  
seven CL:THREAD candle  
“seven candles” (Magyar; cf. Csirmaz and Dékány 2010a: 12)

**ad F₂:** The same noun can take different classifiers (cf. 103).

**ad F₃:** The eighteen available classifiers render the Hungarian system a small one (cf. table 61).

Table 61: The inventory of Hungarian classifiers (cf. Csirmaz and Dékány 2010b: 16; asterisked forms are doubted by my informant).

<table>
<thead>
<tr>
<th>classifier</th>
<th>gloss</th>
<th>categorised items</th>
</tr>
</thead>
<tbody>
<tr>
<td>bokor</td>
<td>“bush”</td>
<td>*potato, raspberry, rose</td>
</tr>
<tr>
<td>csik</td>
<td>“article”</td>
<td>garlic, orange, mandarin</td>
</tr>
<tr>
<td>csérp</td>
<td>“pot”</td>
<td>pot plants</td>
</tr>
<tr>
<td>cső</td>
<td>“tube”</td>
<td>sweet corn, red/green pepper</td>
</tr>
<tr>
<td>darab</td>
<td>“piece”</td>
<td>any noun (according to my informant, only segmentable objects and animals, but not e.g. humans)</td>
</tr>
<tr>
<td>fej</td>
<td>“head”</td>
<td>cabbage, onions, lettuce, kohlrabi, cauliflower, broccoli</td>
</tr>
<tr>
<td>fő</td>
<td>“head”</td>
<td>people in regimented situations (e.g. staff; not for animals)</td>
</tr>
<tr>
<td>gerezd</td>
<td>“clove”</td>
<td>garlic, orange, mandarin</td>
</tr>
<tr>
<td>ív</td>
<td>“sheet”</td>
<td>paper</td>
</tr>
<tr>
<td>*karika</td>
<td>“ring”</td>
<td>sausage</td>
</tr>
<tr>
<td>kötet</td>
<td>“volume”</td>
<td>books, bound volumes</td>
</tr>
<tr>
<td>ríza</td>
<td>“rose”</td>
<td>cauliflower, broccoli</td>
</tr>
<tr>
<td>rúd</td>
<td>“rod”</td>
<td>chitterlings and salami</td>
</tr>
<tr>
<td>szál</td>
<td>“thread”</td>
<td>hair, match, sausage, salami, flower, candle, green onion, cigarette, plank, carrot, Welt</td>
</tr>
<tr>
<td>szem</td>
<td>“eye”</td>
<td>all nuts and berries, tomato, pepper, biscuit, pearl, sand, potato, cucumber</td>
</tr>
<tr>
<td>*tő</td>
<td>“stem”</td>
<td>grape, rose, nursing, any specific plant that has a nursing</td>
</tr>
<tr>
<td>veckő</td>
<td>“loaf”</td>
<td>bread</td>
</tr>
</tbody>
</table>
ad F₄: The classifiers occur once per context (cf. 105).

ad F₅: The classifiers occur within the borders of the NP (cf. 105).

ad F₆: The classifiers may semantically contribute to their host NP (cf. 103).

ad F₇: Though “a selection restriction is operative between the lexical noun and the sortal or group classifier” (Csirmaz and Dékány 2010a: 7), the speaker is able to vary the classifier (at least between a specific and the generic classifier).

13.3.3 Mapping

![Graph showing the mapping of formalisation and transparency values for Hungarian.](image)

Figure 29: Hungarian.

Table 62: The formalisation and transparency value of Hungarian.

<table>
<thead>
<tr>
<th>x-criteria value</th>
<th>y-criteria value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size X₁ 0.75</td>
<td>semantic structure Y₁ 1</td>
</tr>
<tr>
<td>host number X₂ 0.25</td>
<td>rule types Y₂ 1</td>
</tr>
<tr>
<td>locus operandi X₃ 0.5</td>
<td>rule number Y₃ 1</td>
</tr>
<tr>
<td>obligatoriness X₄ 0</td>
<td>categorial independence Y₄ 1</td>
</tr>
<tr>
<td>boundedness X₅ 0</td>
<td>discreetness of markers Y₅ 1</td>
</tr>
<tr>
<td>marking X₆ 0</td>
<td>redundancy Y₆ 0.5</td>
</tr>
<tr>
<td>exhaustivity X₇ 0</td>
<td>flexibility Y₇ 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x-value</th>
<th>y-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.214</td>
<td>0.929</td>
</tr>
</tbody>
</table>

ad X₁: The Hungarian system employs eighteen classes (cf. table 61).
ad X₂: The classifiers occur with quantifiers and demonstratives (cf. 104 and 105).

ad X₃: The classifiers occur NP-internally and independent of the presence of the classified noun (cf. 104).

ad X₄: The classifiers occur optionally in all contexts (cf. 105).

ad X₅: The classifiers occur as morphosyntactically unbound elements (cf. 61).

ad X₆: The classifiers occur once per context (cf. 104).

ad X₇: The specific classifiers are not able to categorise the bulk of the Magyar lexicon. The generic classifier *darab* “piece” can categorise every count noun that does not refer to a human (cf. table 61).

ad Y₁: The semantic structure of the classes is straightforward (cf. table 61).

ad Y₂: The consulted sources and available data suggest a purely semantic and speaker-sided assignment on basis of the classified referent’s shape or configuration (cf. table 61).

ad Y₃: The small class inventory and semantic assignment results in a low number of assignment rules.

ad Y₄: The markers convey class semantics and are not fused with other categories (cf. table 61).

ad Y₅: Each class has its own, distinct marker (cf. table 61).

ad Y₆: The classification is optional (cf. 105), which indicates that the classifiers do not necessarily semantically contribute to the noun phrase in an informative way. There are however examples where this is the case (cf. 103), which results in a mixed redundancy value.

ad Y₇: The same noun can take different classifiers (cf. 103).
13.4 Grammaticalisation

The Hungarian classifier system displays a very low degree of formalisation (value: 0.214) and a high degree of transparency (value: 0.929): Its classifiers take the same form as their source nouns and they occur optionally. There are more or less fixed combinations of specific classifiers and nouns that they are able to categorise, yet several classificatory elements such as szál “thread” and szem “eye” have a more general semantic scope. In addition, there is a general classifier, darab “piece”. All things considered, the Hungarian numeral classifier system shows parallels to the recently emerged Persian classifier system, which suggests the system to be a fairly recent development.

14 Ingush

14.1 Background

Ingush (Galgay, Kistin) is a Caucasian language which is spoken in the North Caucasus (Lewis et al. 2013: s.v.). The main source for this dossier is Nichols 2011. Since the documentation of Ingush is relatively sparse and since its grammar is much alike the one of Chechen (Lewis et al. 2013: s.v., Geiger et al. 1959: 20; the concordial class system is practically the same; cf. Nichols 1996: 21), this dossier additionally builds on Nichols 1996. Ingush employs a concordial class system, which is exceptional in that all types of modifiers that agree with the noun only allow a small percentage of their members to display the agreement markers (Nichols 2011: 114).

nominal classification devices

CLSs

nominal
numeral

possessive

CLS

CCSs

Figure 30: Classification of the Ingush system of nominal classification.
14.2 Classification Typology

14.2.1 Noun Classifiers
Ingush does not employ a noun classifier system (cf. 106 where no classifier occurs).

(106) *Muusaa shie cwea kinashja dieshazh vaagha*
Musa 3SG.REFLEXIVE alone book read:CONVERB.D sit:PRS.V
“Musa is sitting by himself reading a book.” (Ingush; cf. Nichols 2011: 199)

14.2.2 Numeral Classifiers
Ingush does not employ a numeral classifier system (cf. 107 where no classifier occurs).

(107) *cwa sag, qa sag, vi’ sag, ji’ jisha, di’ sahwat*
one person three person four:V four:J sister four:D hour
“one person, three persons, four persons, four sisters, four hours” (Ingush; cf. Nichols 2011: 197f.)

14.2.3 Possessive Classifiers
Ingush does not employ a possessive classifier system (cf. 108 where no classifier occurs).

(108) *sy vosh, sy kuorta, meaga ch’egaljg, desha mawan*
1SG.GEN brother 1SG.GEN head bread:GEN piece:DIMINUTIVE word:GEN meaning
“my brother, my head, (small) piece of bread, the word’s meaning” (Ingush; cf. Nichols 2011: 417)

14.2.4 Multiple Classifier Systems
Ingush does not employ a multiple classifier system (cf. 106 107 and 108 where no classifier occurs).

14.2.5 Concordial Class Systems
Ingush employs a concordial class system, which employs agreement with parts of its adjectives, verbs, and postpositions (cf. Nichols 2011: 431ff.). This system is analysed in section 14.3
14.3 Analysis

14.3.1 Functional Setup

Table 63: The semantic and discourse properties of the Ingush system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification ✓</td>
<td></td>
</tr>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability ×</td>
<td></td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values ×</td>
<td></td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora ✓</td>
<td></td>
</tr>
</tbody>
</table>

ad C₁: Ingush employs separate classes for male and female animates (Nichols 2011: 143) and thus allows the speaker a differentiation of referents by classificatory means. The differentiation occurs in form of derivation (ib.: 145).

ad C₂: The consulted sources do not suggest the Ingush system to be involved in individuation.

ad C₃: Except for humans, the concordial class of nouns in Ingush is arbitrary and invariable (Nichols 2011: 144f.). An attribution of properties thus not possible.

ad C₄: By employing agreement on e.g. verbs, the Ingush system allows its speaker to identify referents by class marking (cf. 109).

(109) "In this play Mariem plays Musa." (Ingush; cf. Nichols 2011: 433)

14.3.2 Formal Setup

Table 64: The formal properties of the nominal classification system of Ingush.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>obligatory assignment of each noun ✓</td>
<td></td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership ✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>single class membership ✓</td>
<td></td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>limited class number ✓</td>
<td></td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once ×</td>
<td></td>
</tr>
<tr>
<td></td>
<td>multiply marked ✓</td>
<td></td>
</tr>
</tbody>
</table>
Appendix

locus operandi (F₅)  
CLS-like: locus operandi: NP-internal  ✗
CCS-like: locus operandi: NP-internal and -external  ✓

semanticity (F₆)  
CLS-like: semantic contribution  ✗
CCS-like: semantically vain  ✓

assignment (F₇)  
CLS-like: chosen freely by the speaker  ✗
CCS-like: mechanic assignment  ✓

ad F₁: All Ingush nouns are assigned to a concordial class (Nichols 2011: 143).

ad F₂: All Ingush nouns belong to one concordial class, except for a handful of denotations of animals, which are respectively assigned to the male or female class according to the real life referent’s biological sex (cf. Nichols 2011: 145).

ad F₃: Ingush employs a small system with six concordial classes (i.e. controller classes) which are labeled by their respective marker’s phonological form as the V- (v-) , J- (j-) , B- (b-) , and D-class (d- ; cf. Nichols 2011: 141). The combination of these four markers (or target classes) results in the six controller classes, so the V-, J-, B-, and D-‘class’ are not concordial classes on their own. In this dossier, each controller class is referred to by a Roman numeral (cf. table 65) wherever possible.

Table 65: The inventory of the Ingush class markers (cf. Nichols 2011: 144).

<table>
<thead>
<tr>
<th>marker (sg./pl.)</th>
<th>class</th>
</tr>
</thead>
<tbody>
<tr>
<td>v-/b-</td>
<td>human male (I)</td>
</tr>
<tr>
<td>j-/b-</td>
<td>human female (II)</td>
</tr>
<tr>
<td>b-</td>
<td>various non-human (III)</td>
</tr>
<tr>
<td>b-/d-</td>
<td>various non-human (IV)</td>
</tr>
<tr>
<td>d-</td>
<td>various non-human (V)</td>
</tr>
<tr>
<td>j-</td>
<td>various non-human (VI)</td>
</tr>
</tbody>
</table>

ad F₄: The same context may feature multiple class marking (cf. [110]).

(110) *Hwoa tolxa-bea*

<table>
<thead>
<tr>
<th>brain</th>
<th>spoil-B.CAUS.CONVERB.ANTERIOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>zubbolazh</td>
<td>beac</td>
</tr>
<tr>
<td>B.FUT.CONVERB.SIMULTANEOUS</td>
<td>B.INFERENTIAL.NEG 3PL.ERG</td>
</tr>
<tr>
<td>cyn.</td>
<td>3SG.GEN</td>
</tr>
</tbody>
</table>

“I guess they must not have hurt his brain.” (Ingush; cf. Nichols 1996: 261)

ad F₅: Ingush marks class membership among others on the verb (cf. [110]) and thus NP-externally.
ad F₆: Except for the distinction of a referent’s biological sex, the semantics of the individual classes are arbitrary (Nichols 2011: 144) and thus a noun’s concordial class membership does not contribute to the noun phrase.

ad F₇: Since the semantics of the concordial classes are arbitrary except for those nouns referring to animates (Nichols 2011: 144), the assignment of nouns to a concordial class is a mechanic one. Loans or unfamiliar words are assigned by the speaker to a concordial class in analogy to generics or synonyms (Nichols 2011: 144).

14.3.3 Mapping

![Diagram](image)

Figure 31: Ingush.

Table 66: The formalisation and transparency value of Ingush.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>0.875</td>
<td>semantic structure</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>0.75</td>
<td>rule types</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>1</td>
<td>rule number</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>X₄</td>
<td>0.666</td>
<td>categorial independence</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>1</td>
<td>discreteness of markers</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>1</td>
<td>redundancy</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>1</td>
<td>flexibility</td>
</tr>
<tr>
<td>x-value</td>
<td>0.899</td>
<td>y-value</td>
<td>0.357</td>
</tr>
</tbody>
</table>
ad X₁: The six classes of the Ingush system render it a small one (cf. table 65).

ad X₂: The Ingush concordial class system employs agreement with parts of its adjectives, verbs, and postpositions (cf. 110 and Nichols 2011: 431ff.).

ad X₃: Ingush class markers occur among others on the verb (cf. 110) and thus NP-externally.

ad X₄: Class marking occurs on adjectives, verbs, and postpositions. However, each of these parts of speech realises class marking only on a fraction of its actual members. Ingush thus displays a ‘semi-obligatory’ classification system.

ad X₅: The markers occur in bound form (cf. 110 and table 65).

ad X₆: The marking may occur multiply in the same context (cf. 110).

ad X₇: The system classifies the full nominal lexicon (Nichols 2011: 143).

ad Y₁: Class I hosts human male referents, class II human female referents (Nichols 2011: 143). Besides a few semantic consistencies within the individual non-human classes, there are no recognisable semantics behind the system (Nichols 2011: 144). The consistencies include implements and tools (J), abstractions (D), and some large animals (B; Nichols 2011: 144).

ad Y₂: New or unfamiliar words are assigned to a gender on the basis of the gender of associated words or generics (Nichols 2011: 144), which falls in the scope of semantic assignment. As class membership is largely arbitrary (ib.), opaque assignment is frequent in Ingush.

ad Y₃: Due to the large portion of opaquely assigned nouns (Nichols 2011: 144), there are many assignment rules.

ad Y₄: The class markers convey class information only (cf. table 65).

ad Y₅: The four class markers serve the marking of six controller classes (cf. the discussion under F₃) and thus are not discrete per class.

ad Y₆: The class markers’ semantics are either vain or mirror information that is available in the classified noun (Nichols 2011: 143ff.).

ad Y₇: The distinction of biological sex is the only niche for flexibility that the consulted sources discuss.
14.4 Grammaticalisation

The Ingush concordial class system indicates a high degree of grammaticalisation. The markers are single consonants (which may be interpreted as a sign for phonological erosion), semantics play a very marginal role in the system, the number of classes is relatively low, nouns are mechanically and inflexibly assigned to a concordial class and the agreement (despite its unusual manifestation on only small parts of the individual types of modifiers) occurs on adjectives, verbs, and a few adpositions. At the same time, the system is fairly intransparent.

15 Jakaltek

15.1 Background

Jakaltek (Popti’) is spoken in the Mesoamerican “Huehuetenango Department near [the] Mexico border, Concepción Huista and Jakaltenango areas” (Lewis et al. 2013: s.v.); it is thus located in Guatemala and partly Mexico. The main works that were consulted for the description of the Jakaltek classifier system are Craig 1977, Craig 1986a and 1986b, and Day 1973. Throughout the literature, Jakaltek is described as a language employing a noun classifier system.

![Figure 32: Classification of the Jakaltek system of nominal classification.](image-url)
15.2 Classification Typology

15.2.1 Noun Classifiers

Jakaltek employs a noun classifier system (cf. [111]). This system is analysed in section 15.3.

(111) y-oj\[3mm\]
cumi’\[3mm\]
x’ahaw

E3.foot CL:GODDESS moon

“moonlight” (Jakaltek; cf. Craig 1986a: 273)

15.2.2 Numeral Classifiers

Day (1973: 59) reports the presence of ‘numeral classifiers’ in Jakaltek, which distinguish three types (object, quantity, and action classifiers; cf. table [67]).

Table 67: So-called numeral classifiers in Jakaltek (cf. Day 1973: 59ff.).

<table>
<thead>
<tr>
<th>type</th>
<th>form</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>object</td>
<td>putzan</td>
<td>“short, thick, solid”</td>
</tr>
<tr>
<td>object</td>
<td>c’olan</td>
<td>“spherical”</td>
</tr>
<tr>
<td>object</td>
<td>lisan</td>
<td>“in form of stripe”</td>
</tr>
<tr>
<td>object</td>
<td>tzol(an)</td>
<td>“in form of row”</td>
</tr>
<tr>
<td>object</td>
<td>molan</td>
<td>“paired (of people or animals)”</td>
</tr>
<tr>
<td>object</td>
<td>b’utan</td>
<td>“in form of solitary grains or gravel”</td>
</tr>
<tr>
<td>object</td>
<td>c’ut'an</td>
<td>“in form of large pinch of powdery substance (falling or fallen)”</td>
</tr>
<tr>
<td>quantity</td>
<td>pulato</td>
<td>“paired (of people or animals)”</td>
</tr>
<tr>
<td>quantity</td>
<td>c’uchub’</td>
<td>“in form of solitary grains or gravel”</td>
</tr>
<tr>
<td>quantity</td>
<td>chehal</td>
<td>“in form of large pinch of powdery substance (falling or fallen)”</td>
</tr>
<tr>
<td>quantity</td>
<td>tG•bra</td>
<td>“lying down hands and feet tied (person or animal)”</td>
</tr>
<tr>
<td>action</td>
<td>way</td>
<td>“paired (of people or animals)”</td>
</tr>
<tr>
<td>action</td>
<td>xew</td>
<td>“in form of solitary grains or gravel”</td>
</tr>
<tr>
<td>action</td>
<td>pil</td>
<td>“in form of large pinch of powdery substance (falling or fallen)”</td>
</tr>
<tr>
<td>action</td>
<td>b’ik’</td>
<td>“lying down hands and feet tied (person or animal)”</td>
</tr>
</tbody>
</table>

These elements are no classifiers in the definition of this study: ‘Action classifiers’ are in fact verbal roots that describe actions (e.g. ox b’ik’ “three actions of swallowing” < b’ik’a “to swallow something”; cf. German drei Schluck “three actions of swallowing ~ sips” < schlucken “to swallow”) and are “never accompanied by any other element of the noun phrase” besides the numeral (cf. Day 1973: 61). The ‘quantity classifiers’ are “nouns which denote containers or units of measure”, while “other nouns used as [numeral classifiers] are productively derived” by -al ~ -il “-fuls” (Day 1973: 60; cf. [112]).

(112) A pulato “plate” > ca pulato chib’e “two plates of meat”

The ‘quantity classifiers’ thus correspond to measure nouns, and not to nominal classifiers. Object numeral modifiers finally are “positional stems with the stem formative -an” (Day 1973: 60). This suffix -an is the regular formative for positional stems and no specific morpheme for a specific category of categorisation (Day 1973: 29).

Next to the abovementioned elements, Jakaltek employs elements labeled ‘plural classifiers’. These constitute a tripartite set of markers, which take the form heb’ “human”, hej “animal”, and the form of a zero-marker for inanimate nouns (cf. 113 and Craig 1986a: 265).

\[(\text{113})\]
\[
\begin{align*}
\text{A} & : \text{ca-}u\text{-n}ni & & \text{heb’} & & \text{naj} & & \text{winaj} \\
& : \text{two-HUMAN} & & \text{PL:HUMAN} & & \text{CL:MAN} & & \text{man} \\
& & \text{“two men”} \\
\text{B} & : \text{ca-c’o}’ni & & (\text{hej}) & & \text{na’} & & \text{nok’} \\
& : \text{two-ANIMAL} & & \text{PL:ANIMAL} & & \text{CL:ANIMAL} & & \text{animal} \\
& & \text{“two animals”} \\
\text{C} & : \text{ca-b’} & & (\text{a}) & & \text{na’} & & \text{žila} \\
& : \text{two-INANIMATE} & & \text{CL:ANIMAL} & & \text{chair} \\
& & \text{“two saddles” (inanimate object made of animal product; cf. Craig 1986a: 265)}
\end{align*}
\]

Quantifying expressions employ the human ‘plural classifier’ obligatorily, whereas the marker hej occurs optionally (cf. Craig 1977: 127). Inanimate nouns may take the non-human marker hej in instances where the ‘number classifier’ is omitted (cf. Craig 1977: 127). Since there is only one obligatorily overtly marked category, which does not employ a single function listed in Contini-Morava and Kilaraksi (2013), the ‘plural classifiers’ are not analysed as nominal classifiers in this dossier. Jakaltek does thus not employ a numeral classifier system.

**15.2.3 Possessive Classifiers**

Jakaltek does not employ a possessive classifier system (cf. [114] where no classifier occurs).

\[(\text{114})\]
\[
\begin{align*}
\text{wa}l & & \text{tina’} & & \text{mach=za} & & \text{ch-o}’ni\text{-}to & & (\text{co}) & & \text{b’et} & & \text{co} & & \text{co’ni\’}b’. \\
& & \text{on.the.other.hand} & & \text{now} & & \text{no.longer} & & \text{we.go} & & (\text{our}) & & \text{to} & & \text{our} & & \text{town}
\end{align*}
\]

“We, now! We won’t go to our town any more.” (Jakaltek; cf. Day 1973: 122)

**15.2.4 Multiple Classifier Systems**

The Jakaltek classifiers are typically described as noun classifiers (e.g. Craig 1977: 133). This naturally allows the classifiers to occur in noun phrases with a variety of nominal modifiers (among them e.g. the possessive; cf. [115]). Note that the classifier in [115] occurs adjacent to the possessive, not the noun.

\[(\text{115})\]
\[
\begin{align*}
\text{xcam} & & \text{na’} & & \text{hin} & & \text{cheh} \\
\text{died} & & \text{CL:ANIMAL} & & \text{my} & & \text{horse}
\end{align*}
\]

“My horse died.” (Jakaltek; cf. Craig 1977: 140)
However, as the classifiers also occur in constructions that lack a modifier (cf. [116]), an analysis as a noun classifier system that allows its markers to attach to the full noun phrase rather than only the noun is not in contradiction with the data. Jakaltek thus does not employ a multiple classifier system.

(116) A  y-oj  naj  pel  
   E3.foot  CL:MAN  Peter 
   “Peter’s foot”

B  no’ sasim  
   CL:ANIMAL  liver 
   “animal liver” (Jakaltek; both Craig 1986a: 273)

15.2.5 Concordial Class Systems

The consulted sources and available data do not suggest that Jakaltek employs a concordial class system.

15.3 Analysis

15.3.1 Functional Setup

Table 68: The semantic and discourse properties of the Jakaltek system of nominal classification.

<table>
<thead>
<tr>
<th>differentiating referents (C_1)</th>
<th>productive derivation, variable classification ✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C_2)</td>
<td>unisation for countability X</td>
</tr>
<tr>
<td>attribution of properties (C_3)</td>
<td>speaker-sided values ✓</td>
</tr>
<tr>
<td>reference identification (C_4)</td>
<td>reference tracking via AGR, anaphora ✓</td>
</tr>
</tbody>
</table>

ad C_1: Being independent elements, the Jakaltek classifiers do not take part in derivational processes. They may serve the differentiation of referents (cf. again [117], where the noun for “chair” is clearly specified by the classifier).

(117) A  ca-b’ (a)  na’  ’ilα   
   two-INANIMATE  CL:ANIMAL  chair 
   “two saddles” (inanimate object made of animal product; Jakaltek; cf. Craig 1986a: 265)

B  ca-b’ (a)  te’  ’ilα   
   two-INANIMATE  CL:PLANT  chair 
   “two chairs” (Jakaltek; compiled after Craig 1986a)

ad C_2: The consulted sources and available data do not suggest the Jakaltek classifiers to be involved in individuation.
ad C₃: A part of the Jakaltek classifier set accounts for the referent’s social status (cf. table 118 on page 473), which allows the speaker to express attitudes about a referent by classificatory means (“reclassification”, cf. Craig 1986a: 269f.).

(118)

A naj cap
CL:MAN Gabriel
“Gabriel (neutral)”
B ya’ cap
CL:RESPECTED.NONDEITY Gabriel
“Gabriel (respectful)”
C ho’ cap
CL:NONEPRESSED.NONINFLRT.MALE.KIN Gabriel
“Gabriel (neutral, akin)” (examples compiled by the description in Craig 1977: 153, Craig 1986a: 269f. and Day 1973: 125f.)

ad C₄: The Jakaltek noun classifiers may be used as carriers of anaphora and thus serve reference identification (cf. 119).

(119)

A xil naj zun wa’ no’ lab’a.
saw CL:MAN Juan CL:ANIMAL snake
“Juan saw the snake.”
B xil naj no’
saw CL:MAN CL:ANIMAL
“He saw it.” (Jakaltek; cf. Craig 1986a: 264)

15.3.2 Formal Setup

Table 69: The formal properties of the nominal classification system of Jakaltek.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness</td>
<td>nouns may remain uncategorised</td>
<td>✓</td>
</tr>
<tr>
<td>flexibility</td>
<td>flexible class membership</td>
<td>✓</td>
</tr>
<tr>
<td>inventory</td>
<td>largish class number</td>
<td>✓</td>
</tr>
<tr>
<td>marking</td>
<td>marked once</td>
<td>✓</td>
</tr>
<tr>
<td>locus operandi</td>
<td>locus operandi: NP-internal</td>
<td>✓</td>
</tr>
<tr>
<td>semanticity</td>
<td>semantic contribution</td>
<td>✓</td>
</tr>
<tr>
<td>assignment</td>
<td>chosen freely by the speaker</td>
<td>✓</td>
</tr>
</tbody>
</table>
**Appendix**

**ad F₁:** Objects that are “perceived by several senses” and that are “made of unidentified substance or of an undifferentiated mixture of substances” (e.g., beer, Coca-Cola, smoke, clouds or garbage (cf. Craig 1986a: 273); both citations Craig 1986a: 276) are not subject to categorisation (ib.).

**ad F₂:** Depending on the situational context, the use of the individual classifiers may vary (cf. 118).

**ad F₃:** The system is subdivided into two semantic subsets, each consisting of twelve classifiers. The first subsystem can be subsumed as a classification in terms of social interaction, the second as a classification in terms of “physical and functional interaction” (cf. Craig 1986a: 266f.; based on the principles in Denny 1976), thus inherent characteristic properties. Both are listed in table 70.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Semantic Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>comam</td>
<td>“male deity”</td>
</tr>
<tr>
<td>comi’</td>
<td>“female deity”</td>
</tr>
<tr>
<td>ya’</td>
<td>“respected nondeity (male or female)”</td>
</tr>
<tr>
<td>unin’</td>
<td>“infant”</td>
</tr>
<tr>
<td>ho’, xo’</td>
<td>“nonrespected noninfant kin (male, female)”</td>
</tr>
<tr>
<td>ho’ ni’an, xo’ ni’an</td>
<td>“noninfant child kin (male, female)”</td>
</tr>
<tr>
<td>naj ni’an, ix ni’an</td>
<td>“noninfant nonkin child (male, female)”</td>
</tr>
<tr>
<td>naj, ix</td>
<td>“nonrespected nonchild nonkin; Ladino; dead person; low deity, human (male, female)”</td>
</tr>
<tr>
<td>metz’</td>
<td>“dog”</td>
</tr>
<tr>
<td>no’</td>
<td>“(made from) animal (except dog); (consisting of) rubber; rainbow”</td>
</tr>
<tr>
<td>ixim</td>
<td>“(consisting of) corn, wheat”</td>
</tr>
<tr>
<td>tz’al’</td>
<td>“thread; woven band”</td>
</tr>
<tr>
<td>tz’añi’</td>
<td>“(consisting of) fiber rope, cord”</td>
</tr>
<tr>
<td>k’ap’</td>
<td>“(consisting of) cloth (except wool)”</td>
</tr>
<tr>
<td>te’’</td>
<td>“(consisting of) plant (except corn, wheat, thread, cloth, rope)”</td>
</tr>
<tr>
<td>ha’</td>
<td>“(consisting of) water”</td>
</tr>
<tr>
<td>ch’en’</td>
<td>“(consisting of) stone, metal, glass, ice; money”</td>
</tr>
<tr>
<td>tz’otz’</td>
<td>“soil, land”</td>
</tr>
<tr>
<td>k’a’</td>
<td>“fire”</td>
</tr>
<tr>
<td>atz’am</td>
<td>“salt”</td>
</tr>
</tbody>
</table>

**ad F₄:** The Jakaltek classifiers occur once per context (cf. Day 1973: 66).

**ad F₅:** The Jakaltek classifiers occur only NP-internally (cf. Day 1973: 66).

**ad F₆:** The Jakaltek noun classifiers may contribute to a noun phrase’s semantics (cf. 117).
ad F₇: The bijective and transparent semantics of the classifier inventory suggest a semantic and speaker-sided assignment (cf. table 70).

15.3.3 Mapping

![Figure 33: Jakaltek.](image)

Table 71: The formalisation and transparency value of Jakaltek.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obliogatoriness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreetness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td>x-value</td>
<td>0.304</td>
<td>y-value</td>
<td>0.893</td>
</tr>
</tbody>
</table>

ad X₁: With a total of 24 classes, the Jakaltek system is a small one (cf. table 70).

ad X₂: The classifiers are reported to occur adjacent to the classified noun or, in the case of anaphorical reference, the position of the deleted noun (Craig 1977: 153). Note however that the classifier in [115] occurs adjacent to the possessive, not the noun.

ad X₄: The classification of nouns is not obligatory in Jakaltek (cf. [120] where no classifier occurs).

(120) A potz’-om txitam
    kill-NOMINALISER pig
    “pig killer”
B tz’a k’a
    excrement fire
    “charcoal”
C sapato plastico
    shoe plastic
    “plastic shoe” (Jakaltek; Craig 1986b: 258ff.)

ad X₅: The classifiers are morphosyntactically independent elements (cf. [119]).

ad X₆: The classifiers occur once per context (cf. Day 1973: 66).

ad X₇: Parts of the nominal lexicon are left unclassified due to the composition of the Jakaltek noun classifier system. Among these, there are objects that are “perceived by several senses” and that are “made of unidentified substance or of an undifferentiated mixture of substances” (e.g. beer, Coca-Cola, smoke, clouds or garbage (cf. Craig 1986a: 273; both citations Craig 1986a: 276). Abstracts are also excluded from nominal categorisation due to the non-concrete nature of the concepts they refer to (cf. howal “war” in [116] A). Other uncategorised nouns are locative nouns (cf. coñob’ “town” in [116] A) and (inalienable) body parts (besides (alienable) edible animal bodyparts are classified; cf. [116] B and C; Craig 1986a: 273).

ad Y₁: The Jakaltek classes are semantically straightforward (cf. table [70]).

ad Y₂: The straightforward semantics of the class inventory allow a semantic assignment (cf. table [70]). The consulted sources and available data do not suggest other involved assignment rule types. Craig (1986a: 276) lists the following four strategies for the “classification of the concrete non-human objects of the world”:

1. Natural classification by material substance (unmarked case).
2. Natural classification by perceptual analogy (limited to rock and corn class).
3. Totemic classification of the objects as members of the spiritual world in the social interaction subsystem (e.g. sun, moon, lightning, crosses, mountains, corn and wind; cf. [117])
4. Absence of classification
ad Y₃: The transparent semantics of the class inventory allow the speaker to master the Jakaltek system with a low number of semantic rules (cf. table 70).

ad Y₄: The classifiers convey classificatory semantics only (cf. table 70).

ad Y₅: Each class has its own, distinct classifier (cf. table 70).

ad Y₆: The classifiers may convey either redundant information, which is already present in the noun (cf. 119), or they may contribute to their host NP (cf. 116 B and 117).

ad Y₇: The physical classifiers are mutually exclusive and may not be used interchangeably. For instance, dogs can be only classified by metz’ “dog”, not by no’ “animal” (Craig 1986a: 281).

15.4 Grammaticalisation

The Jakaltek classifiers display a moderate degree of formalisation (value: 0.304): The classifiers occur once per classificatory context as independent elements, and their lexical origin is traceable in many cases (cf. table 72).


<table>
<thead>
<tr>
<th>classifier</th>
<th>lexical source</th>
</tr>
</thead>
<tbody>
<tr>
<td>unin “infant”</td>
<td>unin “infant”</td>
</tr>
<tr>
<td>(x) mi’an “noninfant child”</td>
<td>mi’an “infant”</td>
</tr>
<tr>
<td>naj “nonrespected nonchild nonkin male”</td>
<td>unaj “man, male”</td>
</tr>
<tr>
<td>ix “nonrespected nonchild nonkin female”</td>
<td>ix “woman, female”</td>
</tr>
<tr>
<td>no’ “animal”</td>
<td>nok’ “animal”</td>
</tr>
<tr>
<td>ixim “corn”</td>
<td>ixim “corn”</td>
</tr>
<tr>
<td>tz’al “thread”</td>
<td>tz’al “thread”</td>
</tr>
<tr>
<td>tz’a’i “rope, cord”</td>
<td>tz’a’i “rope, cord”</td>
</tr>
<tr>
<td>k’ap “cloth”</td>
<td>k’ap “cloth”</td>
</tr>
<tr>
<td>te’ “plant”</td>
<td>te’ “wood, tree, bush, log, stick”</td>
</tr>
<tr>
<td>ha’ “water”</td>
<td>ha’ “water, river, creek”</td>
</tr>
<tr>
<td>ch’em “stone, metal, ice”</td>
<td>ch’em “stone, metal”</td>
</tr>
<tr>
<td>tz’otz’ “soil, land”</td>
<td>tz’otz’ “soil, land”</td>
</tr>
<tr>
<td>k’a “fire”</td>
<td>k’a “fire”</td>
</tr>
<tr>
<td>atz’am “salt”</td>
<td>atz’am “salt”</td>
</tr>
</tbody>
</table>

They display transparent semantics, which are mutually exclusive in the case of the physical class set.
16 Ket

16.1 Background

< Yeniseian (ISO 639-3: ket; cf. Lewis et al. 2013: s.v.)
< isolate (cf. Ruhlen 1987)

Ket is the only extant Yeniseian language and was long thought to be a language isolate. It is spoken in the “Upper Yenisei valley, Krasnoyarski District, Turukhansk, and Baikitsk regions; Sulomai, Bakhta, Verkhneimbatsk, Kellog, Kangatovo, Surguitikha, Vereshchagino, Baklanikha, Farkovo, Goroshikha, and Maduyka villages; eastern Siberia” (cf. Lewis et al. 2013). The Ket language is moribund and gradually replaced by Russian, which is also the major source for loans since the 18th century. This dossier is mainly based on Georg 2007, Vajda 2004, and Werner 1994 and 1997. Ket employs a concordial class system.

nominal classification devices

\[
\text{CLSs} \quad \text{CCSs}
\]

nominative 
numeral 
possessive

Figure 34: Classification of the Ket system of nominal classification.

16.2 Classification Typology

16.2.1 Noun Classifiers

Ket does not employ a noun classifier system (cf. [121] where no classifier occurs).

(121) ki̱de s̕i̱̕-ka is̕e̕ ɣ is
    DEM river-LOC exist fish
    “There are fish in this river.” (Ket; cf. Georg 2007: 315)

16.2.2 Numeral Classifiers

Ket does not employ a numeral classifier system (cf. [122] where no classifier occurs).

(122) s̕i̱̕ k a̕y̕ q, a̕y̕ n e̕γ̕ p
    four tree:PL seven day:PL
    “four trees, seven days” (Ket; cf. Georg 2007: 182)
16.2.3 Possessive Classifiers

Ket does not employ a possessed classifier system (cf. 123 where no classifier occurs).

(123) bû-da, da?n, da¬da?n
he-GEN:M.SG knife GEN:M.SG-knife
“his knife, his knife” (Ket; cf. Georg 2007: 119)

16.2.4 Multiple Classifier Systems

Ket does not employ a multiple classifier system (cf. 121, 122, and 123 where no classifier occurs).

16.2.5 Concordial Class Systems

Ket employs a tripartite concordial class system (cf. Georg 2007: 89ff.). This system is analysed in section 16.3.

16.3 Analysis

16.3.1 Functional Setup

Table 73: The semantic and discourse properties of the Ket system of nominal classification.

<table>
<thead>
<tr>
<th>differentiating referents (C₁)</th>
<th>productive derivation, variable classification</th>
<th>✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability</td>
<td>✗</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
<td>✗</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
<td>✓</td>
</tr>
</tbody>
</table>

ad C₁: Since concordial class is not marked on the noun itself and since the speaker has no influence on the class assignment, Ket does not employ its system in nominal derivation. However, the Ket nominal lexicon distinguishes males and females by means of class marking in predicative constructions (cf. [124]).

(124) A a-yissal, t-yissal, 0-kissal,
SUBJ.M.SG-overnight SUBJ.F.SG-overnight SUBJ.N-overnight
ag-yissal
SUBJ.ANIMATE.PL-overnight
“he overnights, she overnights, it/they [inanimate] overnight, they [animate] overnight” (Ket; cf. Werner 1997: 190)

B daR-a-tavet, daR-i-tavet,
SUBJ.M.SG-OBJ.M.SG-leave SUBJ.M.SG-OBJ.F.SG-leave
daq-tavet, daR-ag-tavet
SUBJ.M.SG-OBJ.N-leave SUBJ.M.SG-OBJ.ANIMATE.PL-leave
“he leaves him, he leaves her, he leaves it/them [inanimate], he leaves them [animate]” (Ket; cf. Werner 1997: 191)

**ad C**₂: The consulted sources and available data do not suggest the Ket system to be involved in individuation.

**ad C**₃: There is neither class variation nor speaker-sided assignment in the Ket concordial class system (Georg 2007: 88f.).

**ad C**₄: By employing agreement on a range of nominal satellites (Werner 1994: 43, Georg 2007: 88), the Ket system serves the means of reference identification (cf. 124).

### 16.3.2 Formal Setup

Table 74: The formal properties of the nominal classification system of Ket.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong> (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility</strong> (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td><strong>inventory</strong> (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td><strong>marking</strong> (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi</strong> (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity</strong> (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td><strong>assignment</strong> (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

**ad F**₁: Every noun is assigned to a concordial classes (cf. Georg 2007: 88).

**ad F**₂: Every noun is assigned to only one concordial class (cf. Georg 2007: 88).

**ad F**₃: Ket distinguishes three concordial classes, a masculine, a feminine, and an inanimate/neuter class. The case markers of each class are listed in table 75.
Table 75: The inventory of the Ket case markers (cf. Werner 1994: 55)

<table>
<thead>
<tr>
<th>case</th>
<th>m. sg.</th>
<th>f./n. sg.</th>
<th>m./f. pl.</th>
<th>n. pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>absolutive</td>
<td>-∅</td>
<td>-∅</td>
<td>-∅</td>
<td>-∅</td>
</tr>
<tr>
<td>genitive</td>
<td>-da</td>
<td>-(d(i)/-t)</td>
<td>-na</td>
<td>-(d(i)/-t)</td>
</tr>
<tr>
<td>dative</td>
<td>-daŋa</td>
<td>-diŋa</td>
<td>-naŋa</td>
<td>-diŋa</td>
</tr>
<tr>
<td>benefactive</td>
<td>-data</td>
<td>-d(i)/-t</td>
<td>-nata</td>
<td>-d(i)/-t</td>
</tr>
<tr>
<td>ablative</td>
<td>-daŋal'</td>
<td>-diŋal'</td>
<td>-naŋal'</td>
<td>-diŋal'</td>
</tr>
<tr>
<td>adessive</td>
<td>-daŋta</td>
<td>-diŋta</td>
<td>-naŋta</td>
<td>-diŋta</td>
</tr>
<tr>
<td>locative</td>
<td>-</td>
<td>-ka/-ga</td>
<td>-</td>
<td>-ka/-ga</td>
</tr>
<tr>
<td>prosecutive</td>
<td>-bes'</td>
<td>-bes'</td>
<td>-bes'</td>
<td>-bes'</td>
</tr>
<tr>
<td>comitative</td>
<td>-as'</td>
<td>-as'</td>
<td>-as'</td>
<td>-as'</td>
</tr>
<tr>
<td>caritative</td>
<td>-an'</td>
<td>-an'</td>
<td>-an'</td>
<td>-an'</td>
</tr>
<tr>
<td>vocative</td>
<td>-</td>
<td>-a/-A</td>
<td>-A</td>
<td>-</td>
</tr>
<tr>
<td>instrumental</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

ad F4: A noun’s concordial class is not expressed on the noun itself, but on a range of nominal satellites and the verb (cf. Georg 2007: 88 and Werner 1994: 43).

ad F5: A noun’s concordial class is marked among others by subject and object markers on the predicate (cf. 124) and thus outside of the noun phrase (cf. Georg 2008: 88).

ad F6: A noun’s concordial class membership does not encode a semantic component (besides a human being’s biological sex) and is generally idiosyncratic (cf. Georg 2007: 89).

ad F7: While those nouns that are referring to humans are generally transparent with respect to their semantics, a noun’s class membership in Ket is mostly idiosyncratic (cf. Georg 2007: 89), which means that a speaker of Ket has to learn a noun’s concordial class and cannot freely assign it.
16.3.3 Mapping

![Graph showing the mapping of transparency value to formalisation value.

Table 76: The formalisation and transparency value of Ket.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td><strong>x-value</strong></td>
<td>0.964</td>
<td><strong>y-value</strong></td>
<td>0.298</td>
</tr>
</tbody>
</table>

**ad X₁:** Its three concordial classes render the Ket system a small one (cf. table 75).

**ad X₂:** The Ket adjectives, numerals, possessives, pronouns, verbs, and adverbs are marked for a noun’s concordial class (cf. Werner 1994: 43 and Georg 2007: 88).

**ad X₃:** The class marking occurs among others on the verb (cf. 124) and thus NP-externally.
ad X₄: As the class markers are fused with other, discoursally obligatorily marked categories such as case and number (cf. table 75), classification in Ket is obligatory.

ad X₅: The markers occur in form of affixes (cf. table 75) or clitics (e.g. in form of the possessive markers; cf. 125, table 77, and Georg 2007: 119).

(125) b=hi?b, k=hi?b, da=hi?b, d=hi?b
    POSS1SG-son POSS2SG-son POSS3SG.M-son POSS3SG.F-son
    “my son, your son, his son, her son” (Ket; compiled after Georg 2007: 119)


<table>
<thead>
<tr>
<th>person</th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>b</td>
<td>na</td>
</tr>
<tr>
<td>2</td>
<td>k</td>
<td>na</td>
</tr>
<tr>
<td>3M</td>
<td>da</td>
<td>na</td>
</tr>
<tr>
<td>3F</td>
<td>d</td>
<td>na</td>
</tr>
<tr>
<td>3N</td>
<td>d</td>
<td>d</td>
</tr>
</tbody>
</table>

ad X₆: As adjectives, numerals, possessives, pronouns, verbs, and adverbs show agree with the noun’s class (cf. Werner 1994: 43 and Georg 2007: 88), multiple marking is possible.

ad X₇: The Ket system classifies the full nominal lexicon (Georg 2007: 88).

ad Y₁: Except for the distinction of a referent’s biological sex, the Ket classes are semantically vain (cf. Georg 2007: 89).

ad Y₂: Only the class membership indicating a referent’s biological sex is semantically transparent and productive. The remaining nominal lexicon is opaquely assigned (cf. Georg 2007: 89).

ad Y₃: The high number of idiosyncratically assigned nouns results in a high number of assignment rules.

ad Y₄: A noun’s class is marked by portmanteau morphemes (cf. the tables 75 and 77).

ad Y₅: Different class make use of the same class markers: The feminine and neuter share the same singular forms and the masculine and feminine class the same plural forms (cf. table 75).
ad Y6: The class membership of the bulk of nouns is idiosyncratic (cf. Georg 2007: 89). Due to the animistic perception of the world (which lasted for the Ket until the 20th century; cf. Werner 1994: 14), we find not only biological males and females in their respective classes, but also objects that are not living in our definition (e.g. o'ks’ m. “(growing) tree” or bejoks m. “mast”; cf. Werner 1994: 32 and 1997: 93). The nominal referent thus is characterised by perceived inherent characteristics, namely animacy. This corresponds to a state of categorisation which we expect to find in classifier systems, but which may also be expected to be preserved in later stages, e.g. when the system grammaticalises and the concordial class membership is being lexicalised. The Ket concordial class system at its synchronic state does not allow its speaker to assign a noun to a concordial class by will, but forms like o'ks’ m. “(living/growing) tree” vs. o'ks’ n. “stick” (cf. Werner 1994: 32) indicate that this process at least must have been productive until late.

In the following, a list of the semantic groupings of each noun class is given. The masculine class comprises male gods and spirits, humans and animals (cf. 126 A), reptiles, worms, insects, trees (cf. 126 B), orbs, meteorological phenomena, months, and seasons (cf. 126 C), certain kinds of animals, birds, and fish (cf. 126 D), and certain tools and instruments (both ritual and common; cf. 126 E and Werner 1994: 23ff.).

C s’el’, ajak, qāRip, q’qdes’ orion thunder january fall “orion, thunder, january, (weather typical for) fall” (Ket; cf. Werner 1994: 26)
D čel, diʔ, boŋtu mammoth eagle herring “mammoth, eagle, herring” (Ket; cf. Werner 1994: 24f.)

The feminine class comprises female gods and spirits, humans and animals (cf. 127 A), certain kinds of animals, birds, and fish (cf. 127 B), reptiles and amphibians, insects, orbs (cf. 127 C), plants (and fruits that entered Ket through...

9Werner (1994: 18f.) a secondary labelling ‘active vs. inactive’, since inanimate objects are found in the animate class (e.g. ul’git m. “whirlwind”); cf. Werner 1997: 93). These secondary labels are not adopted here, since the traditional world view of the Ket was an animistic one (cf. Werner 1994: 14ff.), thus “features and actions, relations, spatial and temporal terms are not less material than that what we [i.e. the westerners; M.P.] merely understand as ‘things’, they are indeed not less alive than our animates [. . .]. The universe is animate, ensouled, and inspired” (Vereno 1960: 45).
Russian), geographical names, body parts (and the seven souls) and dermal diseases (cf. 127 D), and (cult) objects and instruments (cf. 127 E and Werner 1994: 27ff.).

(127) A haʔn', hučä, Dštam, Kalbasem
daughter mare mountain.ruler female.woodghost
“daughter, mare, ruler (f.) of the mountains, female woodghost” (Ket; cf. Werner 1994: 27, 29, 34)
B kuʔs', kiʔl, hal'g
cow raven ruffe
“cow, raven, ruffe” (Ket; cf. Werner 1994: 28)
C tul'n, ʔl', ʔl'm, ʔal'l
lizard frog spider mussle
“lizard, frog, spider, mussle(!)” (Ket; cf. Werner 1994: 28f.)
D qɔʔ, kal'zęs'ka, Jełsk, maʔm, kal'as'
onion-type potato Jeloguj bosom big.boil
“distinct kind of onion [Lilienzwiebel; MP], potato, Jeloguj [affluent of the Yenisei; MP], bosom, big boil” (Ket; cf. Werner 1994: 29f.)
E kal'tag, hil'cm, qo'ndis', kal'l
button spindle glass.pearl twine
“button, spindle, glass pearl, twine” (Ket; cf. Werner 1994: 30)

These semantics are not informative (Georg 2007: 89).

ad Yγ: The synchronic state of Ket mirrors a petrified semantic assignment, which is not always transparent: “A noun’s membership of a certain noun class in the Yenisey languages is only motivated in those cases where the natural sex is distinctly expressed on the lexical level” (Werner 1994: 33). A look at the numerous Russian loans into Ket allows a glance at the currently productive concordial class assignment mechanisms. The bulk of loan refers to social, economical, and cultural concepts that are not genuine to the Ket (Werner 1994: 40). Many of the loans do not follow the Russian example of class assignment, but assign Russian inanimate nouns to the Ket animate class (e.g. names for trees (masculine), fruits, and vegetables (feminine), tools and other instruments (masculine)). The remaining Russian inanimate nouns are assigned to the Ket inanimate class (ib.: 40f.). Russian masculine or feminine nouns are assigned to the respective class in Ket (especially if they denote professions of the biological sex). Diminutives and other nouns in Ket are assigned to the class that hosts the respective diminutive’s category, independently from the Russian source’s gender (ib.: 42f.). Once assigned, a noun does not change its class (Georg 2007: 88).

16.4 Grammaticalisation

Ket displays a strong degree of formalisation (value: 0.964) and a low degree of transparency (0.298). It corresponds closely to the definition of a prototypical concordial class system.
17 Kiowa

17.1 Background

< Kiowa-Tanoan (ISO 639-3: kio; cf. Lewis et al. 2013: s.v.)
< Kiowa-Tanoan < Tanoan < Central Amerind < Amerind (cf. Ruhlen 1987)

Kiowa is primarily spoken in the southwestern parts of Oklahoma “and especially in Caddo, Kiowa, and Comanche counties” (Watkins 1982: 1). This dossier is mainly based on Wonderly et al. 1954, Merrifield 1959, Harbour 2003, 2007, and 2011, and Watkins 1980. Kiowa nouns are inherently marked for number, not for concordial class in a traditional sense (cf. Harbour 2003: 550), which does not naturally suggest an inclusion of Kiowa in the sample. Nevertheless, semantic criteria that are typically involved in nominal classification (such as the distinction of animates and inanimates) seem to play a role in the class distribution and result in a number of nominal classes, which are distinguished by their marking (cf. Wonderly et al. 1954: 3, Harrington 1928: 14). Therefore, Kiowa’s “noun classification system” (Harbour 2003: 550) qualifies for an account within this dossier. This study comes to the conclusion that Kiowa does not employ a productive system of nominal classification.

![Diagram of nominal classification devices](image)

Figure 36: Classification of the Kiowa system of nominal classification.

17.2 Classification Typology

17.2.1 Noun Classifiers

Kiowa does not employ noun classifiers (cf. [128] where no classifier occurs).

(128) A sán 0-kʰóp-dí:  
child 3SG-hurt-be  
“The child is hurt.”

B á:dís  ñlʰétn-gyá  
stick:INV 3INV-break-DETRANSITIVE.PF  
“The stick broke.”
17.2.2 Numeral Classifiers

Kiowa does not employ numeral classifiers (cf. (129) where no classifier occurs).

(129) A zlṓ́́́g yₕ-a-thʰ́́́́n
rock 1SG-SG-find.PF
“I found one rock.”

B zlṓ́́́g ne-n-thʰ́́́́n
rock 1SG-DL-find.PF
“I found two rocks.”

C zlṓ́́́g yₕ-at-thʰ́́́́n
rock 1SG-PL-find.PF
“I found several rocks.”

D ḏló́́́́-hó́́́́-d́̀ thʰ́́́́n
plum INV 1SG:INV find.PF
“I found one plum/several plums.” (Kiowa; cf. Harbour 2003: 554)

17.2.3 Possessive Classifiers

Kiowa does not employ possessive classifiers (cf. (130) where no classifier occurs).

(130) A ṭh₈́ń
head:INV 2/3SG.AGENT:1SG.PATIENT:INV.OBJECT-hurt
“My head hurts.” (Kiowa; cf. Watkins 1980: 131)

B n₃-+t₅́á, á-t₅́á:,
[+author]-wife [author +participant]-wife
“my wife, your wife, his wife” (Kiowa; cf. Watkins 1980: 132)

C n₃-+c₇́á, á-gd₇́á:,
1SG-PL-horse 3SG/PL-horse
“my horse, your horse, his/her/their horse” (Kiowa; cf. Watkins 1980: 137)

17.2.4 Multiple Classifier Systems

Kiowa does not employ a multiple classifier system (cf. (128) (129) and (130)).

17.2.5 Concordial Class Systems

Kiowa integrates cardinality as a categorising criterion in both its noun classification and the category of number: “Under Number, the features yield the cardinalities singular, dual, and plural [ldots]. Under Class, they encode abstract properties of nouns, namely, their natural, or default, distribution in the real world” (Harbour 2011: 582). The distinction of nominal classes that are distinguished by different patterns of marking corresponds to the definition of a concordial class system. This system is analysed in section (17.3) and its investigation shows that Kiowa does not employ a concordial class system.
17.3 Analysis

17.3.1 Functional Setup

Table 78: The semantic and discourse properties of the Kiowa system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification</td>
</tr>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
</tr>
</tbody>
</table>

ad C₁: The consulted sources and available data do not suggest the Kiowa system to be involved in the differentiation of referents.

ad C₂: By its fusion with the category of number, the Kiowa system is involved in an individuational function (cf. the semantics of the individual classes in table 82).

Table 79: The Kiowa concordial classes according to Harbour (2007: 105 and 2011: 578)

<table>
<thead>
<tr>
<th>Class</th>
<th>Class Features</th>
<th>Semantic Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDP</td>
<td>∅</td>
<td>default</td>
</tr>
<tr>
<td>SII</td>
<td>[+singular]</td>
<td>first person only</td>
</tr>
<tr>
<td>SDI</td>
<td>[+augmented]</td>
<td>independently mobile objects</td>
</tr>
<tr>
<td>IDP</td>
<td>[+singular]</td>
<td>vegetation; most non-SDI-implements and body parts</td>
</tr>
<tr>
<td>IDS</td>
<td>[+singular +group]</td>
<td>vegetation occurring in natural collections; implements that may act collectively</td>
</tr>
<tr>
<td>IDI</td>
<td>[+singular -augmented]</td>
<td>hair types; midsize fruit growing in clusters</td>
</tr>
<tr>
<td>SDS</td>
<td>[+group]</td>
<td>non-shape-inductive objects</td>
</tr>
<tr>
<td>SSS</td>
<td>[+augmented +group]</td>
<td>nongranular mass nouns</td>
</tr>
<tr>
<td>PPP</td>
<td>[+augmented -group]</td>
<td>pluralia tantum nouns; granular mass nouns</td>
</tr>
</tbody>
</table>

ad C₃: Properties such as a speaker’s attitude towards a referent cannot be expressed by classificatory means in Kiowa.

ad C₄: A Kiowa noun’s concordial class is marked among others on the demonstrative. Reference identification thus plays a role in the Kiowa system (cf. Harbour 2003: 556).
17.3.2 Formal Setup

Table 80: The formal properties of the nominal classification system of Kiowa.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness</strong> ($F_1$)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility</strong> ($F_2$)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td><strong>inventory</strong> ($F_3$)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td><strong>marking</strong> ($F_4$)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi</strong> ($F_5$)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity</strong> ($F_6$)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td><strong>assignment</strong> ($F_7$)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

**ad $F_1$**: Since the Kiowa concordial class system closely interacts with the category of number and every noun occurs with a certain number, there are no uncategorised nouns in Kiowa.

**ad $F_2$**: Due to Harbour’s (2011) definition of the concordial classes by the features $\pm$singular, $\pm$augmented, and $\pm$group, the individual nouns are assigned to one class and one class only by the inherent properties of their referent (cf. e.g. Harbour 2011: 579).

**ad $F_3$**: The concordial class number is clearly limited, in traditional approaches (four classes; cf. table S1 and Watkins 1980: 102)) as well as in Harbour’s approach (nine classes; cf. table S2 and Harbour 2011: 578).
Table 81: The Kiowa concordial classes (cf. Harbour 2003: 551 and Wonderly et al. 1954: 2ff.).

<table>
<thead>
<tr>
<th>class</th>
<th>singular</th>
<th>dual</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>inherent</td>
<td>inherent</td>
<td>marked</td>
</tr>
<tr>
<td></td>
<td><em>cenbō “a cow”</em></td>
<td><em>cenbō “two cows”</em></td>
<td><em>cenbō-(g) “cows”</em></td>
</tr>
<tr>
<td>II</td>
<td>marked</td>
<td>inherent</td>
<td>inherent</td>
</tr>
<tr>
<td></td>
<td><em>(3lp^b, p^g) “a ribbon”</em></td>
<td><em>(3lp^b, p^g) “two ribbons”</em></td>
<td><em>(3lp^b, p^g) “ribbons”</em></td>
</tr>
<tr>
<td>III</td>
<td>marked</td>
<td>inherent</td>
<td>marked</td>
</tr>
<tr>
<td></td>
<td><em>(áλ-(q) “one apple”</em></td>
<td><em>(áλ “two apples”</em></td>
<td><em>(áλ-(gg) “apples”</em></td>
</tr>
<tr>
<td>IV</td>
<td>inherent</td>
<td>inherent</td>
<td>inherent</td>
</tr>
<tr>
<td></td>
<td><em>(hól-(da “one dress”</em></td>
<td><em>(hól “two dresses”</em></td>
<td><em>(hól “three dresses”</em></td>
</tr>
</tbody>
</table>

Table 82: The Kiowa concordial classes according to Harbour (2007: 105, 2011: 578, and p.c.).

<table>
<thead>
<tr>
<th>class</th>
<th>corresponds to</th>
<th>class features</th>
<th>semantic characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDP</td>
<td>class IVa</td>
<td>∅</td>
<td>default</td>
</tr>
<tr>
<td>SII</td>
<td>class I*</td>
<td>[+singular]</td>
<td>first person only</td>
</tr>
<tr>
<td>SDI</td>
<td>class I</td>
<td>[-augmented]</td>
<td>independently mobile objects</td>
</tr>
<tr>
<td>IDP</td>
<td>class IIa</td>
<td>[-singular]</td>
<td>vegetation; meet non-SDI-implements and body parts</td>
</tr>
<tr>
<td>IDS</td>
<td>class IIb</td>
<td>[-singular +group]</td>
<td>vegetation occurring in natural collections; implements that may act collectively</td>
</tr>
<tr>
<td>IDI</td>
<td>class III</td>
<td>[-singular -augmented]</td>
<td>hair types; midsize fruit frowning in clusters</td>
</tr>
<tr>
<td>SDS</td>
<td>class IVb</td>
<td>[+group]</td>
<td>non-shape-inductive objects</td>
</tr>
<tr>
<td>SSS</td>
<td>class IV*</td>
<td>[+augmented +group]</td>
<td>nongranular mass nouns</td>
</tr>
<tr>
<td>PPP</td>
<td>class IVc</td>
<td>[+augmented -group]</td>
<td>pluralia tantum nouns; granular mass nouns</td>
</tr>
</tbody>
</table>

ad F₄: A noun’s concordial class is marked on the noun itself, the adjective, and the demonstrative. Since these latter two always mirror their respective referent noun’s class value, we are dealing with agreement (cf. 131).

(131) A \(áγ\-\(g\) \(\text{cenbō-}\(g\)) \(\text{cow-INV}\) \(\text{INV}\)

“those cows” (Kiowa; cf. Wonderly et al. 1954: 4)

B \(áλ\-\(\emptyset\) \(\text{apple-INV.SG/PL}\) \(\text{big-INV.SG}\)

“big apple” (Kiowa; cf. Wonderly et al. 1954: 6)
**ad F₅**: Class marking occurs on nouns, adjectives, and demonstratives and thus NP-internally (cf. 131).

**ad F₆**: The Kiowa concordial class system is based on semantic grounds (namely the inherent properties of the referent: animacy, collectiveness, countability; cf. e.g. Harbour 2011: 562), but does not contribute to the NP’s semantics in a sense that they contribute add semantics other than on the quantity and/or degree of individuation of a nominal referent (cf. 81).

**ad F₇**: The inherited Kiowa nominal lexicon displays quite transparent semantics (cf. Harbour 2011 and Watkins 1980: 119), and “[t]he noun classes are productive today. Kiowa speakers make fairly frequent use of unassimilated English nouns, all of which must be assigned to some class for the purposes of determining pronominal prefixes” (Watkins 1980: 118). Animates are assigned to class I, tangible and inanimate objects to class II, and abstract inanimates to class IV (a “natural extension within the system”; cf. Watkins 1980: 118f.). The speakers are bound to assign the nouns within the semantic mechanisms that Kiowa employs and thus do not have some interactive freedom.

### 17.3.3 Mapping

![Figure 37: Kiowa.](image-url)
Table 83: The formalisation and transparency value of Kiowa.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size $X_1$</td>
<td>0.875</td>
<td>semantic structure $Y_1$</td>
<td>1</td>
</tr>
<tr>
<td>host number $X_2$</td>
<td>0.5</td>
<td>rule types $Y_2$</td>
<td>1</td>
</tr>
<tr>
<td>locus operandi $X_3$</td>
<td>0</td>
<td>rule number $Y_3$</td>
<td>1</td>
</tr>
<tr>
<td>obligatoriness $X_4$</td>
<td>0.666</td>
<td>categorial independence $Y_4$</td>
<td>0.666</td>
</tr>
<tr>
<td>boundedness $X_5$</td>
<td>1</td>
<td>discreteness of markers $Y_5$</td>
<td>0</td>
</tr>
<tr>
<td>marking $X_6$</td>
<td>1</td>
<td>redundancy $Y_6$</td>
<td>0</td>
</tr>
<tr>
<td>exhaustivity $X_7$</td>
<td>1</td>
<td>flexibility $Y_7$</td>
<td>0</td>
</tr>
<tr>
<td>x-value</td>
<td>0.720</td>
<td>y-value</td>
<td>0.524</td>
</tr>
</tbody>
</table>

ad $X_1$: Depending on the analysis, Kiowa employs four or nine classes (cf. the tables [81] and [82]). This analysis follows Harbour’s (2011) account and assumes Kiowa to employ nine classes.

ad $X_2$: Kiowa marks a noun’s class membership on the noun, the adjective, and the demonstrative (cf. [131]).

ad $X_3$: The marking occurs NP-internally on the noun, the adjective, and the demonstrative (cf. [131]). Despite class being marked on the demonstrative, the available data does not provide examples for class marking in absence of the classified noun (cf. e.g. the transcription of a Kiowa hunting story in Harbour 2007: 203-206, where).

ad $X_4$: Class and number are intertwined inseparably (cf. e.g. Harbour 2011: 582). Since every noun in discourse has a number value, classification is obligatory.

ad $X_5$: A noun’s class membership is marked by means of the affix -g (cf. table [81]).

ad $X_6$: The multiple occurrence of class marking in the same context is possible (cf. [131]).

ad $X_7$: Every noun, which includes ad hoc loans, is assigned to a class (Watkins 1980: 118).

ad $Y_1$: Since a noun’s class membership is determined by its referent’s degree of individuation (cf. table [82]), the classes are semantically straightforward.

ad $Y_2$: The consulted sources suggest a productive and semantic assignment (cf. e.g. Watkins 1980: 118).
ad Y₃: Due to the low class number and straightforward semantics of the individual class (cf. table 82), the number of rules necessary to master the Kiowa system is low.

ad Y₄: Class and number are intertwined inseparably (cf. the tables 81 and 82).

ad Y₅: All classes are marked by the marker -g (cf. table 81).

ad Y₆: The marker contributes semantically to the noun phrase in that it marks a number of nominal referents that is not inherent to the lexeme (cf. table 81). This contribution falls under the scope of number marking. The class semantics are a byproduct of the grouping of nouns that display the same inherent degree of individuation than their class members, while the individual classes are semantically vain.

ad Y₇: Nouns can be assigned to one class only. This assignment is based on the satisfaction of the features that determine the individual classes (cf. table 82).

17.4 Grammaticalisation

The Kiowa system does not serve nominal classification, but the marking of (inverse) number. The individual classes are a byproduct of noun groupings that have the same degree of individuation. The system shows an advanced, yet intermediate degree of formalisation (value: 0.720) and a medium degree of transparency (value: 0.524). The other sample systems with a formalisation value lower than 0.8 uniformly display a transparency value that is not lower than 0.7.

18 Korean

18.1 Background

< language isolate (ISO 639-3: kor; cf. Lewis et al. 2013: s.v.)

Korean (Hangugmal, Chosŏnmal, Goryeomal) is mainly spoken in Korea (East Asia). This dossier is primarily based on Byon 2008 and 2009, Lee 1989, Martin 1992, Ramstedt 1939, and Sohn 1999. The Korean multiple classifier system largely corresponds to a typical numeral classifier system. A certain amount of semantic potency (in terms of adding semantics to the noun phrase) seems to be only only present in ‘unitisers’, which loosely correspond to count nouns.
Figure 38: Classification of the Korean system of nominal classification.

18.2 Classification Typology

18.2.1 Noun Classifiers

The Korean classifiers occur in quantificational contexts and not adjacent to the bare noun (cf. [132]).

(132) A Waikhiki ka kyengchi ka coh-ta
Waikiki NOM scenery NM good-DECL.SENTENCE
“Waikiki has nice scenery.” (Korean; cf. Sohn 1999: 290)

B ku swunkyeng i totwuk ul cap-ass-e.yo
the police NOM thief ACC caught-PST-POLITE
“The police caught a thief.” (Korean; cf. Sohn 1999: 288)

In addition, Korean employs honorific noun forms, which are typically separate lexical forms which complement a honorific verb form (typically in the role as this verb’s subject, but also in other argument positions; cf. Lee 1989: 57 and Martin 1992: 298). Korean marks the verb with the honorific suffix -(u)si-, which is complemented by a honorific noun (typically with a human referent; cf. Martin 1992: 298f., Lee 1989: 57f., and table [84]). Korean thus does not employ a noun classifier system.

Table 84: Honorific and plain forms of Korean nouns (cf. Lee 1989: 57f.)

<table>
<thead>
<tr>
<th>honorific form</th>
<th>plain form</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>malsim</td>
<td>mal</td>
<td>“word, speech”</td>
</tr>
<tr>
<td>mapi</td>
<td>bah</td>
<td>“food, meal”</td>
</tr>
<tr>
<td>sayham</td>
<td>ilim</td>
<td>“name”</td>
</tr>
<tr>
<td>ogce</td>
<td>mom</td>
<td>“body”</td>
</tr>
<tr>
<td>abapnini</td>
<td>abapyi</td>
<td>“father”</td>
</tr>
<tr>
<td>halmanini</td>
<td>halmani</td>
<td>“grandmother”</td>
</tr>
</tbody>
</table>
18.2.2 Numeral Classifiers

Korean employs a numeral classifier system (cf. [133]), which allows its markers to attach to the adjective and demonstrative (cf. [134]). This system is discussed in section [18.3].

(133) A namu së́y kulwu
     tree three CL:TREE
     “three trees”

B namu sam-cwu
     tree three CL:TREE
     “three trees” (Korean; cf. Martin 1992: 180)

(134) A i/kn/co pun/nom/saalam
     this/the/that CL:PERSON.HONORIFIC/GUY/PERSON
     “this/the/that person/guy/person”

B cakun pun/nom/saalam
     small CL:PERSON.HONORIFIC/GUY/PERSON
     “a small (esteemed) person/guy/person” (Korean; cf. Koo 2008: 65)

18.2.3 Possessive Classifiers

The Korean classifiers occur only in quantificational contexts, not in possessive ones (cf. [135]). Korean thus does not employ a possessive classifier system.

(135) A hyeng un tampay bul/ka phiwu-ko
     brother TOPIC contrast cigarette ACC/NOM smoke-NOM
     wish-PST-POLITE
     “My older brother wanted to smoke.” (Korean; cf. Sohn 1999: 331)

B hal.apeci kkey chayk i mash-a.yo
     grandfather at book NOM many-POLITE
     “My grandfather has many books.” (Korean; cf. Sohn 1999: 284)

18.2.4 Multiple Classifier Systems

The Korean classifiers occur only in quantificational contexts and may attach to the adjective and demonstrative (cf. [133] and [134]). This renders it a multiple classifier system.
18.3 Analysis

18.3.1 Functional Setup

Table 85: The semantic and discourse properties of the Korean system of nominal classification.

<table>
<thead>
<tr>
<th>differentiating references (C₁)</th>
<th>productive derivation, variable classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
</tr>
</tbody>
</table>

ad C₁: According to Lee (1989: 55), “[a]ll classifiers can combine with the noun su ‘number’ to form compound nouns” à la mali-su “number of animals”, jay-su “number of sheets of paper”, or gwansu “number of volumes of books”. Considering Lee’s (1989) label “compound nouns” and the bound status of the otherwise free classifiers, this appears to be a derivational process rather than free classification and thus serves the means of the expansion of the lexicon. In addition, the classification of nouns is variable within certain limitations (e.g. with the general classifier kay and a more specific classifier).

ad C₂: The classifiers occur in quantificational constructions only, which in combination with classifier semantics such as “unit” or “bundle of 20 sheets” (as opposed to the specific classifiers’ individualising semantics) strongly suggests them to take part in the individuation of nouns (cf. 136).

(136)  
A congi sêy yen  
  paper three  CL:PAPER.REAMS  
“three reams (500 sheets) of paper”
B congi sêy yen  
  paper three  CL:20.SHEETS  
“60 sheets of paper”
C congi sêy yen  
  paper three  CL:200.SHEETS  
“600 sheets of paper” (Korean; cf. Sohn 1999: 181)

ad C₃: Since the classifier set contains a number of social status classifiers (cf. 134), attitudes can be expressed speaker-sidedly by the choice of a classifier.

ad C₄: The classifiers may combine with adjectives or demonstratives lacking an overt noun (cf. 134). For this reason, they are involved in reference identification.
Appendix

(137)  A $Eonehakja-2.myeng-i$ $chek-ul$ $sse-ss-ta.$
    linguist-2.CL-NOM book-ACC write-PST-DECL
    “Two (specific or nonspecific) linguists wrote the/a book.”
B $Eonehakja-ga-2.myeng$ $chek-ul$ $sse-ss-ta.$
    linguist-NOM-2.CL book-ACC write-PST-DECL
    “Two (nonspecific) linguists wrote the/a book.”
C $2.myeng-e-eonehakja-ga$ $chek-ul$ $sse-ss-ta.$
    2.CL-GEN-linguist-NOM book-ACC write-PST-DECL
    “Two (nonspecific or specific) linguists wrote the/a book.” (Korean; cf. Kim 2005: 219)

18.3.2 Formal Setup

Table 86: The formal properties of the nominal classification system of Korean.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obligatoriness ($F_1$)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>Flexibility ($F_2$)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>Inventory ($F_3$)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>Marking ($F_4$)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>Locus Operandi ($F_5$)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>Semanticsity ($F_6$)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>Assignment ($F_7$)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad $F_1$: Nouns may remain uncategorised in Korean (cf. 138, where not every noun is classified). Furthermore, abstract nouns may lack their categorisation (Aikhenvald 2000: 334).

(138) na nun cwi lul han mali lul kkoli lul
    1SG TOPIC contrast rat ACC one CL-ANIMAL ACC tail ACC
    call-ass-ta cut-PST-DECL.SENTENCE
    “I cut the tail of one rat.” (Korean; cf. Sohn 1999: 289)

ad $F_2$: At least one classifier (kay “small countable inanimates”) serves as a residue and default classifier as well as an alternative classifier for small
countable items (e.g. *yemphil twu kay* “two pencils” (Aikhenvald 2000: 336) vs. *yemphil twu calwu* “two pencils” (compiled after Aikhenvald 2000 and Martin 1992)).

**ad F.3:** Korean has by far more than a hundred classifiers (cf. Martin 1992: 179ff.), which represents a largish set by the definition of this study. Martin (1992: 179ff.) provides a list of classifiers that is “not quite exhaustive, but […] fairly representative”. It comprises 139 Korean classifiers and 58 Sino-Korean classifiers. A selection of these is given in table 87.

Table 87: The inventory of Korean numeral classifiers (selection; cf. Martin 1992: 179ff.)

<table>
<thead>
<tr>
<th>Korean</th>
<th>Sino-Korean</th>
<th>category</th>
</tr>
</thead>
<tbody>
<tr>
<td>kay</td>
<td></td>
<td>things, items, matches, pencils</td>
</tr>
<tr>
<td>nath</td>
<td></td>
<td>units, items, bits, grains, beans</td>
</tr>
<tr>
<td>kwěn</td>
<td></td>
<td>books, magazines, 20 sheets of Korean paper</td>
</tr>
<tr>
<td>mali</td>
<td></td>
<td>animals, birds, fish</td>
</tr>
<tr>
<td>phil</td>
<td>-phíl</td>
<td>horse</td>
</tr>
<tr>
<td>dho,</td>
<td>-twu</td>
<td>ox</td>
</tr>
<tr>
<td>pun</td>
<td></td>
<td>honored persons</td>
</tr>
<tr>
<td>salam</td>
<td></td>
<td>persons</td>
</tr>
<tr>
<td>myeng</td>
<td></td>
<td>persons (formal)</td>
</tr>
<tr>
<td>cang</td>
<td></td>
<td>flat things, sheets, (news)papers, letters</td>
</tr>
<tr>
<td>twayki</td>
<td></td>
<td>sheets, mats</td>
</tr>
<tr>
<td>chay</td>
<td>-ho, -tong۱</td>
<td>buildings</td>
</tr>
<tr>
<td>tay۱</td>
<td>-tay</td>
<td>vehicles, machines</td>
</tr>
<tr>
<td>calwu</td>
<td></td>
<td>long objects with handles, pencils</td>
</tr>
<tr>
<td>tay۲</td>
<td></td>
<td>cigarettes, pipefuls</td>
</tr>
<tr>
<td>tags</td>
<td></td>
<td>slaps</td>
</tr>
<tr>
<td>pel</td>
<td></td>
<td>suits of clothes/garments, sets of dishes/tablewear, copies of a set of documents</td>
</tr>
<tr>
<td>cwuk</td>
<td></td>
<td>ten garments/dishes</td>
</tr>
<tr>
<td>cepsî</td>
<td></td>
<td>(menu) dishes</td>
</tr>
<tr>
<td>khyel(l)ey</td>
<td></td>
<td>pairs of footwear/gloves</td>
</tr>
<tr>
<td>kulwu</td>
<td>-mun</td>
<td>cannons, big guns</td>
</tr>
<tr>
<td>kwul</td>
<td>-cwú</td>
<td>trees, shrubs</td>
</tr>
<tr>
<td>kawso</td>
<td>-kawso</td>
<td>places, institutions</td>
</tr>
<tr>
<td>tan</td>
<td></td>
<td>bunches (vegetables)</td>
</tr>
<tr>
<td>tapal</td>
<td></td>
<td>bunches (flowers, plants)</td>
</tr>
<tr>
<td>kkokci</td>
<td></td>
<td>bunches, clumps (false hair, seaweed)</td>
</tr>
<tr>
<td>cwal</td>
<td></td>
<td>sheaves, tied bunches, tobacco strings</td>
</tr>
<tr>
<td>mus۱</td>
<td></td>
<td>sheaves (straw)</td>
</tr>
<tr>
<td>tengeli</td>
<td></td>
<td>bundles of chopped firewood</td>
</tr>
<tr>
<td>pen</td>
<td>-hay</td>
<td>times</td>
</tr>
<tr>
<td>swu</td>
<td>-pal</td>
<td>moves (chess/checkers)</td>
</tr>
<tr>
<td>păng</td>
<td></td>
<td>(gun) shots</td>
</tr>
<tr>
<td>swun</td>
<td></td>
<td>rounds (of 5 arrows shot)</td>
</tr>
<tr>
<td>chung</td>
<td>-chung</td>
<td>floors, stories</td>
</tr>
</tbody>
</table>

10If referring to big items, these are understood as a miniature version or a toy (cf. Aikhenvald 2000: 336).
ad F4: The Korean classifiers occur once and overtly in the context of classification (cf. again [136] and [139]).

(139)  A  kāy  sēy  mali
       dog  three  CL:ANIMAL
       “three dogs” (Korean; cf. Martin 1992: 179)
B  catong-cha  sēk  tay1
   car     three  CL:VEHICLE
       “three automobiles”
C  pus  sēk  calwy
       brush three  CL:LONG.OBJECT.WITH.HANDLE
       “three writing brushes”
D  ppyam  sēk  tay3
    slap   three  CL:SLAP
       “three slaps” (Korean; cf. Martin 1992: 180)

ad F5: The Korean classifiers occur within the borders of the NP (cf. again [138] where the classifier mali “animal” occurs in the NP which is headed by cwi “rat”).

ad F6: The bulk of Korean classifiers does not seem to add semantic content to the noun phrase, since its assignment is based on the inherent characteristics of the nominal referent. This information is copied from the head noun rather than imposed on it. An exception are the ‘unitisers’, which express a certain portion or collection of an entity (cf. again [136]), but which are interpreted as a functional equivalent to measure nouns and rather than classifiers within this dossier. However, there is another semantic niche of the Korean classifier system, which displays a semantic contribution to the classifier’s host NP. This niche contains classifiers expressing a person’s social status or value (cf. again [134]).

ad F7: The semantics of the classifiers are transparent (cf. table 87 on page 498) and there is a large set of classifiers available, which enables the speaker to assign them at his or her free will (cf. again [134]).

11 If the function of measure nouns would be treated as nominal classification, languages such as English or German had to be analysed as employing classifiers, too (cf. e.g. German ein Stück Brot, zwei Stück Brot ‘one piece of bread, two piece[s] of bread’. Since this would cause serious problems for a valid definition of classifiers, measure nouns (and their functional equivalents) are analysed as a phenomenon different from nominal classification. However, the functional account for the Korean system acknowledges this individualising function of the Korean system.
18.3.3 Mapping

![Graph showing mapping values]

Figure 39: Korean.

Table 88: The formalisation and transparency value of Korean.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>0</td>
<td>semantic structure</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>0.5</td>
<td>rule types</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>0.5</td>
<td>rule number</td>
</tr>
<tr>
<td>obligatoryness</td>
<td>$X_4$</td>
<td>0.333</td>
<td>categorial independence</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>0.25</td>
<td>discreteness of markers</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>0</td>
<td>redundancy</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>0</td>
<td>flexibility</td>
</tr>
<tr>
<td>x-value</td>
<td>0.226</td>
<td>y-value</td>
<td>0.702</td>
</tr>
</tbody>
</table>

**ad X₁:** Korean employs more than a hundred classifiers (cf. Martin 1992: 179ff.) and thus possesses a large inventory.

**ad X₂:** The classifiers occur with quantifiers (cf. [139]) adjectives, and demonstratives (cf. [134]).

**ad X₃:** The classifiers occur in noun phrases, both ones that feature the classified noun and ones that lack an overt noun (cf. [134]).
ad X₄: Outside quantificational contexts, classification is not obligatory (cf. 138, where not all nouns are categorised). Abstract nouns may remain uncategorised in general (Aikhenvald 2000: 334).

ad X₅: The Korean classifiers occur both in independent and bound form (cf. table 87 on page 498).

ad X₆: The classifiers occur once per context, not multiply.

ad X₇: As abstract nouns are not classified (Aikhenvald 2000: 334), the Korean system is not exhaustive.

ad Y₁: Most of the Korean classes are semantically fairly specific, some of them are a more generic in their semantics (cf. table 87 on page 498). Therefore, Korean is inbetween the state of a mixed and a specific system.

ad Y₂: The consulted sources do not report on other assignment rule types than semantic ones.

ad Y₃: Due to the high number of classes, the number of assignment rules is high as well.

ad Y₄: The classifiers convey classificatory semantics and classificatory semantics only.

ad Y₅: Most Korean classes are marked by classifiers that are reserved for a single class. There are however a few classifiers that refer to more than one class (cf. e.g. tay in table 87 on page 498).

ad Y₆: Most of the available examples do not suggest an informative semantic contribution to the noun phrase by classificatory means (cf. e.g. 139). The most obvious semantic contribution is found in social classification (cf. 134). The Korean system is thus at an intermediate position between semantically redundant and informative classification.

ad Y₇: The classifier (kay “small countable inanimates”) can vary with more specific classifiers for small countable items (e.g. yemphil twu kay “two pencils” (Aikhenvald 2000: 336) vs. yemphil twu calwu “two pencils” (compiled after Aikhenvald 2000 and Martin 1992)).
18.4 Grammaticalisation

There are several indications of a progress of grammaticalisation with respect to the Korean numeral classifier system. One is the semantic extension of the classifiers such as *tay*, which originally denoted traditional vehicles and later got extended to the denotation of artifacts with wheels, which even includes “any electric machinery, and [...] other kinds of machines or instruments, including even the piano” nowadays (Aikhenvald 2000: 311). Another is the general semantic bleaching, which leads to an impoverishment of the classifier system among younger speakers of Korean. This decline is typical for mensural classifiers, which “often refer to culture-specific arrangements and measures, and the obsolescence of cultural knowledge inevitably leads to their loss” (Aikhenvald 2000: 386).

According to Hwang et al. (2008: 155f.), the classifiers can be divided into three groups (‘prototype classifiers’, ‘quasi-classifiers’, and ‘classifier nouns’). ‘Prototype classifiers’ do not have free lexical nouns as a counterpart; ‘quasi-classifiers’ are “used largely as a classifier but can often be used as a common noun (e.g., *-galag, -doe, -majigi*, etc.)” (Hwang et al. 2008: 156). They appear to have entered the classifier system only recently and are still in the process of grammaticalisation (cf. ib.: 155f.). The same accounts for “classifier nouns”, which correspond to count nouns and occur both in classificatory function and as free nouns (ib.).

Korean displays a low degree of formalisation (value: 0.226), which is noticeable mostly in the partial boundedness of its markers and the extended range of morphosyntactic hosts. At the same time, the system is fairly transparent (value: 0.702), as large parts are semantically specific and bijective. The system allows for a certain degree of classificatory flexibility. The high number of assignment rules and the low degree of productive and informative semantic contribution to the noun phrase are the factors that repress a higher degree of classificatory transparency in Korean.

19 Maasai

19.1 Background


Maasai is a language that is spoken in the Rift Valley Province and the Kajiado and Narok districts of Kenya, as well as in Tanzania. This dossier is mainly based on Payne 1998 and Tucker and Tempo Ole Mpaayei 1955, with additional consultation of Dimmendaal 2000, Sim 1980 and van Dorn Lambourneaux 2004. Maasai employs a concordial class system that is semantically
highly productive.

![Diagram of nominal classification devices]

Figure 40: Classification of the Maasai system of nominal classification.

19.2 Classification Typology

19.2.1 Noun Classifiers

Maasai does not employ a noun classifier system (cf. 140 where no classifier occurs).

(140) *Ei-torrono ol-osowuan?*  
be-bad M.SG-buffalo:SG  
“Is the buffalo bad?” (Maasai; cf. Tucker and Tompo Ole Mpaayei 1955: 12)

19.2.2 Numeral Classifiers

Maasai does not employ a numeral classifier system (cf. 141 where no classifier occurs).

(141)  
A *Ol-tuyani obo, Iluyana aare.*  
M.SG-man:SG one M.PL-man:PL two  
“One man, two men.”

B *En-koloŋ nabo, In-koloiŋi are.*  
F.SG-day:SG one F.PL-day:PL two  
“One day, two days.”

C *E-wueji nebo, I-wuejitin are.*  
PLACE.SG-place:SG one PLACE.PL-place:PL two  
“One place, two places.” (Maasai; cf. Tucker and Tompo Ole Mpaayei 1955: 32)

19.2.3 Possessive Classifiers

Maasai does not employ a possessive classifier system (cf. 142 where no classifier occurs).

(142) *Enaai ena alem.*  
POSS.F.SG DEM.CLOSE.F.SG knife  
“This knife [is] mine.” (Maasai; cf. Tucker and Tompo Ole Mpaayei 1955: 12)
Appendix

19.2.4 Multiple Classifier Systems

Maasai does not employ a multiple classifier system (cf. 140, 141 and 142, where no classifier occurs).

19.2.5 Concordial Class Systems

Maasai employs a tripartite concordial class system (cf. 143).

(143) ol-ayioni, M.SG-boy il-ayiok, M.PL-boys en-tito, F.SG-girl in-toyie, F.PL-girls en-títo, PLACE.SG-place i-wuejitin PLACE.PL-places

“boy, boys, girl, girls, place, places” (Maasai; cf. Tucker and Tompo Ole Mpaayei 1955: 3 and 15)

19.3 Analysis

19.3.1 Functional Setup

Table 89: The semantic and discourse properties of the Maasai system of nominal classification.

Table: differentiating referents (C₁) productive derivation, variable classification ✓
individuation (C₂) unitisation for countability ✗
attribute of properties (C₃) speaker-sided values ✓
reference identification (C₄) reference tracking via AGR, anaphora ✓

ad C₁: There are a few productive derivation rules involving the Maasai system: Nominalisations take the feminine gender, except for negative concepts, which take the masculine gender (cf. 144 A and B). The Maasai system also distinguishes the biological sex of the nominal referent and a referent’s size/strength by flexible class marking (cf. 144 C and D).

(144) A em-bútá, F.SG-extraction e-míshírà́r`, F.SG-branding e-řok, F.SG-blackness e-káh, F.SG-hate
“extraction, branding, blackness, hate” (Maasai; cf. Payne 1998: 170)

B ol-lóm, M.SG-envy ol-mená M.SG-contempt
“jealousy, contempt” (Maasai; cf. Payne 1998: 171)

“female/small doctor (or quak), male doctor, sister, very large sister, girl, ‘large shapeless hulk of a woman’, weak brother, brother, wimpy male donkey, male donkey, woman, very respected man” (Maasai; cf. Payne 1998: 166)

ad C2: The consulted sources and available data do not suggest the Maasai system to be involved in individuation.

ad C3: The connotation of the classes with size and strength allow the speaker to attribute values to a nominal referent by classificatory means (cf. 144 C).

ad C4: By marking e.g. relative constructions for concordial class, the Maasai system employs reference identification (cf. 145).

“the man who is running, the cow which is running” (Maasai; cf. Tucker and Tompo Ole Mpaayei 1955: 23)
B obo/nabo/nebo, one:M./F./PLACE aare/are/are, two:M./F./PLACE okuni/uni/uni, three:M./F./PLACE oopishana/naapishana/naapishana, seven:M./F./PLACE ooudo/naaudo/naaudo, nine:M./F./PLACE
“one, two, three, four, seven, nine” (Maasai; cf. Tucker and Tompo Ole Mpaayei 1955: 32)

19.3.2 Formal Setup

Table 90: The formal properties of the nominal classification system of Maasai.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
</tbody>
</table>
assignem**nt** (F7) | CLS-like: chosen freely by the speaker | ✓
| | CCS-like: mechanic assignment

**ad F₁:** Each noun is assigned to a concordial class (cf. Payne 1998: 159).

**ad F₂:** The “vast majority of common nouns [...] can choose either the masculine or the feminine gender prefix” (Payne 1998: 159f.).

**ad F₃:** Maasai employs three concordial classes, the masculine, the feminine, and the place class (cf. Tucker and Tompo Ole Mpaayei 1955: 15).

**ad F₄:** Since the Maasai concordial class system employs agreement on different nominal modifiers, multiple occurrence of concordial class marking is possible (cf. 146).

(146) *ainaŋai naa-shomo*

who:F.PL F.PL-have.gone

“Who (f.) have gone?” (Maasai; cf. Tucker and Tompo Ole Mpaayei 1955: 27)

**ad F₅:** Verbal forms can be marked for concordial class (cf. 146), which means that class marking in Maasai is not restricted to the borders of the noun phrase.

**ad F₆:** A noun’s concordial class membership usually contributes to the noun phrase’s semantics (cf. 144 C and Payne 1998: 171-173).

**ad F₇:** Only a small part of the nominal lexicon displays an entirely fixed concordial class membership (cf. Payne 1998: 169). The nouns possess an inherent class value, but can be assigned to either concordial class according to the speaker’s cognitive construal of a referent (Payne 1998: 168).
19.3.3 Mapping

Table 91: The formalisation and transparency value of Maasai.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>semantic structure</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>rule types</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>rule number</td>
<td>$Y_3$</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>categorial independence</td>
<td>$Y_4$</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>discreteness of markers</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>redundancy</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>flexibility</td>
<td>$Y_7$</td>
</tr>
<tr>
<td>x-value</td>
<td>1</td>
<td>y-value</td>
<td>0.690</td>
</tr>
</tbody>
</table>

**ad X₁:** The Maasai system employs three classes (cf. table 92). The third class contains only the word *ww* 'ejì “place” (Payne 1998: 160).
Table 92: The inventory of the Maasai class markers (cf. Tucker and Tompo Ole Mpaayei 1955: 15).

<table>
<thead>
<tr>
<th>marker (sg./pl.)</th>
<th>class/number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ol-/il</td>
<td>M.SG/PL</td>
</tr>
<tr>
<td>en(k)-/in(k)-</td>
<td>F.SG/PL</td>
</tr>
<tr>
<td>e-/i-</td>
<td>PLACE.SG/PL</td>
</tr>
</tbody>
</table>

ad X2: Several Maasai modifiers are marked for their head noun’s class: Demonstratives, possessives, indefinites, relativisers, ‘genitive linkers’, certain singular interrogatives and numerals (Payne 1998: 160).

ad X3: Class marking may attach to verbal forms (cf. 146) and thus occur NP-externally.

ad X4: Class marking is obligatory. Speakers virtually always give a prefixed noun stem as citation form (Payne 1998: 162).

ad X5: The markers occur as bound affixes (cf. table 92).

ad X6: The system allows multiple marking in the same context (cf. 142).

ad X7: The classification applies to the full nominal lexicon (Payne 1998: 159).

ad Y1: The classes are semantically mixed. As a rule, most nouns with an unmarked, fixed class can be shifted, if the size/strength connotations are applied (Payne 1998: 171f.). The feminine class seems to be the default class (which is e.g. assigned when a noun’s original class is unknown or if a noun is borrowed; Payne 1998: 172).

ad Y2: A noun typically has an inherent class value (Payne 1998: 171f.). Most nouns can be reassigned to either class, if a size or strength connotation is made (cf. 145 C).

ad Y3: Due to the transparent and productive semanticity of the system and the feminine class serving as the default class, a low number of assignment rules is sufficient for the speaker to master the system.

ad Y4: The class markers are fused with the category number (cf. 92).

ad Y5: Each class has its own, distinct markers (cf. table 92).
ad Y₆: Due to the connotation of sex, size, and strength, the markers contribute to their noun phrase in a productive way (cf. 145 C).

ad Y₇: According to Payne (1998: 159f.), the “vast majority of common nouns [...] can choose either the masculine or the feminine gender prefix”.

19.4 Grammaticalisation
The Maasai system displays an advanced degree of formalisation (value: 1) by marking class membership on a wider range of nominal satellites, in- and outside of the noun phrase, and by means of bound markers. Due to a strong association of the masculine class with male sex, strength, and bigger size as opposed to a strong association of the feminine class with female sex, weakness, and littleness, it is highly transparent and semantically productive (value: 0.690).

20 Mandarin

20.1 Background

< Chinese < Sino-Tibetan (ISO 639-3: cmn; cf. Lewis et al. 2013: s.v.)


I want to thank my colleague Jing Lin for contributing to this dossier with her abilities as a linguist and a native speaker of Chinese.
20.2 Classification Typology

20.2.1 Noun Classifiers

The Mandarin classifiers may occur with the bare noun, the numeral, and the demonstrative (cf. 147).

(147) A  naï yī-pí hēi mā
        DEM:REMOTE   one-CL:HORSE black horse
        “That black horse.” (Mandarin; cf. Sun 2006: 160)
B  naï pí hēi mā
        DEM:REMOTE  CL:HORSE black horse
        “That black horse.” (Mandarin; cf. Sun 2006: 159)
C  mài pí hēi mā
        buy CL:HORSE black horse
        “Buy one black horse.” (Mandarin; cf. Sun 2006: 160)

Mandarin does not employ a noun classifier system (cf. 148, where no classifier occurs).

(148) zài jiā kǎo fù, mǔ, zài wài kǎo péngyǒu
        at home rely father mother at outside rely friends
        “At home one relies on parents, outside one relies on friends.” (Mandarin; cf. Sun 2006: 119)

20.2.2 Numeral Classifiers

The Mandarin classifiers occur among others in quantificational contexts. They also occur together with the bare noun and demonstratives (cf. 147), which renders these classifiers are multiple classifiers.

20.2.3 Possessive Classifiers

Mandarin does not employ a possessive classifier system (cf. 149, where no classifier occurs).
Appendix

(149) A  tāmén  duì  zhè-ge  wèntī  de  taolún
they  towards  this-CL  problem  POSS  discussion
“their discussion of this problem” (Mandarin; cf. Wu 2004: 78)
B  zhè  shì  nín-de  chá
DEM.CLOSE  be  you:POLITE-POSS  tea
“This is your tea.” (Mandarin; cf. Po-Ching and Rimmington 2004: 48)

20.2.4 Multiple Classifier Systems

The Mandarin classifiers occur in quantificational and demonstrative constructions and with the noun (cf. 147). The Mandarin system is thus a multiple classifier system. It is analysed in section 20.3.

20.2.5 Concordial Class Systems

Mandarin does not employ a concordial class system (cf. 150, where no class agreement occurs on nominal satellites).

(150) A  hǎo  xuěshēng
good  student
“a good student/good students” (Mandarin; cf. Po-Ching and Rimmington 2004: 59)
B  zhè běn  zìdiǎn  hǎo.
DEM:CLOSE:DEF  dictionary  good
“This dictionary is good.” (Mandarin; cf. Po-Ching and Rimmington 2004: 60)

20.3 Analysis

20.3.1 Functional Setup

Table 93: The semantic and discourse properties of the Mandarin system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Mandarin System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differentiating referents (C₁)</td>
<td>productive derivation, variable classification ✔</td>
</tr>
<tr>
<td>Individuation (C₂)</td>
<td>unitisation for countability ✔</td>
</tr>
<tr>
<td>Attribution of properties (C₃)</td>
<td>speaker-sided values ✔</td>
</tr>
<tr>
<td>Reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora ✔</td>
</tr>
</tbody>
</table>

ad C₁: The Mandarin classifiers are independent elements and do not take part in nominal stem formation. They may serve the purpose of referent differentiation by means of variable classification (cf. 151 on page 512).

(151) A  yì  lún  yuèliáng
one  CL:ROUND  moon
“a full moon”
Appendix

B  yu  wan  yueliang
one  CL:CURVED  moon
“a crescent moon” (Zhang 2007: 48)

**ad C₂:** The Mandarin classifiers serve the function of individuation (Bisang 1999: 4).

**ad C₃:** The classifier set employs a certain variability, e.g. social classification (cf. 152). A speaker-sided attribution of properties is thus possible.

(152)  A  yi  zhang  /  fu  /  zhen  hua
one  CL:FLAT  /  CL:CLOTH  /  CL:ROLE  painting
“a painting”

B  yi  ge  /  zhi  /  tou  yang
one  CL:GENERAL  /  CL:ANIMAL  /  CL:HEAD  sheep
“a sheep”

C  yi  ge  /  wei  /  jiangjun
one  CL:GENERAL  /  CL:POLITE  /  CL:PEOPLE.COUNTER  /  CL:RANK.GENERAL  general
“a general” (Mandarin; cf. Zhang 2007: 53)

**ad C₄:** Chinese classifiers can be employed as carriers of anaphora and thus can serve reference identification (cf. 153).

(153)  Zhe  jian  wuzi  you  liang  zhang  shuzhuo,  na  zhang  shi  wode.
this  CL  room  have  two  CL  desk  that  CL  be  mine
“This room has two desks, that (one) is mine.” (Mandarin; cf. Zhang 2007: 47)

20.3.2 Formal Setup

Table 94: The formal properties of the nominal classification system of Mandarin.

<table>
<thead>
<tr>
<th></th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness (F₁)</strong></td>
<td>nouns may remain uncategorised</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>obligatory assignment of each noun</td>
<td>✗</td>
</tr>
<tr>
<td><strong>flexibility (F₂)</strong></td>
<td>flexible class membership</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>single class membership</td>
<td>✓</td>
</tr>
<tr>
<td><strong>inventory (F₃)</strong></td>
<td>largish class number</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>limited class number</td>
<td>✗</td>
</tr>
<tr>
<td><strong>marking (F₄)</strong></td>
<td>marked once</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>multiply marked</td>
<td>✗</td>
</tr>
</tbody>
</table>
Appendix

<table>
<thead>
<tr>
<th>F</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>F3</td>
<td>locus operandi</td>
<td>locus operandi: NP-internal</td>
</tr>
<tr>
<td>F6</td>
<td>semanticity</td>
<td>semantic contribution</td>
</tr>
<tr>
<td>F7</td>
<td>assignment</td>
<td>chosen freely by the speaker</td>
</tr>
</tbody>
</table>

ad F1: Nouns are only obligatorily categorised in numeral and demonstrative phrases (cf. Li and Thompson 1981: 104 and 154 where the underlined nouns are not categorised).

(154) lù-shàng chè-le chē-huò
road-up happen-PERFECTIVE car-crash

“[There was] a traffic accident on the road.” (Mandarin; cf. Sun 2006: 198)

ad F2: “The choice of classifier is determined by the noun” (Li and Thompson 1981: 105). Except for the generic classifier gē, which can occur in all instances, the combination of a noun and a classifier is “usually considered to be a stable process” in Mandarin. However, spoken Mandarin appears to be more liberal with respect to the otherwise standardised classifier use (cf. Zhang 2007: 55). Erbaugh (1986: 404ff.) reports a great degree of classificatory flexibility across the registers.

ad F3: Li and Thompson (1981: 105) mention “several dozen classifiers”. Zhang (2007: 44) reports of countings that reach the incredible number of 902 classifiers (including verbal classifiers and measure nouns). Peiru et al. (2001: preface) mention that they treat 187 classifiers, which cover only two of four categories of classifier-like elements. The internet offers a list of 162 nominal classifiers (http://en.wikipedia.org/wiki/List_of_Chinese_classifiers#Nominal_classifiers; viewed 27.11.2013, 15:00). The number of classifiers is thus large.

ad F4: The classifiers occur once per context (cf. 147 where only one classifier occurs, also in contexts with more than one potential host).

ad F5: The classifiers occur NP-internally and independent of the presence of the classified noun (cf. 153 where the classifier is used as a carrier of anaphora).

ad F6: The classifiers may semantically contribute to the noun phrase (cf. 151).

20.3.3 Mapping

![Figure 43: Mandarin.](image)

Table 95: The formalisation and transparency value of Mandarin.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>semantic structure</td>
<td>$Y_1$ 0.75</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>rule types</td>
<td>$Y_2$ 0.666</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>rule number</td>
<td>$Y_3$ 0</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>categorial independence</td>
<td>$Y_4$ 1</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>discreteness of markers</td>
<td>$Y_5$ 0.75</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>redundancy</td>
<td>$Y_6$ 0.75</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_7$</td>
<td>flexibility</td>
<td>$Y_7$ 1</td>
</tr>
</tbody>
</table>

| x-value                  | 0.417 |
| y-value                  | 0.702 |

**ad X₁:** This study follows Peiru et al. (2001: preface) in assuming Mandarin to employ at least 187 classifiers.

**ad X₂:** The classifiers combine with quantifiers, demonstratives, and the noun (cf. [147].)
ad X₃: While Erbaugh (1986: 403) “finds that classifiers were, in fact, seldom used as pro-forms in spoken Mandarin”, an anaphoric use of classifiers is possible (cf. 153). The classifiers thus operate NP-internally and independent of the presence of the classified noun.

ad X₄: Quantificational and demonstrative contexts feature an obligatory classifier (Li and Thompson 1981: 104).

ad X₅: The classifiers occur as free and bound elements (cf. 147).

ad X₆: The classifiers occur once per context, even if it contains more than one potential host (cf. 147).

ad X₇: The general classifier ge makes a classification of the full nominal lexicon possible.

ad Y₁: The many specific and straightforward classifiers of Mandarin (e.g. pǐ “horse”) are complemented by more general ones (e.g. the general classifier ge). Beyond this, individual classifiers display semantic bleaching, which leads to them categorising nouns that would not fall under their original scope (cf. 155, where the original shape categorisation does not apply in all cases).

(155)

A yi tiao neiku
    one   CL:LONGISH underwear
“one pair of underwear”

B yi tiao sanjiao ku
    one   CL:LONGISH brief
“one pair of briefs”

C yi tiao youyong ku
    one   CL:LONGISH swimming.trunk
“a swimming trunk”

D yi tiao duan ku
    one   CL:LONGISH shorts
“a pair of shorts” (Mandarin; cf. Zhang 2007: 47)

ad Y₂: The originally semantic assignment is often overridden by conventionalised and analogical assignment (cf. Bisang 1999: 23 and 155).

ad Y₃: The large classifier inventory and frequent cases of opaque assignment lead to a high number of assignment rules (cf. Zhang 2007: 52 and 53).

ad Y₄: The classifiers are not fused with other categories (cf. e.g. 152).
Most of the classifiers denote a single class. However, a number of classifiers referring to more than one class does exist (e.g., zhi “animal (mammals, birds, insects, frogs); one of a pair; three-dimensional, container-/ball-like object”; cf. 156).

(156) A yi zhi laohu, yi zhi niao, yi zhi pangxie
one CL:ANIMAL tiger one CL:ANIMAL bird one CL:ANIMAL crab
“one tiger, one bird, one crab” (Mandarin; cf. Zhang 2007: 55)

B yi zhi yanjing, yi zhi xie
one CL:ONE.OF.PAIR eye one CL:ONE.OF.PAIR shoe
“one eye, one shoe”

C yi zhi lanzi, yi zhi qiqiu
one CL:CONTAINER basket one CL:CONTAINER balloon
“one basket, one balloon” (Mandarin; cf. Zhang 2007: 56)

Classifiers may semantically contribute to the noun phrase (cf. 151) or mirror semantics that are present in the categorised noun (cf. 156 A and 147).

The same noun may take different classifiers without changing its referent (cf. 152 and Bisang 1999: 23).

20.4 Grammaticalisation
The Mandarin system shows distinct traits of formalisation (value: 0.417): Its markers may occur in bound form and attach to a variety of hosts. At the same time, the frequently conventionalised classifier-noun pairings and the large class inventory render the system a comparably intransparent one (value: 0.702).

21 Masri

21.1 Background
< Arabic < South Central Semitic < Central Semitic < Semitic < Afro-Asiatic (ISO 639-3: arz; cf. Lewis et al. 2013: s.v.)
< Arabic < Arabo-Canaanite < Central West Semitic < West Semitic < Semitic < Afro-Asiatic (cf. Ruhlen 1987: 18)

Since Ruhlen (1987: 323) does not list Egyptian Arabic or Masri by its name, but mainly distinguishes the varieties of Arabic by the cardinal directions, another interpretation of its affiliation is possible. This other interpretation would be < South Semitic < West Semitic < Semitic < Afro-Asiatic.
Arabic is a Semitic language comprising several varieties. This dossier focuses on Masri (or Egyptian Arabic), which is spoken in Egypt, Iraq, Israel, Jordan, Kuwait, Libya, Saudi Arabia, Syria, the United Arab Emirates, and Yemen (cf. Lewis et al. 2013: s.v.). It is mainly based on Abdel-Massih et al. 1979 and Spitta-Bey 1880. The Masri fully meets the expectations of a concordial class system.

Figure 44: Classification of the Masri system of nominal classification.

### 21.2 Classification Typology

#### 21.2.1 Noun Classifiers

Masri does not employ noun classifiers (cf. [157] where no classifier occurs).

(157) \( \text{da } \text{walad } \text{kibiir} \)

\( \text{this } \text{boy:M.SG } \text{big:M.SG} \)

“This [is] a big boy.” (Masri; cf. Abdel-Massih et al. 1979: 47)

#### 21.2.2 Numeral Classifiers

Masri does not employ numeral classifiers, and the Masri numerals are not even marked for concordial class (except \( \text{waalid} \) m. “one” and \( \text{waalda} \) f. “one”; cf. Abdel-Massih et al. 1979: 164ff.).

#### 21.2.3 Possessive Classifiers

The Masri possessive construction does not employ classifiers, but may indicate the possessor’s concordial class by pronominal suffixes or the possessive marker \( \text{bitaa}^{\prime} \) m., \( \text{bitaa}^{\prime}a \) f., or \( \text{bituu}^{\prime} \) in the plural (cf. Abdel-Missah et al. 1979: 15ff. and 216ff.).

#### 21.2.4 Multiple Classifier Systems

Masri does not employ multiple classifiers (cf. Abdel-Missah et al. 1979: 15ff., 166ff., and 216ff., and [157] where no classifier occurs with the noun, demonstrative, or adjective).
21.2.5 Concordial Class Systems

Masri employs a concordial class system (cf. 158). This system is analysed in section 158.

(158) fallahi, fallahi-in, fallahi-n, fallahi-naat
farmer:MG | farmer:MP | farmer:FG | farmer:FP

"(male) farmer, (male) farmers, (female) farmer, (female) farmers" (Masri; cf. Abdel-Massih et al. 1979: 211)

21.3 Analysis

21.3.1 Functional Setup

Table 96: The semantic and discourse properties of the Masri system of nominal classification.

| differentiating referents (C₁) | productive derivation, variable classification | ✓ |
| indviduation (C₂) | unitisation for countability | ✗ |
| attribution of properties (C₃) | speaker-sided values | ✗ |
| reference identification (C₄) | reference tracking via AGR, anaphora | ✓ |

ad C₁: Variations in a noun’s class membership are intradialectal; dialect-internally, a noun’s concordial class membership is invariable (Spitta-Bey 1880: 123f.). However, the sex of human referents is distinguished by concordial class marking (cf. 158).

ad C₂: The consulted sources and available data do not suggest the Masri system to be involved in the individuation of nominal referents.

ad C₃: The consulted sources and available data do not suggest the Masri system to be involved in the speaker-sided attribution of properties by classificatory means.

ad C₄: By marking e.g. pronouns for concordial class, the Masri system plays a role in referent identification (cf. 159).

(159) A kitaab-ak, kitaab-ik
book-2SG.M book-2SG.F

“your book (male referent), your book (female referent)"

B kitaab-u, kitab-ha
book-3SG.M book-3SG.F

21.3.2 Formal Setup

Table 97: The formal properties of the nominal classification system of Masri.

<table>
<thead>
<tr>
<th>Feature</th>
<th>CLS-like:</th>
<th>CCS-like:</th>
</tr>
</thead>
<tbody>
<tr>
<td>obliteratedness ($F_1$)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility ($F_2$)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory ($F_3$)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking ($F_4$)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi ($F_5$)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity ($F_6$)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment ($F_7$)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad $F_1$: The Masri nouns all “are either masculine [...] or feminine [...] in gender” (Abdel-Massih et al. 1979: 101; cf. also Kremers 2003: 50).

ad $F_2$: A noun’s class membership is rigid. Cases of variable classification are interdialectal (Spitta-Bey 1880: 123f.).

ad $F_3$: Masri employs two concordial classes (cf. table 98 for an example of pronominal class marking).

Table 98: The inventory of Masri pronominal suffixes (cf. Abdel-Massih et al. 1979: 216f. and 219f.).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Form after consonant</th>
<th>Form after vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>-i/-ni (always after verbs)</td>
<td>-ga</td>
</tr>
<tr>
<td>2SG.M</td>
<td>-ak</td>
<td>-k</td>
</tr>
<tr>
<td>2SG.F</td>
<td>-ak</td>
<td>-ki</td>
</tr>
<tr>
<td>3SG.M</td>
<td>-(h)i</td>
<td>-(h)i</td>
</tr>
<tr>
<td>3SG.F</td>
<td>-ha</td>
<td>-ha</td>
</tr>
<tr>
<td>1PL</td>
<td>-na</td>
<td>-na</td>
</tr>
<tr>
<td>2PL</td>
<td>-kum</td>
<td>-kum</td>
</tr>
<tr>
<td>3PL</td>
<td>-hum</td>
<td>-hum</td>
</tr>
</tbody>
</table>
ad F₄: Masri marks class on pronouns, adjectives, the verb, and optionally on the noun (cf. e.g. Abdel-Massih et al. 1979: 101, 211, 215f., and 261), which allows multiple marking in the same context.

ad F₅: Class marking occurs among others on the verb (cf. 160) and therefore NP-externally.

\begin{align*}
\text{(160) } & \text{Daras.} & \text{Daras-it.} \\
& \text{study:PF3SG.M} & \text{study:PF3SG.F} \\
\end{align*}

“He studied. She studied.” (Masri; cf. Abdel-Massih 1979: 275)

ad F₆: Despite several relics of a semantics-based prestage of the current Masri concordial class system (e.g. female class membership of names of countries and paired body parts), the system as a whole does not employ semantics actively (the only exception being the differentiation of the sex of animates; cf. 160 and e.g. Abdel-Massih et al. 1979: 101).

ad F₇: Besides a few lexical areas (e.g. paired body parts, names of countries, reference to natural sex; cf. Spitta-Bey 1880: 124ff.), there is no obvious pattern of class assignment in Masri.

21.3.3 Mapping

![Figure 45: Masri.](image-url)
Appendix

Table 99: The formalisation and transparency value of Masri.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>1</td>
<td>semantic structure</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>1</td>
<td>rule types</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>1</td>
<td>rule number</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>1</td>
<td>categorial independence</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>1</td>
<td>discreteness of markers</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>1</td>
<td>reducency</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>1</td>
<td>flexibility</td>
</tr>
<tr>
<td>x-value</td>
<td>1</td>
<td>y-value</td>
<td>0.512</td>
</tr>
</tbody>
</table>

ad $X_1$: Masri employs two concordial classes, a masculine and a feminine one (cf. table 98).

ad $X_2$: Masri marks nouns, adjectives, verbs, personal pronouns, possessives for class (cf. e.g. Abdel-Massih et al. 1979: 101, 211, 215f., and 261).

ad $X_3$: Class marking occurs NP-internally and NP-externally independent of the presence of the classified noun (cf. 159 and 160).

ad $X_4$: The class markers are fused with the category number (cf. table 98). Since number is obligatorily marked in discourse, classification is also obligatory.

ad $X_5$: The class markers occur in affixed form (cf. 158).

ad $X_6$: Since adjectives, verbs, pronouns, and possessives are all marked for class, Masri allows multiple class marking in the same context (cf. 157).

ad $X_7$: All Masri nouns are assigned to a class (Abdel-Massih et al. 1979: 101; cf. also Kremers 2003: 50).

ad $Y_1$: The classes’ semantic structure is mixed: Nouns denoting paired body parts, names of countries, and the sex of animate referents have default classes, whereas the remaining nominal lexicon lacks a semantic motivation for their class membership (cf. Spitta-Bey 1880: 124ff.).

ad $Y_2$: Next to a few semantic assignment rules, class assignment is often opaque. In addition, phonological conditions apply: Most feminine nouns end in -a, most masculine nouns do not (cf. Abdel-Massih et al. 1979: 101, 101 A for examples, and 161 B for counterexamples).
ad Y₃: The more recent consulted sources emphasise that the class assignment is intransparent. This means that a high number of opaque assignment rules apply.

ad Y₄: The class markers are fused with the category number (cf. table 98).

ad Y₅: Each class has its own distinct markers (cf. table 98).

ad Y₆: Class membership can be either informative (cf. 161 A) or opaque (cf. 161).

ad Y₇: Class membership is rigid (Spitta-Bey 1880: 123f.).

21.4 Grammaticalisation

The Masri system displays a far progressed degree of formalisation (value: 1). At the same time, it proves to lack a greater semantic basis and is thus fairly intransparent (value: 0.512).

22 Moro

22.1 Background

Moro (or Dhinorong) is a language spoken in the Sudanic “South Kordofan State, Moro hills” (Lewis et al. 2013: s.v.). The situation of documentation of Moro is relatively sparse: After some “informal linguistic work since the 1930s”, a grammar on Moro was published (Black and Black 1971). In 1981, Schadeberg published a “summary account […] as part of a survey of the Heiban group of languages” (all citations taken from Blench 2005: II). Finally, in the late 1990s, a number of field notes and manuscripts were composed by Guest. Blench (2005: II) notes that the insights of Schadeberg 1981 were not taken into account by Guest. This dossier is mainly based on Black and Black 1971, Blench 2005,

![nominal classification devices]

![CLSs](noun CLS numeral CLS possessive CLS multiple CLS)

![CCSs](✓)

Figure 46: Classification of the Moro system of nominal classification.

### 22.2 Classification Typology

#### 22.2.1 Noun Classifiers

Moro does not employ a noun classifier system (cf. 162, where no classifier occurs).

(162) *Iria yerto ındriano ındağin ındıswordo.*

> cows had oxen among three fat

“There were three fat oxen among the cows.” (Moro; cf. Guest 1998: 8)

#### 22.2.2 Numeral Classifiers

Moro does not employ a numeral classifier system (cf. 163, where no classifier occurs).

(163) *Lëmu leto kman labwerto yeuria ındağin.*

> Some Arabs came riding on three camels.


#### 22.2.3 Possessive Classifiers

Moro does not employ a possessive classifier system (cf. 164, where no classifier occurs).

(164) *na igci  gı-tura  gı-avora tyarba*

> and DEM.I I-old I-more.than his.brother

“and he is older than his brothers” (Moro; cf. Guest 1998: 2)

#### 22.2.4 Multiple Classifier Systems

Moro does not employ a multiple classifier system (cf. 162, 163, and 164, where no classifier occurs).
22.2.5 Concordial Class Systems

Moro employs a concordial class system (cf. 165, where class marking occurs on a range of nominal satellites). This system is analysed in section 22.3.

(165) A δ-\\textit{amalk\={a}}-\textit{G}• δ-\\textit{e-t-\={a}} \textit{VIII-camel-DEM.VIII \textit{VIII-REL-small-ADJ \textit{VIII-run-PF}}

“This small camel ran away.”

B γ-\\textit{\={d}nt-\={g}f} γ-\\textit{e-t-\={a}} \textit{VI-dog-DEM.VI \textit{VI-REL-small-ADJ \textit{VI-run-PF}}

“This small dog ran away.” (Moro; cf. Gibbard et al 2009: 108)

22.3 Analysis

22.3.1 Functional Setup

Table 100: The semantic and discourse properties of the Moro system of nominal classification.

| differentiating referents (C₁) | productive derivation, variable classification | X |
| indidication (C₂) | unitisation for countability | X |
| attribution of properties (C₃) | speaker-sided values | X |
| reference identification (C₄) | reference tracking via AGR, anaphora | ✓ |

\textit{ad C₁}: A noun’s concordial class is determined by its initial phoneme and not vice versa; therefore, the Moro system does not play a role in the expansion of the nominal lexicon. The consulted sources and available data do not suggest the Moro system to be involved in variable classification.

\textit{ad C₂}: The consulted sources and available data do not suggest the Moro system to be involved in the individuation of nominal referents.

\textit{ad C₃}: The consulted sources and available data do not suggest the Moro system to be involved in the speaker-sided attribution of properties by classificatory means.

\textit{ad C₄}: By employing agreement on a wide range of nominal modifiers such as the verb (cf. 165), we can assume the Moro system to play a role in reference identification.
22.3.2 Formal Setup

Table 101: The formal properties of the nominal classification system of Moro.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad F₁: The consulted sources do not state whether the system applies to the full nominal lexicon or not. The consulted data does not suggest that there are parts of the nominal lexicon remain unclassified.

ad F₂: The nouns are assigned due to their inherent semantics or phonological properties (cf. Gibbard et al. 2009: 110). There are no reported cases about a flexible class membership of nouns.

ad F₃: Guest (1997a: 1) identified thirteen concordial classes, whereas Gibbard et al. (2009: 107) suggest a total of eighteen classes, consisting of eight main classes with a singular-plural distinction, five ‘unpaired’ classes (hosting mass nouns and verbal nouns), and five ‘minor categories’, which host one to five nouns (which is indicated by bracketed numerals in table 102). These are listed in table 102.
### Table 102: The inventory of Moro class markers (cf. Gibbard et al. 2009: 107 and 109f.)

<table>
<thead>
<tr>
<th>agreement marker (SG, PL)</th>
<th>semantics</th>
<th>glossing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>main class pairings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g- (k-), l-</td>
<td>“people”</td>
<td>I</td>
</tr>
<tr>
<td>g- (k-), n-</td>
<td>uncertain</td>
<td>II</td>
</tr>
<tr>
<td>j- (k-, s-), j- (s-)</td>
<td>uncertain</td>
<td>III</td>
</tr>
<tr>
<td>l-, y-</td>
<td>“round/long things, fruit”</td>
<td>IV</td>
</tr>
<tr>
<td>l-, j-</td>
<td>“animals, body parts, objects”</td>
<td>V</td>
</tr>
<tr>
<td>y-, j-</td>
<td>“small animals”</td>
<td>VI</td>
</tr>
<tr>
<td>h-, r-</td>
<td>“animals, long things”</td>
<td>VII</td>
</tr>
<tr>
<td>d-, j-</td>
<td>uncertain</td>
<td>VIII</td>
</tr>
<tr>
<td><strong>unpaired classes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g- (k-)</td>
<td>“mass” (not further specified)</td>
<td>IX</td>
</tr>
<tr>
<td>j- (k-), s-</td>
<td>“mass” (not further specified)</td>
<td>X</td>
</tr>
<tr>
<td>y-</td>
<td>“liquids, mass nouns”</td>
<td>XI</td>
</tr>
<tr>
<td>d-1</td>
<td>“nature”</td>
<td>XII</td>
</tr>
<tr>
<td>d-2</td>
<td>“infinitives, abstracts”</td>
<td>XIII</td>
</tr>
<tr>
<td><strong>minor categories</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j- (k-), y-</td>
<td>“termite” (1)</td>
<td>XIV</td>
</tr>
<tr>
<td>l-, j- (s-)</td>
<td>“tooth” (1)</td>
<td>XV</td>
</tr>
<tr>
<td>r-, j- (s-)</td>
<td>“goat, etc.” (4)</td>
<td>XVI</td>
</tr>
<tr>
<td>d-, k- (p-)</td>
<td>“derivatives of trees” (5)</td>
<td>XVII</td>
</tr>
<tr>
<td>l-</td>
<td>“honey” (1)</td>
<td>XVIII</td>
</tr>
</tbody>
</table>

Typically, the classes seem to be labeled by their concordial singular and plural affixes, e.g. the “g/l” class, the “g/n” class, etc. (cf. Gibbard et al. 2009: 107 and Guest 1997a: 2).

**ad F4:** Moro allows multiple marking in the same context (cf. [165]).

**ad F5:** Moro marks class membership among others on the verb and thus NP-externally (cf. [165]).

**ad F6:** Since a concordial class is assigned due to inherent characteristics of a noun’s referent or phonological characteristics of its form (cf. Gibbard et al. 2009: 110), the concordial class does not add semantic information to it.

**ad F7:** Though the semantic grounds of the individual classes are neither exhaustive nor compulsory, the individual noun’s phonological properties provide enough information to theoretically allow a speaker to assign nouns (cf. Gibbard et al. 2009: 110ff.). Since this happens within the mechanisms of the system and the speaker has no influence to alter a noun’s class by his or her will, the assignment stays mechanic.
22.3.3 Mapping

![Figure 47: Moro.](image)

Table 103: The formalisation and transparency value of Moro.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>0.75</td>
<td>semantic structure</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>1</td>
<td>rule types</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>1</td>
<td>rule number</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>1</td>
<td>categorial independence</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>1</td>
<td>discreteness of markers</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>1</td>
<td>redundancy</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>1</td>
<td>flexibility</td>
</tr>
<tr>
<td><strong>x-value</strong></td>
<td>0.964</td>
<td><strong>y-value</strong></td>
<td>0.404</td>
</tr>
</tbody>
</table>

ad X₁: Moro employs eighteen concordial classes (cf. table 102).

ad X₂: Several Moro modifiers agree in class with their head noun, namely several tense-aspect-mood forms of the verb as their argument agreement markers, adjectives, numerals (and other quantifiers), demonstratives, possessives, and the instrumental suffix (cf. Gibbard et al. 2009: 107).

ad X₃: The class markers occur among others on the verb and thus NP-externally (cf. 165).
ad X₄: The consulted sources report a number of nominal satellites to agree with their head noun. As agreement is typically obligatory and as the sources and available data do not suggest otherwise, the class marking is obligatory.

ad X₅: The class marking occurs in the form of affixes (cf. table 102).

ad X₆: Moro allows multiple class marking in the same context (cf. 165).

ad X₇: The consulted sources do not state whether the system applies to the full nominal lexicon or not. The consulted data does not suggest that there are parts of the nominal lexicon remain unclassified.

ad Y₁: The class semantics which are indicated in table 102 above are neither compulsory nor exhaustive, they just sum up a trend of semantic groupings within one class (Gibbard et al. 2009: 110). For example, different denotations for human beings are spread over different classes (e.g. ɲerá “girl” in the ɲ/ŋ class, ɓęgę “mother” is in the ɬ/ɬ class, ɓuŋa “poor person” is in the g/l class; cf. Gibbard et al. 2009: 107 and 110), and the human g/l class contains also the noun ɭi “rat” (Gibbard et al. 2009: 110).

ad Y₂: A noun’s concordial class is either semantically determined or by its initial phoneme (cf. table 102). If a noun’s initial consonant differs from the abovementioned initial segments, it is assigned to the concordial class the agreement markers of which are the most similar to it (e.g. sůwéja “(a particular) dance” is assigned to the j class; cf. Gibbard et al. 2009: 110). The ɗ class seems to represent the ‘ultimate’ default class, which hosts all nouns that cannot be assigned by other factors (e.g. maŋwáta “peanut” or kúŋa “ball”). Especially initial labial nouns (which do not have a corresponding class prefix) are assigned to this class and display another peculiarity: They do not form plurals, even if the noun’s semantics allowed it (Gibbard et al. 2009: 110).

ad Y₃: The combination of semantic, phonological, and default assignment rule types result in a higher number of assignment rules.

ad Y₄: The Moro class markers are not fused the category number (cf. 165).

ad Y₅: Different singular and plural classes may make use of the same marker (cf. table 102).

ad Y₆: The classes at their contemporary state are largely not semantically motivated (Gibbard et al. 2009: 110). Related discourse referents can be linked by assignment to the same class (cf. 166); this phenomenon falls into the domain of discourse manipulation rather than representing a case of semantic contribution to the noun phrase.
Appendix

(166) A δ-ωnąd r δ-ωγαw
   ÑII(I)-boat ÑII(I)-water
   “the boat on the water”

B y-ωγαjnu y-σmaje
   XI-parable XI-man
   “a parable about a man” (Moro; cf. Guest 1998: 5)

ad Yγ: Class assignment occurs on semantic and phonological grounds (cf. Gibbard 2009: 110) and does not leave room for class variation (cf. table 102).

22.4 Grammaticalisation

The Moro system displays a fair degree of formalisation (value: 0.964) and an overall lack of transparency (value: 0.404), for instance in the prevailing phonological assignment and the marking by monoconsonantal markers, the recognisable, but unproductive semantic underpinning, the wide range of nominal satellites which are marked for a noun’s class, the marking beyond the borders of the noun phrase.

23 Mundurukú

23.1 Background

< Mundurukú < Tupian (ISO 639-3: myu; cf. Lewis et al. 2013: s.v.)

Mundurukú (also known as Caras-Pretas, Monjoroku, Mundurucu, Paíquize, Pari, or Weijenye) is spoken in 22 villages in Pará, Amazonas, at the middle and upper Tapajós river and the middle Madeira rivers in Central Brazil (cf. Lewis et al. 2013: s.v. and Rodrigues 1999: 109). This dossier is mainly based on Crofts 1973, Gomes 2006 and Gonçalves 1987. Though many references to secondary literature have been made, this literature is based on these works. References from other sources than the mentioned ones have been verified by looking them up in these three works wherever possible. The Mundurukú classifiers occur on the verb, the noun, the numeral, the demonstrative, and the adjective. As class marking in Mundurukú is not analysed as a form of agreement in this study, this renders the system a multiple classifier system.
23.2 Classification Typology

23.2.1 Noun Classifiers

Mundurukú employs a large nominal classification system, which attaches its markers to a range of nominal modifiers (cf. 167). The marking is not analysed as a form of agreement here, which renders the system a multiple classifier system. This system is analysed in section 23.3.

(167) A xepxep-‘a wezik-á two-CL:ROUND potato-CL:ROUND “two potatoes”

B wa-pi-men-na FORMATIVE-CL:STICK-red-ADJ “red (stick, pencil, etc.)”

C ija-bá ako-bá this-CL:LONG banana-CL:LONG “this banana” (MundurukúG’I; cf. Derbyshire and Payne 1990: 260f.)

D Ti dojat paye, o-‘íz-mag ip baseya’a be. water bring when they-CL:WATER-place they basin in “When they brought water, they placed it in a basin.” (Mundurukú; Mithun 1986: 381)

23.2.2 Numeral Classifiers

Mundurukú employs a large nominal classification system, which attaches its markers to a range of nominal modifiers (cf. 167). The marking is not analysed as a form of agreement here, which renders the system a multiple classifier system. This system is analysed in section 23.3.

23.2.3 Possessive Classifiers

Mundurukú employs a large nominal classification system, which attaches its markers to a range of nominal modifiers (cf. 167). The marking is not analysed as a form of agreement here, which renders the system a multiple classifier system. This system is analysed in section 23.3.
23.2.4 Multiple Classifier Systems

Mundurukú employs a large nominal classification system, which attaches its markers to a range of nominal modifiers (cf. 167). The marking is not analysed as a form of agreement here, which renders the system a multiple classifier system. This system is analysed in section 23.3.

23.2.5 Concordial Class Systems

Mundurukú employs a large nominal classification system, which attaches its markers to a range of nominal modifiers (cf. 167). The marking is not analysed as a form of agreement here, which renders the system a multiple classifier system. This system is analysed in section 23.3.

23.3 Analysis

23.3.1 Functional Setup

Table 104: The semantic and discourse properties of the Mundurukú system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification (✓)</td>
</tr>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability (✓)</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values (✗)</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora (✓)</td>
</tr>
</tbody>
</table>

**ad C₁:** The Mundurukú classificatory markers are employed as semantic specifiers (e.g. ako “banana” must combine with a classifier in order to denote a specific part of the concept “banana”: ako-x “banana fruit (CL:LONG)”, ako-dup “banana leaf (CL:LEAF)”, ako-dot “a bunch of bananas (CL:BUNCH)”; cf. Picanço 2003: 4, cited in Gomes 2006: 197). Though Picanço calls the second element of these compounds ‘nouns’ and the process ‘compounding’, these elements are part of Mundurukú’s classifier inventory (e.g. -ba CL:LONG; cf. Rodrigues 1999: 116). This phenomenon appears to represent a productive process (cf. also Gomes 2006: 104, where these constructions are more carefully labeled ‘nominal syntagnata’). Since the Mundurukú noun does not take e.g. plural affixes, which proved the coinage of a lexical item, it is not necessary to assume a compound status in these cases, but rather an instance of flexible categorisation, which serves the differentiation of referents (cf. 168).

(168) A kape-di coffee-CL:LIQUID

“coffee (drink)”
B kape-da coffee-CL:SEED
“coffee (seed)” (Mundurukú; cf. Derbyshire and Payne 1990: 261)

ad C₂: Examples like ako-ba “banana fruit (CL:LONG)”, ako-dup “banana leaf (CL:LEAF)”, ako-dot “a bunch of bananas (CL:BUNCH)” (cf. Picanço 2003: 4, cited in Gomes 2006: 197) strongly suggest that the classificatory markers in Mundurukú serve the means of individuation of otherwise semantically neutral or the unitisation of otherwise unbound lexemes.

ad C₃: According to Gomes (2006: 216f.), Mundurukú does not employ speaker variation or free and flexible class assignment within its nominal categorisation system. A speaker can thus not freely attribute properties to nouns by means of class assignment.

ad C₄: As illustration, the classificatory markers of the Mundurukú system of nominal classification can serve the purpose of reference tracking.

(169) Ti dojot puye. o’-ti-mog ip baseya’a be.
water bring when they-CL:WATER-place they basin in
“When they brought water, they placed it in a basin.” (Mundurukú; Mithun 1986: 381)

23.3.2 Formal Setup

Table 105: The formal properties of the nominal classification system of Mundurukú.

<table>
<thead>
<tr>
<th></th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>
Appendix

ad F₃: [170] illustrates that not all nouns have to be categorised under all circumstances (cf. also Gomes 2006: 209f.), except for lexical areas that are obligatorily categorised (e.g. plant terms; cf. Gomes 2006: 197).

(170) bekkitik ako-ba o'-su-ba-dobuzik
child banana-CL:LONG 3-REF-CL:LONG-find

“The child found the banana.” (Mundurukú; cf. Derbyshire and Payne 1990: 261)

ad F₄: The Mundurukú noun may be categorised by different classifiers (cf. 168).

ad F₅: Mundurukú employs a large inventory of about 120 classifiers (cf. Aikhenvald 2000: 355) or (according to Aikhenvald 2000: 160) about 100 forms. Derbyshire and Payne (1990: 261; also based on Gonçalves 1987) count 125 classifiers. The bulk of them (“at least 96”; cf. Gonçalves 1987: 24-29 and Aikhenvald 2000: 355; 98 according to Derbyshire and Payne 1990: 261) is derived from body part terms, 13 forms refer to plants, six forms refer to natural elements, and nine forms refer to “culturally significant items” (Derbyshire and Payne 1990: ib.).

ad F₆: Mundurukú seemingly employs class agreement by allowing more than one class marker to occur on adjacent elements (cf. 167 A and C). In this study, Mundurukú is analysed as a system that allows multiple classification (“two [round ones], potato [round ones]”) and not multiple marking (“two CL:ROUND potatoes CL:ROUND”), i.e. each element carrying a class marker is analysed as a classificatory context of its own. This analysis is supported by examples such as 167 B and 168 which clearly show a referential potential of the class markers, and not an instance of class agreement.

ad F₇: Mundurukú allows its class markers to attach to the verb (cf. 169). The system’s locus operandi is thus not restricted to the borders of the noun phrase.

ad F₈: The class markers may semantically contribute to their host noun phrase (cf. 168).

ad F₉: The classifier adds a semantic component to the noun phrase. Different classifiers may occur on the same noun stem and modify the semantic content of this noun. Due to the large inventory with its straightforward semantics, the speaker is the most likely assigning agent.
23.3.3 Mapping

![Graph showing mapping values]

Figure 49: Mundurukú.

Table 106: The formalisation and transparency value of Mundurukú.

<table>
<thead>
<tr>
<th>x-criteria value</th>
<th>y-criteria value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size X₁ 0.125</td>
<td>semantic structure Y₁ 0.75</td>
</tr>
<tr>
<td>host number X₂ 1</td>
<td>rule types Y₂ 1</td>
</tr>
<tr>
<td>locus operandi X₃ 1</td>
<td>categorical independence Y₃ 1</td>
</tr>
<tr>
<td>boundedness X₄ 0.666</td>
<td>discreteness of markers Y₄ 1</td>
</tr>
<tr>
<td>marking X₅ 0</td>
<td>redundancy Y₅ 0.75</td>
</tr>
<tr>
<td>exhaustivity X₆ 0</td>
<td>flexibility Y₆ 0.666</td>
</tr>
</tbody>
</table>

| x-value 0.542 | y-value 0.738 |

ad X₁: The inventory of Mundurukú comprises more than 120 classes (cf. e.g. Derbyshire and Payne 1990: 261), which makes the system a large one.

ad X₂: The Mundurukú classifiers are incorporated in the verb (cf. 169 and 170) and appear as affixes on the numeral, the noun (nominalised verbs (com verbo nominalizado), inalienable nouns (com nomes também inalienáveis), and ‘descriptive’ nouns (com nomes descritivos); cf. Gomes 2006: 185ff.), the adjective, and the demonstrative (cf. 167).

ad X₃: The Mundurukú classifiers can occur on the verb (cf. 169 and 170). Their scope is thus not reduced to the noun phrase.
ad X₄: Except for certain lexical areas such as plant terms (Gomes 2006: 197), nominal classification in Mundurukú is not obligatory in all contexts (cf. 170 where not all nouns are categorised).

ad X₅: The class markers occur as affixes or infixes (cf. 167).

ad X₆: As argued above, this study analyses Mundurukú as a language that does not employ multiple marking. Instances of seeming class agreement are considered adjacent noun phrases here.

ad X₇: The consulted sources do not comment on the exhaustiveness of the Mundurukú system. The class inventory provided in Gonçalves (1987: 24-29) comprises a list of semantically highly specific classifiers, which mostly refer to the spatial domain of the body. Due to the lack of more general classifiers and only a small set of classifiers encoding natural elements and objects (e.g. di² “water, liquid”, i³pi³ “earth, soil”), vegetation (e.g. dup² “leaf”, dit² “blossom”), and cultural elements and objects (e.g. ka⁴ “hut”, e³ “way”), nouns such as abstracts or semantic fields such as weather phenomena cannot be categorised.

ad Y₁: The bulk of classes is bijective and semantically highly specified (e.g. da³bi² “vagina”, bi³di² “saliva”; Gonçalves 1987: 25). Only a few classes display a semantic extension (e.g. (’)a⁴ “head, front, round thing”, bu² “hand, finger, long object, round and flexible”; Gonçalves 1987: 24f.). This makes the Mundurukú inventory an intermediate one between mixed classes and bijective ones.

ad Y₂: The consulted sources report on semantic assignment rules only.

ad Y₃: Due to the large class inventory, the number of assignment rules is high.

ad Y₄: The class markers convey only classificatory semantics. The markers are therefore categorially independent.

ad Y₅: The bulk of classificatory markers marks one class and one class only. A small number of markers is reported to denote more than one class (e.g. a³ a³ “voice, language” and “upper part of the trunk”; Gonçalves 1987: 24); these cases can however be argued to form a semantically extended category.

ad Y₆: The class markers can convey informative semantics (cf. 168) or mirror redundant semantic information, which is already present in the categorised noun (cf. 169).
ad Y\textsubscript{17}: The same noun can take different class markers (cf. [168]), which renders the classification a flexible one.

23.4 Grammaticalisation

Derbyshire and Payne (1990: 261) argue in the case of Mundurukú for an originally verbal classifier system that only later exceeded its range of morphosyntactic hosts. As Aikhenvald (2000: 227) points out, this claim cannot be made on historical or typological evidence. The system displays a medium degree of formalisation (value: 0.542) with a larger number of NP-internal and -external classificatory hosts, which are marked by bound markers. The system is nevertheless fairly transparent (value: 0.738), as its classes are semantically highly specified and flexibly marked by categorically independent and discrete markers, which in many cases semantically contribute to their host noun phrase. Only the high number of rules reduces the systemic transparency notably. The class inventory presented in Gonçalves (1987: 24-29) however displays at least the possibility that complex markers are of a compositional character, which would enhance the system’s transparency. At this point, the compositionality of the Mundurukú class inventory is a conjecture and has to be confirmed by a more in-depth study, which includes a larger database and native speaker consultation.

24 Nasioi

24.1 Background

24.2 Classification Typology

24.2.1 Noun Classifiers

Nasioi attaches its class markers to a whole variety of morphosyntactic hosts, including the noun (cf. 171A). Therefore, Nasioi is not analysed as a noun classifier system here, but as a multiple classifier system.

(171) A nto-na-u’\textsubscript{CAP} \quad bee-ru’-pi\textsubscript{CAP}
water-DERIV-CL:UNIT.LIQUID \quad three-CL:UNIT.LIQUID-PL
a-ru’-daang\textsubscript{CAP}
this-CL:UNIT.LIQUID
“These three inland lakes.” (Nasioi; cf. Aikhenvald 2000: 219)

B tamp-a-u’\textsubscript{CAP}
good:DERIV-CL:TARO
“a good taro” (Nasioi; cf. Hurd 1977: 132)

C n-ce-ka-na-va\textsubscript{CAP}
us-DL-INTENS-DERIV-CL:HOUSE
“(both) our house” (Nasioi; cf. Hurd 1977: 155)

D A-u\textsubscript{DEM} \quad marasini-n-u \quad bore’\textsubscript{DEM}
DEM-CL:TABLET \quad medicine-DERIV-CL:TABLET \quad head
siipa-koo-n-u\textsubscript{DEM}
sick-part-DERIV-CL:TABLET
“This pill is for headaches.” (Nasioi; cf. Hurd 1977: 140)

E Teni \quad toire’\textsubscript{ART.F} \quad tareuri-ma\textsubscript{ART.F-PR.S}
ART.F children \quad care.for-PRS.HAB-DERIV-CL:F
bau’uri-ma\textsubscript{ART.F-PR.S}
feed-PRS.PROGR-DERIV-CL:F
“The lady who cares for the children (is) the one who is feeding them.” (Nasioi; cf. Hurd 1977: 144 and Aikhenvald 2000: 162)

24.2.2 Numeral Classifiers

Nasioi attaches its class markers to a whole variety of morphosyntactic hosts, including the numeral (cf. 171A). Therefore, Nasioi is not analysed as a numeral classifier system here, but as a multiple classifier system.
24.2.3 Possessive Classifiers
Nasioi attaches its class markers to a whole variety of morphosyntactic hosts, including the possessive (cf. [171] C). Therefore, Nasioi is not analysed as a numeral classifier system here, but as a multiple classifier system.

24.2.4 Multiple Classifier Systems
Nasioi attaches its class markers to a whole variety of morphosyntactic hosts (cf. [171] C). This system is analysed in section 24.3.

24.2.5 Concordial Class Systems
Nasioi employs a nominal classification system that marks class membership on a variety of morphosyntactic hosts (cf. [171]). Despite formally resembling a system that employs agreement marking, the Nasioi system is analysed as a multiple classifier system in this study, not a concordial class system.

24.3 Analysis
24.3.1 Functional Setup

Table 107: The semantic and discourse properties of the Nasioi system of nominal classification.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents</td>
<td>productive derivation, variable classification ✓</td>
</tr>
<tr>
<td>individuation</td>
<td>unitisation for countability ✓</td>
</tr>
<tr>
<td>attribution of properties</td>
<td>speaker-sided values ×</td>
</tr>
<tr>
<td>reference identification</td>
<td>reference tracking via AGR, anaphora ✓</td>
</tr>
</tbody>
</table>

**ad C1:** There is no information available on class variation in Nasioi (cf. Serzisko 1982: 114). However, the Nasioi class markers are involved in noun formation (cf. [172]) and thus serve the differentiation of referents.

(172)  *manii'-koo-∅-nai’*

> women-part-DERIV-CL: COMPETING SIDES

“dispute over women” (Nasioi; cf. Hurd 1977: 140)

**ad C2:** Due to the semantic resemblance of the system to individuating classifier systems and since e.g. *biam* “banana” can be specified by a range of specific classifiers like *-voonku* “clumps of banana plants”, *-vareng* or *-u* “stalk of bananas”, *-noo’* “single banana”, and *-ne’* or *-rone’* “banana leaf” (cf. Hurd 1977: 118 and 138), the system appears to serve individuation.
### 24.3.2 Formal Setup

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness ((F_1))</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility ((F_2))</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory ((F_3))</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking ((F_4))</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi ((F_5))</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity ((F_6))</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment ((F_7))</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

**ad C₃:** Though the Nasioi system employs a class for social categorisation, it seems not to offer a possibility for the speaker to attribute positive or negative properties to a referent (cf. Hurd 1977: 115f.).

**ad C₄:** The Nasioi system’s agreement markers occur e.g. on the pronoun, which allows for reference identification by classificatory means (cf. [171] C).

**ad F₁:** Nouns are not always categorised in Nasioi (cf. [173] where *toire’ “children” remains uncategorised).

(173) *nikanamonoto* *toire’* *mamau’-po’-na-nona*  
our.village children many-POSS-DERIV-CL:VILLAGE  
“Our village has many children.” (Nasioi; cf. Aikhenvald 2000: 140)

**ad F₂:** There is no information on classificational flexibility in Nasioi (Serzisko 1982: 114), except for a general understandability of specific classifier replacement by a general classifier (Hurd and Hurd 1966: 20).

**ad F₃:** Nasioi employs between 40 and 50 referential classes “according to the size, shape or nature of the object being counted” (Capell 1969: 115, cited in
Appendix

Serzisko 1982: 113f.; cf. also Hurd and Hurd 1966: 20; “several dozen agreement classes” according to Aikhenvald 2000: 79) that include more than 100 ‘classifiers’ (Aikhenvald 2000: 219, Foley 1986: 83, Terrill 2002: 293). Hurd (1977: 162ff.) lists 118 of these elements, which he calls case-gender-number markers (CGNs; Hurd 1977: 114 and Terrill 2002: 292). They are grouped in 22 semantic categories (a default one and 21 specific ones), a selection of which are listed in figure 109, thus we find “more concord classes than genders”. Some of these ‘classifiers’ include measures such as -vu’ “piece”, -marang “part”, or -mare’ “year” (cf. Hurd 1977: 122 and Serzisko 1982: 114).


<table>
<thead>
<tr>
<th>category</th>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>social</td>
<td>-varang</td>
<td>“generation, peer group”</td>
</tr>
<tr>
<td></td>
<td>-vuntu’</td>
<td>“village, community, kingdom”</td>
</tr>
<tr>
<td></td>
<td>-ni</td>
<td>“feminine, woman-one”</td>
</tr>
<tr>
<td>body part</td>
<td>-rata</td>
<td>“eye”</td>
</tr>
<tr>
<td></td>
<td>-romi’</td>
<td>“ear”</td>
</tr>
<tr>
<td></td>
<td>-ri’</td>
<td>“bone, tooth, tool, direction”</td>
</tr>
<tr>
<td>animals &amp; food</td>
<td>-rempo’</td>
<td>“litter of pigs, all offspring of one cow”</td>
</tr>
<tr>
<td></td>
<td>-u</td>
<td>“animal, coin, pineapple, sweet potato, sore”</td>
</tr>
<tr>
<td></td>
<td>-maku</td>
<td>“pineapple, possum”</td>
</tr>
<tr>
<td>trees &amp; wood</td>
<td>-vare’</td>
<td>“tree”</td>
</tr>
<tr>
<td></td>
<td>-ve</td>
<td>“tree branch, horizontal timber, beam”</td>
</tr>
<tr>
<td></td>
<td>-ma’</td>
<td>“dead stick of wood, fire(wood), smoke column”</td>
</tr>
<tr>
<td>bananas</td>
<td>-vareng/-u</td>
<td>“stalk of bananas”</td>
</tr>
<tr>
<td></td>
<td>-noo’</td>
<td>“single banana; arrow, spear, sago shingle”</td>
</tr>
<tr>
<td></td>
<td>-(ro)ne’</td>
<td>“banana leaf”</td>
</tr>
</tbody>
</table>

The number of 40-50 referential classes is received when one subtracts the “about fifty classes of numbers or count nouns” (Hurd and Hurd 1966: 20). The number of 118 classificatory markers represents the class inventory in this dossier.

ad F4: Examples such as [171]A show a class marking that formally resembles class agreement. This study analyses the Nasioi class marking not as an instance of multiple marking by means of agreement, but as adjacent and independent contexts of classification (cf. water-units, three-units, these-units).

ad F5: Class marking occurs among others on the verb (cf. [171] E) and thus NP-externally.

ad F6: The Nasioi classificatory elements may “stand for nouns, modify nouns, and be modified by nouns” (Hurd 1977: 111, cited in Terrill 2002: 293) and thus contribute semantically to their nominal referent (cf. e.g. [172]).
Appendix

ad F7: There is no information on an ad hoc categorisation of the Nasioi classificatory elements available (Serzisko 1982: 114); the systemic semantics however suggest that the speaker is able to carry out class assignment on semantic grounds.

24.3.3 Mapping

![Figure 51: Nasioi.](image)

Table 110: The formalisation and transparency value of Nasioi.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size X1</td>
<td>0.125</td>
<td>semantic structure Y1</td>
<td>0.75</td>
</tr>
<tr>
<td>host number X2</td>
<td>1</td>
<td>rule types Y2</td>
<td>1</td>
</tr>
<tr>
<td>locus operandi X3</td>
<td>1</td>
<td>rule number Y3</td>
<td>0</td>
</tr>
<tr>
<td>obligatory X4</td>
<td>1</td>
<td>categorial independence Y4</td>
<td>1</td>
</tr>
<tr>
<td>boundedness X5</td>
<td>1</td>
<td>discreteness of markers Y5</td>
<td>0.5</td>
</tr>
<tr>
<td>marking X6</td>
<td>0</td>
<td>redundancy Y6</td>
<td>1</td>
</tr>
<tr>
<td>exhaustivity X7</td>
<td>1</td>
<td>flexibility Y7</td>
<td>0.666</td>
</tr>
</tbody>
</table>

x-value 0.732  
y-value 0.702

ad X1: With an inventory of 118 classes (cf. Hurd 1977: 162ff.), the Nasioi system is a large one.

ad X2: The classifiers occur on the adjective, article, pronoun, and numeral (Serzisko 1982: 113), furthermore demonstratives and the verb, in possessive

ad X₃: Class marking occurs among others on the verb and thus NP-externally (cf. 171 E).

ad X₄: Nouns in discourse may remain uncategorised (cf. 173, where toire’ “children” remains uncategorised).

ad X₅: The markers occur in bound form (cf. 171).

ad X₆: This study analyses multiple occurrences of class markers as cases of multiple, adjacent contexts of classification (cf. 171 A ⊃ water-units, three-units, these-units)

ad X₇: Due to the presence of a general class (Hurd and Hurd 1966: 20), the full nominal lexicon can be categorised.

ad Y₁: The bulk of Nasioi classes is semantically specific; a part of the class inventory is of a more generic nature (cf. table 109).

ad Y₂: The consulted sources and available data suggest no other assignment than a semantic one.

ad Y₃: Due to the high number of classes, the number of assignment rules is also high.

ad Y₄: The class markers convey classificatory semantics only (cf. 171 and table 109).

ad Y₅: Not all markers denote a single class (cf. e.g. -u “stalk of bananas; Chinese taro, sweet potato, pineapple; coins; sores, animals”; Hurd and Hurd 1977: 118).

ad Y₆: The markers typically add informative semantic content to the noun phrase (cf. 171 and 172).

ad Y₇: Despite a general availability of variation between a more general and a more specific class marker, there is no information available on the flexibility of classification in Nasioi (Serzisko 1982: 114).
24.4 Grammaticalisation

There is no direct information available on the diachrony of the Nasioi system of nominal classification, but a seemingly typologically parallel system in Aiwoo and in Buin (Terrill 2002: 292f.) suggests that it is deeply rooted in Papua. The Nasioi system as such displays incipient properties of a concordial class system by a seemingly inflexible class assignment and a marking that formally resembles agreement. However, the huge class inventory and highly productive semantics display its classifier-like nature. A system at this stage with such distinct characteristics of both classifier and concordial class systems is possibly representing a case of a classifier system grammaticalising into a concordial class system.

25 Ngan’gityemerri

25.1 Background

Ngan’gityemerri (or Moil) is a Murrinh-Patha language of the Daly phylum in Australia. Ruhlen (1987: 362) does not list Ngan’gityemerri, but assumes a Moil branch as a subfamily of Daly languages. Moil is the alternate name for Ngan’gityemerri (cf. Lewis et al. 2013: s.v.), and the specific dialect investigated by Reid (Ngan’giwumirri) is “highly consistent with noun class features of its sister dialect Ngan’gikurunggurr, which is listed as a Moil language in Ruhlen (1987). The main source for this dossier is Reid (1997). Ngan’gityemerri employs a multiple classifier system.

![Diagram of nominal classification devices](image-url)

Figure 52: Classification of the Ngan’gityemerri system of nominal classification.
25.2 Classification Typology

25.2.1 Noun Classifiers

Ngan’gityemerri employs a classification system that attaches its markers among others to the noun (cf. 174).

(174) A \( \text{wa}= \text{yedi} \ \text{wa}= \text{wunu} \ \text{wa}= \text{lenggir} \)
M man M that M bad
“that bad man”

B \( \text{wayedi} \ \text{wunu} \ \text{walingir} \)
M man M that M bad
“That man is bad.” (Ngan’gityemerri; cf. Reid 1997: 169)

Since the markers also attach to a variety of other morphosyntactic hosts (cf. 175) and 176, the system of Ngan’gityemerri is treated as a multiple classifier system in this dossier.

(175) A \( \text{wur}= \text{yuku} \)
F good
“a good woman”

B \( \text{wur}= \text{kmyi} \)
F this
“this woman”

C \( \text{wur}= \text{ngayi} \)
F mine
“my woman”

D \( \text{wur}= \text{wungume} \)
F one
“a certain woman”

E \( \text{wur}= \text{ngin} \)
F anaphor
“that (previously mentioned) woman” (Ngan’gityemerri; cf. Reid 1997: 169)

(176) A \( \text{tyen} \ -\text{ga} \)
what ANIMAL
“What kind of animal is this?” (Ngan’gityemerri; cf. Reid 1997: 170)

B \( \text{gagu} \ \text{minbe} \ -\text{ga} \ -\text{wurru} \ \text{yenim-} \ \text{way} \ -\text{me} \)
ANIMAL NEW ANIMAL UNSATIS 3SGS:AUX empty hand
“He has no beef, he is emptyhanded.” (Ngan’gityemerri; cf. Reid 1997: 171)

C \( g= \text{ngagan} \)
ANIMAL whatchamacallit
“that whatchamacallit animal” (Ngan’gityemerri; cf. Reid 1997: 172)

25.2.2 Numeral Classifiers

Ngan’gityemerri does not employ a numeral classifier system (cf. 177, where the quantificational phrase does not feature a numeral classifier).

(177) e- \( \text{menggin} \ \text{mipurr} \ \text{da-} \ \text{nguru} \ \text{wagarri} \ \text{dem-} \ \text{baty} \)
ANIMAL goanna man BODY penis two 3SGS:AUX hold
“A goanna, a male one, has two penises.” (Ngan’gityemerri; cf. Reid 1997: 180)
25.2.3 Possessive Classifiers

The classifiers occur among others in possessive contexts (cf. 175 C). As the same set of classifiers combines with a range of other nominal satellites, this system is analysed as a multiple classifier system, not a possessive classifier system.

25.2.4 Multiple Classifier Systems

Ngan’gityemerri employs a multiple classifier system (cf. 175 and 176). This system is analysed in section 25.3.

25.2.5 Concordial Class Systems

Ngan’gityemerri does not employ a concordial class system. Though it is argued that Ngan’gityemerri employs class agreement, the class marking on nominal modifiers is generally optional (Reid 1997: 168). Instances of multiple marking in the same context are thus analysed as adjacent noun phrases, which represent a classificational context of their own.

25.3 Analysis

25.3.1 Functional Setup

Table 111: The semantic and discourse properties of the Ngan’gityemerri system of nominal classification.

| differentiating referents (C₁) | productive derivation, variable classification ✓ |
| indviduation (C₂) | unitisation for countability ✗ |
| attribution of properties (C₃) | speaker-sided values ✓ |
| reference identification (C₄) | reference tracking via AGR, anaphora ✓ |

ad C₁: The markers are used to coin descriptive labels and nominalise even phrases, which is a productive process in Ngan’gityemerri (cf. 178 and Reid 1997: 222ff.).

(178)  
A wa= wedimuy ngagadi-tye wa=  
M little.boy 1SGS:go-PST M  
[deme-ngi-pi-yiri-tye-yedi]  
3SGS:AUX-1SGDO-head-numb-PST-3SGS:go  
“As a young boy, I was (a male who was) habitually shy.”

B wur= [pagu wudumbun-ngirr-bal]  
F meat 3SGS:AUX-1PLEX-cook  
“The woman who cooked meat for us.” (Ngan’gityemerri; cf. Reid 1997: 206)
“There are two white women (security guards) both standing there (with) a thing (metal detector) lest you have something like a knife.” (Ngan’gityemerri; cf. Reid 1997: 205)

In addition, the markers can be assigned variably in order to distinguish referents (cf. again 179).

(179) A syiri yawurr
   STRIKERS stick
   “stick for fighting”
B kini yawurr
   DIGGING stick
   “stick for digging”
C yenggi yawurr
   FIRE stick
   “stick for burning” (Ngan’gityemerri; cf. Reid 1997: 178)

ad C2: The consulted sources and available data do not suggest the Ngan’gityemerri system to be involved in individuation.

ad C3: The attribution of properties by classificatory means is possible: For instance, by assigning a human noun to the animal category, the speaker of Ngan’gityemerri imposes that this being has spiritual properties (cf. 180).

(180) gagu pope dini-gumu-tye Alice Springs[. . .]
   ANIMAL pope 3SGS:sit-while-PST Alice.Springs
   “While the pope was in Alice Springs[. . .]” (Ngan’gityemerri; cf. Reid 1997: 184)

ad C4: Ngan’gityemerri allows its nominal classification system to take part in reference tracking; “there is preference, once a specific noun is contextually established, to omit the specific noun, and subsequently refer to it by its generic” (Reid 1997: 167; cf. 181).

(181) yawurr men’gin-ya
   TREE/THING 3SGS:AUX-pick.up
   “He picked up the [didgeridoo].” (Ngan’gityemerri; cf. Reid 1997: 167)
25.3.2 Formal Setup

Table 112: The formal properties of the nominal classification system of Ngan’gityemerri.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like:</th>
<th>CCS-like:</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>flexible class membership</td>
<td></td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad F₁: Some nouns cannot be marked for their referential class. These constitute an own concordial class in Reid’s classification (Reid 1997: 172).

ad F₂: The Ngan’gityemerri nouns frequently display various gender markers in combination with one and the same root; Reid (1997: 198) calls this phenomenon “crossclassing” and attributes it a “fairly widespread” status.

ad F₃: Ngan’gityemerri distinguishes fifteen marked classes, the markers of which can occur in free or bound form (cf. table [113]), and an unmarked class (“mostly terms for natural objects such as sun, rocks, ground, stars, clouds etc.”; cf. Reid 1997: 172).
Appendix

Table 113: Concordial class marking in Ngan’gityemerri (cf. Reid 1997: 173).

<table>
<thead>
<tr>
<th>class</th>
<th>head marking</th>
<th>AGR marking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>free</td>
<td>dependent</td>
</tr>
<tr>
<td>1 male</td>
<td>(mipurr)</td>
<td>wu= ~ Ø-</td>
</tr>
<tr>
<td>2 female</td>
<td>(fahmi)</td>
<td>wur= ~ Ø-</td>
</tr>
<tr>
<td>3 group</td>
<td>awa=</td>
<td>empha=</td>
</tr>
<tr>
<td>4 animal</td>
<td>gagu</td>
<td>a- ~ Ø-</td>
</tr>
<tr>
<td>5 vegetable</td>
<td>migi</td>
<td>mi- ~ Ø-</td>
</tr>
<tr>
<td>6 body parts</td>
<td>da-</td>
<td>ū- ~ a-</td>
</tr>
<tr>
<td>7 canines</td>
<td>wu-</td>
<td>emphyu=</td>
</tr>
<tr>
<td>8 trees/things</td>
<td>yawurr</td>
<td>yerr= ~ Ø-</td>
</tr>
<tr>
<td>9 bamboo spears</td>
<td>yawul</td>
<td>Ø-</td>
</tr>
<tr>
<td>10 fire</td>
<td>yenggi</td>
<td>yenggi ~ Ø</td>
</tr>
<tr>
<td>11 strikers</td>
<td>syiri</td>
<td>syiri ~ Ø</td>
</tr>
<tr>
<td>12 canegras spears</td>
<td>kurum</td>
<td>kurum ~ Ø</td>
</tr>
<tr>
<td>13 drinks</td>
<td>kuru</td>
<td>kuru ~ Ø</td>
</tr>
<tr>
<td>14 woomeras</td>
<td>tyin</td>
<td>tyin ~ Ø</td>
</tr>
<tr>
<td>15 digging sticks</td>
<td>kini</td>
<td>kini ~ Ø</td>
</tr>
</tbody>
</table>

ad F₄: Next to clitics and affixes, Ngan’gityemerri also employs optional free form class markers (cf. [182] and table [113], which is a highly unusual feature for an agreement system (and indicates the different ages of the markers in the system; cf. Reid 1997: 211 and 219).

(182) A (syiri) magulfu (syiri) marrgu
      strike cylindrical.fighting.stick strike new
dem-wursty-dim
      AUX:3SG.SUBJ-make-sit:2SG.SUBJ

“He is making a new cylindrical fighting stick.” (Ngan’gityemerri; cf. Reid 1997: 177)

B (yawurr) ganbi kide
      TREE didgeridoo where

“Where is the didgeridoo?” (Ngan’gityemerri; cf. Reid 1997: 167)

In general, “gender marking tends to be obligatory on head nouns”, while “agreement marking [...] is not strictly obligatory” (Reid 1997: 168). Since optional agreement marking occurs nowhere else in the sample, the seeming occurrences of multiple marking (e.g. [174] A) are interpreted as instances of multiple classification in this study (cf. the [male] man, that [male], the [male] bad one).

ad F₅: Among others, class marking may occur on the verb (cf. [183]) and thus NP-externally.
Appendix

(183) ganbi yerr= buy mem -garri kide
didgeridoo CL:TREE light 3SGS.do leg where

“Where is that white digeridoo?” (Ngan’gityemerri; cf. Reid 1997: 194)

ad F6: The class markers may semantically contribute to their host noun phrase (cf. 179).

ad F7: The freedom of class assignment appears to be but rather limited; the ‘cроссclassed’ nouns seem to be lexicalised in many cases (which is why different speakers came up with different folk etymologies for certain terms; cf. Reid 1997: 200). The most ‘liberal’ and speaker-sided assignment possibilities appear to be situated in the realm of metaphor (cf. [184]), while the other class variations are lexicalised and/or related to other circumstances such as the discourse-active linguistic register (cf. Reid 1997: 200f.).

(184) ngatya ngikin fi-ngumu-pe ya mi = dityunggurr-wurr
father shit puff-do:1SG.SUBJ-FUT hey VEG short-UNSATIS

“Hey daddy, let me have a puff of that shit... Oh, it’s too short now.”
(Ngan’gityemerri; cf. Reid 1997: 201)

25.3.3 Mapping

![Figure 53: Ngan’gityemerri.](image_url)
Table 114: The formalisation and transparency value of Ngan’gityemerri.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>semantic structure</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>rule types</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>rule number</td>
<td>$Y_3$</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>categorial independence</td>
<td>$Y_4$</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>discreteness of markers</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>redundancy</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>flexibility</td>
<td>$Y_7$</td>
</tr>
</tbody>
</table>

| x-value | 0.512 |
| y-value | 0.905 |

**ad X₁:** With fifteen marked classes and an unmarked one (cf. table 113), the system is a small one.

**ad X₂:** The classificatory markers combine among others with adjectives, pronouns, anaphoric demonstratives, deictics, or numerals, where they can also serve as nominalisers (cf. 175).

**ad X₃:** The class markers may occur on the verb (cf. 183) and thus NP-externally.

**ad X₄:** Not all nouns are categorised in all instances (cf. 182, where the classification occurs optionally).

**ad X₅:** The markers occur in both bound and free form (cf. table 113).

**ad X₆:** Ngan’gityemerri does not employ multiple marking by means of agreement (cf. the discussion under F₁).

**ad X₇:** The unmarked sixteenth class (cf. Reid 1997: 172) cannot partake in the classification system.

**ad Y₁:** The class semantics are straightforward (cf. table 113).

**ad Y₂:** The system is semantically highly productive (cf. e.g. 178 and 179) and a small number of semantic assignment rules is sufficient to master the system (cf. table 113). Especially a number of ‘crossclasses’ nouns seem to be lexicalised (cf. Reid 1997: 200), which means that opaque assignment is also present in Ngan’gityemerri.

**ad Y₃:** The consulted sources and available data suggest that the number of assignment rules in Ngan’gityemerri is small.
ad $Y_4$: The class markers convey classificatory semantics only (cf. table 113).

ad $Y_5$: Each class has its own, distinct markers (cf. table 113).

ad $Y_6$: The class markers typically add informative semantics to the noun phrase (cf. e.g. 178 and 179).

ad $Y_7$: The same noun can take different class markers (cf. e.g. 179).

25.4 Grammaticalisation

The system of Ngan’gityemerri displays an intermediate degree of formalisation (value: 0.512) by employing a mixed set of bound and unbound markers, which tend to occur obligatorily on the head noun and optionally on nominal modifiers. Since they are commonly used as nominalisers, this study argues that marked elements represent a noun phrase of their own. At the same time, the system is highly productive with respect to its semantics. The small class inventory with its bijective classes render it a highly transparent system (value: 0.905).

26 North Ambrym

26.1 Background


North Ambrym is a language of Vanuatu spoken in the “Malampa Province, North Ambrym island” (Lewis et al 2013: s.v.). It employs a possessive classifier system. This dossier is mainly based on Franjieh (2012, 2014, 2015, in prep.).
26.2 Classification Typology

26.2.1 Noun Classifiers

North Ambrym does not employ a noun classifier system (cf. [185] where no classifier occurs).

(185) A Yafa kirine mwenan yanarr
CONJ REC.PST3DL give piglet IND go.TR3SG
“And they gave a piglet to him.” (North Ambrym; cf. Franjieh 2012: 95)
B Liseseu rro gili tomo
Liseseu CONT dig.TR rat
“Lisepsep was digging out the rat.” (North Ambrym; cf. Franjieh 2012: 102)

26.2.2 Numeral Classifiers

North Ambrym does not employ a numeral classifier system (cf. Franjieh 2012: 194 and [186] where no classifier occurs).

(186) A Rrun ne teere wa te ru
custom.story ASS child PART NONRECENT.PST two
B Ma geyene bu be lim
REC.PST3SG pay.tr castrated.pig COP five
“He paid five pigs for it.” (North Ambrym; cf. Franjieh 2012: 85)

26.2.3 Possessive Classifiers

North Ambrym employs a possessive classifier system (cf. [187]). This system is analysed in section 26.3.

(187) A Mwene-ng ol
CL:RESIDUE-1SG coconut
“my (coconut as) copra”
Appendix

26.2.4 Multiple Classifier Systems

The classifiers occur in possessive contexts only (cf. 185, 186, and 187). Therefore, North Ambrym does not employ a multiple classifier system.

26.2.5 Concordial Class Systems

North Ambrym does not employ a concordial class system (cf. Franjieh 2012: 173, footnote 2, and 190).

26.3 Analysis

26.3.1 Functional Setup

Table 115: The semantic and discourse properties of the North Ambrym system of nominal classification.

<table>
<thead>
<tr>
<th>differentiating referents (C₁)</th>
<th>productive derivation, variable classification ✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability ✓</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values ✓</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora ✓</td>
</tr>
</tbody>
</table>

ad C₁: The classifiers combine with the marker for person and number and thus do not take part in productive derivation. Certain nouns may combine with a variety of classifiers, which also contribute to their NP’s semantics in order to specify the referent (cf. 187).

ad C₂: The consulted sources and available data do not suggest the North Ambrym system to be involved in individuation.

ad C₃: The consulted sources and available data do not suggest the North Ambrym system to be involved in the speaker-sided attribution of properties by classifiericatory means.

ad C₄: The head noun of a classifier construction in North Ambrym may be omitted (if it is recoverable from the context; Franjieh 2012: 237); the remaining classifiers thus take part in reference identification.
26.3.2 Formal Setup

Table 116: The formal properties of the nominal classification system of North Ambrym.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

**ad F1:** The classificatory markers occur are only in possessive constructions obligatory (cf. \[188\]) where no class marker occurs.

\[188\] narr vya nane lon skul 1SG.NONRECENT.PST go yesterday in school
“I went yesterday to school.” (North Ambrym; cf. Franjieh 2012: 145)

**ad F2:** While classifier assignment is rigid at large, a few nouns can take different classifiers (cf. \[187\]).

**ad F3:** North Ambrym employs a set of five classifiers (cf. table \[117\]).

Table 117: The inventory of North Ambrym classifiers (cf. Franjieh 2014: 5).

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>animals, tools, units of time, some trees and kin terms</td>
</tr>
<tr>
<td>ma</td>
<td>liquids, containers of liquids, buildings, holes, mats</td>
</tr>
<tr>
<td>bo</td>
<td>fire, firewood</td>
</tr>
<tr>
<td>lo</td>
<td>baskets</td>
</tr>
<tr>
<td>mwena</td>
<td>residue</td>
</tr>
</tbody>
</table>
ad F₄: The classifier occur once, not multiply (cf. e.g. [187]).

ad F₅: The classifiers occur NP-internally (Franjieh 2012: 237).

ad F₆: The classifiers may contribute semantically to their host NP (cf. [187]).

ad F₇: Since the classifier assignment is rigid (Franjieh in prep.: 8ff.) and semantically very distinct, the speaker is bound to the system’s semantic assignment mechanisms and lexicalised class members.

26.3.3 Mapping

![Figure 55: North Ambrym.](image)

Table 118: The formalisation and transparency value of North Ambrym.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size X₁</td>
<td>0.875</td>
<td>semantic structure Y₁</td>
<td>0.5</td>
</tr>
<tr>
<td>host number X₂</td>
<td>0</td>
<td>rule types Y₂</td>
<td>0.666</td>
</tr>
<tr>
<td>locus operandi X₃</td>
<td>0</td>
<td>rule number Y₃</td>
<td>1</td>
</tr>
<tr>
<td>obligatoriness X₄</td>
<td>0.333</td>
<td>categorial independence Y₄</td>
<td>1</td>
</tr>
<tr>
<td>boundedness X₅</td>
<td>0.25</td>
<td>discreetness of markers Y₅</td>
<td>1</td>
</tr>
<tr>
<td>marking X₆</td>
<td>0</td>
<td>redundancy Y₆</td>
<td>0.5</td>
</tr>
<tr>
<td>exhaustivity X₇</td>
<td>1</td>
<td>flexibility Y₇</td>
<td>0.333</td>
</tr>
<tr>
<td>x-value</td>
<td>0.351</td>
<td>y-value</td>
<td>0.714</td>
</tr>
</tbody>
</table>
ad X₁: With a total of five classes (cf. table [117]), the North Ambrym system is a small one.

ad X₂: The classifiers occur in possessive constructions (cf. [187]), not elsewhere (cf. [185] [186] and [188]).

ad X₃: The class markers occur in possessive constructions and thus NP-internally (cf. Franjieh 2012: 237).

ad X₄: The classifiers occur obligatorily in indirect possessive constructions (cf. Franjieh 2012: 232).

ad X₅: The markers occur in bound form (cf. [187]) and as free form markers (cf. [189]).

(189) amkumku a CL:ANIM John

“John’s truck” (North Ambrym; cf. Franjieh 2012: 237)

ad X₆: The markers occur once, not multiply (cf. [187]).

ad X₇: Due to the presence of a residue marker (cf. table [117]), all nouns can be classified.

ad Y₁: There are both specific and general classes present (cf. table [117]), which makes the class structure a mixed one.

ad Y₂: Classification in North Ambrym is largely rigid (Franjieh in prep.: 8ff.) and thus lexicalised. Nevertheless, a number of nouns allows a flexible semantic assignment (cf. [187]), and the bulk of classified nouns is semantically transparent despite its rigid class membership.

ad Y₃: The semantic transparency of class membership results in a small number of assignment rules.

ad Y₄: The class markers convey classificatory semantics only (cf. table [117]).

ad Y₅: Each class has its own, distinct marker (cf. table [117]).

ad Y₆: Class markers typically are not informative (cf. [189]). Cases like [187] where a flexible class assignment and a productive semantic contribution occurs, are rather infrequent (Franjieh 2012: 385).
ad Y₇: The bulk of nouns cannot change its class (Franjieh in prep.: 8ff.).

### 26.4 Grammaticalisation

The North Ambrym relational classifier system corresponds to the formal definition and general understanding of a possessive classifier system (value: 0.351), though its assignment is not free and its inventory is comparatively small. The small classifier inventory and the rigidity and lack of manipulability of the classifier assignment indicate that the North Ambrym relational classifier system has a fairly low degree of transparency (value: 0.714).

### 27 Nùng

#### 27.1 Background

< Central Tai < Kam-Tai < Tai-Kadai (ISO 639-3: nut; cf. Lewis et al. 2013: s.v.)
< Central Tai < Kam-Tai < Lakkia-Kam-Tai < Be-Kam-Tai < Li-Kam-Tai < Daic < Austro-Tai < Austric (cf. Ruhlen 1987)

Nùng (or Nùng Fan Sling) is spoken in the Southeast Asian Cao Bang and Lang Son provinces, Ho Chi Minh City, Dong Nai, Lam Đồng, and Dak Lak, thus mainly in Vietnam. There are also communities in Canada, Laos, and the USA (cf. Lewis et al. 2013: s.v.). The sparse literature on the grammar of Nùng provides confusing information on its multiple classifier system. The main publications on Nùng classifiers are Conklin 1981, Freiberger 1976, Saul 1965, and a cooperation of the latter two, Saul and Wilson 1980. Conklin (1981: 110f) analysed the findings in Freiberger 1976 and Saul 1965 and came to the conclusion that both based their study on the same data, though the findings of both differ. Conklin (1981: 110) points to “serious discrepancies in transcriptions” between the works on Nùng. Since all the mentioned authors and works showed differences in the transcription of Nùng, the examples were gathered from Saul/Wilson 1980 wherever possible.

![Figure 56: Classification of the Nùng system of nominal classification.](image-url)
27.2 Classification Typology

27.2.1 Noun Classifiers

The Nùng classifiers occur in various contexts, and also adjacent to nominals (cf. 190). As the same classifiers also occur in other contexts, the Nùng system is analysed as a multiple classifier system within this dossier.

(190) mùhn láo tú phi lái
      he    fear CL:ANIMATE spirit much

“He is very afraid of evil [sic] spirits.” (Nùng; cf. Saul and Wilson 1980: 26)

27.2.2 Numeral Classifiers

The Nùng classifiers occur in various contexts, including quantificational ones (cf. 32.4). As the same classifiers also occur in other contexts, the Nùng system is analysed as a multiple classifier system within this dossier.

(191) lẹọ háhn tú mà lyhc
      then kill two CL:ANIMATE dog,child

“Then kill two puppies.” (Nùng; cf. Saul and Wilson 1980: 14)

27.2.3 Possessive Classifiers

The Nùng classifiers occur in various contexts, including possessive ones (cf. 192). As the same classifiers also occur in other contexts, the Nùng system is analysed as a multiple classifier system within this dossier.

(192) A tú mà cáu phao lái nhếhng hồn tú mà mưhn
      CL:ANIMATE horse my run much fast than CL:ANIMATE
      horse your

“My horse runs faster than your horse.” (Nùng; cf. Saul and Wilson 1980: 76)

B ñọn tẹ̀ng cáu chékp dúg kí pi lọ
      CL:INANIMATE stomach I pain able several year emphatic

“My stomach ached for several years.” (Nùng; cf. Saul and Wilson 1980: 26)

27.2.4 Multiple Classifier Systems

The Nùng classifiers occur in various contexts (cf. 190, 32.4, and 192). The system is thus analysed as a multiple classifier system in section 27.3

27.2.5 Concordial Class Systems

Nùng does not employ a concordial class system (cf. e.g. 192 where no class agreement marking occurs on the nominal satellites).
27.3 Analysis

27.3.1 Functional Setup

Table 119: The semantic and discourse properties of the Nùng system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Nùng System Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents</td>
<td>productive derivation, variable classification ✓</td>
</tr>
<tr>
<td>individuation</td>
<td>unitisation for countability ✓</td>
</tr>
<tr>
<td>attribution of properties</td>
<td>speaker-sided values ×</td>
</tr>
<tr>
<td>reference identification</td>
<td>reference tracking via AGR, anaphora ✓</td>
</tr>
</tbody>
</table>

**ad C₁:** The Nùng classifiers are independent elements and as such do not take part in noun formation processes. A certain variability at least between a more general and a more specific classifier seems possible (cf. [193]).

\[(193)\] A lẹọ pọy hahn cáh cá dähm  
then go see CL:GENERAL crow black  
“Then he went and saw a black crow.” (Nùng; cf. Saul and Wilson 1980: 26)

B lẹọ pọy hahn tũ cá dähm  
then go see CL:ANIMATE crow black  
“Then he went and saw a black crow.” (Nùng; compiled after Saul and Wilson 1980)

**ad C₂:** The Nùng classifiers occur obligatorily in quantificational phrases (Conklin 1981: 113 and Saul and Wilson 1980: 27) and specify the nominal semantics there (cf. [32.4]). Therefore, the system most likely partakes in individuation.

\[(194)\] tũ tể oc dähm høy ma  
CL:ANIMATE that out CL:INANIMATE egg come  
“She came out of the egg.” (Nùng; cf. Saul and Wilson 1980: 64)

**ad C₃:** The consulted sources and available data do not suggest the Nùng system to be involved in the speaker-sided attribution of properties by classificatory means.

**ad C₄:** The Nùng system allows for referent identification, since the classifiers can substitute the head noun (cf. [194] [195] and Saul and Wilson 1980: 25).

\[(195)\] óhng tũ sńńń pọy  
CL:HUMAN CL:ANIMATE elephant go  
27.3.2 Formal Setup

Table 120: The formal properties of the nominal classification system of Nùng.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness</td>
<td>nouns may remain uncategorised</td>
<td>obligitory assignment of each noun</td>
</tr>
<tr>
<td>flexibility</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad F1: Nouns may remain uncategorised in Nùng (cf. [196] where no classifier occurs).

(196) A ké cõ vahng té muhn ñhíng ma
man father boy that he then come
"Then that boy’s father came." (Nùng; cf. Saul and Wilson 1980: 16)

B vahng pi bao muhn hết cahn sláy
boy older brother his do person priest
"His older brother is a sorcerer." (Nùng; cf. Saul and Wilson 1980: 71)

ad F2: The system allows for a certain flexibility: “Cáh is the most general classifier. It may occur before animate beings [...] and before inanimate objects [...]” (Saul and Wilson 1980: 26). At the same time, animates can be categorised by tú and inanimates can be categorised by áhn.

ad F3: Nùng employs four general classifiers and seven specific classifiers for inanimate objects; these are listed in table [121]
Table 121: The Nùng classifiers (cf. Saul and Wilson 1980: 25ff.).

<table>
<thead>
<tr>
<th>classifier</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>áhn</td>
<td>“inanimate”</td>
</tr>
<tr>
<td>tū</td>
<td>“animate”</td>
</tr>
<tr>
<td>ơng</td>
<td>“human”</td>
</tr>
<tr>
<td>cāh</td>
<td>general classifier</td>
</tr>
<tr>
<td>bāh</td>
<td>“leaf-like object”</td>
</tr>
<tr>
<td>cōn</td>
<td>“lump-shaped object”</td>
</tr>
<tr>
<td>hōng</td>
<td>“clothing”</td>
</tr>
<tr>
<td>mōs</td>
<td>“bean-like object without shell”</td>
</tr>
<tr>
<td>phōhn</td>
<td>“bedding”</td>
</tr>
<tr>
<td>tēu</td>
<td>“stick or string-like object”</td>
</tr>
<tr>
<td>maht</td>
<td>“fabric, board, brick, tile, or raindrop”</td>
</tr>
</tbody>
</table>

ad F₄: The Nùng classifiers occur once within the noun phrase (cf. Conklin 1981: 111). The only exception is the optional cooccurrence of the general classifier cāh and tū “animate” (cf. 197). This represents a case of double classification and not multiple class marking within the noun phrase.

(197) cāh  tū  pá  du  chúng  áhn  hō
     CL:GENERAL  CL:ANIMATE  fish at insiden  CL:INANIMATE  house
   nĩ,  muhn  va
   FOC  he  say

   “The fish inside the house said…” (Nùng; cf. Saul and Wilson 1980: 16)

ad F₅: The classifiers occur NP-internally (Conklin 1981: 111).

ad F₆: The classifiers do not add semantics to the noun phrase, but mirror semantic properties of the nominal referent without changing their host NP’s meaning (cf. 193 and Saul and Wilson 1980: 26).

ad F₇: In general, the classifiers are assigned by transparent semantic mechanisms (cf. table 121). The use of classifiers may vary (mostly due to the occurrence of the general classifier cāh, but also humans may be referred to by the classifier tū “animate”), thus the speaker has influence on the occurrence of the individual classifiers (cf. e.g. Saul and Wilson 1980: 26).
27.3.3 Mapping

Figure 57: Nùng.

Table 122: The formalisation and transparency value of Nùng.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td></td>
<td></td>
</tr>
<tr>
<td>obligatoriness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td>x-value</td>
<td>0.408</td>
<td>y-value</td>
<td>0.893</td>
</tr>
</tbody>
</table>

ad X₁: The eleven classes of the Nùng system render it a small one (cf. table 121).

ad X₂: The classifiers occur obligatorily in noun phrases that contain a numeral and a classifiable noun; they are inserted between these (cf. Conklin 1981: 113 and Saul and Wilson 1980: 27). They also occur in contexts lacking quantification, where they typically occur next to the classified noun (cf. 192) or replace it (cf. 195).

ad X₃: The classifiers occur NP-internally and independent of the presence of the classified noun (cf. 195 and Conklin 1981: 111).
ad X₄: The classifiers are optional in contexts that contain powers of ten (cf. Conklin 1981: 113 and Saul and Wilson 1980: 27). If they occur in a context without numeral, the numeral “one” is understood (ib.; cf. 198 A). The occurrence differs with respect to unspecific quantifiers (Saul and Wilson 1980: 28): With measures and lái “much, many”, the classifier is omitted (cf. 198 B - D); with ki “several”, the classifier may occur (also with instances of powers of ten; cf. 198 E).

(198)

A məhn khâu hchn áhŋ bōpyyyy du
he enter side CL:INANIMATE mine go stay
“He entered the side of the mine and stayed.” (Nùng; cf. Saul and Wilson 1981: 28)

B slám áhŋ vět
three CL:INANIMATE spoon
“three spoons”

C slám vět náhmm
three spoon water
“three spoonfuls of water” (Nùng; cf. Saul and Wilson 1980: 29)

D məhn kihn lái məhn lái
he eat many potato many
“He ate many potatoes”

E mi kí pă̄ (áhbm) kō mày
have several hundred CL:INANIMATE house burn
“There were several hundred houses burned.” (Nùng; cf. Saul and Wilson 1980: 28)

ad X₅: The classifiers occur as free elements (cf. 197).

ad X₆: The classifiers occur once per classificatory context (Conklin 1981: 111).

ad X₇: The system applies to the full nominal lexicon due to the presence of a general classifier (cf. table 121).

ad Y₁: The class inventory is mixed in structure: Some classes are fairly specific, others are more general (cf. table 121).

ad Y₂: The consulted sources and available data suggest no other assignment than a semantic and speaker-sided one.

ad Y₃: Due to the small number of semantically straightforward classes (cf. table 121), a small number of semantic assignment rules is sufficient to master the Nùng classification system.
ad Y₄: The markers convey classificatory semantics only (cf. table 121).

ad Y₅: Each class has its own, distinct marker (cf. table 121).

ad Y₆: The consulted sources and available data do not suggest the classifiers to semantically contribute to their host NP. Instead, the markers mirror semantic properties of the nominal referent (cf. 104).

ad Y₇: The same noun may take different classifiers, e.g. a general or specific one (Saul and Wilson 1980: 26). Different classifiers categorising the same noun can even cooccur (cf. 107).

27.4 Grammaticalisation
The Nùng numeral classifier system matches in its core a prototypical classifier system, though it deviates from such one in a few characteristic aspects. The small set size, the lack of semantic contribution to the noun phrase, and the expansion of the classifiers’ potential hosts indicate a modest formalisation process (value: 408). At the same time, the small and semantically straightforward class inventory make the system fairly transparent (value: 0.893).

28 Sumerian
28.1 Background
< language isolate (ISO 639-3: sux)
< language isolate (cf. Ruhlen 1987)

Sumerian is an isolate, extinct language of Southern Mesopotamia, which is not related to, but in close contact with Akkadian. It employs a concordial class system. This dossier is mainly based on Edzard 2003, Foxvog 2011, Hayes 1990 and 1999, Jagersma 2010, and Thomsen 1984.

![nominal classification devices](image)

Figure 58: Classification of the Sumerian system of nominal classification.
Appendix

28.2 Classification Typology

28.2.1 Noun Classifiers

Sumerian does not employ a noun classifier system (cf. [199] where no classifier occurs).

(199) **engar =e šu =e ∅ =ba**

farmer ERG hand DIR VOCALIC PREFIX 3N.OBJECT(indirect)

=ba 3N.AGENT approach 3N.SUBJECT/OBJECT(direct)

“The farmers received this.” (Sumerian; cf. Jagersma 2010: 104)

28.2.2 Numeral Classifiers

Sumerian does not employ a numeral classifier system (cf. [200] where no classifier occurs).

(200) **alan ur.namma.k =ak limmu =be, igi min-ʔ=t=be, ša.g kišib**

statue Urnamma GEN four DEM eye two-NOM DEM heart

namin =ak

seal forty GEN

“the four statues of Ur-Namma, both eyes, among [‘at the heart of’; MP] forty sealed documents” (Sumerian; cf. Jagersma 2010: 244f. and 248f.)

28.2.3 Possessive Classifiers

Sumerian does not employ a possessive classifier system (cf. [201] where no classifier occurs).

(201) **A lugal ki =∅ G•∅**) =ane

king place ABSOLUTIVE measure.out-NONFINITE his

=e ERGATIVE

“his beloved master”

B **G•j =ane =ʔa**

house his LOC

“into his house” (Sumerian: Jagersma 2010: 214)

C **geme2 =ane =∅**

female.slave his ABSOLUTIVE

“his slave woman”

D **kišib urdu dam damu =ane =ak**

seal Urdudam son his GEN

“seal of Urdudam, his son” (Sumerian; cf. Jagersma 2010: 215)

E **ka.g damu =ane =enē =ak =ʔa**

mouth child her PL GEN LOC

“in the mouth of her children” (Sumerian; cf. Jagersma 2010: 212)
28.2.4 Multiple Classifier Systems

Sumerian does not employ a multiple classifier system (cf. 199, 200, and 201, where no classifier occurs).

28.2.5 Concordial Class Systems

Sumerian employs a concordial class system (cf. Jagersma 2010: 101 and Foxvog 2011: 23ff.). This system is analysed in section 28.3.

28.3 Analysis

28.3.1 Functional Setup

Table 123: The semantic and discourse properties of the Sumerian system of nominal classification.

<table>
<thead>
<tr>
<th>differentiating referents (C1)</th>
<th>productive derivation, variable classification</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuation (C2)</td>
<td>unitisation for countability</td>
<td>x</td>
</tr>
<tr>
<td>attribution of properties (C3)</td>
<td>speaker-sided values</td>
<td></td>
</tr>
<tr>
<td>reference identification (C4)</td>
<td>reference tracking via AGR, anaphora</td>
<td>✓</td>
</tr>
</tbody>
</table>

ad C1: Nouns themselves are not marked for concordial class (Jagersma 2010: 101). The semantics of the system do not allow a variable classification of referents (except for slaves, which may be assigned to either the common or the neuter gender; cf. 202 on page 566). The Sumerian system seems not to be involved in derivational noun formation.

(202) 1 saḫḫ.ī,mī sar.uru.a =∅ mu =ane =∅ =(?a)m
one female.slave Šarrûa ABS name her.C ABS be:3SG.SUBJECT
niḡ.sīm =beE price its:N

“One female slave, Šarrûa [is] her name, its price […].” (Sumerian; cf. Jagersma 2010: 103)

ad C2: The consulted sources and available data do not suggest the Sumerian system to be involved in individuation.

ad C3: The consulted sources and available data do not suggest the Sumerian system to be involved in the speaker-sided attribution of properties by classificatory means.
Among others, pronominals are marked for class (cf. 203). The Sumerian system thus is allowed in reference identification in discourse.

(203) \[
\begin{array}{llllll}
gem\text{ne}_2 & = & \emptyset & \text{nin} & = & \text{d(a)} \\
\text{female.slave} & \text{ABS} & \text{her} & \text{COMITATIVE} & \text{VOCALIC.PREFIX} \\
\text{-mu} & \text{n} & \text{-da} & \text{-sd} & \emptyset \\
\text{VENTIVE} & \text{3SG.COMMON} & \text{with} & \text{be.equal} & \text{3SG.SUBJECT/OBJECT} \\
\text{-?} & \text{=} & \text{?} & \text{=} & \text{?am} \\
\text{NOM} & \text{ABS} & \text{be:3N.SG} \\
\end{array}
\]

“It was [the case] that a female slave was equal to her mistress.” (Sumerian; cf. Jagersma 2010: 199)

28.3.2 Formal Setup

Table 124: The formal properties of the nominal classification system of Sumerian.

<table>
<thead>
<tr>
<th>Feature (F)</th>
<th>CLS-like (Concordial)</th>
<th>CCS-like (Case)</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F1)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility (F2)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory (F3)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking (F4)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi (F5)</td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and external</td>
</tr>
<tr>
<td>semanticity (F6)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment (F7)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad F1: All Sumerian nouns are assigned to a concordial class (Jagersma 2010: 102).

ad F2: Besides the mentioned cases of nouns that denote slaves or groups of people (cf. 202 and 203), every Sumerian noun is assigned to one concordial class only (Jagersma 2010: 102): It “does not betray its class” (Edzard 2003: 29).
Appendix

ad F₃: Sumerian employs two concordial classes (cf. Jagersma 2010: 101 and Foxvog 2011: 23f.), a common and a neuter class (cf. table 125 for a list of the pronominal class markers).

Table 125: The inventory of the Sumerian pronominal class markers (cf. Jagersma 2010: 102 and 212).

<table>
<thead>
<tr>
<th>3rd sg.</th>
<th>common</th>
<th>neuter</th>
</tr>
</thead>
<tbody>
<tr>
<td>verbal person-prefix</td>
<td>n-</td>
<td>b-</td>
</tr>
<tr>
<td>possessive pronoun</td>
<td>=anc</td>
<td>=be</td>
</tr>
<tr>
<td>interrogative pronoun</td>
<td>a-ba</td>
<td>a-na</td>
</tr>
</tbody>
</table>

ad F₄: Sumerian allows multiple marking in the same context (cf. 202, where a slave is marked by both common and neuter class markers).

ad F₅: Among others, the class markers are attached to the verb and thus operate NP-externally (cf. 203).

ad F₆: While semantics is of importance in the process of concordial class assignment (cf. Jagersma 2010: 102), a noun’s concordial class membership does not add semantics to the noun’s lexical entry, but rather mirrors lexical properties (‘humanness’) of the noun.

ad F₇: The Sumerian noun is assigned to its concordial class ‘on a strictly semantic basis’ (Jagersma 2010: 102).
28.3.3 Mapping

![Graph](image)

Figure 59: Sumerian.

Table 126: The formalisation and transparency value of Sumerian.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>1</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>0.75</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>1</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>1</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>0.75</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>1</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>1</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td>x-value</td>
<td>0.929</td>
<td>y-value</td>
<td>0.738</td>
</tr>
</tbody>
</table>

**ad X₁:** Its two classes render the Sumerian system a small one (cf. table 125).

**ad X₂:** Some nominal cases occur are expressed with only one of the classes (e.g. the directive case with the neuter class and the dative marker with the common class (Edzard 2003: 29, Jagersma 2010: 102). Only common nouns take a plural marker (cf. 204) where only the noun diĝir receives plural marking).

(204) A /diĝir / gal-gal /lagas =a$kI| =enê  ê
god big-big Lagash GEN PL ERG

“the great gods of Lagash” (Sumerian; cf. Jagersma 2010: 106)
Furthermore, the markers occur in fused form with pronominal verb prefixes, possessives, and interrogatives (cf. table 125).

**ad X₃:** Among others, the class markers attach to the verb and thus occur NP-externally (cf. 203).

**ad X₄:** The class markers are fused with obligatory categories such as pronominal marking on the verb (cf. table 125). Therefore, classification in Sumerian is obligatory.

**ad X₅:** The markers occur in bound and fused form (cf. table 125).

**ad X₆:** Sumerian allows multiple marking (cf. 202, where a slave is marked for both common and neuter class).

**ad X₇:** All Sumerian nouns are assigned to a concordial class (Jagersma 2010: 102).

**ad Y₁:** The class semantics are straightforward: The common class comprises proper names of and generic terms for humans and gods as well as kinship and professional terms, and the neuter class comprises all other nouns such as plants and animals, all sorts of items, geographical terms and abstractions, etc. (cf. Jagersma 2010: 103).

**ad Y₂:** Only semantic assignment rules determine the class assignment in Sumerian. The abovementioned assignment rules are complemented by only two exceptions: Terms for groups of persons are commonly, but not always assigned to the neuter class (except for a group of persons who “are important as separate individuals” (Jagersma 2010: 103). The words for slaves, saḫ “male slave” and geme₂ “female slave”, may show agreement with markers from both classes (cf. 202). This is rather an intertextual than an intratextual phenomenon (ib.).

**ad Y₃:** The two classes are straightforward and know only two exceptions, the ‘epicene’ class membership of slave terms and the neuter as the default class for group terms (Jagersma 2010: 103). The number of assignment rules in Sumerian is thus low.

**ad Y₄:** The class marking is fused with the category number (cf. 204 and the discussion under X₃).
ad Y₅: Each class possesses its own, distinct markers (cf. table 125).

ad Y₆: The class marking mirrors semantic properties of the nominal referent and thus is redundant (cf. 202, where a class reassignment from the common to the neuter class does not influence the semantics).


### 28.4 Grammaticalisation

Sumerian is a largely formalised concordial class system (value: 0.929). Surprisingly, the origin of its markers is recognisable: While Jagersma (2010: 105) notes that he “cannot offer a plausible reconstruction of what preceded the attested uses of the [class markers; MP] /n/ and /b/ in Sumerian”, Foxvog (2011: 24) assumes their origin to lie in deictics, “/n/ designating near-deixis “this one here”, and /b/ far-deixis “that one there””. The small class inventory with its highly transparent semantics makes the system a transparent one (value: 0.738).

### 29 Tainae

#### 29.1 Background

Tainae (or Ivori) is an Angan language that is spoken in the “Gulf Province, Ivori-Swanson district” (Lewis et al. 2013: s.v.) on Papua New Guinea. This dossier is based on Carlson 1991, the only comprehensive grammar on Tainae. The Tainae system eludes a straightforward assignment to one of the subtypes of nominal classification devices, as it displays fairly mixed properties of both classifier and concordial class systems. It is typically identified as a concordial class system (cf. Carlson 1991: 23), while this study treats Tainae as a multiple classifier system due to the seeming lack of agreement.
29.2 Classification Typology

29.2.1 Noun Classifiers

Tainae attaches its class markers to the noun (cf. 205). Since they also occur in other kinds of constructions (cf. 206 and 207), the Tainae system is analysed as a multiple classifier system in this dossier.

(205)  
\[
F_{-ai} \quad \text{popti-} \quad \text{haminim} \\
\text{road-LONG} \quad \text{enter-LONG} \quad \text{gone}
\]

“The way to enter was gone.” (Tainae; cf. Carlson 1991: 118)

29.2.2 Numeral Classifiers

Tainae attaches its class markers to quantifiers (cf. 206). Since they also occur in other kinds of constructions (cf. 205 and 207), the Tainae system is analysed as a multiple classifier system in this dossier.

(206)  
\[
\text{Padae-ta} \quad \text{mpa-po} \quad \text{ha-e-yo} \quad f_{-ai-ka} \\
\text{exit-1PL} \quad \text{outside-static} \quad \text{stone-PL-OBJ} \quad \text{one-LONG-two} \\
\text{ki-nk-i} \quad f_{-ai-ka} \quad \text{otiku-w} \quad \text{mpi-}n\text{u} \\
3\text{SG-POSS-EPENTHESES} \quad \text{one-LONG-two} \quad \text{dirt-PL} \quad \text{one-only}
\]

“We went outside and he took two stones and one [clump of] dirt.” (Carlson 1991: 117)

29.2.3 Possessive Classifiers

Tainae attaches its class markers to possessives (cf. 207). Since they also occur in other kinds of constructions (cf. 205 and 206), the Tainae system is analysed as a multiple classifier system in this dossier.

(207)  
\[
\text{Tikwamk-ai} \quad \text{ni-nk-ai} \quad \text{nampi-ti} \\
\text{charcoal-LONG} \quad 1\text{SG-POSS-LONG} \quad \text{give.me-2SG.IMP}
\]

“Give me my pencil!” (Tainae; cf. Carlson 1991: 40)
29.2.4 Multiple Classifier Systems

Tainae employs a multiple classifier system (cf. 205, 206, and 207). This system is analysed in section 29.3.

29.2.5 Concordial Class Systems

Tainae employs a nominal classification system. This system is typically identified as a concordial class system (cf. Carlson 1991: 23). This study analyses the Tainae system as a multiple classifier system due to the lack of evidence for the presence of agreement, the highly flexible classification according to an objects shape, and the overall optionality of the system.

29.3 Analysis

29.3.1 Functional Setup

Table 127: The semantic and discourse properties of the Tainae system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Tainae System</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification</td>
</tr>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
</tr>
</tbody>
</table>

ad C₁: Class markers may be involved in derivational noun formation of e.g. verb roots (cf. 205 and Carlson 1991: 28 and 118). The Tainae system differentiates the sex of referents by concordial class marking (cf. 208 on page 573).

(208) A at-a
grandparent-M
“grandfather”
B at-epi
grandparent-F
“grandmother” (Tainae; cf. Carlson 1991: 24)

In addition, the Tainae system allows a productive variable classification of shape (cf. 209 and Carlson 1991: 27f).

(209) A skyppu-ina
leaf-FLAT
“leaf”
B skyppu-ina
leaf-CYLINDRICAL
“leaf (roll)”
Appendix

C  ikypw-itl
leaf-THIN
“leaf (strip)”

D  ikypw-ai
leaf-LONG
“(folded) leaf (rod)”

E  ikypw-ipi
leaf-FLUID
“leaf (ball)” (Tainae: Carlson 1991: 27f.)

**ad C**₂: The consulted sources and available data do not suggest the Tainae system to be involved in individuation.

**ad C**₃: According to Carlson (1991: 23 and 27), the Tainae speakers are not aware of the system’s semantic criteria; a speaker-sided attribution of properties to referents by classificatory means is thus highly unlikely.

**ad C**₄: Among others, most pronouns are marked for concordial class and thus allow reference identification by classificatory means (cf. [210] and Carlson 1991: 34).

(210)  nimi-taka  q  te-mti-ki,...
tell.3OBJ-SG.DR 3SG.M this-say.3-SG.REMOTE.PAST

“When he said that to him, he said this,...” (Tainae; cf. Carlson 1991: 134)

### 29.3.2 Formal Setup

Table 128: The formal properties of the nominal classification system of Tainae.

<table>
<thead>
<tr>
<th></th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obligatoriness (F₁)</strong></td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td><strong>flexibility (F₂)</strong></td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td><strong>inventory (F₃)</strong></td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td><strong>marking (F₄)</strong></td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td><strong>locus operandi (F₅)</strong></td>
<td>locus operandi: NP-internal</td>
<td>locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td><strong>semanticity (F₆)</strong></td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
</tbody>
</table>
Appendix

575

assignment (F7)  CLS-like: chosen freely by the speaker ✓
                 CCS-like: mechanic assignment ✗

ad F1: Every noun is assigned to a concordial class in Tainae. This is not explicitly mentioned in the consulted source, but can be derived from the description (cf. e.g. Carlson 1991: 23).

ad F2: The individual concordial classes of Tainae cover certain semantic fields, which suggests that at least a part of them is not able to shift from one concordial class to another. If this study’s reanalysis of the so-called ‘plural’ marker -e as a generic/default class marker is correct, this means that all nouns actually are able to shift their class. Furthermore, there are entities the shape of which can be manipulated, which triggers a change in referential class (cf. 209 and Carlson 1991: 27f.).

ad F3: Tainae employs eleven noun classes, two of which are frozen (Carlson 1991: 23). The markers and numeral forms are given in table 129. The frozen classes are marked by an asterisk.

Table 129: The inventory of Tainae noun class markers (cf. Carlson 1991: 23).

<table>
<thead>
<tr>
<th>marker</th>
<th>numeral “one” + N</th>
<th>semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>fonu wako “one man”</td>
<td>masculine</td>
</tr>
<tr>
<td>epi</td>
<td>fepinu atepe “one grandmother”</td>
<td>feminine</td>
</tr>
<tr>
<td>I</td>
<td>finu si “one possum”</td>
<td>animate</td>
</tr>
<tr>
<td>iwa</td>
<td>fiwanu hakina “one bamboo tube”</td>
<td>cylindrical</td>
</tr>
<tr>
<td>ina</td>
<td>fnanu alina “one ear”</td>
<td>“flat”</td>
</tr>
<tr>
<td>iti</td>
<td>ftiina pikiti “one vine”</td>
<td>“thin, flexible”</td>
</tr>
<tr>
<td>ipi</td>
<td>mpina takipi “one blood”</td>
<td>“fluid”</td>
</tr>
<tr>
<td>ipa</td>
<td>mpanu kstaipa “one bush knife”</td>
<td>“tool”</td>
</tr>
<tr>
<td>*u</td>
<td>fonu fayu “one hand”</td>
<td>“indeterminate”</td>
</tr>
<tr>
<td>*a</td>
<td>fonu oga “one girl”</td>
<td>“girl; rain”</td>
</tr>
</tbody>
</table>

ad F4: Despite Carlson (1991: 23) arguing that a range of nominal modifiers “all are in concord with their head noun or referent”, there is no convincing evidence that proves the existence of agreement in Tainae. This study defined agreement among others as a the copying of information from the head noun and marking it on other elements. Examples such as 209 clearly show the referential character of the class markers, which cannot add semantic content to the noun phrase and display copied semantic content from the head noun at the same time. Therefore, Tainae is not assumed to employ agreement here; the sample study discusses the case of Tainae more in detail.
ad F₅: “The numeric system, adjectives, copular constructions, demonstratives and nominalized verbs all are in concord with their head noun or referent” (Carlson 1991: 23). The only potentially NP-external elements that are marked for class are the verb and the copula. Class marking on the verb nominalises the verb (cf. 205). There is no evidence that allows the justified assumption that the copula construction represents an NP-external constituent (cf. 211). Therefore, the Tainae system operates NP-internally.

ad F₆: The markers may add semantic content to their host NP (cf. 208 and 209).

ad F₇: “Unlike the gender system of Indo-European languages, which for the most part seem arbitrary, Tainae’s productive noun classes are based more self evidently on semantic grounds, and even non-native speakers can hazard a reasonable guess as to the noun class of any particular noun” (Carlson 1991: 23). The system is productive and new loanwords are assigned to their respective concordial class without problem by the speakers despite their unawareness of the system’s underlying semantics (ib.: 23 and 27).

29.3.3 Mapping

Figure 61: Tainae.
Table 130: The formalisation and transparency value of Tainae.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>semantic structure</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>rule types</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td>rule number</td>
<td>$Y_3$</td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>categorial independence</td>
<td>$Y_4$</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>discreteness of markers</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>redundancy</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>flexibility</td>
<td>$Y_7$</td>
</tr>
<tr>
<td>x-value</td>
<td>0.893</td>
<td>y-value</td>
<td>0.929</td>
</tr>
</tbody>
</table>

ad $X_1$: The eleven classes of Tainae render it a small system (cf. table 129).

ad $X_2$: The class markers occur on numerals, adjectives, copular constructions, demonstratives, nouns, and nominalised verbs (Carlson 1991: 23).

ad $X_3$: As the copula cannot be convincingly argued to be an NP-external element (cf. 211), the marking occurs NP-externally. The occurrence of class marking is independent of the presence of the classified noun (cf. 210).

ad $X_4$: The Tainae system serves as a pragmatic tool for focussing discourse referents: Nouns are typically only marked for a specific class if "a specific person or thing is in mind, and not when just talking about something generally" (Carlson 1991: 27). If reference to a thing in general is made, the so-called 'plural' marker -$i/-e$ is attached (cf. 212).

(212) mi nai wink-x ne-ate naoni
QUOTE 1PL garden-CL:RESIDUE do-SG about
"This is about our garden work." (Tainae; cf. Carlson 1991: 27)

This study reanalyses the so-called 'plural marker' as a residue marker, since examples such as 212 do not deliver satisfactory evidence for plural reference, and since instances of nouns categorised by specific markers can also be interpreted as generic forms which refer to more than one real life referent (cf. 213).

(213) Mi apiho-iwa-mi-aiki wink-e te-ye
SR.SEQ troll-SEG.M-OBJ-3.POSS work-PL this-SUBJ
"This is what trolls do." (Tainae; cf. Carlson 1991: 45)

If the reanalysis of the so-called ‘plural marker’ as a generic classifier is correct, classification in Tainae is obligatory.

ad $X_5$: The markers occur in bound form (cf. 209).
Appendix

**ad X₆**: This study argues that Tainae does not allow multiple class marking in the same context (cf. the discussion under F₄).

**ad X₇**: If the reanalysis of the so-called ‘plural marker’ as a generic classifier is correct (cf. the discussion under F₄), the Tainae system applies to the full nominal lexicon.

**ad Y₁**: The individual classes possess straightforward semantics (cf. table 129).

**ad Y₂**: The consulted sources and available data do not suggest assignment rules other than semantic ones to be involved in class assignment. The consulted sources do not report on exceptions from the semantic assignment rules.

**ad Y₃**: The straightforward class structure and low number of classes result in a small number of rules.

**ad Y₄**: The markers convey class information and class information only (cf. table 129).

**ad Y₅**: Each class possesses its own, distinct marker (cf. 129).

**ad Y₆**: In general, unspecified discourse referents are marked by the residue marker -i/-e, which thus does not convey informative semantic content. A change in class may trigger a change in the noun phrase’s reference (cf. 209). Class marking may also convey information of a referent’s biological sex, which can be either informative in the case of epicene nouns or redundant in the case that the information on the referent’s sex is available from the lexical item itself (cf. 214 for a number of examples of both redundant and informative class marking).

(214)  

\[
\begin{array}{llll}
\text{A} & \text{papa-}epi & \text{at-}epi & \text{Pikinika-}epi \\
& \text{cross.cousin-3SG.F} & \text{grandparent-3SG.F} & \text{Pikinini-3SG.F} \\
& \text{“female cross cousin, grandmother, Pikinini”} \\
\text{B} & \text{oy-}o & \text{s-}o & \text{Nesa-}o & \text{pat-}o \\
& \text{boy-3SG.M} & \text{wild.pig-3SG.M} & \text{Nesao-3SG.M} & \text{hawk-3SG.M} \\
& \text{“boy, wild pig, Nesao (“Razor”), hawk” (Tainae; cf. Carlson 1991: 24)}
\end{array}
\]

**ad Y₇**: The same noun can take various classifiers (cf. 208 and 209). As nouns are only marked by specific class markers if the referent is discourse-specific and otherwise take the marker -i/-e (Carlson 1991: 27), all nouns are at least able to vary between this general marker and a specific marker.
29.4 Grammaticalisation

The Tainae system of nominal categorisation falls between the cracks of the prototypical systems of nominal categorisation. Some of its properties (the size of its inventory, the multiple marking, the employed agreement, the opaque origin of its markers, and the obligatory assignment of each noun to a concordial class) suggest it to be more a concordial class system, whereas others (the flexible class membership, the semanticity of its markers, their NP-internal occurrence, and the assignment by the speaker) suggest it to be more classifier-like.

The Tainae system is the only one from the sample that is analysed as a classifier system, but displays a formalisation value of more than 0.8 (value: 0.893). At the same time, the system is highly transparent (value: 0.929). This ‘mismatch’ is due to the reanalysis of the so-called ‘plural marker’ as a generic classifier, which occurs with all nouns when “a specific person or thing is in mind, and not when just talking about something generally” (Carlson 1991: 27). Due this reanalysis, the classification becomes exhaustive and obligatory in all contexts. The Tainae case suggests that the individual measures of the mapping instrument are further refined with more unusual systems like the one of Tainae.

30 Tamil

30.1 Background

Tamil is a Dravidian language originally spoken in Southern India (federal state Tamil Nadu) and Sri Lanka. There are also speaking communities in Bahrain, Belize, Canada, Denmark, Fiji, Germany, Malaysia, Mauritius, the Netherlands, Qatar, Réunion, Singapore, South Africa, Thailand, the United Arab Emirates, the United Kingdom and the United States (cf. Lewis et al. 2013: s.v.). Among the Dravidian languages, it is the diachronically best documented language covering a span of more than 2000 years (Andronov 2003: 21). Five grammars and two comparative handbooks constitute the main base of this dossier: Andronov 2003 and 2004, Asher 1982, Krishnamurti 2003, Lehmann 1989, Rhenius 1836, and Schiffman 1999. Tamil employs a concordial class system.
30.2 Classification Typology

30.2.1 Noun Classifiers
Tamil does not employ a noun classifier system (cf. 215 where no classifier occurs).

\[(215)\] kumaar kattiʼy-aal paza-tt-ai veft-in-aan
Kumar:M knife:N-INSTR fruit-OBL-ACC cut-PAST-3SG.M
“Kumar cuts the fruit with a knife.” (Tamil; cf. Lehmann 1989:36)

30.2.2 Numeral Classifiers
Tamil does not employ a numeral classifier system (cf. 216 where no classifier occurs).

\[(216)\] kumaar aintu mani-kku vara-v-aan
Kumar:M five day:DAT.N work do-PRS-3SG.M
“Kumar works for five days.” (Tamil; cf. Lehmann 1989:33)

30.2.3 Possessive Classifiers
Tamil does not employ a possessive classifier system (cf. 217 where no classifier occurs).

\[(217)\] kumaar enkal viit-tu-kku va-nt-aan
Kumar:M we:OBL house:DAT.N come-PAST-3SG.M
“Kumar came to our house.” (Tamil; cf. Lehmann 1989:32)

30.2.4 Multiple Classifier Systems
Tamil does not employ a multiple classifier system (cf. 215 216 and 217 where no classifier occurs).
30.2.5 Concordial Class Systems

Tamil employs a concordial class system (cf. 215, 216 and 217, where class marking occurs e.g. on the verb. This system is analysed in section 30.3

30.3 Analysis

30.3.1 Functional Setup

Table 131: The semantic and discourse properties of the Tamil system of nominal classification.

| differentiating referents (C₁) | productive derivation, variable classification ✓ |
| indification (C₂) | unitisation for countability ✓ |
| attribution of properties (C₃) | speaker-sided values ✓ |
| reference identification (C₄) | reference tracking via AGR, anaphora ✓ |

ad C₁: There is no information available on the Tamil system to be involved in free, variable classification. Nouns can distinguish a referent’s biological sex by concordial class marking (cf. 218 and e.g. Lehmann 1989: 14) and rational/nonrational referents can be distinguished by assignment of the same noun to the respective classes (Andronov 2004: 58f.); the Tamil system thus allows a productive derivation for the purpose of referent differentiation.

(218) A maan˙a
ющую-3SG.M
v-an
˙
hab-3SG.F
maan˙a
˙

student-3SG.M
student-3SG.F

“male student, female student” (Tamil; cf. Lehmann 1989: 14)
B maan˙a
˙

maan˙a
˙

student-3SG.HONORIFIC.EPICENE/3PL.EPICENE student-3F-PL

student-3SG.M

“student(s) (epicene honorific singular or epicene plural), female students”
(Tamil; cf. Lehmann 1989: 14 and 20)

ad C₂: The consulted sources and available data do not suggest the Tamil system to be involved in individuation.

ad C₃: The system contains a honorific category, which allows the speaker to attribute properties to a referent (cf. 218 B).

ad C₄: Among other constituents, pronouns are marked for concordial class (cf. 219 and Andronov 2011: 147). The Tamil system thus allows reference identification by classificatory means.

(219) geun geun-ai vera-ker-an

geun-ai

vera-ker-an

3SG.M 3SG.M-acc hate-PRS-3SG.M

“He hates himself.” (Tamil; cf. Lehmann 1989: 96)
30.3.2 Formal Setup

Table 132: The formal properties of the nominal classification system of Tamil.

<table>
<thead>
<tr>
<th>Property</th>
<th>CLS-like:</th>
<th>CCS-like:</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F1)</td>
<td>nouns may remain uncategorised</td>
<td>obligatory assignment of each noun</td>
</tr>
<tr>
<td>flexibility (F2)</td>
<td>flexible class membership</td>
<td>single class membership</td>
</tr>
<tr>
<td>inventory (F3)</td>
<td>largish class number</td>
<td>limited class number</td>
</tr>
<tr>
<td>marking (F4)</td>
<td>marked once</td>
<td>multiply marked</td>
</tr>
<tr>
<td>locus operandi (F5)</td>
<td>locus operandi: NP-internal</td>
<td>NP-internal and -external</td>
</tr>
<tr>
<td>semanticity (F6)</td>
<td>semantic contribution</td>
<td>semantically vain</td>
</tr>
<tr>
<td>assignment (F7)</td>
<td>chosen freely by the speaker</td>
<td>mechanic assignment</td>
</tr>
</tbody>
</table>

ad F₁: “Tamil nouns fall into two lexical classes. The superior class comprises nouns which denote human beings and anthropomorphic personages of literature. The rest [...] belong[s] to the inferior class” (Andronov 2004: 57).

ad F₂: Since a noun’s gender is determined by properties of its referent, it cannot change its gender as a rule. (Andronov 2004: 57).

ad F₃: The two basic referential classes of Tamil are traditionally referred to as the ‘superior’ and ‘inferior’ class (e.g. Andronov 2004; the ‘superior’ class denoting the masculine and feminine gender and the ‘inferior’ class denoting the neuter gender) or comparable judgemental terms such as ‘high caste’ and ‘low caste’ (Rhenius 1836). This is the notion of the traditional grammars (cf. Krishnamurti 2003: 205), whereas a modern approach would rather label them ‘human/animate’ and ‘nonhuman/-animate’ (Krishnamurti 2003: 205) or ‘rational’ (which equals “human”) and ‘nonrational’ (Asher 1982: 136). Since the latter terms appear to be the most unloaded of undesirable connotations, they were adopted for this dossier. Wherever possible, reference to a referential class was made by “masculine”, “feminine” and “neuter”.

ad F₄: Tamil theoretically allows for multiple marking in the same environment by marking e.g. demonstratives and certain nouns for class.
ad F₅: Class marking occurs among others on the verb and thus NP-externally (cf. 219).

ad F₆: Nominal gender is determined by the respective noun’s semantics and vice versa affects a noun’s semantics, as noun stems with different class markers demonstrate (cf. 220).

(220)  
A  ēlai  
poor.man:M  
“poor man”
B  ēlai  
poor.woman:F  
“poor woman”
C  ēlai  
stupidity:N  
“stupidity” (Tamil; cf. Andronov 2004: 59)

ad F₇: The assigning agent is the speaker: Tamil subdivides the rational class into two subclasses, which are usually referred to as the masculine and feminine class, which reflects the biological sex of the typical referents of the respective class members (humans, deities and other “anthropomorphic personages”; Andronov 2004: 57f.; cf. 221).

(221)  
A  veelekkaaran,  paṇakkaaran,  piccekkaaran  
worker:M  rich:M  beggar:M  
“male worker, rich man, beggarman”
B  veelekkaari,  paṇakkaari,  piccekkaari  
worker:F  rich:F  beggar:F  
“female worker, rich woman, beggarwoman” (Tamil; cf. Asher 1982: 136)

Nonrational nouns denoting humans and nouns denoting small children with respect to their biological sex change their gender to the rational class in the respective construction (cf. 222 and Andronov 2004: 58).

(222)  
A  kulantu  atu-ata  
child:N  weep:FAST-3SG.N  
“The child wept.”
B  kulantu  vicīlam  tan  takappanūrīlam  varukir-ăl  
“The girl [literally: child (n.)] Vishalam comes [f.] near her father.”  
(Tamil; cf. Andronov 2004: 58)

The nonrational (or neuter) class hosts the rest of the nominal lexicon and includes nouns denoting small children, animals, inanimate objects, and abstractions (cf. 223 A and Andronov 2004: 58f.). Just as members of the nonrational class change their gender to the rational class when denoting a rational entity, members of the rational class become nonrational when referring to a non-rational entity (cf. 223 B and Andronov 2004: 58f.). Furthermore, “[s]emantic change or metaphorization can lead to corresponding modifications in the noun gender” (Andronov 2004: 59).
Appendix

A kulantai, puli, nerup, amp
child tiger fire love
“child, tiger, fire, love” (Tamil; compiled after Lehmann 1989)

B pakatūr parvaikkul alakāy irunt-atu
“Bahadur [a proper name (m.) of a dog] looked [n.] fine.”
(Tamil; cf. Lehmann 1989: 58)

30.3.3 Mapping

![Graph showing formalisation and transparency values for Tamil]

Figure 63: Tamil.

Table 133: The formalisation and transparency value of Tamil.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th></th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td>rule number</td>
<td>Y₃</td>
</tr>
<tr>
<td>obligatoryness</td>
<td>X₄</td>
<td>categorial independence</td>
<td>Y₄</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreetness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
</tbody>
</table>

| x-value       | 0.964 | y-value       | 0.917 |

ad X₁: The Tamil system consists of a masculine, a feminine, and a neuter class (cf. the discussion under F₇).
ad X₃: The class markers attach among others to the predicate and thus occurs NP-externally (cf. 224).

ad X₄: The so-called ‘persona nouns’, predicates, and demonstratives require a class suffix in the third person (cf. 224 and 225).

ad X₅: The Tamil markers occur in bound form (cf. 225).

ad X₆: Most nominal satellites are not marked for class. I was not able to find a concrete example of multiple marking in Tamil. However, since some nouns are overtly marked for class (cf. 221) and the demonstrative is, too (cf. 224), multiple marking is theoretically possible.

ad X₇: “Tamil nouns fall into two lexical classes. The superior class comprises nouns which denote human beings and anthropomorphic personages of literature. The rest [ . . . ] belong[s] to the inferior class” (Andronov 2004: 57).

ad Y₁: The classes are semantically straightforward (cf. the discussion under F₇).

ad Y₂: Due to the small class inventory and the straightforward class structure, there are only semantic assignment rules necessary to master the system.

14Persona nouns are nominal stems that take personal suffixes that “distinguish the categories of person, number, and, in the 3rd person, gender” (Andronov 2004: 146).
Appendix

ad Y₃: Due to the small class inventory and the straightforward class structure, a small number of assignment rules is sufficient to master the system.

ad Y₄: The class markers are fused with the category number (cf. C and D).

ad Y₅: Each class has its own, distinct markers (cf. e.g. C and D).

ad Y₆: Class marking can be either redundant (cf. where a male person is marked for the masculine class) or informative (cf. C and D).

ad Y₇: A noun cannot change its class as a rule (Andronov 2004: 57).

30.4 Grammaticalisation

The Tamil system displays a high degree of formalisation (value: 0.964) and at the same time a high degree of transparency (value: 0.917). The most eyecatching property of the system is its straightforward and seemingly exceptionless transparent class assignment. Furthermore, only a small number of nominal satellites is marked for class, which results in a low visibility of the Tamil system.

31 Tlingit

31.1 Background

< language isolate (ISO 639-3: tli; cf. Lewis et al. 2013: s.v.)
< Tlingit < Continental Na-Dene < Na-Dene (cf. Ruhlen 1987)

Tlingit is a Northern American language spoken around Yakutat in Alaska, Carcross Tagish, and Canada (cf. Lewis et al. 2013: s.v.). Aikhenvald (2000: 121) lists Tlingit as a numeral classifier language. This claim is based on Sherzer (1976: 74), where Tlingit (among others) is listed as a language in which numerals are “classified according to the form or shape of objects” or “other specifications of objects” (Sherzer 1976: 74). Dryer/Haselmath (2013: s.v.) and Kinkade (2001) also list Tlingit as a numeral classifier language. For this reason, it is treated as such in this dossier; however, the system is not productive anymore. This dossier is mainly based on Boas 1917, Crippen 2010, Edwards 2009, Eggleson 1013, Krauss 1968, and Swanton 1911.
31.2 Classification Typology

31.2.1 Noun Classifiers

Tlingit does not employ a noun classifier system (cf. (226) where no classifier occurs).

(226) *Caqabha’dt gəkʷ-’t a-wu-go’qʷ du-axa’yî.*

bowman canoe-into CL-COMPLETE.ACTION-push his-paddle-POS

“The bowman pushed his paddle [down] into the canoe.” (Tlingit; cf. Swanton 1911: 201, 204)

31.2.2 Numeral Classifiers

Tlingit is said to employ a numeral classifier system that distinguishes a human and a non-human class (cf. e.g. Aikhenvald 2000: 121). This system, which is not productive anymore, is analysed in section 31.3.

31.2.3 Possessive Classifiers

The Tlingit possessive construction consists of the possessive pronoun preceding the noun and the possessive suffix -(y)i, -wu or wo, the distribution of which is determined by the phonological ending of the possessed noun (Swanton 1911: 169f.). If the possessed noun “is a term of relationship or part of the body, or one of few other terms” (Swanton 1911: 169f.), the possessive suffix is omitted. Therefore, Tlingit does not employ a possessive classifier system (cf. also (226) where no classifier occurs in the possessive construction).

31.2.4 Multiple Classifier Systems

Tlingit does not employ classifiers in possessive, numeral, or other contexts (cf. (226). The supposed numeral classifiers are not a productive mean of nominal classification and are analysed in section 31.3.

Figure 64: Classification of the Tlingit system of nominal classification.
31.2.5 Concordial Class Systems

Tlingit does not employ a concordial class system (cf. Swanton 1911: 167ff. and 226, where no class marking occurs on nominal satellites).

31.3 Analysis

31.3.1 Functional Setup

Table 134: The semantic and discourse properties of the Tlingit system of nominal classification.

| differentiating referents (C₁) | productive derivation, variable classification | ✓ |
| individuation (C₂)             | unitisation for countability                   | ❌ |
| attribution of properties (C₃)  | speaker-sided values                           | ❌ |
| reference identification (C₄)   | reference tracking via AGR, anaphora            | ✓ |

ad C₁: The marking occurs on the numeral only and thus does not take part in productive derivational processes. A variable classification of the same entity is not possible, since the class assignment is inflexible.

ad C₂: The consulted sources and available data do not suggest the Tlingit system to be involved in individuation.

ad C₃: Though the semantics would theoretically allow to assign a human to the non-human category to refer to him or her in a pejorative manner, this is not suggested in any of the consulted sources or by the available data.

ad C₄: The markers copy the semantic value \([±\text{human}]\) from the head noun to the numeral. This is a property of agreement in the definition of this study. Class agreement serves among others the purpose of reference tracking in discourse. Though numerals often cooccur with the noun that they quantify, reference tracking by classificatory means is theoretically possible in Tlingit.

31.3.2 Formal Setup

Table 135: The formal properties of the nominal classification system of Tlingit.

| obligatoriness (F₁) | CLS-like: nouns may remain uncategorised | ✓ |
|                    | CLS-like: obligatory assignment of each noun | ❌ |
Appendix

<table>
<thead>
<tr>
<th>flexibility (F₂)</th>
<th>CLS-like: flexible class membership</th>
<th>CCS-like: single class membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory (F₃)</td>
<td>CLS-like: largish class number</td>
<td>CCS-like: limited class number</td>
</tr>
<tr>
<td>marking (F₄)</td>
<td>CLS-like: marked once</td>
<td>CCS-like: multiply marked</td>
</tr>
<tr>
<td>locus operandi (F₅)</td>
<td>CLS-like: locus operandi: NP-internal</td>
<td>CCS-like: locus operandi: NP-internal and -external</td>
</tr>
<tr>
<td>semanticity (F₆)</td>
<td>CLS-like: semantic contribution</td>
<td>CCS-like: semantically vain</td>
</tr>
<tr>
<td>assignment (F₇)</td>
<td>CLS-like: chosen freely by the speaker</td>
<td>CCS-like: mechanic assignment</td>
</tr>
</tbody>
</table>

**ad F₁:** Since only the numerals differentiate humans and non-humans formally (Edwards 2009: 17), unquantified entities remain uncategorised.

**ad F₂:** The semantics of the classification are straightforward (human vs. non-human; Edwards 2009: 17) and do not suggest a classificational flexibility.

**ad F₃:** Tlingit employs two classes, a human and a non-human one (Edwards 2009: 17).

**ad F₄:** The regular set of Tlingit numerals take the form *tléix, déix, nás’k,* etc. and count objects and animals. If humans are counted, an additional infix *-ná-* enters the numeral between the stem and its suffix, which results in the numeral form *tléináx, dáxnáx, nás’gínáx* (Edwards 2009: 17).

**ad F₅:** Only numerals take the [+human] infix *-ná-* (Edwards 2009: 17). The class marking thus only occurs within the borders of the noun phrase.

**ad F₆:** The consulted sources and available data suggest that the class marking mirrors the referent’s humanness and does not add semantic content to the noun phrase.

**ad F₇:** The classificatory semantics are straightforward and transparent. The speaker is able to carry out the class assignment, but is strictly bound to the systemic semantics and has no reported potential for class variation.
31.3.3 Mapping

![Graph showing the mapping of transparency value to formalisation value.]

Figure 65: Tlingit.

Table 136: The formalisation and transparency value of Tlingit.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>X₁</td>
<td>semantic structure</td>
<td>Y₁</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>host number</td>
<td>X₂</td>
<td>rule types</td>
<td>Y₂</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>locus operandi</td>
<td>X₃</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>obligatoryness</td>
<td>X₄</td>
<td>categorical independence</td>
<td>Y₄</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>boundedness</td>
<td>X₅</td>
<td>discreteness of markers</td>
<td>Y₅</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>marking</td>
<td>X₆</td>
<td>redundancy</td>
<td>Y₆</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>X₇</td>
<td>flexibility</td>
<td>Y₇</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>x-value</td>
<td>0.643</td>
<td>y-value</td>
<td>0.786</td>
</tr>
</tbody>
</table>

**ad X₁**: Tlingit employs two classes, a human and a non-human one (Edwards 2009: 17).

**ad X₂**: The numeral is the only element that takes the [+human] infix -ná- (Edwards 2009: 17).

**ad X₃**: As only the numeral takes the supposed class marking (Edwards 2009: 17), the marking occurs NP-internally. The consulted sources and available data did not provide an example for the occurrence of the marker -ná- in the absence of a classified noun, yet this possibility can be derived from the description of the marking as obligatory (Edwards 2009: 17).
Appendix

ad X₄: The numeral is described as taking the marker -ná-, if it refers to a human (Edwards 2009: 17). As the consulted sources do not further comment on the marking, it can be concluded that this marker occurs obligatorily.

ad X₅: The marking occurs in form of the presence or absence of the infix -ná- (Edwards 2009: 17).

ad X₆: As only the numeral is marked by -ná- (Edwards 2009: 17), multiple marking is not possible in Tlingit.

ad X₇: Human nouns are assigned to the class marked by -ná-, all other nouns to the second class (Edwards 2009: 17). The classification is thus exhaustive.

ad Y₁: The class semantics are straightforward: Human nouns are assigned to one class, all other nouns to the other class (Edwards 2009: 17).

ad Y₂: The consulted sources and available data suggest only semantic assignment rules to be present in the Tlingit system.

ad Y₃: The speaker of Tlingit needs to know two semantic assignment rules to master the system: Humans are assigned to one class, the remaining nominal lexicon to the other class (Edwards 2009: 17). There are no reported exceptions from these rules.

ad Y₄: There is only one marker, -ná- “human”. This marker is not fused with other categories (cf. Edwards 2009: 17).

ad Y₅: The human class is marked by the infix -ná-, the residue class remains unmarked (Edwards 2009: 17). Each class therefore possesses its own, distinct marking.

ad Y₆: The marker mirrors the humanness of the nominal referent (Edwards 2009: 17) and thus does not semantically contribute to its noun phrase.

ad Y₇: The underlying semantics of the classes are mutually exclusive and thus do not allow a flexible classification.
31.4 Grammaticalisation

The supposed numeral classifier system is quite unusual for a classifier system: It employs only two categories, one of which is marked by infixes on numerals. The system seems to exert no influence on the noun phrase’s semantics, and the classification is rigid. In its contemporary state, there is no evidence for the assumption that the Tlingit numeral marking corresponds to a nominal classification system. The system rather represents a rich way to quantify nouns, though it might very well represent an impoverished and petrified form of an earlier productive numeral classifier system.

32 Tohono O’odham

32.1 Background

Tohono O’odham (or Papago, Nevome) is an Uto-Aztecan language spoken in South Central Arizona and Mexico (cf. Lewis et al. 2013: s.v.). The Uto-Aztecan languages are claimed to employ possessive classifiers (cf. Aikhenvald 2000: 147), which are inherited from Proto Uto-Aztecan (cf. Langacker 1977: 90). This dossier is mainly based on Langacker 1977, Saxton 1983 and Zepeda 1983.

32.2 Classification Typology

32.2.1 Noun Classifiers

Tohono O’odham does not employ a noun classifier system (cf. where no classifier occurs).
(227) gook a-t ab wa’i hi-hi-X g
   two MOOD-TENSE LOC only go-REDUPL-PERFECTIVE ART
   či-č-aj man1-REDUPL2
   “Only two men came.” (Tohono O’odham; cf. Saxton 1983: 201)

32.2.2 Numeral Classifiers
Tohono O’odham does not employ a numeral classifier system (cf. 227, where no classifier occurs).

32.2.3 Possessive Classifiers
Tohono O’odham employs a possessive classifier system (Aikhenvald 2000: 127), which distinguishes only two categories, animate and inanimate. This system is analysed in section 32.3.

32.2.4 Multiple Classifier Systems
Tohono O’odham employs a set of classifiers, which occurs only in possessive constructions (Aikhenvald 2000: 127). Therefore, Tohono O’odham does not employ a multiple classifier system.

32.2.5 Concordial Class Systems
Tohono O’odham does not employ a concordial class system (cf. 228, where no class marking occurs on nominal satellites).

(228) A s ˙am hu hiba’i hima d ˙gi’i
   QUOTATIVE LOC REMOTE sometime one EQUATIONAL big
   ban coyote
   “One time, it is said, there was a big coyote.” (Tohono O’odham; cf. Saxton 1983: 263)

   B pi o sa’i ap d o’odham
   NEG MOOD DEGREE good EQUATIONAL person
   “He is not at all a good person.” (Tohono O’odham; cf. Saxton 1983: 168)
32.3 Analysis

32.3.1 Functional Setup

Table 137: The semantic and discourse properties of the Tohono O’odham system of nominal classification.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
<th>Tohono O’odham</th>
</tr>
</thead>
<tbody>
<tr>
<td>differentiating referents (C₁)</td>
<td>productive derivation, variable classification</td>
<td>✘</td>
</tr>
<tr>
<td>individuation (C₂)</td>
<td>unitisation for countability</td>
<td>✘</td>
</tr>
<tr>
<td>attribution of properties (C₃)</td>
<td>speaker-sided values</td>
<td>✘</td>
</tr>
<tr>
<td>reference identification (C₄)</td>
<td>reference tracking via AGR, anaphora</td>
<td>✓</td>
</tr>
</tbody>
</table>

ad C₁: Tohono O’odham distinguishes a class for animates from a class for inanimates (Saxton 1983: 186). Since a nominal referent can be either animate or inanimate, a variable classification is not possible.

ad C₂: The consulted sources and available data do not suggest the Tohono O’odham system to be involved in individuation.

ad C₃: The consulted sources and available data do not suggest the Tohono O’odham system to be involved in the speaker-sided attribution of properties.

ad C₄: The class marking may serve the tracking of referents in discourse (cf. 230, where the class marker allows the identification of the intended referent).

(229) has-çu soi-ga-j $g$ huan

what-thing CL:ANIMATE-POSS:ALI-GEN ART Juan

“What kind of animal does Juan have?” (Tohono O’odham; cf. Saxton 1983: 187 and Aikhenvald 2000: 127)

32.3.2 Formal Setup

Table 138: The formal properties of the nominal classification system of Tohono O’odham.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
<th>Tohono O’odham</th>
</tr>
</thead>
<tbody>
<tr>
<td>obligatoriness (F₁)</td>
<td>CLS-like: nouns may remain uncategorised</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>CCS-like: obligatory assignment of each noun</td>
<td>✘</td>
</tr>
<tr>
<td>flexibility (F₂)</td>
<td>CLS-like: flexible class membership</td>
<td>✘</td>
</tr>
<tr>
<td></td>
<td>CCS-like: single class membership</td>
<td>✓</td>
</tr>
<tr>
<td>inventory (F₃)</td>
<td>CLS-like: largish class number</td>
<td>✘</td>
</tr>
<tr>
<td></td>
<td>CCS-like: limited class number</td>
<td>✓</td>
</tr>
</tbody>
</table>
Appendix

marking \((F_4)\)

<table>
<thead>
<tr>
<th></th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>marked once</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>multiply marked</td>
<td></td>
<td>✗</td>
</tr>
</tbody>
</table>

locus operandi \((F_5)\)

<table>
<thead>
<tr>
<th></th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>locus operandi: NP-internal</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>locus operandi: NP-internal and -external</td>
<td></td>
<td>✗</td>
</tr>
</tbody>
</table>

semanticity \((F_6)\)

<table>
<thead>
<tr>
<th></th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantic contribution</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>semantically vain</td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

assignment \((F_7)\)

<table>
<thead>
<tr>
<th></th>
<th>CLS-like</th>
<th>CCS-like</th>
</tr>
</thead>
<tbody>
<tr>
<td>chosen freely by the speaker</td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td>mechanic assignment</td>
<td>✔</td>
<td></td>
</tr>
</tbody>
</table>

ad \(F_1\): Nouns may remain uncategorised (cf. 227 and 228, where no classifier occurs).

ad \(F_2\): A nominal referent can be either animate or inanimate, not both. Therefore, a classificatory flexibility is not given in Tohono O’odham.

ad \(F_3\): Tohono O’odham distinguishes an animate and an inanimate class (Saxton 1983: 186), which renders the system a small one.

ad \(F_4\): The marking occurs once per classificatory context (cf. 230).

\[(230)\]

\[
g \quad \text{gogs} \quad \\text{soi}-\text{ga}-\\text{j} \quad \text{g} \quad \text{huan} \quad \text{g} \quad \text{huan} \quad \text{gogs}
\]

\[\text{ART} \quad \text{dog} \quad \text{CL:ANIMATE-POSS:ALI-GEN} \quad \text{ART} \quad \text{Juan} \quad \text{ART} \quad \text{Juan} \quad \text{dog} \]

\[\text{soi-}\text{ga} \quad \text{CL:ANIMATE-POSS:ALI}\]

“the dog of Juan, Juan’s dog” (Tohono O’odham; cf. Saxton 1983: 187)

ad \(F_5\): The marking occurs on the possessive (cf. 230) and thus NP-internally.

ad \(F_6\): The class marker mirrors its referent’s animacy and thus does not semantically contribute to the noun phrase (cf. 230).

ad \(F_7\): The semantic assignment criteria are of a straightforward and transparent nature (cf. Saxton 1983: 186). Therefore, a noun’s natural referential class does not have to be learned, but can be identified by the speaker, who is further bound to the semantic assignment rules of the system.
32.3.3 Mapping

![Graph showing mapping between formalisation and transparency values.]

Figure 67: Tohono O’odham.

Table 139: The formalisation and transparency value of Tohono O’odham.

<table>
<thead>
<tr>
<th>x-criteria</th>
<th>value</th>
<th>y-criteria</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>inventory size</td>
<td>$X_1$</td>
<td>semantic structure</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>host number</td>
<td>$X_2$</td>
<td>rule types</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>locus operandi</td>
<td>$X_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>obligatoriness</td>
<td>$X_4$</td>
<td>categorical independence</td>
<td>$Y_4$</td>
</tr>
<tr>
<td>boundedness</td>
<td>$X_5$</td>
<td>discreteness of markers</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>marking</td>
<td>$X_6$</td>
<td>redundancy</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>exhaustivity</td>
<td>$X_7$</td>
<td>flexibility</td>
<td>$Y_7$</td>
</tr>
</tbody>
</table>

| x-value             | 0.429 | y-value             | 0.786 |

**ad X₁:** The inventory size of two classes renders the Tohono O’odham system is a small one (cf. Saxton 1983: 186).

**ad X₂:** The markers attach to the possessive (cf. Saxton 1983: 186).

**ad X₃:** Since the markers occur only with the possessive (cf. Saxton 1983: 186), they operate NP-internally. The presence of the classified noun is not required (cf. 230).

**ad X₄:** The marking by the classifiers $sɔi$- “animate” and $iŋ$ “inanimate” is fully optional (Saxton 1983: 186).
Appendix

ad X₅: The markers occur in bound form (cf. 231).

(231)  
A  huan  gogs  sas-ga  
Juan  dog  CL:ANIMATE-POSS:ALI  
“Juan’s dog” (Tohono O’odham; cf. Langacker 1977: 91)

B  huan  uus  iin-ga  
Juan  tree  CL:INANIMATE-POSS:ALI  
“Juan’s tree” (Tohono O’odham; cf. Langacker 1977: 92)

ad X₆: The markers occur once per classificatory context (cf. 231).

ad X₇: The categories [animate] and [inanimate] apply to the full nominal lexicon.

ad Y₁: The class structure is transparent and straightforward: Nouns with animate referents constitute one class, nouns with inanimate referents the other class (cf. Saxton 1983: 186).

ad Y₂: The consulted sources and available data do not suggest another assignment than a semantic one.

ad Y₃: Due to the small class inventory and the transparent and bijective semantics of the individual classes, the speaker of Tohono O’odham needs to know only two semantic assignment rules to master the system.

ad Y₄: The markers convey classificatory information only (cf. 231).

ad Y₅: Each class has its own, distinct marker (cf. 231).

ad Y₆: The markers mirror the animacy of their nominal referent (cf. 231), and thus convey redundant semantics.

ad Y₇: The class semantics are mutually exclusive, as a nominal referent can be either animate or inanimate. Therefore, a classificatory flexibility is not given in Tohono O’odham.

32.4 Grammaticalisation

The possessive classifiers of Tohono O’odham are fully optional and thus may represent an incipient (or decaying) classifier system with a lower degree of formalisation (value: 0.429). The minimalistic and straightforward class inventory of two classes, which do not semantically contribute to their noun phrase, makes the system fairly transparent (value: 0.786).


AUDRING Jenny (in print): “Calibrating complexity: How complex is a gender system?” In Language Sciences (special issue). Manuscript provided by the author.


BARDDAL Jóhanna 2013: Constructional and Used-Based Approaches to Historical Linguistics. Constructional Approaches to Grammaticalization. Handout of the presentation at the LOT Winterschool 2013 (17.01.2013).


CORBETT Greville G. 2003: “ Agreement: the range of the phenomenon and the principles of the Surrey Database of Agreement.” In Transactions of the Philologial Society 101:2:


CSIRMAZ Aniko & DÉKÁNY Éva 2010a: “Hungarian is a classifier language.” Online version from https://www.academia.edu/4769418/Hungarian_is_a_classifier_language_with_Aniko_Csirmaz_(viewed_17.02.2014, 12:08).


ERHART Adolf 1993: *Die indogermanische Nominalflexion und ihre Genese*. Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck.


FISCHER Susann 2007: “Ian Roberts u. Anna Roussou, Syntactic change. A minimalist approach to grammaticalization.” In *Beiträge zur Geschichte der deutschen Sprache und
Literatur (PBB) 129:3, 469–476.


FRANJIEH Michael James 2016: “Possessive classifiers in North Ambrym: relational classifiers, possessed classifiers or noun classes?” Manuscript, University of Newcastle.


Bibliography


JACOBSON Steven A. 1979: *A Grammatical Sketch of Siberian Yupik Eskimo as Spoken on St. Lawrence Island, Alaska*. Alaska Native Language Center, Fairbanks.


LICHTENBERK Frantisek 1983a: “Relational classifiers.” In Lingua 60, 147–176.


LICHTENBERK Frantisek 2009: “Attributive possessive constructions in Oceanic.’In Mc- 
249–292.

LIGHTFOOT David W. 1979: Principles of Diachronic Syntax. Cambridge University Press, 
Cambridge.


LIU Shiru 1965: Weijin nanbei chao liangci yanjiu (“A study of the classifiers in the Weijin 
Nanbei period”). Zhonghua Shuju, Beijing.

LUCY John A. 1992: Grammatical categories and cognition. Cambridge University Press, 
Cambridge.

341.

considerations.” In Folia Linguistica 45:2, 435–464.

LYNCH John 1996: “Proto Oceanic Possessive Marking.” In LYNCH John & PAT Fa’afo: Oceanic 
Pacific Linguistics, Canberra. 93–110.

Cambridge.

MACE John 2003: Persian Grammar. For Reference and Revision. Routledge Curzon, Lon-
don/New York.


MARNITA Rina 1996: Classifiers in Minangkabau. MA Thesis. Australian National Univer-
sity, Canberra.


MARTINS Samuel A. 1994: Análise da morfosintaxe da língua Dâw (Maku-Kamã) e sua 
classificação tipológica. MA Thesis, Universidade Federal de Santa Catarina, Florianópolis, 
Brazil.

MATTEI-MULLER Marie-Claude 1979: “Cuestiones de categorización en las construcciones 
posesivas del panar.” In Antropológica 38, 3–14.


Bibliography


PASSEY Matthias Benjamin 2016: “Relational classifiers—a phantom category.” Manuscript, University of Amsterdam.


Bibliography


Index

A-Má-Lo, 266
Abau, 266
Acehnese, 181, 266
acquisition, 58, 65, 128, 142, 272, 300, 324, 336, 338, 341
Akatek, 182, 265
Aleut, 108, 109
Anêm, 266
analogy, 3, 34, 35, 52, 54, 60, 61, 63, 219
anaphora, 11, 16, 18, 25, 36, 37, 106, 116, 132, 155, 158, 166, 202, 206, 210, 229, 268, 282, 283, 287, 310, 315
Ancient Greek, 1, 2, 61
Anjom, 181, 265
Anindilyakwa, 266
anomaly, 3
Araki, 181, 266
argument marking, 35–38
Arrente, 266
Assamese, 181, 266
assigning agent, 5, 11, 18, 23, 64, 67, 72, 92, 94, 106, 116, 121, 124, 125, 127–129, 140, 141, 144, 147, 151, 154, 167, 172, 177, 216–222, 227, 230, 231, 235, 244, 246, 249, 256, 257, 276, 340
Ata, 266
Athabaskan, 260, 261
attribution, 11, 18, 92, 104, 105, 113, 116, 173, 340
Austrian, 210
automatisation, 5, 51, 64, 67, 89, 94, 187, 262, 277, 341
Awtuw, 161, 162
Babungo, 265
Bahwana, 182, 266
Baining, 265
Baniwa, 28, 31
Bantu, 100, 111
Basque, 108
Bavarian, 210
Bengali, 181, 266
bijectivity, 98, 271–273, 277, 278, 290, 292, 294, 323, 331
body part terms, 29, 30, 58, 109, 131, 132, 146, 150, 153, 218, 223, 335
Bukiyip, 44, 235, 266
Burmese, 13, 24, 79, 182, 266
Caddo, 39
canonical typology, 93, 119, 157, 262, 264, 267, 268
Carib, 40
Carrier, 261
Chalcatongo Mixtec, 59
Chamalal, 266
Chambri, 266
Chechen, 266
Chenapian, 266

Index
## Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>class terms, classificatory host, classifiers proper, cliticisation, closed systems, construction, context, core classifiers, core function, core property, count/mass distinction, creoles, Cubeo, conventionalisation, crèoles, Danes, Dâw, Danes, decay, decategorisation, deictic classifiers, degrammaticalisation, deponent pathways, diegueno, diversification, divergent layering, diversity value, documentation, discourse, discourse reference, Dongo-Ko, Duyari, Dutch, divergent layering, diversity value, Dutch, divergent layering, Dutch, Dutch, Dutch, Dutch, Dutch</td>
<td>621</td>
</tr>
</tbody>
</table>

**Note:** The term entries include page numbers where each term is mentioned or defined within the document.
Index


Ejagham, 266
Emmi, 181, 265
epicene, 12, 142, 212, 221, 222
erosion, 30, 56, 57, 59, 60, 70, 71, 74, 94, 96, 235, 269, 327, 337
Eskimo, 40, 41, 108, 109
Etruscan, 108
etymology, 26, 29, 31, 45, 47, 49, 59, 75, 98, 105, 135, 143, 154
exhaustivity, 118, 159, 163, 265, 271, 272, 282–286, 288, 289, 301, 311, 316, 322, 326–329
first language, 64, 272, 300, 336, 341
folk etymology, 153
formal liaison, 65, 74, 150
French, 52, 53, 56, 57, 59, 62, 63, 71, 81, 83, 84, 86, 211
Fula, 266
functional typology, 11, 19, 264
Gaagudju, 266
generics, 30, 47, 99
Gilyak/Nivkh, 108
Godié, 266
goodness, 56, 63, 65, 66, 71, 72, 74, 75, 85, 94, 95, 119, 121, 192, 252, 263, 325, 332
grammaticalisation cycles, 52, 57, 60
Greek, 83
grey area, 45, 78, 94, 95, 119, 120, 122, 125, 157, 158, 161, 163, 164,
Index

167, 241, 242, 286, 303, 304
Guaicuruan, 40
Gunbarlang, 266
Gunwinggu, 87
Gurr-Goni, 266
Hawaiian, 31
hadrology, 125, 177
hinge systems, 2, 73, 124, 144, 160, 302
Hupa, 260, 261
Hurrian, 108
idiosyncracy, 217, 230, 247, 252, 257, 260, 272, 278, 299, 319, 323, 324
inalienability, 31, 34, 109, 132, 135, 136, 187, 252
incorporation, 27, 30, 38, 39, 78, 165, 166, 264
indicativity, 11
individuation, 5, 11, 14, 15, 18, 23, 26, 30, 44, 46, 104, 105, 113, 114, 116, 127, 143, 184, 186, 213, 231, 285, 293, 340
information structure, 65
inheritance, 3, 83, 136, 273, 324, 326, 329
injunctivity, 98
Italian, 53, 62, 63, 83, 84, 334
Itonama, 265
Iwam, 266
Japanese, 182, 184, 256, 266
Jarawa, 265
Kana, 78
Kanjokal Mayan, 182, 265
kataphora, 16, 36
Kayardild, 265
Ket, 108, 110–113, 120, 121, 123, 126, 139, 158, 164, 170, 175, 178, 180, 181, 187, 193, 194, 203, 205, 211, 216, 223, 228, 230, 232,
Index

240, 242, 243, 253, 265, 289, 296, 298, 300, 304, 348
Khinalug, 266
Kílívila, 21, 26, 32–34, 135, 181, 256, 265, 266, 317
kinship terms, 30, 109, 193, 219, 262
kinship terms, 132, 165, 166, 220, 261
Khíowa, 48, 110, 112, 113, 120–122, 126, 139–143, 158, 164, 165, 169, 170, 266
Koko Bera, 265
Kugu-Muminh, 265
Kuseian, 265
Kuuk Thaayorre, 266
Kwakwala, 261
Kwaza, 266
Kókó Bera, 265
Kugu-Muminh, 265
Kuseian, 265
Kuuk Thaayorre, 266
Kwakwala, 261
Kwazá, 266
Lakondê, 266
language type, 14, 42, 101, 128, 182, 246, 247, 339
lateral conversion, 61
Latin, 2, 19, 44, 47, 52, 56, 60–62, 81–84, 86, 207, 334
Lelemi, 266
Lenakel, 181, 266
lexical entry, 48, 183, 200, 246, 331
lexical items, 12–14, 46, 51, 54, 55, 58, 64, 69, 75–77, 79, 80, 82, 85, 86, 128, 133, 136, 141–143, 183
lexical restriction, 35, 38
linguistic area, 27, 35, 40, 41, 182, 261
linguistic economy, 11, 64, 64, 184, 186, 217, 219, 229, 278, 317, 318, 338, 341
locative classifiers, 30–41
Lokono, 40
Lolovoli, 181, 266
Lonwolwol, 181, 265
Máíhìkì, 11, 12
Maimandê, 266
Mal, 21, 26, 87, 181, 266
Mam, 182, 266
Manam, 36, 37, 135, 181, 265
Mandan, 40, 41
Marind, 266
Masri, 110, 112, 113, 120, 121, 123, 126, 139, 158, 164, 170, 175, 178, 180, 181, 187, 193, 194, 203, 205, 210, 211, 216, 223, 228, 230, 232, 240, 242, 243, 252, 253, 265, 272, 289, 296, 298, 300, 304, 348
Mawng, 132, 133, 247
Mawung, 265
Mayali, 99, 209
Mba-Ne, 266
measure nouns, 18, 26, 30, 72, 74, 181–184
mensural classifiers, 26, 69, 85
Meroitic, 108
Mian, 266
Index

Middle English, 325, 327–329, 331–333, 337
Middle High German, 329, 332
Minangkabau, 55, 68, 74, 77, 83
minor classifier types, 9, 27, 31, 39
Mohawk, 38, 39
Moro, 111–113, 120, 121, 123, 126, 139, 158, 160, 164, 170, 175, 178, 180, 182, 187, 193, 194, 196, 199, 204, 205, 211, 216, 223, 228, 230, 232, 240, 242, 243, 253, 266, 289, 296, 300, 304, 348
morphology, 1, 40, 41, 43, 44, 47, 52, 55, 58, 61, 71, 74–76, 78, 81, 93, 94, 96, 121, 159, 160, 217, 235, 243, 273, 293, 330, 337
Mota, 135
Motuna, 35, 265, 266
Mparutwe Arrerute, 266
multiple classification, 164, 169, 194, 203, 205, 228, 236, 249, 255, 258, 270, 287
multiple classifiers, 22, 27, 35, 36, 112, 114, 134, 138, 159, 179, 184, 199, 233, 236, 259, 267, 270
Murrinhpata, 181, 266
Na-Dene, 108, 110, 260
Nahai/Nihali, 108, 109
naturalism, 3, 4
Ndali, 265
Ndunga-Le, 265
Ngandi, 77
Ngangumartu, 265
Index


nominalisation, 29, 83, 132, 203, 208, 209, 220, 221, 237


noun class s, 2–5, 9, 10, 12, 14–18, 20, 21, 26, 42, 43, 47, 68, 69, 85, 88, 100, 137, 139, 240

noun classifiers, 22, 26, 27, 29, 59, 77, 78, 82, 87, 98, 104, 114, 167, 177, 196, 200, 218


numeral classifiers, 5, 14, 22, 26, 28, 29, 34, 55, 58, 74, 75, 78, 87, 89, 102, 104, 109, 111, 114, 122, 127, 128, 134, 159–161, 172, 179, 181–184, 186, 189, 224, 261, 269

Nynorsk, 330


Occitan, 83, 84

Oceanic, 28, 34, 110, 136

Old English, 63, 326, 327, 329–333, 337

Old French, 56, 63

Old High German, 101, 329, 332

Olgolo, 97, 98

Omani, 266

omission, 57, 59, 69, 77, 98, 109, 206, 287, 299

opacity, 84, 273, 274, 277, 278, 291, 293, 294, 300, 317, 321, 323, 324, 329–331


order of deviation, 163, 164, 167, 241, 242, 254, 256, 302–305

Paamese, 181, 266

Pajtjamalh, 181, 265

Panare, 22, 27, 87, 266

Paumari, 266

Persian, 29, 74, 75, 161, 162, 181, 189–192, 265


pidgins, 107, 109, 110

Pilagā, 39, 266

Polish, 80, 83, 85, 89

Polynesia, 219

Ponapean, 78

Ponca, 41

portmanteau morphemes, 43–45, 166, 203, 274, 275, 278

Portuguese, 83, 84


pragmatics, 20, 52, 55, 64, 65, 77, 96, 214

predicativity, 11

Index

proficiency, 324, 333, 336, 341
Proto-Australian, 77, 99, 100
Proto-Germanic, 101, 326, 329–333, 341
Proto-Indo-European, 94, 100, 101, 142, 326, 329, 330, 332, 333
Proto-Oceanic, 136
Puluwat, 29, 181
quantification, 14, 20, 26, 72, 111, 122, 160, 161, 170, 190, 260, 261, 265, 269, 272, 283, 309–311
quantifiers, 22, 26, 28, 35, 74, 171, 179, 188, 239, 261, 267, 282, 283
reassignment, 11, 12, 17, 25, 46, 57, 59, 84, 196, 210, 211, 220, 222, 225, 226, 292, 320, 340, 341
reconstruction, 94, 97–101, 136, 322, 324
reference tracking, 5, 10, 11, 16, 18, 19, 25, 36, 46, 48, 92, 106, 116, 132, 143, 155, 173, 206, 248, 340
referentiality, 10, 12, 38, 55, 76, 80, 81, 92, 99, 105, 106, 134, 143, 145, 167, 169, 201–204, 208, 217, 233, 271, 273, 277, 287
remotivation, 71, 72, 102, 140, 247, 249, 252, 253, 257, 260, 280, 297, 319, 322–325, 329, 331, 333
repeaters, 23, 30, 39, 41, 148
research history, 2–5
Rikvani Andi, 265
Romance, 52, 53, 60–62, 334
Romanian, 83, 84
Rotokas, 266
Russian, 83
Sabanē, 266
Sakao, 181, 265
sampling, 106, 107
Sarawak, 181, 266
scope, 11, 19, 20, 31, 71, 73, 81, 82, 161, 166, 172, 187, 236, 247, 257, 259, 261, 262, 272, 302, 313, 321, 335, 338
second language, 64, 142, 324, 336, 341
Sesotho, 266
Siberian, 109
Siouan, 40, 41
social status, 15, 212, 317
sortal classifiers, 26
Spanish, 83, 84, 109
specificity, 11, 64, 96, 100, 183, 184, 235, 258, 274, 276, 277, 335
Squamish, 266
structuralism, 4, 5, 55, 58, 61, 65, 77, 107, 247
Swedish, 85, 330
synchrony, 4, 47, 60, 65, 75, 77, 101, 106, 119, 133, 143, 160, 166, 177, 236, 237, 243, 244, 246, 252, 254, 261, 262, 273, 277, 305, 307, 323, 326, 328, 333
Index


Tape, 181, 266
Tariana, 22
Tashkent Uzbek, 182, 265
Telugu, 265
Terêna, 38
Thai, 23, 39, 182, 266
Tibetan, 15, 108, 111, 182, 265


transparency threshold, 299, 316, 317

Tsafiki, 265

Tsowa-Tush, 266
Tswana, 266
Tzeltal, 266

unidirectionality, 52, 54, 58, 59, 63, 64, 281

Vanuatu, 32, 135
Vedie Sanskrit, 61, 100
verb classifiers, 4, 9, 29, 31, 35, 58
verbal classifiers, 35, 36, 38, 253, 260
Vietnamese, 182, 266
visibility, 195, 233, 299, 300, 324, 333, 334, 341

Wambaya, 266
Wardaman, 77
Warekena, 266
Wati, 265
West Germanic, 308, 325–327, 329, 332
Wik-Mungknh, 265
Wogamusin, 265
Wolof, 266

Xóõ, 266

Yagua, 182, 266
Yankunytjatjara, 99
Yanomami, 78
Yanyula, 266
Yanyuwa, 77
Yidiny, 1, 182, 266
Yimas, 266
Yir-Yoront, 266
Yucatec Maya, 13, 24, 146
Yuchi, 40
Yucuna, 82, 181, 266
Yupik, 109
Yurok, 261

Zamboangueño, 109, 110
Zande, 98, 101, 266
Zulu, 87
In this thesis, I study the typological structure and pathways of developmental change of nominal classification systems. Nominal classification is a wide-spread cross-linguistic phenomenon, which scholars investigate since almost 2500 years (cf. chapter 1 for an overview on the research history). In a nutshell, nominal classification systems are a grammatical device that categorises nouns and their referents with respect to their semantic characteristics, stem-forming morphemes, or phonological shape.

Chapter 2 presents the two basic types of nominal classification systems, namely classifier systems, which are frequently found in Pacific, East Asian, and South Asian languages, and concordial class systems, which are commonly found e.g. in Indo-European and Bantu languages. The difference between these two types is that classifier systems are fairly interactive in allowing the speaker to assign a variety of classifiers to a given noun or referent, semantically productive in that the inserted classifier contributes to the meaning of its host noun phrase, and morphosyntactically independent in that classifiers are typically clitics or unbound elements, (cf. 1, where all these conditions apply).

(1) Nùng (cf. Saul and Wilson 1980: 14)

\[ \text{lẽo hảhn tú má mạ lụhc} \]
then kill two CL:ANIMATE dog.child

"Then kill two puppies."

Moreover, classifiers are a delimited phenomenon in that they occur once and overtly within the borders of the noun phrase, and in that they are bound to given morphosyntactic contexts: Mandarin for instance does not allow to count objects by the mere combination of a quantifier and a noun (one snake, two books, three bullets,...), but requires a classifier to be inserted in these contexts (cf. 2). Except for quantificational contexts, Mandarin allows demonstratives
and sometimes nouns to combine with a classifier, while other elements are not able to do so.

(2) Mandarin (cf. Zhang 2007: 45f.)

A yi tiao she
one CL:LONGISH snake
“one snake”

B liang ben shu
two CL:VOLUME book
“two books”

C yi ke zidan
three CL:SMALLISH bullet
“three bullets”

There are four main classifier types, which are labeled according to their morphosyntactic operative context: Possessive classifiers, which occur in possessive constructions, numeral classifiers, which occur in quantificational constructions, noun classifiers, which occur in combination with the bare noun, and multiple classifiers, which insert their classifiers in more than a single kind of construction. This study does not consider the types of so-called ‘verb classifiers’, ‘relational classifiers’, deictic classifiers, and locative classifiers.

Concordial class systems in contrast are fairly automatised in that nouns typically belong to a fixed, single class; they are semantically unproductive, since the class of the noun in most cases does not manipulate the meaning of its noun phrase, and they are morphosynthetically dependent in that a noun’s class is marked by means of agreement affixes on a language-specific range of nominal satellites (cf. 3, where all these conditions apply).

(3) Italian

La ART:F.SG mi-a POSS1SG-F.SG ragazza-POSS1SG-F.SG è scaltr-a S G and educated-F.SG
clever-F.SG and educated-F.SG
“My girl is clever and educated.”

Class agreement marking on nominal satellites is obligatory, which means that class markers may occur multiply in the same context. Concordial class marking is thus not limited to a single occurrence per classificatory context, like it is the case for classifiers.

Classifiers and concordial class systems constitute the respective extreme points of a continuum of nominal classification. The gradations of the continuum are created by grammaticalisation, a concept of gradual linguistic change that involves the loss of morphosyntactic independence and categorial characteristics, semantic bleaching, and morphologisation. The properties of classifier systems correspond to a low degree of grammaticalisation, and the properties of concordial class systems correspond to a high degree of grammaticalisation.
A common hypothesis states that a classifier system may undergo grammaticalisation and turn into a concordial class system. I refer to this process as the “grammaticalisation hypothesis”. This process is not observed or documented in real world languages, and, therefore, the present study examines in detail and from a typological perspective how likely this phenomenon is to occur. For this purpose, I study a sample of forty languages, which are projected onto a grammaticalisation scale. If the grammaticalisation hypothesis holds, an uninterrupted increase of grammaticalisation should be observed in the sample.

Chapter 3 presents a theoretical background of grammaticalisation and applies it to the typology of nominal classification devices. As a result, seven properties, which apply to both classifier systems and concordial class systems and are widely used in descriptive analyses of nominal classification, are isolated as a measure for a system’s degree of grammaticalisation. These properties are listed in figure 1.

Figure 1: The basic properties of systems of nominal classification.

Chapter 4 renders the aim and focus of this study and presents its methodology. It puts forward research questions, working hypotheses and predictions and discusses the feasibility of studying a diachronic phenomenon, the grammaticalisation hypothesis, by means of synchronic data, the language sample. It also presents the sampling method and sample composition and provides arguments for the inclusion or exclusion of candidate systems for the sample.

Chapter 5 studies grammaticalisation in the sample from the perspective of the systemic types of nominal classification. When the sample languages are ordered according to an increase in their degree of grammaticalisation, classifier-like systems group together at the less grammaticalised end of the continuum and concordial class-like systems group together at the more grammaticalised end of the spectrum. The full sample displays a recognisable grammaticalisation cline, which is, however, blurred by two coherent groups of deviations among the classifier-like systems. These deviations mostly occur in systems that show a conventionalisation of classifier-noun pairs, while they maintain a low degree of grammaticalisation in their other properties. Moreover, only the Korean sys-
system corresponds to the definition of an ideal classifier system, whereas there are six systems that fully correspond to the definition of an ideal concordial class system. The concordial class-like systems in the sample generally correspond better to the grammaticalisation cline, too.

Besides classifier systems and concordial class systems, two more phenomena are investigated in chapter 5. First, ‘hybrid systems’ are discussed more in detail. These are systems that seemingly employ defining properties of classifier systems (e.g. a large class inventory, a productive semantic contribution to the noun phrase, a speaker-sided and flexible assignment) and concordial class systems (multiple marking, NP-internal and -external marking) at the same time. Therefore, they elude a straightforward identification as an either classifier-like or concordial class-like system. Furthermore, they have been argued to represent the most likely ‘hinge’ between classifier and concordial class systems and thus might represent the ‘missing link’ in the grammaticalisation hypothesis. The sample features four of these systems, the one of Nasioi, Bora, Mundurukú, and Ngan’gityemerri; these show a coherent grouping towards the very lowly grammaticalised end of the spectrum when ordered with respect to degree of grammaticalisation of their properties. This observation does not suggest an analysis of these systems as ‘hinges’ between classifier systems and concordial class systems.

Second, the ‘grey area’ between classifier and concordial class systems is investigated. The grey area is not further specified and can be derived from the conceptuation of the continuum of nominal classification as defined by its two most extreme poles, classifiers and concordial classes. The grey area consists of intermediate systems and is located at the intersection of classifiers and concordial classes. Therefore, it is of great interest for the study of the grammaticalisation hypothesis, because it represents the area in the continuum where a system may shift from a classifier-like type to a concordial class-like type. I put forward three orders of deviation, which are defined on basis of the observations from the sample data and which represent helpful gradations of the grey area: The whole sample shows three properties that are highly consistent with the ideal definition of the respective system. These are the obligatoriness of classification, the potential to feature a class marker multiply or not, and the potential to feature a class marker outside of the borders of the noun phrase or not. A system that shows a deviation in one or more of these ‘core properties’ is a first order deviation; first order deviations are the narrowest form of the grey area. The supposed ‘hybrid systems’ are the only sample-internal systems that show both classifier-like and concordial class-like properties here, and thus are the first order deviations in the sample. Systems that show a consistent classifier-like or concordial class-like setup of properties, but deviate in their flexibility of classification or their assignment from this setup, are second order deviations. First and second order deviations together constitute the wider notion of the grey area. Finally, there are systems that show a consistent classifier-like or concordial class-like setup of properties, but employ individual criteria that are partly classifier-like and concordial class-like
at the same time (e.g. if both speaker- and system-sided assignment occur). These are third order deviations, which together with second and first order deviations constitute the widest available notion of the grey area.

Chapter 6 studies grammaticalisation in the sample from the perspective of the sample-internal distribution of the individual properties. Most of the properties correspond well to the expected grammaticalisation cline; this includes the size of the class inventory, the obligatoriness of classification, the potential to multiply mark class membership in the same context, and the potential to feature class marking NP-internally and -externally. The remaining two properties, the flexibility of classification and the either speaker- or system-sided class assignment, show a less coherent distribution. The classifier-like systems of the sample show a preference for a flexible and speaker-sided classification, and the concordial class-like systems of the sample show a preference for a rigid and system-sided assignment. However, both systemic types feature a number of deviations from this preference, and thus blur the grammaticalisation cline in the sample. The property of the semantic contribution of a system proves to be highly varied throughout the sample. Nominal classification systems are reported to employ various functions in the domain of semantics and discourse; to capture this rich functionality in a rather vague criterion such as “semantic contribution” thus is too wide a notion to represent a useful measure for the grammaticalisation of nominal classification systems. The property is therefore dropped as a measure in the further study.

In chapter 6, I provide an alternative analysis for the seemingly multiple marking in the supposed ‘hybrid systems’ as an actual instance of adjacent multiple classification. This means that a class value in those systems is not copied from the noun and expressed in form of agreement on elements surrounding the noun, but that each element carrying a class marker represents an individual context of classification. This allows the reanalysis of the ‘hybrid systems’ as multiple classifier-like systems. Furthermore, I demonstrate that the core properties show strong correlations in the sample—systems with optional classification do not mark class multiply in the same context, and restrict their locus operandi to the borders of the noun phrase. Nasioi, Bora, Mundurukù, and Ngan‘gityemerri are the only exception in that they allow NP-external class marking. Furthermore, the ‘non-core properties’ of classificatory flexibility and assignment show correlating traits: The more flexible the classification of a system is, the more it is likely to involve the speaker as assigning agent, and vice versa.

The analysis of the sample data in chapter 5 and 6 suggests a modification of the notion of grammaticalisation in nominal classification devices: The sample systems show two main lines of development, which may occur independent of each other. On the one hand, there are systems that show a grammaticalisation of their formal means (“formalisation”); they are characterised by an expansion of the host of the class marker and the locus operandi, the loss of morphosyntactic independence of their markers and their phonological erosion, and a decreasing class inventory. On the other hand, there is a larger number
of classifier-like systems that does not alter its means of formal expression, but shows an increase in conventionalised and inflexible class-noun pairings and a decrease in the semantic contribution to the noun phrase ("conventionalisation"). The concepts of formalisation and conventionalisation are taken up in chapters 7 and 8 and applied to a refined mapping of nominal classification systems. Other strategies of development are less frequent, but also present in the sample. Systems may expand their functionality beyond the scope of nominal classification; Great Andamanese for instance maintains a nominal classification function when its class markers attach to nouns and adjectives, but serves the function of verbal classification (i.e. the semantic specification of the concept that is expressed by the predicate) when its markers attach to adverbs and verbs. Finally, a system may be given up; in this case, formal remains may indicate the earlier presence of a nominal classification system. I have argued that this is the case for Tlingit and Chukchi, where modifiers of human referent nouns carry an overt marker, while there is no recognisable nominal classification system or function.

Chapter 7 introduces a mapping instrument for nominal classification devices, which is based on both earlier research and the observations from the language sample. The sample study shows that there are systems at all stages of formalisation, and that systems at any given point of formalisation display differences in their semantic transparency and productivity. Therefore, the mapping account distinguishes the dimension of formalisation from the dimension of transparency. These dimensions are defined on the basis of two packages of seven properties. Each property possesses two or more measure points, which can be projected on a numerical scale of grammaticalisation. The mean value of the seven properties that render the degree of formalisation reveals the formalisation value of a system, and the one of the seven properties that render the degree of conventionalisation reveal the transparency value of a system. These two values allow a straightforward visual presentation of a system in a scatter plot, where systems can be directly compared both synchronically or at different stages of their history. The application of the mapping to the language sample results in a clear grouping of systems, as figure 2 demonstrates.
Figure 2: Comparative mapping of the sample systems: Distribution of systemic types (+: classifier system; •: concordial class system).

Figure 2 shows that systems with a formalisation value up to 0.73 are those systems that were earlier analysed as classifier-like; they have a transparency value of 0.7 or more. They are highly varied with respect to their degree of formalisation. Systems with a formalisation value higher than 0.89 are those that were earlier analysed as concordial class-like and show little variation in their degree of formalisation, but a large variation in their transparency value, which roughly spans from 0.2 to 0.93. The classifier-like systems are separated from the concordial class-like systems by a gap that covers 16 per cent of the formalisation scale. The only exception is the system of Tainae, which is located at the top of the highly formalised systems, yet was analysed as a classifier-like system in chapter 6. The supposed hinge systems of Nasiol, Bora, Mundurukú, and Ngan’gityemerri are at an actual hinge position in the two-dimensional mapping of the sample; they occur among the the more grammaticalised classifier-like systems. The distribution of the sample systems in figure 2 has an important implication for the grammaticalisation hypothesis: If only systems with a high formalisation value are able to lose their transparency to a considerable degree, a ‘unitary’ grammaticalisation approach, which assumes formalisation and conventionalisation to occur at the same time, cannot be maintained. Instead, systems of nominal classification seem to undergo ‘split’ grammaticalisation, which means that formalisation precedes a potential conventionalisation process.

Chapter 8 investigates the diachrony of the systems of Mandarin, Dyirbal, and Western Germanic. The respective systems are mapped at different stages of their development and at different linguistic registers. This diachronic and sociolinguistic mapping allows us to observe the processes of formalisation, con-
ventionalisation, and systemic semantic remotivation. The case studies suggest that a considerable loss in transparency is related to the visibility of a system in discourse, i.e. that only a concordial class-like system, which allows for multiple classificatory marking in the same context, can refrain from entertaining an accessible and productive semanticity without entering the risk of decaying or undergoing semantic remotivation. Beyond this, the linguistic and cultural proficiency of the younger generation of speakers plays an important role in the reanalysis of the system of Dyirbal. Furthermore, the high prestige of English as compared to Dyirbal contributes not only in the remodelling of the nominal classification system, but also influences other parts of the grammar.

In conclusion, this study has shown that grammaticalisation takes a ‘split’ form in nominal classification: The means of formal expression of nominal classification systems grammaticalise independently of their semantic and interactive components. An account for nominal classification which distinguishes these two developmental vehicles yields much more detailed and expressive results than a ‘unitary’ grammaticalisation approach. Moreover, I have shown that the grammaticalisation hypothesis is a scenario which is possible, but highly unlikely to ever occur. Besides the lack of a ‘missing link’ between classifier-like and concordial class-like systems in the sample, the grammatical changes that a system requires to develop concordial class-like affixed agreement, for instance, are not available in many Asian classifier languages of an isolating type. A language that grammaticalises its classifiers into a concordial class-like system requires in-depth grammatical shifts that span far beyond nominal classification; for instance, many classifier languages employ their markers as semantic specifiers of otherwise underspecified lexical nouns. If the classificatory markers shift from their specifying function into a function as ‘mere’ agreement markers, the concept of nouns in these languages had to reform, and the lexical content of nouns had to become more specified. While a scenario with a number of in-depth grammatical changes that are required by the grammaticalisation hypothesis is not impossible, it is improbable.
Samenvatting: De typologie en diachronie van nominale classificatie

In dit proefschrift bestudeer ik de typologische structuur en ontwikkeling van nominale classificatiesystemen. Nominale classificatie is een wijdverspreid fenomeen dat al ruim 2500 jaar bestudeerd wordt (zie hoofdstuk 1 voor een overzicht van de onderzoeksgeschiedenis hiervan). Nominale classificatiesystemen zijn het grammaticale verschijnsel waarbij de zelfstandige naamwoorden van een taal, of hun referenten, op basis van hun semantische eigenschappen, stamvormende morfemen of fonologische vorm worden gecategoriseerd.

In hoofdstuk 2 worden de twee types van nominale classificatiesystemen voorgesteld, namelijk classificeerdersystemen (die vaak in Oost-Azië en Zuidoost-Azië voorkomen) en concordantieklassesystemen (die bijvoorbeeld in de meeste Indo-Germaanse en Bantoetalen voorkomen). Het verschil tussen deze twee types is dat classificeerdersystemen vrij interactief zijn, omdat de spreker een classificeerder aan een naamwoord (of de referent hiervan) toewijst. Ze zijn verder semantisch productief doordat de classificeerder betekenis aan het naamwoordelijk zinsdeel toevoegt. Ze zijn tevens morfosyntactisch onafhankelijk omdat de klassemarkeerder in de vorm van ofwel een clitic of van een vrijstaand element optreedt (zie 1, waar al deze condities van toepassing zijn).

(1) Nùng (zie Saul and Wilson 1980: 14)

lẹọ báhn tá má luhc
dan dood twee CL.LEVEND hond.kind
“Dood dan twee welpjes.”

Bovendien zijn classificeerders beperkt in hun optreden; in elke context van classificatie verschijnen ze eenmalig en slechts binnen de grenzen van het naamwoordelijk zinsdeel. Verder zijn classificeerders aan bepaalde morfosyntactische constructies gebonden: in het Mandarijn kan men bijvoorbeeld niet zomaar tellen door middel van de samenvoeging van een telwoord en een naamwoord
Samenvatting in het Nederlands

(één slang, twee boeken, drie patronen...), maar is de toevoeging van een classificeerder vereist (zie 2). Buiten kwantificerende contexten treden classificeerders in het Mandarijn alleen op in combinatie met aanwijzende voornaamwoorden en soms ook met zelfstandige naamwoorden.

(2) Mandarijn (zie Zhang 2007: 45f.)

A yi tiao she
een ČL:LANG.OBJECT slang
“een slang”

B liang ben shu
twee ČL:VOLUME boek
“two books”

C yi ke zidan
drie ČL:KLEIN.OBJECT patroon
“drie patronen”

Er zijn vier hoofdtype: Bonom-classificeerders, deze zijn vernoemd naar de morfosyntactische context waarin ze optreden: possessieve classificeerders komen voor in bezittelijke constructies voor, numerale classificeerders in kwantificerende constructies, naamwoordelijke classificeerders treden op naast het zelfstandige naamwoord, en multipele classificeerders in meer dan één soort constructie. Deze studie houdt geen rekening met zogenaamde werkwoordclassificeerders, relatiefclassificeerders, aanwijzende classificeerders en locatieve classificeerders.

Concordantieklassesystemen werken daarentegen vrij mechanisch: naamwoorden zijn een onveranderlijk onderdeel van één klasse, en de klassen oefenen geen semantische invloed uit op het naamwoordelijk zinsdeel. De klassemarkeerders zijn morfosyntactisch afhankelijk en treden op in de vorm van congruerende affixe met een taalspecifieke selectie van naamwoordelijke satellieten (zie 3, waarop al deze condities van toepassing zijn).

(3) Italian

La mi-a ragazz-à è scatlr-a e colt-a.
ART:F.SG POSSISG-F.SG meisje-F.SG is slim-F.SG en educated-F.SG
“Mijn vriendin is slim en gecultiveerd.”

De voor een taal vastgelegde selectie van congruerende naamwoordelijke satellieten congrueert verplicht met het zelfstandige naamwoord. Dit betekent dat klassenmarkeerders meermaals binnen dezelfde constructie kunnen optreden. In tegenstelling tot classificeerdersystemen zijn concordantieklassesystemen dus niet beperkt in het aantal tegelijkertijd optredende markeerders.

Classificeerders en concordantieklassesystemen bepalen de eindpunten van een nominaal classificatiecontinuüm. Dit continuüm wordt door grammaticalisatie bij elkaar gehouden; grammaticalisatie is een concept van graduele taalverandering dat bestuwt uit het verlies van morfosyntactische autonomie en categorische eigenschappen, semantische erosie en morfologisatie. De eigenschappen
van classificeerdersystemen komen overeen met een lage mate van grammaticalisatie, en de eigenschappen van concordantieklassessystemen komen overeen met een hoge mate van grammaticalisatie.

Er wordt aangenomen dat classificeerders grammaticalisatie kunnen ondergaan en uiteindelijk in concordantieklassessystemen kunnen veranderen. Ik noem deze veronderstelling de “grammaticalisatiehypothese”. Deze is niet gedocumenteerd; daarom onderzoekt deze studie gedetailleerd en vanuit een typologisch perspectief hoe waarschijnlijk het is dat dit fenomeen soms plaatsvindt. Daartoe bestudeer ik een sample van veertig talen, dat ik op een grammaticalisatieschaal plaats. Als de grammaticalisatiehypothese klopt, toont het sample een continue toename van grammaticalisatie.

Hoofdstuk 3 biedt een theoretische achtergrond van grammaticalisatie en past deze toe op de typologie van nominale classificatiesystemen. Het resultaat is een bundel van zeven eigenschappen, die zowel in classificeerdersystemen als in concordantieklassessystemen terug te vinden zijn en die vaak in beschrijvende analyses van nominale classificatie gebruikt worden. Door middel van deze bundel wordt de gradatie van grammaticalisatie van een nominaal classificatiesysteem gemeten. De eigenschappen staan opgesomd in afbeelding 1.

Afbeelding 1: De basiseigenschappen van nominale classificatiesystemen.

Hoofdstuk 4 beschrijft de doelen en de focus van deze studie en legt het theoretisch kader uit. Hierin worden de onderzoeksvragen, hypothesen en voorverduidelijkingen opgesomd en de haalbaarheid besproken van het bestuderen van een diachroon fenomeen (de grammaticalisatiehypothese) aan de hand van synchrone data (het sample). In hoofdstuk 4 worden tevens de samenstelling van het sample en de redenen voor de opname of verwerping van specifieke talen besproken.

In hoofdstuk 5 wordt de grammaticalisatie in het sample vanuit het perspectief van de twee types van nominale classificatiesystemen bestudeerd. Wanneer de samplletalen op volgorde van hun gegrammaticaliseerde eigenschappen geordend worden, wordt er een duidelijke verdeling zichtbaar van classificeerderachtige systemen richting het minder gegrammaticaliseerde einde van het spectrum, en een verdeling van concordantieklassseachtige systemen richting het
Samenvatting in het Nederlands

meer gegrammaticaliseerde einde van het spectrum. Het gehele sample toont een zichtbaar toenemend grammaticalisatiepatroon; dit wordt slechts door twee coherente groepen van afwijkende systemen verstoord. Deze afwijkingen treden vooral op in systemen, waarin zich een conventionalisering van paren van clas-

ificeerders en naamwoorden voordoet, terwijl de overige eigenschappen een lage graad van grammaticalisatie vertonen. Bovendien is er maar één systeem, namelijk het Koreaanse clasificeerdersysteem, dat aan de voorwaarden van een ideaal classificeerdersysteem voldoet, terwijl er zes systemen in het sample aanwezig zijn die volledig met de definitie van een ideaal concordantieklassesys-

teem overeenkomen. De concordantieklassesystemen in het sample zijn in het algemeen in betere overeenstemming met het grammaticalisatiepatroon dan de clas-

ificeerdersystemen in het sample.

Naast classificeerdersystemen en concordantieklassesystemen worden nog twee fenomenen in hoofdstuk 5 besproken: hybride systemen en het grijze gebied tussen classificeerdersystemen en concordantieklassesystemen. Hybride systemen hebben eigenschappen die bepalend zijn voor classificeerdersystemen (bijvoorbeeld een grote klasseninventaris, een productieve semantische bijdrage aan het naamwoordelijk zinsdeel, een flexibele klasstoewijzing door de spreker) en concordantieklassesystemen (meervoudige markering binnen en buiten de grenzen van het naamwoordelijk zinsdeel). Hierdoor onttrekken ze zich aan een makkelijke identificatie als classificeerderachtig systeem of concor-

dantieklassachtig systeem. Bovendien worden deze systemen beschouwd als het meest waarschijnlijke ‘scharnier’ tussen classificeersystemen en concor-

dantieklassesystemen en zouden ze de missing link van de grammaticalisasiehy-

pthese kunnen zijn. Het sample bevat vier hybride systemen, die van Nasioi, Bora, Mundurukú, en Ngan’gityemerri. Deze vertonen een coherente groep aan het weinig gegrammaticaliseerde einde van het spectrum, als men het sample ordent naar mate van grammaticalisatie. Deze observatie is in tegenspraak met de aanname dat deze systemen ‘scharnieren’ tussen classificeerdersystemen en concordantieklassesystemen zijn.

Het grijze gebied tussen classificeerdersystemen en concordantieklassesystemen is tot dusver niet gespecificeerd. Het kan afgeleid worden van de conceptu-

atie van nominale classificatie die door ideale classificeerdersystemen en concor-

dantieklassesystemen als respectievelijk eindpunt van het continuüm bepaald wordt. Het grijze gebied bestaat uit intermediaire systemen en ligt bij het overgangspunt tussen classificeerdersystemen en concordantieklassesystemen. Daarom is het van groot belang voor het bestuderen van de grammaticalisasiehypothese, omdat dit het gebied binnen het continuüm vertegenwoordigt waarin een systeem van een classificeerderachtig type verandert in een concor-

dantieklassachtig type. Ik stel in dit proefschrift drie gradaties van afwijki-

ng van een ideaal nominaal classificatiesysteem voor: het gehele sample toont drie eigenschappen, die sterk overeenkomen met de definitie van het ideale type van het respectievelijke systeem. Deze eigenschappen zijn de verplicht-

ing van classificatie, het potentieel om meerdere klassenmarkeerders binnen dezelfde context te plaatsen, en het potentieel om klassenmarkeerders buiten
het naamwoordelijke zinsdeel te plaatsen. Een systeem dat afwijkungen binnen één of meer van deze ‘kerneigenschappen’ vertoont, is een eerstegraads afwijk-
ing; eerstegraads afwijkungen vormen de meest beperkte variant van het grijze gebied. De zogenoemden hybride systemen zijn de enige binnen het sample die zowel classificeerderachtige als concordantieklasseachtige kerneigenschappen vertonen, en zijn dus de eerstegraads afwijkungen van het sample. Tweedegraads afwijkungen zijn systemen die consistent classificeerderachtig of concordantieklasseachtig zijn, maar afwijken in hun flexibiliteit met betrekking tot classificatie of klassentoewijzing. Eerstegraads afwijkungen en tweedegraads afwijkungen vormen samen de ruimere variant van het grijze gebied. Tenslotte zijn er systemen die volledig classificeerderachtig of concordantieklasseachtig zijn, maar individuele eigenschappen vertonen die deels classificeerderachtig en deels concordantieklasseachtig zijn (bijvoorbeeld als klassentoewijzing door zowel het systeem als de spreker optreedt). Dit zijn derdegraads afwijkingen; samen met eerstegraads afwijkungen en tweedegraads afwijkungen vormen ze de ruimste variant van het grijze gebied.

Hoofdstuk 6 bestudeert grammaticalisatie in het sample vanuit het perspectief van de sample-interne verdeling van de individuele eigenschappen. De meeste eigenschappen voldoen aan het verwachte stijgende grammaticalisatiepatroon; dit geldt voor de grootte van de klasseninventaris, verplichte classificatie, het potentieel om meerdere klassemrkkeershers binnen dezelfde context te plaatsen, en het potentieel om klassemrkkersders buiten het naamwoordelijke zinsdeel te plaatsen. De overige twee eigenschappen, de flexibiliteit van classificatie en de klassentoewijzing door de spreker of het systeem, vertonen een minder coherente verdeling. De classificeerderachtige sampletalen laten een voorkeur zien voor een flexibele klassentoewijzing door de spreker, en de concordantieklasseachtige sampletalen laten een voorkeur zien voor een rigide klassentoewijzing. Desalniettemin vertonen beide systeemtypes een aantal afwijkingen van deze voorkeur en verstoren ze op deze manier het grammaticalisatiepatroon. De semantische contributie aan het naamwoordelijke zinsdeel varieert sterk in het sample. Nominaire classificatiesystemen worden over het algemeen als een grammaticaal verschijnsel beschouwd, dat een aantal verschillende functies in de domein van semantiek en discourse grammatica vervult; deze functionele rijklom middels een vaag criterium zoals “semantische contributie” te evalueren blijkt in de samplestudie niet mogelijk. Deze eigenschap wordt daarom in de verdere studie niet meer als meetinstrument gebruikt.

Ik stel in hoofdstuk 6 een alternatieve analyse voor de ogenschijnlijke aanwezigheid van meervoudige klassemarkering binnen dezelfde context in de hybride systemen; de meervoudige klassemarkering is in feite een geval van meervoudige classificatie. Dit betekent dat deze systemen niet de klassewaarde van een naamwoord kopiëren en door middel van congruentie op naamwoorddongevende elementen vertonen, maar dat elk element met een klassemrkkeerder een eigen classificatiecontext vormt. Deze analyse laat een heranalyse van de ‘hybride systemen’ als multipel classificeerdersystemen toe. Bovendien laat ik zien dat de kerneigenschappen sterk met elkaar correleren: systemen met
Samenvatting in het Nederlands

een niet-verplichte classificatie laten geen meervoudige klassemarkering binnen dezelfde context toe en beperken de klassemarkering tot het naamwoordelijke zinsdeel. Nasioi, Bora, Mundurukú, en Ngan’gityemerri zijn de enige uitzonderingen, omdat ze een klassemarkering buiten het naamwoordelijke zinsdeel toelaten. Ook de eigenschappen van de flexibiliteit van classificatie en de klasseverwijzing staan met elkaar in verband: hoe flexibeler de classificatie van een systeem is, hoe meer de spreker medeverantwoordelijk is voor de klasseverwijzing, en andersom.

De analyse van de sampledata in de hoofdstukken 5 en 6 stelt een aanpassing van het concept van grammaticalisatie in nominale classificatiesystemen voor: de samptalen laten twee hoofdstromingen van verandering zien die onafhankelijk van elkaar kunnen optreden. Enerzijds bestaan er systemen die een grammaticalisatie van hun formele uitdrukkingmiddelen laten zien ("formalisatie"): deze bestaat uit de expansie van de woordtypes die een classifier eisen en de expansie van de *locus operandi*. Verder raken de markeringen hun morfosyntactische onafhankelijkheid kwijt, worden fonologisch versleten, en de klasseninventaris neemt af. Anderzijds bestaan er systemen die hun formele uitdrukkingmiddelen niet veranderen, maar een toename van geconventionaliseerde en rigide paren van classifier en naamwoord tonen; verder neemt het potentieel voor een semantische contributie aan het naamwoordselement af ("conventionalisatie"). Het onderscheid tussen formalisatie en conventionalisatie wordt in de hoofdstukken 7 en 8 opgenomen en toegepast op een uitgewerkte *mapping* voor nominale classificatiesystemen. Andere veranderingsstrategieën zijn zeldzamer, maar ook in het sample te vinden.

Systemen kunnen hun functie uitbreiden tot meer dan nominale classifiering; Groot-Andamanees bijvoorbeeld blijft naamwoorden klasseren, indien de classifiers samen met een bijvoeglijk of zelfstandig naamwoord optreden, maar vertoont functies van verbaalclassifiering (de semantische specificatie van het door het werkwoord uitgedrukte concept), indien de classifiers samen met een bijwoord of werkwoord optreden. Een systeem kan ook opgegeven worden; indien dit het geval is, kunnen eventueel overgebleven resten van klassenmarkeringen de voormalige aanwezigheid van een nominale classificatiesysteem aanduiden. Ik redeneer dat dit het geval is in Tlingit en Chukchi, waar modifieders van naamwoorden die naar een mens verwijzen middels een affix gemarkereerd worden, terwijl er geen herkenbaar systeem van functie van nominale classifiering achter zit.

Hoofdstuk 7 introduceert een *mapping*-instrument voor nominaleclassificatiesystemen, dat zich baseert op zowel eerder onderzoek als de observaties van het sample. De samptalen laten zien dat er systemen van alle gradaties van grammaticalisatie zijn, en dat systemen op elk mogelijke punt in hun formalisatie een onderscheid in hun semantische transparantie en productiviteit vertonen. Daarom maakt de mapping een onderscheid tussen de dimensie van formalisatie en de dimensie van de systemische transparantie. Deze dimensies zijn gedefinieerd op basis van twee bundels van zeven eigenschappen. Elke eigenschap omvat twee of meer meetpunten, die op een numerieke grammatica-
Samenvatting in het Nederlands

643

lisatieschaal geprojecteerd kunnen worden. De gemiddelde waarde van de zeven eigenschappen, die de gradatie van de formalisatie van een systeem bepalen, levert de formalisatiewaarde van dit systeem op, en de gemiddelde waarde van de zeven eigenschappen, die de gradatie van conventionalisatie bepalen, levert de transparantiewaarde van dit systeem op. Deze twee waardes maken een eenvoudige visuele presentatie van een systeem in een spreidingsdiagram mogelijk, waar systemen direct vergeleken kunnen worden, zowel synchronisch als diachronisch. De toepassing van de mapping op het sample heeft een duidelijke groepering van het systeem tot gevolg (zie afbeelding 2).

Afbeelding 2: Vergelijkende mapping van het sample: Verdeling van systeemtypes (+: classificeerdersystemen; •: concordantieklassesystemen).

Afbeelding 2 laat zien dat systemen met een formalisatiewaarde tot en met 0.73 overeenkomen met systemen die eerder als classificeerderachtig geanalyseerd werden, en dat deze een transparantiewaarde van 0.7 of meer hebben. Ze variëren zeer met betrekking tot hun formalisatiewaarde. Systemen met een formalisatiewaarde boven de 0.89 komen overeen met de eerder als concordantieklaasachtig geanalyseerde systemen; deze vertonen weinig variatie in hun formalisatiewaarde, terwijl hun transparantiewaarde van 0.2 tot 0.93 reikt. De classificeerderachtige systemen zijn gescheiden van de concordantieklaasachtige systemen door een lacune van ongeveer 16 procent van de formalisatieschaal. De enige uitzondering is het systeem van Tainae, dat de positie als meest transparante uitschieter van de hoog geformaliseerde systemen inneemt, maar in hoofdstuk 6 als classificeerderachtig systeem geanalyseerd wordt. De potentiële scharniersystemen van Nasioi, Bora, Mundurukú, en Ngan’gityemerri nemen in de tweedimensionale mapping daadwerkelijk een scharnierpositie in; zij treden aan de rechterkant van de classificeerderachtige systemen op. De
Samenvatting in het Nederlands

verdeling van de systemen in afbeelding 2 heeft een belangrijk gevolg voor de grammaticalisatihypothese: als alleen systemen met een hoge grammaticalisatiewaarde in staat zijn om een groot deel van hun transparantie te verliezen, kan een ‘globaal’ grammaticalisatieconcept, dat gelijktijdige formalisatie en conventionalisatie veronderstelt, niet bewaard blijven. In plaats daarvan vindt een ‘gespleten’ grammaticalisatieproces plaats, wat inhoudt dat formalisatie vóór een mogelijke conventionalisatie uitgevoerd wordt.

Hoofdstuk 8 beschrijft de diachronie van de systemen van het Mandarijn, het Dyirbal en het Westgermaans. De systemen worden op verschillende punten van hun ontwikkeling en in verschillende registers gemapt. Deze diachronische en sociolinguïstische mapping maakt het mogelijk om de processen van formalisatie, conventionalisatie, en systematische semantische remotivatie te observeren. De case studies suggereren dat een aanmerkelijke daling in transparantie gerelateerd is aan de zichtbaarheid van een systeem. Dat wil zeggen dat alleen een concordantieklasseachtig systeem, waar meervoudige klassemarkeringen in dezelfde context mogelijk zijn, een toegankelijke en productieve semantiek kan opgeven zonder het risico lopen opgegeven of semantisch opnieuw gemootiveerd te worden. Bovendien dragen de linguïstische en culturele competentie en het taalvermogen van de jongere spreker generatie van het Dyirbal bij aan de heranalyse van het systeem. Daarbij draagt het hogere aanzien van het Engels in vergelijking met het Dyirbal niet alleen bij aan de herstructurering van het classificatiesysteem, maar ook aan andere onderdelen van de grammatica.

Samenvattend kan worden vastgesteld dat grammaticalisatie in nominale classificatiesystemen in een ‘gespleten’ vorm optreedt: De formele uitdrukkingsmiddelen grammaticaliseren onafhankelijk van hun semantische en interactieve onderdelen. Een onderzoek van nominale classificatie, dat deze twee ontwikkelingsstrajecten onderscheidt, levert veel gedetailleerdeer en veelzeggendere resultaten op dan een ‘globaal’ grammaticalisatieonderzoek. Bovendien laat ik zien dat het door de grammaticalisatihypothese voorgestelde scenario puur technisch mogelijk is, maar toegankelijke en productieve semantiteit kan opgeven zonder het risico lopen opgegeven of semantisch opnieuw gemootiveerd te worden. Naast het gebrek aan een echte missing link tussen classificerderachtige en concordantieklasseachtige systemen in het sample zijn hiervoor grammaticale veranderingen nodig, die vaak niet aanwezig zijn; een classificerderachtig systeem zou bijvoorbeeld concordantie door middel van affixen ter beschikking moeten hebben, wat in vele Aziatische talen met een isolerend taaltype niet het geval is. Een taal die haar classifieerders tot een concordantieklasseachtig systeem grammaticaliseert, eist grammaticale verschuivingen die zich ver van nominale classificatie afspelen. Vele classificerderdalen gebruiken hun markeerders voor een semantische toelichting van anders ondergespecificeerde lexicaal naamwoorden. Als de markeerders van hun specificatiefunctie verschuiven naar een functie als pure concordantiemarkeerders, moet het concept van een naamwoord in deze taal veranderen, en de lexicaal inhoud van naamwoorden moet gespecificeerd worden. Een scenario waarin een aantal fundamentele grammaticale veranderingen plaatsvindt (die door de grammaticalisatihypothese geëist worden) is niet onmogelijk, maar zeer onwaarschijnlijk.
Matthias Benjamin Passer was born on 28 September, 1984, in Burghausen an der Salzach, Bavaria. After the completion of his secondary education, he studied general and historical-comparative linguistics with a supplementary focus on Indo-European studies at the Paris-Lodron-Universität in Salzburg. He graduated in 2008 cum laude on an etymological and comparative study of nautical terms in Old English and other older Germanic languages. Subsequently, he managed the library of the linguistic department in Salzburg and studied the development of grammatical gender in both Proto-Indo-European and Proto-Germanic.

In 2010, he joined the research project Information Structure in Older Indo-European Languages at the Friedrich-Schiller-Universität Jena as an academic researcher. He presented the results of this project and his own research at international conferences in Innsbruck and Salzburg (Austria), Jena (Germany), Thessaloniki (Greece), Logroño (Spain), and Los Angeles (United States), and attained workshops in Berlin (Germany) and Nijmegen (the Netherlands).

In 2012, Matthias obtained the position of a PhD researcher at the University of Amsterdam under the supervision of prof. dr. Kees Hengeveld and prof. dr. Hedde Zeijlstra. The project was entitled The Nature of Nominal Classification and focussed on an in-depth investigation of the grammaticalisation of nominal classification devices. During the four years of the project’s course, he presented at workshops and international conferences in Salzburg (Austria), Leuven (Belgium), Rouen (France), Jena, Leipzig, and Göttingen (Germany), Amsterdam, Leiden, Nijmegen, and Groningen (the Netherlands), Oslo (Norway), Poznań and Warsaw (Poland), Oxford, and Guildford (United Kingdom). In autumn 2013, he spent one month as a visiting researcher at the Uniwersytet im. Adama Mickiewicza at Poznań, Poland.

Besides his research activities, he was involved in teaching undergraduate courses for exchange students during the last five semesters of his research project and published several articles. Furthermore, he attended seven summer and winter schools for PhD students in the Netherlands and Belgium. The
present dissertation contains the results of his PhD research at the University of Amsterdam.