Parts of speech and dependent clauses: A typological study

van Lier, E.H.

Publication date
2009

Document Version
Final published version

Citation for published version (APA):
Linguistic typology studies the various ways in which languages map form onto meaning. Thus, it identifies the different grammatical constructions that languages use to express particular functional distinctions. The present research is concerned with the typology of two construction types: parts of speech and dependent clauses. These constructions are studied in a sample of 50 languages from all-over the world. The study aims to investigate the relationship between the functional possibilities of the parts of speech and dependent clauses attested in each of these languages.

The parts of speech classes and dependent clauses of individual languages are categorized according to their ability to express one or more of the communicative functions of predication, reference, and modification. In addition, dependent clauses are classified according to their internal morpho-syntactic properties, distinguishing between balanced dependent clauses, which are structurally similar to independent clauses, and deranked dependent clauses, which have properties in common with nominal, adjectival, or adverbial constructions.

The results of this study show that the degree of functional flexibility as displayed by a language’s parts of speech classes constrains the degree of flexibility of its deranked dependent clauses, but not its balanced dependent clauses. In particular, the deranked dependent clauses of a language hardly ever show a higher degree of functional flexibility than its parts of speech classes. Rather, deranked dependent clauses have either an equal or a smaller range of functional possibilities as compared to the parts of speech classes on which they are structurally modeled.

The findings are interpreted from a functionalist perspective: They shed light on the way in which languages establish maximal functional transparency, by dividing the workload of assigning specific functions to specific structures over the lexical, morphological, and syntactic devices available in their grammatical systems.

This study is of relevance for linguistic typologists who work in the functionalist framework and who are interested in lexical and morpho-syntactic categorization.
Parts of Speech and Dependent Clauses
A typological study
Parts of Speech and Dependent Clauses
A typological study

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college van promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 25 september 2009, te 12.00 uur

door

Eva Helena van Lier
geboren te Den Haag
Promotiecommissie

Promotor
prof. dr. P.C. Hengeveld

Overige leden
prof. dr. J.B. den Besten
dr. S. Cristofaro
dr. J. Don
prof. dr. O.C.M. Fischer
prof. dr. H. de Hoop
dr. A. Malchukov

Faculteit der Geesteswetenschappen
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>vii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xix</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Parts of Speech</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Functional-typological attempts at defining parts of speech</td>
<td>8</td>
</tr>
<tr>
<td>2.3 Hengeveld's syntactic-pragmatic approach</td>
<td>12</td>
</tr>
<tr>
<td>2.3.1 Slots and fillers</td>
<td>12</td>
</tr>
<tr>
<td>2.3.2 The cross-linguistic perspective</td>
<td>15</td>
</tr>
<tr>
<td>2.3.2.1 General introduction</td>
<td>15</td>
</tr>
<tr>
<td>2.3.2.2 The Parts of Speech Hierarchy (Hengeveld 1992)</td>
<td>17</td>
</tr>
<tr>
<td>2.3.2.3 Intermediate systems (Hengeveld, Rijkhoff &amp; Siewierska 2004)</td>
<td>19</td>
</tr>
<tr>
<td>2.3.2.4 More diversity (Hengeveld &amp; Van Lier 2008, forthcoming)</td>
<td>22</td>
</tr>
<tr>
<td>2.3.3 Summary</td>
<td>33</td>
</tr>
<tr>
<td>2.4 Critique on the syntactic approach</td>
<td>33</td>
</tr>
<tr>
<td>2.4.1 Introduction</td>
<td>33</td>
</tr>
<tr>
<td>2.4.2 Croft’s Universal-Typological Theory of Parts of Speech</td>
<td>34</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Lexical categories within and across languages: The generality and the subclass problems</td>
</tr>
<tr>
<td>2.4.4</td>
<td>Summary, outlook</td>
</tr>
<tr>
<td>2.5</td>
<td>Defining lexical flexibility</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2.5.2</td>
<td>The criterion of compositionality: flexibility and semantic shift</td>
</tr>
<tr>
<td>2.5.2.1</td>
<td>Evans and Osada’s proposal and some alternatives</td>
</tr>
<tr>
<td>2.5.2.2</td>
<td>An integrative approach to semantics in flexible languages</td>
</tr>
<tr>
<td>2.5.3</td>
<td>The criterion of exhaustiveness: How flexible should a flexible language be?</td>
</tr>
<tr>
<td>2.5.4</td>
<td>The criterion of equivalent combinatorics: Bi-directional flexibility</td>
</tr>
<tr>
<td>2.5.5</td>
<td>Lexical flexibility: A summary</td>
</tr>
<tr>
<td>2.6</td>
<td>Summary</td>
</tr>
<tr>
<td>3</td>
<td>Dependent Clauses</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.2</td>
<td>Dependent clauses and propositional functions</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Slots and fillers</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Structural coding in dependent clauses: Subordinating conjunctions and dependent verb forms</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Rigid and flexible dependent clauses</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Summary</td>
</tr>
<tr>
<td>3.3</td>
<td>Behavioural potential in dependent clauses: De-categorization and re-categorization</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Verbal and nominal feature hierarchies</td>
</tr>
<tr>
<td>3.3.2.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.3.2.2</td>
<td>The verbal feature hierarchy</td>
</tr>
<tr>
<td>3.3.2.3</td>
<td>The nominal feature hierarchy</td>
</tr>
<tr>
<td>3.3.2.4</td>
<td>Summary, outlook</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Feature mixing in DCs across languages: Cross-linguistic generalizations and functional motivations</td>
</tr>
<tr>
<td>3.3.3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.3.3.2</td>
<td>The typology of subordination (Cristofaro 2003)</td>
</tr>
<tr>
<td>3.3.3.3</td>
<td>The typology of nominalization (Malchukov 2004)</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Summary, outlook</td>
</tr>
<tr>
<td>3.4</td>
<td>A two-step typology of dependent clauses</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary</td>
</tr>
</tbody>
</table>
4 Sampling, Research Questions, Method

4.1 Introduction

4.2 The language sample

4.3 Research Questions

4.3.1 Introduction

4.3.2 Global functional matching

4.3.3 Global matching including differentiation for structural DC type

4.3.4 Specific functional matching

4.3.5 Specific matching including differentiation for structural DC type

4.3.6 Summary

4.4 Method

4.5 Outlook

5 Parts of Speech in the Languages of the Sample

5.1 Introduction

5.2 Overview of the language data

5.3 Results in relation to implicational map of Parts of Speech

5.3.1 Introduction

5.3.2 Predicted and attested 'pure' systems

5.3.3 Predicted and attested 'intermediate' systems

5.3.4 Predicted but not attested systems

(and attested but not predicted systems)

5.3.5 Summary

5.4 The generality and the subclass problems in practice

5.4.1 The generality problem

5.4.2 The subclass problem

5.4.2.1 Fuzzy boundaries

5.4.2.2 Restricted classes

5.5 Non-verbal predication

5.6 Identifying flexibility

5.6.1 Introduction

5.6.2 Languages with contentives

5.6.2.1 Semantic compositionality

5.6.2.2 Exhaustiveness

5.6.2.3 Equivalent combinatorics

5.6.3 Non-verbs, nominals, and modifiers

5.6.3.1 Non-verbs

5.6.3.2 Nominals
5.6.3.3 Modifiers 183
5.6.4 Summary 188
5.7 Conclusion, outlook 185

6 Dependent Clauses in the Languages of the Sample 187
   6.1 Introduction 187
   6.2 DC constructions and propositional functions 187
      6.2.1 Introduction 187
      6.2.2 Rigid and flexible DCs: Attested types 188
         6.2.2.1 Preliminary considerations 188
         6.2.2.2 Language data 196
      6.2.3 Discussion 204
         6.2.3.1 Attested and predicted types 204
         6.2.3.2 Predicted but not attested types: 211
            The problem of predicate clauses
      6.2.4 Summary 216
   6.3 DC constructions and behavioural potential 216
      6.3.1 Introduction 216
      6.3.2 Basic Data 217
         6.3.2.1 Formal parameters 217
         6.3.2.2 Language data 220
      6.3.3 A three-way typology of DCs according to behavioural potential 229
         6.3.3.1 Introduction 229
         6.3.3.2 Preliminary considerations 230
         6.3.3.3 Language data 235
      6.3.4 Behavioural potential of different structural DC types 242
   6.4 An integrated DC typology: Functional possibilities of three structural DC types 251
   6.5 Summary, outlook 259

7 Results 261
   7.1 Introduction 262
   7.2 Global functional matches:
      Flexibility versus rigidity in the PoS and DC domains
         7.2.1 Global matching without differentiating for structural DC types 262
            7.2.1.1 Introduction 262
            7.2.1.2 Global match for flexible constructions 262
            7.2.1.3 Global match for rigid constructions 264
7.2.1.4 Summary 264
7.2.2 The parameter of structural DC type 264
7.2.2.1 Introduction 264
7.2.2.2 Global match for flexible constructions, differentiating for structural DC type 265
7.2.2.3 Global match for rigid constructions, differentiating for structural DC type 269
7.2.2.4 Summary 277

7.3 Specific functional matches: 278
Types and amounts of flexibility/rigidity in the PoS and DC domains
7.3.1 Introduction 278
7.3.2 Specific matches for flexible constructions 278
7.3.2.1 Introduction 278
7.3.2.2 Flexible match 1: Lexical contentives/non-verbs and contentive/multi-functional clauses 280
7.3.2.3 Flexible match 2: Nominals and nominal clauses 284
7.3.2.4 Flexible match 3: Modifiers and modifier clauses 287
7.3.2.5 Summary 288
7.3.3 Specific matches for rigid constructions 288
7.3.3.1 Introduction 289
7.3.3.2 Rigid match 1: Nouns and complement clauses 289
7.3.3.3 Rigid match 2: Adjectives and relative clauses 291
7.3.3.4 Rigid match 3: Manner adverbs and adverbial manner clauses 295
7.3.3.5 Summary 297

7.4 Summary, Conclusion 297

7.5 Language data 298

8 Discussion – Flexibility and Functional Transparency 313
8.1 Introduction 313
8.2 Theoretical background: 314
Trade-off effects between flexible and rigid constructions
8.3 Applying the complexity hypothesis to PoS and deranked DCs 320
8.3.1 Hypothesis 320
8.3.2 Results 322
8.3.2.1 Distribution of deranked DCs in languages with flexible PoS systems 322
8.3.2.2 Distribution of deranked DCs in languages with rigid PoS systems 331
I like languages. At least that much was clear to me when I finished high school. The problem was that I didn't know yet why I liked them so much. And since there didn't seem to be any particular language that I liked better than any other, I decided to start studying my own language: Dutch. During my first classes in Dutch syntax I discovered, to my great excitement, that there was a real academic discipline devoted to my gut feeling about languages; it was called Linguistics. Moreover, I realized that the existence of Linguistics implied the existence of Linguists. Very soon I knew I wanted to belong to this group of people – that I wanted to be a linguist. I hope this book shows that I have taken an important step on my way to becoming one.

One day in August 2004, when it was pouring with rain and I was in a particularly bad mood, Kees Hengeveld called me to say that I got the job: I was going to be a PhD candidate at the Department of Theoretical Linguistics of the University of Amsterdam. Three days later my desk was waiting for me at the Herengracht 338. Some years earlier, when I was still a student, the huge, high-ceilinged, large-windowed room on the ground floor of this beautiful building had been renovated. I remember clearly that I peeped inside the room and thought: If only one of these desks could once be mine…and here it was!

The other desks in the room belonged to three wonderful persons: Rafael, Annerieke, and Gerry. They were the firsts of many more people that have made my time as a PhD candidate so thoroughly enjoyable. I wish
to express my deep gratitude to all of them.

First, I want to thank Kees, for the enthusiasm, dedication, and patience with which he helped me to structure lines of thought, improve analyses, and gain confidence as a researcher along the way. He did this not only in his role of supervisor of my MA, MPhil, and PhD theses, but also as a colleague in our various joint projects. I am very grateful for all these years of collaboration, during which I learned innumerable cool things about innumerable cool languages, as well as a few lessons about myself. Among the latter, there is one thing Kees taught me that I will always remember: that it is essential to invest energy in those things in (linguistic) life that make one’s heart beat faster.

I also wish to thank Hans den Besten, Sonia Cristofaro, Jan Don, Olga Fischer, Helen de Hoop, and Andrej Malchukov for accepting the invitation to be a member of my committee. Special thanks go to Jan Don, who helped me with editing the manuscript in its final stages, and who has more generally been a wonderful colleague with whom I have collaborated intensively and with great pleasure. Another thank you goes to Andrej Malchukov, who took the time for interesting and helpful discussions during my stay at the Max Planck Institute in Leipzig.

But there are many other people who have made valuable contributions to the present work: Dik Bakker, whose extensive e-mails helped me to overcome occasional fits of doubt and confusion; Elena Maslova, for her generous and crystal clear comments on the interpretation of quantitative typological data; Rob Schoonen, for his kind advice on statistical issues; Michael Cysouw, Martin Haspelmath, and Jaklin Kornfilt for inspiring discussions at the Max Planck in Leipzig; John Peterson and Nick Evans for helping me out on the fascinating languages they know so much about. I also wish to thank all the members of the parts-of-speech research group of the Amsterdam Center for Language and Communication for their sharing of expertise and spirit.

More generally, a great amount of people have spiced up my linguistic life in the past years. First, I wish to thank my fellow students of the Linguistics Research Master in Amsterdam for a very stimulating year: It’s good to be among like-minded language-freaks. Also, many thanks go to all the people who contribute to the warm bath called the Linguistics Department of the University of Amsterdam. I’m going to miss you all. Finally, I feel grateful to have met so many young fellow-typologists, with whom I have spent good times in Leipzig and at conferences all-over Europe. All these people have
confirmed, time and again, my feeling that linguists are inspiring, open-minded, and great fun!

Some linguists are particularly dear to me. First, a special thank you goes to Rafael – whose jokes are even better than his looks – for sharing his Cofán with me, for asking difficult typological questions, and for helping me to believe, notwithstanding gloomy moods, that our work matters after all. Another big thank you goes to Esther and Margot, for lighting up the Bungehuis in their very own lovely ways and for all the shared lunches, dinners, ice-creams, and borrels (and their burning-up during runs along the Amstel). I also want to thank Josje Verhagen and Suzanne Dikker, for great collaboration projects during our student years, and for the many drinks and dinners over which we shared our experiences in and outside of linguistic careers, keeping each other’s spirits high. A very special thank you is for Gerry Wanders, for reading the manuscript of my dissertation with so much care.

Of course, there are also many non-linguists to thank, in fact too many to list them all. Nonetheless, I want to mention Floor and Anne, for their patience with all my struggles, as well as for celebrating the good times with me. And Gijs, for never tiring of asking puzzling questions about my work. Finally, I want thank Claartje, Sanne and Annet for making this book look so pretty.

Very particularly, I want to thank my paranimf, Lotte, for adopting me after the Herengracht 388 was sold, and for being the best possible roomy, colleague, and friend ever since. I am grateful for her sharing; not just of working space and coffee, but also of classical music, intellectual frustrations and successes, panic attacks, statistical know-how, sore muscles, broken hearts, a great many laughs, and more time with me than anyone else. Lotte, you’re a star!

I also want to thank my other paranimf. Joost, thank you for being my big brother, and for giving me your unconditional love and support.

I want to thank my parents. Papa, thanks for being my father in a way that makes me feel proud to be your daughter, and lucky to have your courageous embrace of life as a guiding example. Mama, thanks for making me remember you the way I do.

Finally, I want to thank Sander. My love, I doubt whether there is any other
linguist’s partner who understands linguistics as well as you do. But that’s just one of the countless reasons why you are – and have been for so many years – an essential part of my life.

Amsterdam, June 2009
### List of Figures

<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>Parts of Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Syntactic slots as distinguished in Hengeveld (1992)</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Syntactic slots and lexical categories</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>The functional potential of parts of speech classes</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Possible parts of speech systems according to Hengeveld 1992</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>The flexible intermediate PoS system of Turkish</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>The rigid intermediate PoS system of Alamblak</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>Parts of speech systems, including intermediate types</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>The implicational map of parts of speech</td>
</tr>
<tr>
<td>Figure 2.9</td>
<td>The PoS system of Santali</td>
</tr>
<tr>
<td>Figure 2.10</td>
<td>The PoS system of Garo</td>
</tr>
<tr>
<td>Figure 2.11</td>
<td>Properties of semantic classes</td>
</tr>
<tr>
<td>Figure 2.12</td>
<td>Nouns, adjectives and verb as prototypical combinations of propositional act functions and semantic classes</td>
</tr>
<tr>
<td>Figure 2.13</td>
<td>The noun-adjective distinction in Japanese</td>
</tr>
<tr>
<td>Figure 2.14</td>
<td>The conceptual structure of kiss.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3</th>
<th>Dependent Clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1a</td>
<td>The space of propositional functions</td>
</tr>
<tr>
<td>Figure 3.1b</td>
<td>Propositional functions and dependent clause types</td>
</tr>
</tbody>
</table>
### Chapter 5

#### Parts of Speech in the Languages of the Sample

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>The PoS system of Warao</td>
<td>141</td>
</tr>
<tr>
<td>5.2</td>
<td>The PoS system of Kayardild</td>
<td>148</td>
</tr>
<tr>
<td>5.3</td>
<td>The noun-adjective distinction in Japanese (2)</td>
<td>156</td>
</tr>
</tbody>
</table>
## List of Tables

### Chapter 3  Dependent Clauses

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>DC types and their internal formal properties</td>
<td>108</td>
</tr>
</tbody>
</table>

### Chapter 4  Sampling, Research Questions, Method

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>The language sample</td>
<td>113</td>
</tr>
<tr>
<td>4.2</td>
<td>DC types and their internal formal properties</td>
<td>116</td>
</tr>
<tr>
<td>4.3</td>
<td>Functional connection between PoS and deranked DCs</td>
<td>117</td>
</tr>
<tr>
<td>4.4</td>
<td>A 2x2 contingency table</td>
<td>118</td>
</tr>
<tr>
<td>4.5a</td>
<td>Maslova test no. 1: Change DC parameter to PoS=DC</td>
<td>119</td>
</tr>
<tr>
<td>4.5b</td>
<td>Maslova test no. 2: Change PoS parameter to PoS=DC</td>
<td>120</td>
</tr>
<tr>
<td>4.6</td>
<td>A symmetrical dependency; original distribution</td>
<td>121</td>
</tr>
<tr>
<td>4.7a</td>
<td>Distribution for PoS and PoS = DC (Maslova-test no. 1)</td>
<td>122</td>
</tr>
<tr>
<td>4.7b</td>
<td>Distribution for DC and PoS = DC (Maslova-test no. 2)</td>
<td>122</td>
</tr>
<tr>
<td>4.8</td>
<td>A one-sided asymmetrical dependency; original distribution</td>
<td>123</td>
</tr>
<tr>
<td>4.9a</td>
<td>Distribution for PoS and PoS = DC (Maslova-test no. 1)</td>
<td>123</td>
</tr>
<tr>
<td>4.9b</td>
<td>Distribution for DC and PoS = DC (Maslova-test no. 2)</td>
<td>124</td>
</tr>
<tr>
<td>4.10</td>
<td>A two-sided asymmetrical dependency; original distribution</td>
<td>125</td>
</tr>
<tr>
<td>4.11a</td>
<td>Distribution for PoS and PoS = DC (Maslova-test no. 1)</td>
<td>125</td>
</tr>
<tr>
<td>4.11b</td>
<td>Distribution for DC and PoS = DC (Maslova-test no. 2)</td>
<td>126</td>
</tr>
</tbody>
</table>

### Chapter 5  Parts of Speech in the Languages of the Sample

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>PoS classes in the sample languages</td>
<td>132</td>
</tr>
</tbody>
</table>

XIII
Table 5.2  Verbs, nouns and nominals in Japanese  156
Table 5.3  Verbs and Adjectives in Mandarin Chinese  158

Chapter 6  Dependent Clauses in the Languages of the Sample
Table 6.1  DCs in the sample languages and their possibilities concerning the expression of propositional functions  197-204
Table 6.2  Basic data on behavioural potential of all DC constructions  220-229
Table 6.3  DC types and their internal formal properties  230
Table 6.4  Three-way classification of DCs  236-242
Table 6.5  Behavioural potential of type 1 balanced DCs  242-246
Table 6.6  Behavioural potential of type 2 D-SENT DCs  246-249
Table 6.7  Behavioural potential of type 3 D-ALT DCs  250-251
Table 6.8  Functional possibilities of type 1 balanced DCs  252-255
Table 6.9  Functional possibilities of type 2 D-SENT DCs  255-257
Table 6.10  Functional possibilities of type 3 D-ALT DCs  257-258

Chapter 7  Results
Table 7.1  Frequencies for languages with/without flexible PoS and DCs  263
Table 7.2  Frequencies for language with/without rigid PoS only and rigid DCs only  264
Table 7.3  Frequencies for languages with/without flexible PoS and flexible DCs of type 1  266
Table 7.4  Frequencies for languages with/without flexible PoS and flexible DCs of type 2/3  266
Table 7.5a Maslova test no. 1: change DC parameter to PoS=DC  266
Table 7.5b Maslova test no. 2: Change PoS parameter to PoS=DC  267
Table 7.6  Frequencies for languages with/without flex PoS and flex DCs of type 2  268
Table 7.7  Frequencies for languages with/without flex PoS and flex DCs of type 3  268
Table 7.8a Maslova test no. 1: change DC parameter to PoS=DC  269
Table 7.8b Maslova test no. 2: Change PoS parameter to PoS=DC  269
Table 7.9  Frequencies for languages with/without rigid PoS only and rigid DCs only  270
Table 7.10a Maslova test no. 1: change DC parameter to PoS=DC  280
Table 7.10b Maslova test no. 2: change PoS parameter to PoS=DC  281
Table 7.11  Frequencies for languages with/without contentives/non-verbs and contentive clauses/multi-functional clauses of type 1 282
Table 7.12  Frequencies for languages with/without lexical contentives/non-verbs and contentive clauses/multi-functional clauses of type 3 283
Table 7.13a Maslova test no. 1: change DC parameter to PoS=DC 283
Table 7.13b Maslova test no. 2: change PoS parameter to PoS=DC 283
Table 7.14  Frequencies for languages with/without nominals and nominal clauses 284
Table 7.15  Frequencies for languages with/without nominals and nominal clauses of type 1 285
Table 7.16  Frequencies for languages with/without nominals and nominal clauses of type 2/3 286
Table 7.17  Frequencies for languages with/without nominals and nominal clauses of type 2 286
Table 7.18  Frequencies for languages with/without nominals and nominal clauses of type 3 286
Table 7.19  Frequencies for languages with/without modifiers and modifier clauses 287
Table 7.20  Frequencies for languages with/without nouns and complement clauses 290
Table 7.21  Frequencies for language with/without nouns and complement clauses of type 1 290
Table 7.22  Frequencies for language with/without nouns and complement clauses of type 2/3 291
Table 7.23  Frequencies for language with/without nouns and complement clauses of type 2 291
Table 7.24  Frequencies for language with/without nouns and complement clauses of type 3 291
Table 7.25  Frequencies for languages with/without adjectives and relative clauses 292
Table 7.26  frequencies for languages with/without adjectives and relative clauses of type 1 293
Table 7.27  Frequencies for languages with/without adjectives and relative clauses of type 2/3 293
Table 7.28  Frequencies for languages with/without adjectives and relative clauses of type 2 294
Table 7.29  Frequencies for languages with/without adjectives and relative clauses of type 3 294
Table 7.30  Frequencies for languages with/without (derived) manner adverbs and adverbial manner clauses 295
Table 7.31  Frequencies for languages with/without (derived) manner adverbs and adverbial manner clauses of type 1 296
Table 7.32  Frequencies for languages with/without (derived) manner adverbs and adverbal manner clauses of type 2/3  297

Table 7.33  Availability and type of flexible DCs in languages with one or more flexible PoS classes  298-299

Table 7.34  Availability and type of flexible DCs in languages with rigid PoS classes only  299-300

Table 7.35  Availability and type of contentive/multi-functional clauses in languages with lexical contentives/non-verbs  300

Table 7.36  Availability and type of contentive/multi-functional clauses in languages without lexical contentives/non-verbs  301-302

Table 7.37  Availability and type of nominal clauses in languages with lexical nominals  302

Table 7.38  Availability and type of and nominal clauses in languages without lexical nominals  302-303

Table 7.39  Availability of modifier clauses in languages with (derived) lexical modifiers  303-304

Table 7.40  Languages without (derived) lexical modifiers, but with modifier clauses  304

Table 7.41  Availability and type of complement clauses in languages without nouns  304

Table 7.42  Availability and type of complement clauses in languages with nouns  304-305

Table 7.43  Availability and type of relative clauses in languages without adjectives, but with a flexible lexical strategy for the function of modifier in a referential phrase  306

Table 7.44  Availability and type of relative clauses in languages with an open class of simple adjectives  306-307

Table 7.45  Availability and type of relative clauses in languages without an open class of simple adjectives, but with derived adjectives (and a small class of simple adjectives)  307

Table 7.46  Availability and type of relative clauses in languages without an open class of simple/derived class adjectives, but with a small class of adjectives  308

Table 7.47  Availability and type of relative clauses in languages without a lexical strategy for the function of modifier in a referential phrase  308

Table 7.48  Availability and type of adverbal manner clauses in languages without an open class of simple/derived manner adverbs, but with a flexible lexical strategy for the function of modifier in a predicate phrase (and in some cases also a small class of rigid manner adverbs)  309
Table 7.49  Availability and type of adverbial manner clauses in languages with an open class of simple manner adverbs 310
Table 7.50  Availability and type of adverbial manner clauses in languages without an open class of simple manner adverbs, but with derived manner adverbs (and in some cases a small class of simple manner adverbs) 310
Table 7.51  Availability and type of relative clauses in languages without an open class of simple/derived class manner adverbs, but with a small class of manner adverbs 311
Table 7.52  Availability and type of adverbial manner clauses in languages without a lexical strategy for the function of modifier in a predicate phrase 311

Chapter 8  Discussion: Flexibility and Functional Transparency
Table 8.1  Languages with a flexible PoS class, and the distribution of deranked DCs that appear in at least one of the same function(s) 323–324
Table 8.2  The principle of increasing complexity: Flexible PoS systems and deranked DCs 327
Table 8.3  Rigid PoS classes in languages with flexible PoS, and the distribution of deranked DCs that appear in the same function as these rigid PoS classes. 332–333
Table 8.4  Rigid PoS classes in languages with rigid PoS only, and the distribution of deranked DCs that appear in the same function 334–335
Table 8.5  Flexible balanced DCs in languages with flexible PoS systems 337
Table 8.6  Flexible balanced DCs in languages with rigid PoS systems 341
List of Abbreviations

AB  = ablative
ABS = absolutive
ACC = accusative
ACT = active
ADV = adverbial marker/subordinator
AFF = affirmative
AFR = aforementioned
AG  = agent
ALL = allative
ANAPH = anaphoric proform
ANIM = animate
A.OBL = associating oblique case
AOR = aorist
APPL = applicative
ART = article
ASSOC = associative
ATTR = attributive
AUX = auxiliary
AV  = actor voice
BEN = beneficiary
CA  = continuative aspect
CAUS = causative
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCM</td>
<td>conjunction class marker</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
</tr>
<tr>
<td>CM</td>
<td>challengeable marker</td>
</tr>
<tr>
<td>CNTR</td>
<td>contrastive focus</td>
</tr>
<tr>
<td>C.OBL</td>
<td>complementizing oblique case</td>
</tr>
<tr>
<td>COM</td>
<td>completion</td>
</tr>
<tr>
<td>COMP</td>
<td>complementizer</td>
</tr>
<tr>
<td>COND</td>
<td>conditional</td>
</tr>
<tr>
<td>CONJ</td>
<td>conjunction</td>
</tr>
<tr>
<td>CONN</td>
<td>connector</td>
</tr>
<tr>
<td>CONS</td>
<td>consequential</td>
</tr>
<tr>
<td>CONT</td>
<td>contemporative mood</td>
</tr>
<tr>
<td>CONV</td>
<td>converb</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>CR</td>
<td>correlative</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>DECL</td>
<td>declarative marker</td>
</tr>
<tr>
<td>DEIC</td>
<td>deictic element</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>DEFOC</td>
<td>de-focused</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
</tr>
<tr>
<td>DEP</td>
<td>dependent form</td>
</tr>
<tr>
<td>DIM</td>
<td>diminutive</td>
</tr>
<tr>
<td>DIR</td>
<td>directional</td>
</tr>
<tr>
<td>DIST</td>
<td>distal</td>
</tr>
<tr>
<td>DISTR</td>
<td>distributive</td>
</tr>
<tr>
<td>DNMLZ</td>
<td>denominalizer</td>
</tr>
<tr>
<td>D.SO</td>
<td>distal extension: point of view of source</td>
</tr>
<tr>
<td>DTR</td>
<td>de-transitivizer</td>
</tr>
<tr>
<td>DU</td>
<td>dual</td>
</tr>
<tr>
<td>DUR</td>
<td>durative</td>
</tr>
<tr>
<td>DYN</td>
<td>dynamic</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>EMOT</td>
<td>emotive</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EV.COP</td>
<td>evidential copula</td>
</tr>
<tr>
<td>EXCL</td>
<td>exclusive</td>
</tr>
<tr>
<td>FEM</td>
<td>feminine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Meaning</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>FIN</td>
<td>finite</td>
</tr>
<tr>
<td>FOC</td>
<td>focus</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>GENR</td>
<td>general TAM particle</td>
</tr>
<tr>
<td>GER</td>
<td>gerund</td>
</tr>
<tr>
<td>HAB</td>
<td>habitual</td>
</tr>
<tr>
<td>HUM</td>
<td>human</td>
</tr>
<tr>
<td>IDEOPH</td>
<td>ideophone</td>
</tr>
<tr>
<td>IMMED</td>
<td>immediate</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative</td>
</tr>
<tr>
<td>INANIM</td>
<td>inanimate</td>
</tr>
<tr>
<td>INCEP</td>
<td>inceptive</td>
</tr>
<tr>
<td>INCL</td>
<td>inclusive</td>
</tr>
<tr>
<td>IND</td>
<td>indicative</td>
</tr>
<tr>
<td>INDF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INESS</td>
<td>inessive</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INSTR</td>
<td>instrumental</td>
</tr>
<tr>
<td>INTERJ</td>
<td>interjection</td>
</tr>
<tr>
<td>INTR</td>
<td>intransitive</td>
</tr>
<tr>
<td>I.OBJ</td>
<td>indirect object</td>
</tr>
<tr>
<td>IOV</td>
<td>indirect object version</td>
</tr>
<tr>
<td>IPFR</td>
<td>imperfective</td>
</tr>
<tr>
<td>IRR</td>
<td>irreals</td>
</tr>
<tr>
<td>ITER</td>
<td>iterative</td>
</tr>
<tr>
<td>LG</td>
<td>ligature</td>
</tr>
<tr>
<td>LK</td>
<td>linker</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>L.PRON</td>
<td>logophoric pronoun</td>
</tr>
<tr>
<td>LV</td>
<td>locative voice</td>
</tr>
<tr>
<td>M</td>
<td>middle voice</td>
</tr>
<tr>
<td>M.ABL</td>
<td>modal ablative case</td>
</tr>
<tr>
<td>MASC</td>
<td>masculine</td>
</tr>
<tr>
<td>M.LOC</td>
<td>modal locative case</td>
</tr>
<tr>
<td>MOD</td>
<td>modality marker</td>
</tr>
<tr>
<td>MODIF</td>
<td>modifier</td>
</tr>
<tr>
<td>M.PROP</td>
<td>modal proprietive case</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>N</td>
<td>non-past low tone verbal prefix</td>
</tr>
<tr>
<td>NC</td>
<td>non-control</td>
</tr>
<tr>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>NEUT</td>
<td>neuter</td>
</tr>
<tr>
<td>NMLZ</td>
<td>nominalizer</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>N.FIN</td>
<td>non-finite</td>
</tr>
<tr>
<td>N.PST</td>
<td>non-past</td>
</tr>
<tr>
<td>NV</td>
<td>neutral version</td>
</tr>
<tr>
<td>OBJ</td>
<td>object marker</td>
</tr>
<tr>
<td>OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>ONOM</td>
<td>onomatopoeic form</td>
</tr>
<tr>
<td>OPT</td>
<td>optative</td>
</tr>
<tr>
<td>OR</td>
<td>oriented</td>
</tr>
<tr>
<td>ORIG</td>
<td>origin</td>
</tr>
<tr>
<td>OV</td>
<td>object version</td>
</tr>
<tr>
<td>PART</td>
<td>partitive</td>
</tr>
<tr>
<td>PASS</td>
<td>passive</td>
</tr>
<tr>
<td>PAT</td>
<td>patient</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PHTM</td>
<td>phrase terminal marker</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>PN</td>
<td>proper name</td>
</tr>
<tr>
<td>PO</td>
<td>potential object (verbal extension)</td>
</tr>
<tr>
<td>POSS</td>
<td>possessor</td>
</tr>
<tr>
<td>POSSD</td>
<td>possessed</td>
</tr>
<tr>
<td>POT</td>
<td>potential</td>
</tr>
<tr>
<td>PREF</td>
<td>prefix</td>
</tr>
<tr>
<td>PRET</td>
<td>preterite (simple past)</td>
</tr>
<tr>
<td>PREV</td>
<td>preverb</td>
</tr>
<tr>
<td>PROGR</td>
<td>progressive</td>
</tr>
<tr>
<td>PRON</td>
<td>pronoun</td>
</tr>
<tr>
<td>PROS</td>
<td>prosecutive</td>
</tr>
<tr>
<td>PROX</td>
<td>proximate</td>
</tr>
<tr>
<td>PRS</td>
<td>present</td>
</tr>
<tr>
<td>PRT</td>
<td>particle</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>PTC</td>
<td>participle</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>PV</td>
<td>patient voice</td>
</tr>
<tr>
<td>Q</td>
<td>interrogative</td>
</tr>
<tr>
<td>QUOT</td>
<td>quotative</td>
</tr>
<tr>
<td>RDP</td>
<td>reduplication</td>
</tr>
<tr>
<td>REAL</td>
<td>realis</td>
</tr>
<tr>
<td>REF</td>
<td>referential</td>
</tr>
<tr>
<td>REFL</td>
<td>reflexive</td>
</tr>
<tr>
<td>REL</td>
<td>relative marker</td>
</tr>
<tr>
<td>RELC</td>
<td>relational case</td>
</tr>
<tr>
<td>REM</td>
<td>remote</td>
</tr>
<tr>
<td>RES</td>
<td>resultative</td>
</tr>
<tr>
<td>RCP</td>
<td>reciprocal</td>
</tr>
<tr>
<td>RSP</td>
<td>resumptive pronoun</td>
</tr>
<tr>
<td>SBJV</td>
<td>subjunctive</td>
</tr>
<tr>
<td>SEQ.CONV</td>
<td>sequential converb</td>
</tr>
<tr>
<td>SEMTR</td>
<td>semi-transitivizer</td>
</tr>
<tr>
<td>SF</td>
<td>stem forming formative</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SIM</td>
<td>adverbial marker of simultaneity</td>
</tr>
<tr>
<td>SPEC</td>
<td>specifier</td>
</tr>
<tr>
<td>SS</td>
<td>same subject</td>
</tr>
<tr>
<td>SR</td>
<td>source</td>
</tr>
<tr>
<td>STAT</td>
<td>stative</td>
</tr>
<tr>
<td>SUB</td>
<td>subordinator</td>
</tr>
<tr>
<td>SUBJ</td>
<td>subject marker</td>
</tr>
<tr>
<td>SUFF</td>
<td>suffix</td>
</tr>
<tr>
<td>SUPERESS</td>
<td>superessive</td>
</tr>
<tr>
<td>SUPERL</td>
<td>superlative</td>
</tr>
<tr>
<td>SV</td>
<td>subjective version</td>
</tr>
<tr>
<td>SW</td>
<td>switch reference</td>
</tr>
<tr>
<td>TENT</td>
<td>tentative</td>
</tr>
<tr>
<td>Ths</td>
<td>thematic suffix</td>
</tr>
<tr>
<td>TOP</td>
<td>topic</td>
</tr>
<tr>
<td>TR</td>
<td>transitivizer</td>
</tr>
<tr>
<td>TRL</td>
<td>translative case</td>
</tr>
<tr>
<td>UV</td>
<td>undergoer voice</td>
</tr>
<tr>
<td>VAL</td>
<td>validator</td>
</tr>
<tr>
<td>VE</td>
<td>ventive prefix</td>
</tr>
</tbody>
</table>
Languages are formal systems used for human communication. Three basic communicative functions, or communicative acts, may be distinguished: Reference, predication, and modification. The act of reference identifies a referent, i.e. an entity that one wants to talk about, and establishes a cognitive file for that referent. The act of predication ascribes something to a referent, i.e. it reports that this referent is involved in some state of affairs. The act of modification enriches either reference or predication, by expressing an additional feature of either a referent or a state of affairs (Searle 1969, Croft 2001: 66).

The functions of reference, predication, and modification may be fulfilled through linguistic forms of various types. The most basic way to do this is probably by means of a phrase consisting of a single lexical element. The lexical elements of any particular language can be divided into different groups or classes, called parts of speech. Prototypically, the class of nouns is used for communicative acts of reference (cf. car in example (1a)); verbs are used for predication (cf. wash in (1b)); adjectives for the modification of referential expressions (cf. red in (1c)); and adverbs for the modification of predications (cf. quickly in (1d)).

(1) a. I wash the car.
   b. I wash the car.
   c. I wash the red car.
   d. I wash the car quickly.
However, not all languages follow this particular pattern, in which every part of speech class is specialized for the expression of a single function. Alternatively, languages may display part of speech classifications in which two or more functions can be expressed by members of the same class. Hockett (1958: 225) compares such functionally different parts of speech to “athletic squads, trained in different ways to play much the same game.” Some languages train specialists, while others try to make an “all-round player or triple-threat man of every member in the squad.” These two types of techniques may also be combined within a single language system, “producing some specialists but also good numbers of double-threat and triple-threat men.” Adopting a terminology proposed by Hengeveld (1992), I will call parts of speech that are functionally specialized rigid, and those that can express multiple functions flexible.

Apart from simple lexical items, however, more complex linguistic structures may also be used to express the functions of reference, predication, and modification within an utterance. One possibility is the use of a clause-like construction within a larger utterance. Traditionally, this linguistic situation is referred to as subordination. Consider the example in (2), where the subordinate or dependent clause ‘John will come home soon’ is used as a referring expression that functions as the object of the predicate of the main or matrix clause ‘Mary hopes’, to which it is linked by means of the subordination marker that.

(2) Mary hopes [that John will come home soon].

Like parts of speech, certain types of dependent clause constructions are rigid, i.e. can fulfil only a single type of function, while others are flexible, i.e. may express multiple functions. For instance, the English that-construction, illustrated in (2) above, is flexible because it can also be used for modification, as is shown in (3), where the dependent clause ‘that we bought together yesterday’ expresses an additional property of the referent of ‘book’, which is the direct object within the matrix clause ‘I gave her the book’.

(3) I gave her the book [that we bought together yesterday].

Apart from their possibilities regarding the expression of communicative function(s), dependent clauses – being complex constructions – can
also be studied from the perspective of their internal morpho-syntactic properties. Usually, the morpho-syntactic properties of dependent clauses in a particular language are described in comparison to the properties of ordinary independent clauses in that language. Consider the difference between examples (4a) and (4b): The dependent clause in (4a), marked with the subordinator that, expresses future tense (will move) and the subject Paul is encoded in the same way as it would be in an independent clause. This type of dependent clause, which structurally resembles an independent clause, is called balanced. In contrast, the dependent clause in (4b) is marked by the special -ing form of the predicate, it cannot express tense, and it has a subject that is encoded as a possessor (Paul’s). This second type of dependent clause, the structure of which deviates to some extent from the structure of an independent clause, is called deranked (Stassen 1985).

(4)  a.  John regrets [that Paul will move to Finland].
   b.  John regrets [Paul’s moving to Finland].

The aim of the present research is to investigate the relationship between the functional flexibility or rigidity displayed by a particular language’s parts of speech classes on the one hand, and by its balanced/deranked dependent clauses on the other hand. More specifically, this study tries to discover to what extent it is possible to predict, on the basis of a language’s parts of speech system – that is, its major part of speech classes and their respective potential to express particular functions – what the functional potential of its dependent clause constructions will be.

This goal is approached by means of an investigation of the parts of speech classes and dependent clause constructions of a genealogically and geographically balanced sample of 50 languages. For every major part of speech class and for every relevant dependent clause construction in each of the sample languages it is determined which function(s) it may express. These typological data pertaining to parts of speech and dependent clauses are then combined in order to identify dependency relations between the functional behaviour of lexical and clausal constructions in languages. The results of these analyses are interpreted within the framework of functional linguistics, which assumes that grammars, i.e. linguistic forms and structures, are ultimately explainable in terms of their function: the encoding of human communication.

This book is organized as follows. The first part, to which Chapters 2, 3, and 4 belong, provides the theoretical and methodological background
of the research. Chapters 2 and 3 are devoted to the theory and typology of this study’s two *dramatis personae*: parts of speech and dependent clauses, respectively. First, in Chapter 2, I give an overview of possible approaches to parts of speech classification as proposed in the functional-typological literature. Two (relatively) recent approaches are discussed and compared in detail: Hengeveld’s pragmatic-syntactic theory of parts of speech and Croft’s universal-typological theory of parts of speech. The former serves as a point of departure for the present study, because it defines parts of speech exclusively in terms of the communicative function(s) that they can express (Hengeveld 1992, Hengeveld et al. 2004, Hengeveld & Rijkhoff 2005, Hengeveld & Van Lier 2008, 2009). The mono-dimensionality of this approach has been the target of its most important critic: Croft (2000, 2001, 2005). The latter proposes to define parts of speech in terms of a broader set of criteria, including their semantic meaning and their ability to express particular morphological categories. Croft’s approach yields a typology in which part of speech classes are not directly comparable across languages. On the other hand, the cross-linguistic comparability of parts of speech defined according to Hengeveld’s approach is achieved at the cost of ignoring specific sets of linguistic facts. In addition, the difference between the Hengeveldian and the Croftian approaches has important repercussions for the debate about so-called *flexible languages*, i.e. languages that presumably lack all parts of speech distinctions, most significantly a basic distinction between nouns and verbs. At the end of Chapter 2 I review this particular debate, and propose an approach to lexical flexibility that aims to integrate the insights of several earlier studies (see also Don & Van Lier, forthcoming).

In Chapter 3 I turn to the theoretical and typological treatment of dependent clauses, which is approached from two perspectives. First, it is shown that dependent clause constructions, like parts of speech, can be defined in terms of the communicative function(s) that they are able to express. Second, as mentioned above, dependent clauses are complex constructions, which can also be classified according to their internal morpho-syntactic properties. This second perspective, which is most common in the existing literature, focuses on the fact that dependent clause constructions often display a certain mixture of formal features of independent clauses on the one hand, and features of lexical or phrasal constructions on the other hand (cf. examples (4a-b) above). I briefly review the functional-typological literature concerning these two types of features and the various ways in which they may be combined in
actual dependent clause constructions. Two recent studies on the typology of the internal structure of dependent clauses, by Cristofaro (2003) and Malchukov (2004), are discussed in more detail. These formal features of dependent clauses, as I will show, are expected to influence the hypothesized functional dependency relation between these constructions and parts of speech classes. Therefore, at the end of Chapter 3, I present a classification system of dependent clause constructions that takes into account both their functional properties and their internal formal properties.

Chapter 4 rounds off the theoretical and methodological part of the book. In this chapter, I first present the composition of the language sample investigated for this research. Second, I formulate and operationalize the specific hypotheses that are tested on this sample. These hypotheses take the form of predictions about dependency relations between the functional patterns displayed by the part of speech classes of particular languages, and the functional patterns of their (different structural types of) dependent clause constructions. Finally, Chapter 4 presents the method that will be used to test these predictions.

The second part of the book contains the actual typological data. First, Chapters 5 and 6 separately present the classifications of part of speech classes and dependent clause constructions in the languages of the sample. In addition, a number of issues related to the analyses of both types of constructions are raised, and linked to the theoretical discussions of Chapters 2 and 3, respectively. Subsequently, in Chapter 7, the two data sets of Chapters 5 and 6 are combined in order to test the specific predictions about their interrelationship, as formulated in Chapter 4. This analysis yields a set of statistically significant correlations, which allow for the formulation of a number of typological generalizations. It is shown that these generalizations pertain in particular to languages with part of speech classes displaying a high degree of flexibility or multi-functionality (as opposed to languages with less or no lexical flexibility), and to the functional behaviour of deranked (as opposed to balanced) dependent clauses.

In the third and final part of the book I further discuss the results of the study and present its conclusions. First, in Chapter 8, the findings of Chapter 7 are reconsidered, taking a specific functionalist perspective that makes reference to the overall complexity of language systems and the various ways in which this level of complexity can be attained. To quote Hockett once more, “all languages have about equally complex jobs to do.” (Hockett 1958: 180) This ‘job’ – the encoding of human communication – can indeed be considered
universal. However, languages differ in the way in which they divide the labour of expressing the total amount of functional complexity over the available grammatical resources, including lexicon, morphology, and syntax. It seems that flexibility or multi-functionality in one area of the grammar must be counterbalanced or ‘traded off’ by functional specificity or rigidity in another area. In general, the functional flexibility of linguistic units tends to gradually decrease when they become structurally more complex. When flexibility is nonetheless retained in complex constructions, this requires functional disambiguation at the highest level of phrase-structure, either by means of ordering restrictions or through the use of function-indicating elements. In general, it will be shown in Chapter 8 that an explanatory framework based on functional transparency provides a complete account of the findings of this research. Finally, Chapter 9 summarizes the main points of the study as a whole and presents its overall conclusions.
2.1. Introduction

*Parts of Speech* (henceforth PoS) is the traditional term used to refer to the major classes of words or lexemes that are distinguished in a language (Schachter & Shopen 2007: 1). The issue of lexical classification or categorization has a long and troublesome history in linguistics. Generally, the nature of categorization in human language has been and continues to be the subject of a debate that goes back to Aristotle and has raised the interest of philosophers, psychologists, and linguists alike. In the present chapter, I will not attempt to give an overview of the various conceptions of categorization that have been on and off stage at various periods in the history of these sciences. Rather, I will limit myself to a brief discussion of the problems that have been at stake in defining lexical categories within the sub-discipline of linguistics called *functional typology*. This is the topic of section 2.2.

Subsequently, in section 2.3, I go on to discuss in more detail one specific functional-typological approach to lexical categorization, namely the theory proposed by Hengeveld (1992 a, b; Hengeveld et al. 2004; Hengeveld & Rijkhoff 2005; Hengeveld & Van Lier 2008, 2009). I will explain that

---

1 For a recent overview of this type, see Aarts 2006.
2 For other overviews of approaches to lexical categorization, including approaches that do not belong to the functional-typological framework, the reader is referred to Bisang (f.c.), Evans (2000), Haspelmath (2001), and Rijkhoff (2007).
the Hengeveldian approach provides a suitable basis for the present study, because it defines PoS classes in terms of the *propositional function(s)* they can fulfil in an utterance. These PoS definitions will serve as a basis for the typology of PoS systems presented in Chapter 5.

However, Hengeveld’s approach is certainly not without problems. In section 2.4 I discuss a number of objections that have been raised against this theory, most elaborately by Croft (2000, 2001, 2005). In general terms, these problems involve the tension between the language specific nature of lexical categories on the one hand, and their cross-linguistic comparability on the other hand. As such, they touch upon a fundamental discussion in linguistic typology concerning the (non-)existence of universal linguistic categories (see Haspelmath 2007, Newmeyer 2007).

Section 2.5 is specifically devoted to lexical flexibility, which for the time being I will define as the possibility to use the members of one or more lexeme class(es) of a language in more than one propositional function without needing any morpho-syntactic adaptation. First, I will review the discussion about lexical flexibility in recent literature. Subsequently, I present an approach to lexical flexibility that attempts to integrate insights of various earlier studies (see also Don & Van Lier, forthcoming). Finally, section 2.6 provides a summary of this chapter.

### 2.2. Functional-typological attempts at defining parts of speech

As the binary term implies, a functional-typological approach to the study of linguistics involves two basic ingredients: typology and functionalism. First, typology deals with cross-linguistic comparison, under the assumption that the structural variation displayed by languages across the word is not random, but is rather constrained by a set of principles that are called typological *universals*. Second, functionalism refers to the assumption that these universals of language structure can be explained in terms of language use, i.e. in terms of the semantic and pragmatic meanings expressed by these structures in human communication.

Within the functional-typological literature, various approaches to lexical categorization have been proposed. These can be characterised in terms of the type of criteria on which the definitions of PoS classes are primarily based.

---

1 The term *propositional function* has been introduced by Croft (2001). Its use in the present study will be clarified in the course of this chapter.

2 For a more detailed characterization of the functional–typological approach the reader is referred to Croft (2003), or Cristofaro (2003: 5–9).
According to Givón (2001: 49) there are three different types of such criteria:

(i) **Semantic criteria**: the kind of *meanings* that tend to be coded by words of a particular class.

(ii) **Morphological criteria**: the kind of *bound morphemes* that tend to be affixed to words of a particular class.

(iii) **Syntactic criteria**: the typical *syntactic slot* that words of a particular class tend to occupy.

To these, a fourth type may be added, namely:

(iv) **Pragmatic criteria**: the kind of *discourse function* that tends to be coded by words of a particular class. (cf. Sasse 1993b:196)

Although most approaches acknowledge the relevance of all four types of criteria for the definition of lexical categories, there are clear differences in emphasis. First, there is the so-called *notional approach*, which takes the semantic meaning of lexemes as the most important property of PoS classes, i.e. the types of entities that groups of lexemes denote. Thus, nouns are defined as lexemes denoting *first-order entities* or objects and individuals. In contrast, verbs denote *second-order entities*, i.e. “events, processes, states-of-affairs, etcetera, that are located in time and which are said to 'occur' or 'take place' rather than to 'exist'” (Lyons 1977: 443). Langacker (1987) provides a notional definition of nouns and verbs from the perspective of Cognitive Grammar. According to his definitions, nouns designate bounded *regions*, typically in the domain of space, and are therefore cognitively perceived as static and holistic. Verbs, on the other hand, designate *processes*, and are cognitively perceived as sequences of states evolving through time.

The obvious problem of the notional approach, as has been pointed out by many scholars (Dik 1989; Lehmann 1990, among others), is that the same type of entity may be expressed by lexemes that belong to different lexical classes, both within and across languages. The English word *destruction*, for instance, is a noun, even though it denotes an event rather than an object. Furthermore, property concepts such as ‘small’ are in some languages lexicalized in the form of a separate class of adjectives, while in other languages these concepts are expressed by verbs or nouns.

A second possibility is the *discourse approach*, advanced by Hopper and Thompson (1984, 1985). It defines the differences between PoS classes
in terms of their prototypical discourse function. Thus, a prototypical verb expresses event-occurrence, whereas a prototypical noun introduces a participant into the discourse. Whenever a lexeme does not express its prototypical discourse function, it will show reduced categoriality, meaning that it will have fewer possibilities to take morphological marking. However, as in the case of the notional approach, the discourse approach makes no principled distinction between universal discourse functions and the language-specific lexical categories that express them.

In an attempt to overcome the problem of the notional and discourse-based approaches, focus shifted more towards the morphological characteristics of PoS. For the classical languages, it had been noted long ago that lexical classes could be identified by means of the particular morphological categories for which they are specified. Typical nominal categories would include grammatical gender, number, definiteness and case distinctions. Verbs rather express categories like tense, mood, aspect, and person agreement. And adjectives take morphological markers expressing degrees of comparison and may display gender, number and/or case agreement with the head they modify.

As Croft (1991: 86) observes, the morphological categories mentioned above are indeed “tailored to” a specific lexical class, because they express bits of meaning that are relevant to the semantic meaning or the pragmatic-syntactic function prototypically expressed by members of that class. Thus, aspecual distinctions are relevant to verbs, which prototypically express actions; number is relevant to nouns, which prototypically refer to countable objects; and comparative marking is relevant to adjectives, which typically express gradable property concepts. Similarly, agreement has the function of showing the relationships between typically verbal predicates and their typically nominal argument(s), or between typically nominal heads and their typically adjectival modifiers.

However, morphological criteria are problematic too, because they are no more than “diagnostic” (Beck 2002:14) or “symptomatic” (Langacker 2002: 60) for lexical categories. For one thing, not all PoS classes in all languages express the same type and amount of morphological categories. Therefore, morphological criteria do not provide a suitable basis for cross-linguistic

---

5 As far as (manner) adverbs are concerned, it seems that there are no typical inflectional classes associated with this lexeme class. This is probably related to the fact that the status of manner adverbs as a major word class is somewhat problematic in the first place (see Haspelmath 2001: 16544).
generalizations. In addition, morphological categories are problematic as *intra*-linguistic criteria as well, because words belonging to the same lexical class often do not display the full range of morphological categories belonging to that class.

Schachter (1985) tries to come up with more integrated definitions of PoS classes. He argues that, apart from morphological criteria, the possible *syntactic* function(s) of lexemes should be taken into account as well.6 However, this combination of formal criteria still yields definitions of PoS classes that are language-specific rather than universal. At the same time, Schachter uses notional labels to refer to lexeme classes, which according to him, “appropriately reflect universal semantic considerations” (Schachter 1985: 4). This suggests that some kind of common prototypical meaning of the members of a lexeme class should be acknowledged and play a part in a universal theory of lexical categorization.

Anward et al. (1997: 171) also argue in favour of an approach to PoS definition that integrates semantic, morphological, and syntactic, as well as phonological criteria. However, they clearly state that “given that such properties of constituents differ from one language to another at least in part, parts-of-speech classes will need to be set up separately for each language [...]”. Again, this approach has obvious repercussions for the cross-linguistic comparability of PoS classes.

To sum up, there are several possible approaches to the definition of PoS classes, each with its own problems. These problems can be broadly divided into two types: either universal tendencies are defined without accounting for language-specific categories, or the other way around. In the next sections, I will discuss in more detail two recent functional-typological theories of PoS classification that deal with these problems in very different ways: Hengeveld’s syntactic-pragmatic approach (section 2.3) and Croft’s Universal-Typological approach (section 2.4). These sections, together with the present one, should make clear that there is no single answer to the question of whether and how the four types of criteria – pragmatic, semantic, syntactic, and morphological – should be combined into definitions of lexical categories within and across languages. As Bisang (fc.) puts it: “There is no final solution appearing on the horizon”.

---

6 See also Schachter & Shopen (2007), who explicitly distinguish between notional and grammatical criteria for defining PoS.
2.3 Hengeveld’s syntactic-pragmatic approach

2.3.1 Slots and fillers

Hengeveld (1992) developed his PoS theory as part of his typological study of non-verbal predication. He defines PoS classes in terms of the syntactic slot or slots that their members can occupy in an utterance, without any further morpho-syntactic measures. Hengeveld rules out semantics as a component of PoS definitions, because, as explained in the previous section, denotational meaning is not a water-tight criterion for lexical category membership. The expression of morphological categories is used as an analytic device only, which can help to identify possible candidates for lexeme classes in specific languages. It does not, however, play a role in the actual definition of PoS classes.

Working in the framework of Functional Grammar (henceforth FG; Dik 1989, 1997), Hengeveld distinguishes four possible syntactic slots in which lexemes may be inserted. Two of these slots are obligatorily filled in their respective phrases, namely:

(i) The head of a predicate phrase and
(ii) The head of a referential phrase

The other two slots are optionally filled in each of the two phrase types. They are:

(iii) A modifier within a predicate phrase and
(iv) A modifier within a referential phrase.

The four slots can thus be tabulated as in Figure 2.1:

<table>
<thead>
<tr>
<th></th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicate phrase</td>
<td>(i)</td>
<td>(iv)</td>
</tr>
<tr>
<td>Referential phrase</td>
<td>(ii)</td>
<td>(iii)</td>
</tr>
</tbody>
</table>

Figure 2.1: Syntactic slots as distinguished in Hengeveld (1992)

On the basis of their distribution over these four syntactic slots, the

\footnote{As I will explain shortly, in more recent work (Hengeveld & Van Lier 2008, 2009) the syntactic slots are re-interpreted as propositional functions.}
definitions of the major PoS classes verb, noun, adjective, and manner adverb are formulated as in (i):

(i)
(i) A *verb* is a lexeme that can be used, without further morpho-syntactic measures, as the head of a predicate phrase only (slot (i));
(ii) A *noun* is a lexeme that can be used, without further morpho-syntactic measures, as the head of a referential phrase (slot (ii));
(iii) An *adjective* is a lexeme that can be used, without further morpho-syntactic measures, as a modifier in a referential phrase (slot (iii));
(iv) A *manner adverb* is a lexeme that can be used, without further morpho-syntactic measures, as a modifier in a predicate phrase (slot (iv)).


In Figure 2.2 the four syntactic slots are represented once again; this time with the four lexeme classes appearing in the relevant slots (from Hengeveld et al. 2004: 530):

<table>
<thead>
<tr>
<th>Predicat phrase</th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referential phrase</td>
<td>noun</td>
<td>adjective</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predicate phrase</th>
<th>verb</th>
<th>manner adverb</th>
</tr>
</thead>
</table>

*Figure 2.2: Syntactic slots and lexical categories*

The definitions of the PoS classes in (i) give rise to a number of issues that require some clarification. First, there is quite a bit more to say about the proviso ‘*without further morpho-syntactic measures*’. For a proper understanding of what constitute such a measure, I find it useful to refer to a distinction made by Croft (1991, 2001), between *structural coding and behavioural potential*. The first term, *structural coding*, refers to those morpho-syntactic markers that specifically indicate the syntactic or propositional function of a linguistic unit (a lexeme, in the case at hand). The other type of marking, *behavioural potential*, involves the expression of morpho-syntactic categories that are semantically, syntactically and/or pragmatically relevant to the particular function in which a linguistic unit is used, but do not indicate that

---

8 In Croft’s earlier work (1991: 58) *structural coding* is termed “*function indicating morpho-syntax*”.
function as such (cf. Croft 2001: 66). Thus, structural coding can be thought of as a condition for using a linguistic (lexical) unit in a certain function, whereas behavioural potential is the consequence of using it in that function. Consider for example the Dutch noun *vlees* ‘meat’. If this lexeme is to express the function of modifier in a referential phrase, it needs the affix *-ig*. This affix is an instance of structural coding. However, when the resulting form *vlezig* ‘meaty’ is used to modify a definite neuter noun like *het gezicht* ‘the face’, then it takes the gender agreement suffix *-e* (*het vlezig-e gezicht*). This suffix is an instance of behavioural potential. It is crucial that in Hengeveld’s definitions of PoS classes only structural coding counts as further morphosyntactic measures. In contrast, behavioural potential corresponds to what has been termed ‘morphological criteria’ in section 2.2, and does not play a role in the definitions in (i) above.

Second, it is noteworthy that Hengeveld’s definition of verbs includes a special ‘only’-condition, which is absent in the definitions of the other three PoS. This is because in many languages nouns, adjectives, and adverbs can have an additional function next to their defining function as specified in (i) (ii)-(iv): They can also be used as the head of a predicate phrase9. This means that verbs are the only PoS class whose exclusive, distinguishing function it is to express the head of a predicate phrase; hence the ‘only’-condition in the definition of this lexeme class.

Cross-linguistically, the use of a non-verbal lexeme as the head of a predicate phrase (i.e. non-verbal predication) may trigger various formal expression strategies. Following Hengeveld’s typology of non-verbal predication, two main types of strategies are distinguished: those with a copula and those without one. In the first case, the copula counts as structural coding. The second strategy, without a copula, can be further sub-divided into two types, termed the zero-1 and the zero-2 strategy, respectively. In the case of a zero-1 strategy, the non-verbal predicate shows the same behavioural potential as a verbal predicate in the language under analysis. Thus, any categories that are expressed on a verb, such as TAM and person agreement, should also appear on a non-verbal predicate. In contrast, in the case of a zero-2 strategy the non-verbal predicate does not show the same behavioural potential as a verbal predicate, but rather remains uninflected and is simply juxtaposed to its argument10.

---

9 This is then termed a non-verbal predication, which was in fact the main topic of Hengeveld’s 1992 study. See further below.
A third point concerning the details of Hengeveld’s PoS definitions is that the class of adverbs includes manner adverbs only. This is because other types of adverbs do not modify the head of a predicate phrase but rather another unit. This may be a larger unit, such as a state of affairs, a proposition, an illocution, or an even larger stretch of discourse. Alternatively, it may be another modifying unit, such as an adjective or another adverb.

Taking into account the issues discussed above, the functions of the four PoS classes (verb, noun, adjective, and manner adverb) can be represented as in Figure 2.3, in which ‘+’ stands for the defining use of a lexical class (as mentioned in the left-hand column); ‘-’ for excluded uses; and ‘+/−’ for potential additional uses. The figure shows that verbs are uniquely used as the head of a predicate phrase, while nouns, adjectives, and manner adverbs may have additional uses as the head of a predicate phrase, next to their respective defining uses as the head of a referential phrase, a modifier in a referential phrase, and a modifier in a predicate phrase.

<table>
<thead>
<tr>
<th></th>
<th>Head of Predicate phrase</th>
<th>Head of Referential phrase</th>
<th>Modifier in Referential phrase</th>
<th>Modifier in Predicate phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbs</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Nouns</td>
<td>+/−</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Adjectives</td>
<td>+/−</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Manner adverbs</td>
<td>+/−</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
</tbody>
</table>

_Figure 2.3: The functional potential of parts of speech classes_

2.3.2 The cross-linguistic perspective

2.3.2.1 General introduction

Individual languages across the world do not all have systems of lexical categorization of the type represented in Figures 2.2 and 2.3 above, in

10 In many cases where a zero–2 strategy is used, a copula can optionally be added. The copula often remains un-expressed in ‘default’ cases, i.e. involving present tense and third person singular. In other, non-default cases, TAM and person distinctions are expressed on the copula. See also Chapter 5, section 5.
11 See Hengeveld & Mackenzie (2008, Chapter 3) for examples.
12 Regarding this latter point, the modification of lexical units other than verbs, Hengeveld & Van Lier (2008: 760–762) show that the definition of the class of lexemes functioning as a modifier in a predicate phrase must make reference to the PoS class of the head of that predicate phrase. That is to say, when the predicate phrase has a verbal head, it takes different types of adverbial modifiers (namely manner adverbs) than when it has a nominal, adjectival, or adverbial head (namely reference/referent modifiers or degree adverbs).
which there are four PoS classes, each of which is mapped neatly onto a single defining function. Rather, many languages display different types of mappings of their lexical categories onto the set of possible functions. Hengeveld’s syntactic approach to lexical categorization provides a method to compare these types of mapping systems across languages, called *parts of speech systems* (henceforth PoS systems).

According to Hengeveld’s typology, PoS systems come in two basic types: *rigid* and *flexible* ones. Rigid PoS systems are characterized by the fact that all lexical classes are rigid, i.e. specialized for the expression of a single function. When lexical items do appear in other functions, this must be marked by means of further morpho-syntactic measures or structural coding (with the possible exception of the predicative function, cf. the +/- signs in Figure 2.3). Rigid PoS systems may display all four lexeme classes that are represented in Figures 2.2 and 2.3. Alternatively, however, they may lack one or more PoS classes. This means that one or more function(s) cannot be expressed lexically, so that the relevant language must resort to an alternative, non-lexical strategy. In Garo NPs, for instance, the translational equivalent of an English adjective is always expressed by means of a relative clause, as illustrated in (2):

\begin{quote}
\textit{Garo} (Hengeveld et al. 2004: 531)\textsuperscript{13}
\end{quote}

\textit{(2) da'r-gipa} \textit{mande}
be.big-\textsc{ptc} man
‘the big man’ (lit. the man who is big/the man who ‘big-s’)

Thus, in a rigid PoS system all major lexeme classes are functionally specialized for the expression of a single function, even though there may not be a PoS class available for every function.

In contrast, flexible PoS systems are characterized by the fact that one or more of the available lexeme classes is flexible, meaning that its members can express two or more functions, without further morpho-syntactic measures. In Dutch, for instance, the translational equivalents of English adjectives

\textsuperscript{13} The suffix –\textit{gipa} in this example is not a lexical derivational suffix, as can be seen from example (i), in which the participial form marks a relative clause construction (Burling 2004: 301):

\begin{quote}
\textit{(i) [nok-\textit{o} p\textit{iu-a-ko} nik-gipa] metra}
\textit{house-\textsc{loc} child-\textsc{acc} see-\textsc{ptc} woman}
‘the woman who saw the child at the house’
\end{quote}
and manner adverbs belong to a single class of flexible lexemes that can be used without further measures as modifiers in both referential and predicate phrases. This is shown in examples (3) and (4), respectively.

### Dutch

(3) \[een \ mooi \ meisje\] Referential Phrase

\[
\begin{array}{c}
\text{a beautiful girl} \\
\text{‘a beautiful girl.’}
\end{array}
\]

(4) \[Het \ meisje \ zingt \ mooi\] Predicate Phrase

\[
\begin{array}{c}
\text{The girl sings beautiful} \\
\text{‘That girl sings beautifully.’}
\end{array}
\]

Notice that in the English translation of (4) the adjective beautiful does require a further measure in order for it to be usable in the function of modifier in a predicate phrase, namely the suffix -ly\(^{14}\).

2.3.2.2 The Parts of Speech Hierarchy (Hengeveld 1992)

Hengeveld (1992) shows that the two basic types of PoS systems, flexible and rigid, come in several variants across languages. He claims that this variation can be systematically described in terms of a Parts of Speech (PoS) Hierarchy, which is given in (5):

(5) Head of \(\subseteq\) Head of \(\subseteq\) Modifier in \(\subseteq\) Modifier in

\[
\begin{array}{c}
\text{Predicate} \\
\text{Referential} \\
\text{Referential} \\
\text{Predicate}
\end{array}
\]

\[
\begin{array}{c}
\text{phrase} \\
\text{phrase} \\
\text{phrase} \\
\text{phrase}
\end{array}
\]

Proceeding through this hierarchy from left to right, the chance decreases that a language has a separate, functionally specialized lexical class to express that particular function. For a language with a rigid PoS system this means that if it lacks a lexeme class for just a single function, this will be the rightmost function in the hierarchy: modifier in a predicate phrase. In other words, this language will have no class of manner adverbs. In addition, the

\(^{14}\) In Dutch, modifiers in referential phrases show gender agreement with the head noun in definite contexts. Thus, if in a phrase like (4a) the definite article *het* (rather than indefinite *een*) would be used, then *mooi* would carry the agreement marker \(\text{-e}\). However, as mentioned above, gender agreement does not count as ‘further measures’, since it constitutes behavioural potential rather than structural coding.
hierarchy is implicational. As such, it predicts that, if a rigid PoS system lacks a lexeme class for a specific function in the hierarchy, then it will also lack lexeme class(es) for any function(s) further to the right in the hierarchy. For example, a language without a lexeme class to express the function of modifier in a referential phrase will also lack a lexical class for the function of modifier in a predicate phrase. In other words, a language without adjectives is predicted to also lack manner adverbs.

For languages with flexible PoS systems, the hierarchy predicts that they are most likely to have a flexible lexeme class that can express the two rightmost functions: modifier in a referential phrase and modifier in a predicate phrase (as in the Dutch examples (3) and (4) above). Moreover, if in a flexible PoS system a lexeme class can express two particular functions in the hierarchy, it is predicted that this class will also be able to express any function(s) further to the right in the hierarchy, as well as any function(s) lying in between. For example, if a language has a class of lexemes that can be used as the head of a referential phrase and as a modifier in a predicate phrase, then it is predicted that this class will also be used for modification in a referential phrase. In other words, if a language has a class of lexemes that can function as nouns and as manner adverbs, then these lexemes should also be usable as adjectives.

In its original formulation (as given in (5) above), the PoS Hierarchy predicts seven different PoS systems, which are represented in Figure 2.4 below (Hengeveld 1992: 69; Hengeveld et al. 2004: 537). Three PoS systems are flexible, i.e. they include one lexeme class that can be used in multiple functions. The other four systems are rigid, i.e. they involve functionally specialized classes only. The labels for the flexible PoS classes are taken from Hengeveld et al. (2004): A contentive is a member of a maximally flexible class that can be used in all four functions; a non-verb can be used as the head and modifier in a referential phrase, and as a modifier in a predicate phrase; and a modifier is a lexeme that can be used as a modifier in both predicate and referential phrases.
2.3.2.3 Intermediate systems (Hengeveld, Rijkhoff & Siewierska 2004)

In more recent work by Hengeveld et al. (2004), the typology in Figure 2.4 is expanded in order to accommodate languages with so-called intermediate PoS systems (as first proposed by Smit (2001)). Such intermediate systems exhibit features of two ‘neighbouring’ PoS systems, i.e. two PoS systems that are on contiguous rows in Figure 2.4.

In a flexible intermediate PoS system, the lexeme class that is positioned rightmost in the hierarchy consists of two sub-classes: one containing basic members and the other derived members. The sub-class of derived items can express one function less than the subclass of basic items. In particular, the left-most function that can be expressed by the basic items cannot be expressed by the derived items. For example, Turkish has a PoS system with classes of basic verbs and basic non-verbs. The non-verbs, the right-most lexical class in the system, serve as the input category for derivation with the suffix -CA, which creates a class of modifiers, i.e. derived items that can be used as modifiers in referential and predicate phrases, but not as the head of a referential phrase. This is represented in Figure 2.5, in which the shading indicates the derived nature of the modifier class.
In (6a–c) the functional possibilities of the Turkish non-verb güzel are illustrated: in (6a) it functions as the head of a referential phrase, in (6b) as a modifier in a referential phrase, and in (6c) as a modifier in a predicate phrase.

Turkish (Göksel & Kerslake 2005: 49)

(6)  a. güzel-im
    beauty-1POSS
    ‘my beauty’

    b. güzel bir kopek
    beauty ART dog
    ‘a beautiful dog’

    c. güzel konuştu
    beauty s/he.spoke
    ‘S/he spoke well’

Examples (7a-b) show the distribution of CA-derived modifiers: in (7a) erkek-çe functions as a modifier in a referential phrase, and in (7b) yiy-çe modifies a predicative head.

Turkish (Kornfilt 1997: 92)

(7)  a. Erkek-çe bir ses
    man-MODIF a.voice
    ‘a manly voice’

    b. Hasan gömleği-i yiy-çe yıka-di
    Hasan shirt-ACC good-MODIF wash-PST
    ‘Hasan washed the shirt well.’

A rigid PoS system is classified as intermediate when its rightmost lexical category is a small, closed class, rather than a large, open one. For example, Alamblak has large open classes of rigid verbs and nouns, and a reasonably large class of adjectives (45 items). Its class of manner adverbs, however, comprises only 9 items (Bruce 1984: 87, 88). This is represented in Figure 2.6, in which the darker shading is used to indicate the small, closed nature of the manner adverb class.
The case of Alamblak points to the lack of a sharp distinction between ‘large, open’ and ‘small, closed’ word classes. The class of adjectives in Alamblak is much smaller than the classes of nouns and verbs, but also substantially larger than the class of manner adverbs. Also the notion of openness – the ease with which new members are added to a lexeme class – seems to be of a relative, rather than an absolute nature. As a result, there may be borderline cases, the classification of which inevitably involves a degree of arbitrariness. I will return to this issue in some more detail in Chapter 5 (section 5.4.2.1), where the PoS systems of the individual sample languages are classified.

The intermediate PoS systems predicted by the PoS Hierarchy in (6) are included in Figure 2.7, in which the cells with light shading are derived classes, and darker shaded cells represent small, closed lexical classes (cf. Hengeveld et al. 2004: 409).
2.3.2.3  More diversity (Hengeveld & Van Lier 2008, 2009)

Even more recent typological investigations make clear that the set of PoS systems predicted by the PoS Hierarchy, including the ‘extended’ set represented in Figure 2.7 above, is too restricted to account for the attested cross-linguistic variation (Hengeveld & Van Lier 2008, 2009). In particular, there are languages that display intermediate PoS systems that are not contiguous in Figure 2.4 above. Moreover, the ‘extra’ classes of these intermediate systems do not necessarily consist of derived members in flexible PoS systems, nor are they always small, closed classes in rigid PoS systems (cf. Figure 2.7 above).

Accommodating these findings, Hengeveld’s earlier model has recently been adapted in order to give it higher coverage, without seriously diminishing its accuracy. In other words, the new model predicts a higher amount of different types of PoS systems, and as such accounts for a larger range of variation attested in the world’s languages, without predicting too many PoS systems that have not (or at least not yet) been attested in any language.

Hengeveld and Van Lier (2008, 2009) argue that the four syntactic slots

<table>
<thead>
<tr>
<th>Parts of speech system</th>
<th>Head of Predicate phrase</th>
<th>Head of Referential phrase</th>
<th>Modifier in Referential phrase</th>
<th>Modifier in Predicate phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexible 1</td>
<td>contentive</td>
<td>non-verb</td>
<td>non-verb</td>
<td>non-verb</td>
</tr>
<tr>
<td>Flexible 1/2</td>
<td>contentive</td>
<td>non-verb</td>
<td>non-verb</td>
<td>non-verb</td>
</tr>
<tr>
<td>Rigid 2</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>manner adverb</td>
</tr>
<tr>
<td>Rigid 2/3</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>manner adverb</td>
</tr>
<tr>
<td>Rigid 3</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>–</td>
</tr>
<tr>
<td>Rigid 3/4</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>–</td>
</tr>
<tr>
<td>Rigid 4</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>–</td>
</tr>
<tr>
<td>Rigid 4/5</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>manner adverb</td>
</tr>
<tr>
<td>Rigid 5</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>–</td>
</tr>
<tr>
<td>Rigid 5/6</td>
<td>verb</td>
<td>noun</td>
<td>adjective</td>
<td>–</td>
</tr>
<tr>
<td>Rigid 6</td>
<td>verb</td>
<td>noun</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Rigid 6/7</td>
<td>verb</td>
<td>noun</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Rigid 7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*Figure 2.7: Parts of speech systems, including intermediate types*
distinguished in Figure 2.1 above should be re-interpreted as *propositional functions*\(^\text{15}\). These functions constitute dimensions of a space defined in terms of two parameters, each of which represents a dominance relation between two functional values. In addition, the two parameters stand in a dominance relation with respect to one another. Together, these hierarchical relations yield a set of predictions concerning possible mappings of PoS categories onto the space of propositional functions. The details of this model are presented in the remainder of the present sub-section.

The first and primary parameter involves the distinction between *predication* and *reference*\(^\text{16}\). These two functions are conceived of as the basic communicative functions, with which a speaker builds up the communicative content he or she wishes to express. It is assumed that the function of predication is privileged in relation to the function of reference, since the communicative act of referring presupposes that something is predicated of a referent. For instance, in a noun phrase like ‘*a yellow car*’ the properties ‘*car*’ and ‘*yellow*’ are predicated of the entity being referred to. The hierarchical relation between predication and reference is represented in (8):

\[\text{(8) Predication } \sqsubset \text{ Reference}\]

The second and also secondary parameter involves the distinction between *heads* and *modifiers*. Heads are obligatory and therefore primary in relation to optional modifiers\(^\text{17}\). This dominance relation is also shown by the fact that the lexical class of a modifier is dependent on the lexical class of its head (see note 12). The hierarchical relation between heads and modifiers is represented in (9):

\[\text{(9) Head } \sqsubset \text{ Modifier}\]

\(^{15}\) Croft (2001) considers modification as a propositional function on a par with predication and reference (see section 2.4.2 of the present chapter). In his earlier work (Croft 1991: 52), however, modification is defined as an “*accessory function to reference and predication*”. This latter definition is in accordance with the treatment of the modifier function in the Hengeveldian model.

\(^{16}\) Hengeveld and Van Lier (2008) use the term *ascription* where I use *predication*. The former term is used in the framework of *Functional Discourse Grammar* (FDG), which is the follow-up of Functional Grammar (FG), to distinguish between a propositional or interpersonal function and a semantic or representational function (see Hengeveld & Mackenzie 2008).

\(^{17}\) cf. Bloomfield’s (1962: 194-196) definition of *a head*: the element of a construction that can function like the whole construction.
In addition, as already mentioned, the two hierarchical relations in (8) and (9) are in turn ranked with respect to one another. In view of the fact that there are appositional languages that do not use modification at all, while there are no languages that do not display the predication-reference distinction (even though they may not do so at the level of lexical classification, but rather at the level of morpho-syntactic constructions), it is expected that the predication-reference parameter is primary in relation to the head-modifier parameter. This is represented in (10):

(10)  \(((\text{Predication/Reference}) \sqsubseteq (\text{Head/Modifier}))\)

On the basis of these three hierarchical relations, three implicational constraints can be formulated, as listed in (11)-(13):

(11)  Predication \sqsubseteq Reference

a.   If a language has a rigid class of lexemes that can be used as the head of a referential phrase, it must also have a rigid class of lexemes that can be used as the head of a predicate phrase.

b.   If a language has a flexible class of lexemes that can be used as the head of a referential phrase (but not as the head of a predicate phrase, since then the restriction becomes irrelevant), it must also have a flexible or rigid class of lexemes that can be used as the head of a predicate phrase.

(12)  Head \sqsubseteq Modifier

a.   If a language has a rigid class of lexemes that can be used as the modifier within a phrase, it must also have a rigid class of lexemes that can be used as the head of that phrase.

b.   If a language has a flexible class of lexemes that can be used as the modifier within a phrase (but not as the head of that phrase, since then the restriction becomes irrelevant), it must also have a flexible or rigid class of lexemes that can be used as the head of that phrase.

(13)  \(((\text{Predication/Reference}) \sqsubseteq (\text{Head/Modifier}))^{18}\)

If a language has distinct (rigid or flexible) classes of lexemes for
heads and modifiers within any phrase, then it must also have distinct (rigid or flexible) classes of lexemes for heads of predicate versus referential phrases.

The parameters in (8)-(10) and the constraints in (11)-(13) are accounted for in the layered implicational map in Figure 2.8 below. This figure shows that the Predication-Reference parameter ranks higher than the Head-Modifier parameter (as indicated by the symbol \( \cap \) in the centre), and that the Head-Modifier parameter applies in the domains of both Predication and Reference, which does not exclude the possibility that these domains share a single lexeme class.

\[ \begin{array}{ccc}
\text{Predication} & \subseteq & \cap \\
\cap & \cap & \cap \\
\text{Reference} & \subseteq & \cap \\
\end{array} \]

*Figure 2.8: The implicational map of parts of speech*

The map in Figure 2.8 predicts all seven systems that were also consistent with the original parts of speech hierarchy in (5). These systems are represented in (14)-(20) below. In addition, the new model predicts another 10 systems, which would have been counterexamples to the original model. These are represented in (21)-(30). In what follows, I briefly discuss each of the systems predicted by the map in Figure 2.8. Note that in all representations of PoS

18 It is explicitly not the case that the Head-Modifier distinction must apply within the predicative domain before it can apply in the referential domain. Thus the combination of (8)/(11) and (9)/(12) should not be read as (PredHead \( \subseteq \) PredModifier) \( \subseteq \) (RefHead \( \subseteq \) RefModifier). Thanks to Michael Cysouw for pointing this out.
systems the flexible PoS classes have dark shading, while rigid classes are in a lighter shade. When there is no PoS class available to express a particular function, then the cell corresponding to this function remains unshaded and is marked with a bar.

The system in (14) is consistent with the new model because it involves no lexical distinctions at all. All functions can be expressed by a single class, called contentives (cf. Figures 2.4 and 2.7 above).

(14)

The system in (15) is consistent because (i) it has a flexible lexeme class (non-verbs) that can be used as the head of a referential phrase and a rigid lexeme that can be used as the head of a predicate phrase (verbs) (cf. (11b)); (ii) it has a flexible class of lexemes that can be used as a modifier within a predicate phrase and a rigid class of lexemes that can be used as the head of that phrase (cf. (12b)); and (iii) it has distinct classes for heads and modifiers in the predicate phrase as well as distinct classes for heads of predicative versus referential phrases (cf. (13)).

(15)

The system in (16) is consistent because it involves a predication-reference distinction for heads (verbs and nouns) but not for modifiers. This is in accordance with constraints (11a), (12b) and (13).

(16)
The system in (17) is consistent with the model because it involves head-modifier distinctions in both the predicative and the referential domain. Thus it has *verbs, nouns, adjectives, and manner adverbs*.

\(17\)

<table>
<thead>
<tr>
<th></th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
<td>Manner adverb</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
<td>Adjective</td>
</tr>
</tbody>
</table>

The system in (18) differs from the one in (17) only in that it lacks a lexeme class for the function of modification in predicate phrases (*manner adverbs*), while the head function in this domain is expressed by *verbs*. This does not violate any of the constraints.

\(18\)

<table>
<thead>
<tr>
<th></th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
<td>–</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
<td>Adjective</td>
</tr>
</tbody>
</table>

The system in (19) again involves a predication-reference distinction in the head domain, but does not have lexical classes for either of the two modifier functions. This system conforms to (11a), since it has rigid classes of heads in both the reference and the predication domain (*verbs and nouns*), while constraints (12) and (13) are not applicable.

\(19\)

<table>
<thead>
<tr>
<th></th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
<td>–</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
<td>–</td>
</tr>
</tbody>
</table>

The system in (20), finally, has only one lexical class, namely one that fulfils the function of head of a predicate phrase (*verbs*). This does not violate any of the constraints.
These were the PoS systems predicted by both the original PoS hierarchy in (5) and the implicational map in Figure 2.8. As mentioned, the latter model predicts another 10 possible PoS systems, which would have presented counterexamples to the former model. These systems are represented in (21)–(29) below.

The system in (21) is in accordance with constraint (11b), since it has a class of flexible lexemes that can be used as referential heads (nominals) and a rigid class for predicative heads (verbs). It is also in accordance with (12a) in the predication domain, where it has rigid modifiers (manner adverbs) and rigid heads. Finally, it conforms to (13), since it has distinct classes for heads and modifiers in the predicate phrase, as well as distinct classes for heads in predicative versus referential phrases.

\[
\begin{array}{|c|c|}
\hline
\text{Head} & \text{Modifier} \\
\hline
\text{Predication} & \text{Verb} & \text{Manner adverb} \\
\text{Reference} & \text{Nominal} & \text{Nominal} \\
\hline
\end{array}
\]

The system in (22) is in accordance with constraint (11b), since it has a class of flexible lexemes that can be used as the head of a referential phrase (nominals), and a class of rigid lexemes for the function of head of a predicate phrase (verbs). Constraint (12) and (13) are not applicable.

\[
\begin{array}{|c|c|}
\hline
\text{Head} & \text{Modifier} \\
\hline
\text{Predication} & \text{Verb} & \text{–} \\
\text{Reference} & \text{Nominal} & \text{–} \\
\hline
\end{array}
\]

The system in (23) involves a lexical distinction between predication and reference, but not between heads and modifiers. The PoS classes in this
system are termed *predicatives* and *nominals* (see also Hengeveld & Van Lier 2009). The system is in accordance with (11b). Constraints (12) and (13) are not applicable.

(23)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicative</td>
</tr>
<tr>
<td>Reference</td>
<td>Nominal</td>
</tr>
</tbody>
</table>

The system in (24) conforms to constraint (11b), since it has a flexible class of lexemes that can be used predicatively (*predicatives*), while not having a class of lexemes that can be used referentially. Constraints (12) and (13) do not apply.

(24)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicative</td>
</tr>
<tr>
<td>Reference</td>
<td>–</td>
</tr>
</tbody>
</table>

The system in (25) involves a single flexible class of *heads*, i.e. lexemes that can be used as the head of both predicative and referential phrases. This system does not violate any of the constraints.

(25)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>–</td>
</tr>
<tr>
<td>Reference</td>
<td>Head</td>
</tr>
</tbody>
</table>

The system in (26) also involves a single flexible class, but one that can be used in all functions except modifier in a predicate phrase. This system does not violate any of the constraints\(^{19}\). For want of a good candidate, no special term

\(^{19}\) Note that the *lack* of a lexeme class for the function of modifier in a predicate phrase does not mean that there is a lexical head-modifier distinction within the functional domain of predication.
is proposed for this type of PoS class. It is simply indicated as ‘flexible class A’ (*Flex A*); its functional distribution can be read off the representation.

(26)

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head: Flex A</td>
<td>Modifier: –</td>
</tr>
</tbody>
</table>

The system in (27) is similar to the one in (26), except that the flexible class can now be used in all functions except modifier in a referential phrase. Again, no special term is introduced for this type of flexible lexeme class; it is indicated as *Flex B*.

(27)

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head: Flex B</td>
<td>Modifier: –</td>
</tr>
</tbody>
</table>

The system in (28) involves a rigid class of verbs and a flexible class of lexemes, termed *Flex C*, which can be used as the head of a referential phrase and a modifier in a predicate phrase, but not as a modifier in a referential phrase. As such, it is in accordance with constraint (11b), with (12b) in the predication domain, and with (13).

(28)

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head: Verb</td>
<td>Modifier: Flex C</td>
</tr>
<tr>
<td></td>
<td>Flex C</td>
</tr>
</tbody>
</table>

The system in (29) involves two ‘cross-wise’ flexible lexeme classes: One class (*Flex C*; see above) that can be used as the head of a referential phrase and as a modifier in a predicate phrase, while the other class, termed *Flex D*, can be used as the head of a predicate phrase and as a modifier in a referential phrase. This system does not violate any constraint.
Finally, the system in (30) has rigid lexeme classes for heads and modifiers in the predication domain (*verbs* and *manner adverbs*), while in the reference domain it has a rigid lexeme class for heads only (*nouns*). This system is in accordance with constraints (11a), (12a) and (13).

In sum, the implicational map model in Figure 2.8 predicts 17 possible PoS systems (of which 7 were also predicted by the original PoS hierarchy), which are expected to occur in actual languages. In addition however, the model predicts that a large number of logically possible PoS systems will not occur in any language. In particular, 34 out of 51 logically possible systems are excluded by the model, as they violate one or more of the constraints in (11), (12) and (13) above. For expositional reasons, these excluded systems are listed separately in Appendix II, where I specify for each system which constraint(s) it violates. In Chapter 5 the complete set of predicted and excluded PoS systems will be compared with the PoS systems that are (un-) attested in the sample languages.

At this point, it should be noted that the PoS systems presented above do not include any intermediate systems. It is expected, however, that the predicted systems can indeed be supplemented with additional PoS classes, as long as this does not violate the constraints in (11)-(13). As I mentioned earlier, these ‘extra’ classes need not be contiguous to the rightmost large basic class of a system, in terms of the original PoS hierarchy. Moreover, and again contrary to the predictions of the original hierarchy, extra classes are not necessarily derived, small and/or closed. Rather, they can also be large, open classes consisting of simple members. This will be illustrated in full.
detail in Chapter 5. However, to give a flavour of the attested cross-linguistic variation in PoS systems, I provide some examples below.

First, languages with flexible PoS systems may have extra classes that are large and open rather than derived. Santali, for instance, combines a maximally flexible PoS class of contentives with a large, open class of rigid verbs (Neukom 2001: 17). This is represented in Figure 2.9. Note that this system would have been a counterexample to the original theory, which predicts that an ‘intermediate’ system with contentives can only involve a class of derived non-verbs (cf. Figure 2.7).

![Figure 2.9: The PoS system of Santali](image)

Second, consider Garo, a language with a rigid PoS system. It has large open classes of verbs and nouns, no adjectives, a small class of simple manner adverbs, and a large class of verb-derived manner adverbs (Burling 2004). This is represented in Figure 2.10. The PoS system of Garo would have been a counterexample to the original PoS hierarchy, since there is no lexical class available for the function that is positioned between the functions of modifier in a predicate phrase and head of a referential phrase. Garo’s PoS system does not fit into the extended typology of PoS systems in Figure 2.7 either, since the latter does not take into account the possibility of derived classes in rigid systems.

![Figure 2.10: The PoS system of Garo](image)

These examples give a glimpse of the diversity encountered in the realm of PoS systems across languages, showing that the range of variation is much wider than Hengeveld’s original typology suggests.
2.3.3 Summary
In this section I have presented Hengeveld’s theory of lexical categorization, from its original formulation to its most recent version. I explained that the Hengeveldian approach has two crucial characteristics. The first one is the exclusion of semantics as a criterion for the definition of PoS. The second important component involves the ‘further morpho-syntactic measures’ that do or do not accompany the use of a particular group of lexemes in a particular functional environment. Such further measures include structural coding only; behavioural potential, or morphological criteria, are not taken into account.

Hengeveld’s approach avoids (but does not provide a solution to) the problems associated with semantic and morphological criteria for PoS definitions, as pointed out in section 2.2. In particular, the Hengeveldian method allows for a straightforward cross-linguistic comparison of the various ways in which languages may map groups of lexemes onto a set of propositional functions. However, this convenient situation comes at a cost. In the next section I discuss the problems associated with Hengeveld’s method, and the alternatives that have been proposed for it.

2.4 Critique on the syntactic approach

2.4.1 Introduction
The most elaborate and critical alternative to Hengeveld’s approach is Croft’s *Universal-Typological Theory of Parts of Speech* (2000, 2001). The crucial difference between the two approaches is that Croft’s theory, unlike Hengeveld’s, takes into account semantic and morphological criteria. According to Croft, leaving out these two types of linguistic data implies missing out on the cross-linguistic generalizations that apply to them.

First, in section 2.4.2, I start with a brief outline of Croft’s theory. Subsequently, in section 2.4.3, I discuss in more detail the specific objections that Croft has raised against Hengeveld’s approach. In fact, these objections are not confined to Hengeveld’s PoS theory. Rather, they target any theory of linguistic categories that fails to take into account the full range of linguistic evidence, and that fails to acknowledge that such a complete distributional analysis inevitably leads to the conclusion that cross-linguistically identical categories do not exist.20

20 For a recent general assessment of this fundamental issue in linguistic typology, the reader is referred to Haspelmath (2007) and Newmeyer (2007).
Finally, one particular aspect of Croft’s critique on Hengeveld’s method concerns the role of semantics in languages with flexible word classes. Since this is a complex issue that has been the subject of a rather extensive discussion in recent literature, I treat it separately in section 2.5.

### 2.4.2 Croft’s Universal-Typological Theory of Parts of Speech

Croft (2000, 2001) distinguishes three functions that lexemes may express: predication, reference, and modification. These are termed propositional act functions. Although this resembles the Hengeveldian model, in the latter the modifier function is treated as a dimension of the head-modifier distinction, which cross-cuts the predication-reference distinction, rather than as a basic function in itself (cf. note 15).

A much more important difference between the two theories is that Croft divides the lexemes that may fulfil the three propositional act functions into three different semantic classes: objects, properties, and actions. These semantic classes are in turn defined in terms of the following four semantic properties (Croft 2001: 87):

1. **Relationality**: whether a definition of a concept inherently requires reference to another concept;
2. **Stativity**: whether the concept represents a state or a process;
3. **Transitoriness**: whether the concept represents a transitory state or process or an inherent or permanent state of the entity in question;
4. **Gradability**: whether the entity is graded along a scalar dimension.

The definitions of the semantic classes in terms of these properties are presented in Figure 2.11 below (taken from Croft 2001: 87)

<table>
<thead>
<tr>
<th></th>
<th>relationality</th>
<th>stativity</th>
<th>transitoriness</th>
<th>gradability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objects</strong></td>
<td>non-relational</td>
<td>state</td>
<td>permanent</td>
<td>non-gradable</td>
</tr>
<tr>
<td><strong>Properties</strong></td>
<td>relational</td>
<td>state</td>
<td>permanent/</td>
<td>gradable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>transitory</td>
<td></td>
</tr>
<tr>
<td><strong>Actions</strong></td>
<td>relational</td>
<td>process</td>
<td>transitory</td>
<td>non-gradable</td>
</tr>
</tbody>
</table>

*Figure 2.11: Properties of semantic classes*

In principle, lexemes belonging to any of the three semantic classes (objects, properties, and actions) may express any of the three propositional act functions (reference, modification, and predication). However, the
assumption of the prototype approach (first formulated in Croft 1991) is that the semantic classes of objects, properties, and actions are the *typological prototypes* of referring, modifying, and predicating constructions, respectively. Croft thus defines PoS as prototypical or *unmarked* combinations of semantic class and propositional function. A prototypical noun denotes an object and functions referentially; a prototypical adjective denotes a property and functions as a modifier; and a prototypical verb denotes an action and functions predicatively. This is represented in Figure 2.12 (adapted from Table 2.3 in Croft 2001: 88). Any combination of propositional act function and semantic class other than the three prototypical ones represented in Figure 2.12 is claimed to be typologically *marked*.

<table>
<thead>
<tr>
<th></th>
<th>reference</th>
<th>modification</th>
<th>predication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objects</strong></td>
<td>Unmarked nouns</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Properties</strong></td>
<td></td>
<td>Unmarked adjectives</td>
<td></td>
</tr>
<tr>
<td><strong>Actions</strong></td>
<td></td>
<td>Unmarked verbs</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 2.12: Nouns, adjectives and verbs as prototypical combinations of propositional act functions and semantic classes*

Importantly, although Croft defines PoS as typological prototypes, rather than as language-specific categories, this by no means implies that he regards PoS-typology as a pointless enterprise. On the contrary, the conceptual notion of typological markedness (versus unmarkedness or prototypicality), as it applies to the combinations of propositional functions and semantic categories, correlates cross-linguistically with formal phenomena. In

---

21 In an earlier version of his theory (1991), Croft uses different labels for some of the semantic properties listed in Figure 2.11. Pustet (2000, 2003) draws heavily on Croft’s work, but she proposes a slightly different semantic system. Apart from the property of *transitivity/valency*, which is the same as Croft’s (1991), she introduces a new semantic parameter, namely *energy release* (renamed *dynamicity* in Pustet 2003) and leaves out Croft’s parameter of *gradability*. The latter is apparently integrated, together with *perfectivity* and *processuality* (corresponding to Croft’s *stativity* and *persistence/transitoriness*), into the notion of *transcience*. Just for the sake of clarity, this rather confusing terminological situation is represented in the schema below:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>valency</td>
<td>-</td>
<td>relatiology</td>
<td>transitivity</td>
<td>transitivity/valence</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>energy release</td>
<td>dynamicity</td>
<td></td>
</tr>
<tr>
<td>persistence</td>
<td>transitoriness</td>
<td>perfectivity</td>
<td>-</td>
<td>transcience</td>
</tr>
<tr>
<td>stativity</td>
<td>stativity</td>
<td>processuality</td>
<td>transcience</td>
<td></td>
</tr>
<tr>
<td>gradability</td>
<td>gradability</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
particular, Croft identifies two implicational universals, which constrain cross-linguistic structural variation in terms of both structural coding and behavioural potential.

The first universal, pertaining to structural coding, predicts that a typologically marked combination of pragmatic function and semantic class must be encoded by at least as many morphemes as an unmarked combination. The second universal makes the opposite prediction for behavioural potential: It states that the range of behavioural potential displayed in the case of an unmarked combination of pragmatic function and semantic class must be at least as wide as the range displayed in the case of a marked combination (Croft 2001: 90, 91). Notably, both universals allow for the possibility of a typologically marked combination showing exactly the same amount of morpho-syntactic measures as an unmarked combination, including zero-marking. What is excluded is a situation in which a marked combination shows less structural coding or more behavioural potential than an unmarked combination.

In sum, Croft’s PoS theory defines a universal conceptual space that is made up of a pragmatic (propositional) and a semantic (denotational) dimension. Onto this conceptual space, a set of language-specific lexical categories is mapped, which are identified on the basis of a distributional analysis of formal criteria, including both structural coding and behavioural potential. Croft’s theory and Hengeveld’s theory thus resemble one another to the extent that both are concerned with form-function mappings (or ‘construction-filler interactions’, as Croft (2005: 437) calls them). However, their approaches differ in two fundamental respects. First, Croft’s functional space includes a semantic dimension, which is lacking in the Hengeveldian model. Second, Croft is interested not only in structural coding – as Hengeveld is in further morpho-syntactic measures –, but also takes into account all other distributional properties of lexemes, i.e. their behavioural potential.

According to Croft, the theoretical repercussions of using only a subset of morpho-syntactic criteria are serious. He argues that constructions or

---

22 As far as structural coding is concerned, this would amount to Hengeveld’s definition of flexibility.

23 Note that the amount of marking is measured in terms of the number of morphemes (Croft 2001: 90, 91). This is not a very straightforward method, though, since it is not clear how to treat for instance portmanteau morphemes. Therefore, it would probably be better to count the number of semantic categories expressed on the lexeme, either by bound (agglutinative or fusional) morphemes or by analytic grammatical particles.
categories (including PoS classes) identified in this way are “epiphenomenal” and that “one will not find language universals that way” (Croft 2005: 436). In what follows, I discuss these claims in some more detail.

2.4.3 Lexical categories within and across languages: The generality and the subclass problems

Croft’s main objection to Hengeveld’s PoS theory is that it is selective about which distributional tests are taken into account (structural coding) and which are not (behavioural potential). According to Croft, Hengeveld’s approach amounts to “methodological opportunism”, which can be avoided only by taking into account the complete set of distributional facts in each individual language. This, he argues, will reveal “covert categories, that is, categories not obligatorily flagged by overt structural coding” (Croft 2001: 75; 2005: 434)\(^{24}\).

However, the approach advocated by Croft raises two potential problems, which are in close relation to each other, and which have been termed the subclass problem and the generality problem (Haspelmath 2001). The former problem is concerned with intra-linguistic lexical categorization; the latter with inter-linguistic comparability of lexical categories.

Starting with the subclass problem, if one takes into account the full set of distributional facts about groups of lexemes in a specific language, this gives rise to a “myriad” of different types of sub-categorizations, without any means to decide which of these categories are to be considered ‘PoS classes’ and which ones are not (Croft 2001: 78ff, Haspelmath 2001: 16540). This problem is also aptly described by Schachter and Shopen:

“It must be acknowledged […] that there is not always a clear basis for deciding whether two distinguishable open classes of words that occur in a language should be identified as different parts of speech or as subclasses of a single part of speech. The reason for this is that the open parts of speech classes must be distinguished from one another on the basis of a cluster of properties […]. Typically there is some overlap, some sharing of features, as well as some differentiation. […] What this means is that there may in some cases be considerable arbitrariness in the identification of two open word classes as distinct parts of speech.

\(^{24}\) Croft (2001: 73) also points out that a PoS theory like Hengeveld’s, which lacks a semantic dimension, is bound to overlook regularities in the relevant domain, such as those applicable to semantic shifts in flexible lexemes. This issue will be discussed further in section 2.5.
rather than subclasses of a single part of speech.” (Schachter & Shopen 2007: 4; emphasis in the original)

Croft (2001: 81-83, following Uehara 1998) illustrates this point with an example from Japanese, which I will discuss here in some detail. In Japanese, the boundary between nouns and adjectives is not clear-cut. Traditionally, three categories are distinguished: ‘nouns’, ‘adjectives’, and ‘nominal adjectives’. ‘Nouns’ can function without structural coding as the head of a referential phrase. When they are used as the head of a predicate phrase they need a copula, and when they are used as a modifier in a referential phrase they take a genitive particle no. ‘Nominal adjectives’ behave in the same way as nouns when functioning as the head of a referential phrase or the head of a predicate phrase, but they take the linking particle na when used in modifier function. ‘Adjectives’ can not be used as heads in referential phrases. In contrast, they can be used without structural coding as referential modifiers and as heads of predicate phrases. In both functions, the behavioural potential of adjectives consists of tense-marking: –i for non-past and –katta for past (see Hinds 1986: 44)25.

However, a number of other lexical classes may be distinguished in Japanese with distributional patterns that seem to fall somewhere in between the patterns of ‘nouns’, ‘nominal adjectives’, and ‘adjectives’. First, there is a class of lexemes that sometimes behave like nouns and sometimes like nominal adjectives. In particular, when functioning as modifiers, these items can appear either with the genitive marker no, like nouns (see (31a)), or with the linker na, like nominal adjectives (see (31b)):

Japanese (Croft 2001: 82)
(31) a. heiwa no sisya
   peace GEN messenger
   ‘a messenger of peace’

   b. heiwa na kuni
   peace(ful) LK country
   ‘a peaceful country’

25 In Chapter 5 I will interpret the fact that ‘adjectives’ inflect for tense when used as modifiers as evidence that these lexemes are in fact verbs that may be used as dependent predicates in (zero-marked) relative clauses. At present, however, I will follow the argument as presented by Croft.
The members of another such 'in-between class' behave partly like nominal adjectives and partly like adjectives. One part occurs either with a copula when used as the head of a predicate phrase, or without a copula and with tense-marking (see (32a)). In modifier function these items can take either the linker \textit{na} or zero structural coding with tense inflection (see (32b)).

\begin{itemize}
\item[(32)]
\begin{enumerate}
\item \textit{ataka-}i / \textit{ataka da}
\begin{itemize}
\item warm-\texttt{PRS} / warm \texttt{COP}
\end{itemize}
\end{enumerate}
\end{itemize}

\begin{itemize}
\item[(32)] \begin{enumerate}
\item \textit{ataka-}i \ \textit{bi} / \textit{ataka na} \ \textit{bi}
\begin{itemize}
\item warm-\texttt{PRS} \ \texttt{day} / warm \texttt{Lk} \ \texttt{day}
\end{itemize}
\end{enumerate}
\end{itemize}

'\textit{It is warm}.'

'a warm day''

The second part of this subclass displays alternative coding strategies only in the modifier function (\textit{na} versus zero structural coding and tense marking, as in (33a), cf. (32b)). When used as the head of a predicate phrase the relevant lexemes always occur without a copula and with tense marking (see (33b) and cf. (32a)).

\begin{itemize}
\item[(33)]
\begin{enumerate}
\item \textit{tiisa-}i \ \textit{hon} / \textit{tiisa na} \ \textit{hon}
\begin{itemize}
\item small-\texttt{PRS} \ \texttt{book} / small \texttt{Lk} \ \texttt{book}
\end{itemize}
\end{enumerate}
\end{itemize}

'a small book'

'\textit{It is small}.'
These distributional facts are summarized in Figure 2.13:

<table>
<thead>
<tr>
<th>PoS class</th>
<th>Function/Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Head of Pred Phrase</td>
</tr>
<tr>
<td>Noun</td>
<td>da (COP)</td>
</tr>
<tr>
<td>Noun/nominal adjective</td>
<td>da (COP)</td>
</tr>
<tr>
<td>Nominal adjective</td>
<td>da (COP)</td>
</tr>
<tr>
<td>Nominal adjective/adjective 1</td>
<td>da (COP)/Ø + Tense</td>
</tr>
<tr>
<td>Nominal adjective/adjective 2</td>
<td>Ø + Tense</td>
</tr>
<tr>
<td>Adjective</td>
<td>Ø + Tense</td>
</tr>
</tbody>
</table>

*Figure 2.13: The noun–adjective distinction in Japanese*

The crucial question is now whether one should interpret all these different distributional patterns as evidence for as many PoS classes in Japanese. If so, then a Hengeveldian typology of PoS systems would be split up into a potentially unlimited number of sub-classifications. For Croft’s theory, however, this problem is irrelevant: From his point of view, it is simply not necessary to decide whether the classes identified above are nouns or adjectives, or both, or neither. They just represent a set of language-specific formal categories mapped onto a universal functional space.

A direct consequence of the subclass problem is the generality problem: lexical categories like the Japanese ones identified in Figure 2.13 are not comparable with lexical categories attested in any other individual language. Since languages differ in terms of the number and type of relevant distributional criteria, a lexeme class that is defined on the basis of such criteria in one language can never be identical to a lexeme class in another language, which expresses different morpho-syntactic distinctions. Again, this is a non-issue in terms of Croft’s prototype approach. The semantic dimension of his theory accounts for cross-linguistic generalizations about the ways in which individual languages cut up the universal functional space cake: the less prototypical the combination of semantic meaning and propositional function, the more structural coding will be required, and the less behavioural potential will be displayed.

The Hengeveldian approach, in contrast, ignores fuzzy boundaries between lexeme classes, such as between nouns and adjectives in Japanese. According to this approach, the relevant facts for Japanese are that it has
(i) a class of lexemes that can be used without structural coding as the head of a referential phrase (i.e. Hengeveld’s definition of nouns), and (ii) a class of lexemes that can be used without structural coding as a modifier in a referential phrase (i.e. Hengeveld’s definition of adjectives). Nouns and adjectives (and other PoS) defined along these lines are not universal in the way that Croft’s prototypes are universal, and they are not language-specific in the Croftian sense either (i.e. they are not based on a complete distributional analysis of a the particular language system of Japanese). However, these PoS are cross-linguistically comparable, even though they can be considered identical across languages only as far as Hengeveld’s narrow (or should one say ‘broad’) definitions go. Adding more analytical primitives, as Croft does, inevitably leads to more categorial differentiation.

### Summary, outlook

In sum, Croft’s method of identifying lexical categories gives rise to the subclass problem and the generality problem, which concern the intra-linguistic definition and inter-linguistic comparability of PoS classes. However, Croft avoids these problems by defining PoS as prototypes rather than as language-specific categories, and by identifying cross-linguistic generalizations about patterns of relative formal markedness, which make reference to both structural coding and behavioural potential. In contrast, Hengeveld’s approach to lexical categorization avoids the subclass and generality problems by looking at a restricted set of distributional data, namely at the presence versus absence of structural coding accompanying the use of particular groups of lexemes in particular propositional functions.

In section 2.5 I turn to another problematic issue that has been raised in relation to Hengeveld’s PoS theory. This issue concerns the absence of a semantic dimension in relation to the problem of defining *lexical flexibility*. In particular, since the Hengeveldian theory does not take into account the semantic denotation of lexemes, it has nothing to say about the semantic interpretation of supposedly flexible lexemes in the various functions that these items may express. However, claims about the existence of lexical flexibility (in general as well as with reference to particular languages) have often been challenged on semantic grounds: Even though presumed flexible lexemes do not change formally when used in various functions, they may have different (if related) meanings in each of those functions. Scholars disagree about the origin and theoretical significance of such semantic shifts. In addition to the problem of semantic interpretation, a number of other
issues have entered the discussion about lexical flexibility. The following section reviews this discussion.

2.5 Defining lexical flexibility

2.5.1 Introduction
This section addresses the debate surrounding so-called flexible languages. As a point of departure for the discussion, I use the three criteria for lexical flexibility proposed by Evans and Osada (2005). The first and most complicated criterion, to be discussed in section 2.5.2, is concerned with semantic (non-)compositionality, i.e. with the semantic interpretation of presumably flexible lexemes in different functional environments. In section 2.5.3 I take up the criterion of exhaustiveness, which relates to the problem of quantitative measurement of lexical flexibility. Finally, in section 2.5.4, I discuss the criterion of equivalent combinatorics, which pertains to the issue of bi-directional lexical flexibility.

2.5.2 The criterion of compositionality: flexibility and semantic shift

2.5.2.1 Evans and Osada’s proposal and some alternatives
Presumably flexible lexemes, i.e. lexemes that can be used without further measures in two or more propositional functions, receive slightly different semantic interpretations, depending on the particular function in which they appear. Although these meanings are (closely) related, particular meaning shifts may differ in terms of their degree of regularity: while some are apparently very predictable, others seem to be rather more idiosyncratic.

Evans and Osada’s (2005) criterion of semantic compositionality states that, for a lexeme to be ‘truly’ flexible,

“[A]ny semantic differences between the uses of a putative ‘fluid’ [i.e. ‘flexible’, EvL] lexeme in two syntactic positions must be attributable to the function of that position.” (Evans & Osada 2005: 367)

A corollary of this criterion is that the meaning shift, or what Evans and Osada call the coercion effect, should be the same for any member of a flexible lexeme class that is used in a particular function, apart from possible semantic interactions attributable to behavioural potential that belongs to the function, such as for instance aspect markers. According to Evans and
Osada, if the compositionality criterion is not satisfied, i.e. if the semantic interpretation of a flexible lexeme in different functional environments is not entirely compositional, then what looks like flexibility should in fact be analyzed as ‘rampant’ zero-conversion: large-scale derivation from one lexical category to another without any formal reflection.

Similarly, Croft (2000, 2001, 2005) argues that cases of non-compositional semantic shifts are not instances of flexibility. Rather, he argues, they are instances of polysemy, i.e. they involve two or more related meanings associated with a single form. Since the relation between these various meanings is not predictable, each of them must be stored individually in the lexicon, and must be labelled as appropriate in a particular functional environment. This amounts to the same kind of categorization that is found in languages with rigid PoS systems.

As we have seen, Hengeveld’s original PoS theory makes no reference to semantics, and therefore it does not account for meaning shifts in flexible lexemes. In Croft’s words: “Hengeveld ignores what happens to a lexical root’s meaning when used in more than one function” (Croft 2000: 69). This point is illustrated with examples from Tongan, which were adduced by Hengeveld (1992: 66) as evidence for analyzing Tongan as a flexible language. According to Hengeveld, the examples in (34) demonstrate that the lexeme *si‘i* is a flexible lexeme, because it can be used without further measures as the head of a predicate phrase (34a), the head of a referential phrase (34b), and a modifier in a referential phrase (34c). In contrast, Croft argues that this analysis is problematic, since it does not account for the different meanings of *si‘i* that are associated with each of its various uses.

(34)  a.  [Na‘e si‘i]  ‘ae  ako  
    PST  small  ABS  school:DEF  
    ‘The school was small.’

    b.  ‘i  [‘ene  si‘i]  
    in  3SG.POSS  childhood:DEF  
    ‘in his/her childhood’

Note in passing that *ako* presumably also represents an instance of lexical flexibility.
In response to Croft’s critique, Hengeveld and his colleagues argue that there is no need to assume multiple meanings for flexible lexemes like si’i (Hengeveld et al. 2004, Hengeveld & Rijkhoff 2005). In other words, they regard flexible lexemes as monosemous rather than polysemous. Meaning shifts, they argue, come about because the use of a flexible lexeme in a particular function activates certain aspects of the lexeme’s general meaning. This view is clearly stated in the following quotation:

“[E]ach flexible lexeme has a single (vague) sense. By placing the flexible lexeme in a particular slot or by providing it with certain morphological markers\textsuperscript{27}, the speaker highlights those meaning components of the flexible lexemes that are relevant for a certain lexical (verbal, nominal, etc.) function. Thus we contend that the meaning of a flexible lexeme always remains the same, and that morpho-syntactic and other contextual clues signal to the addressee how to interpret this lexeme in an actual utterance. In other words, it is the use of a vague lexeme in a certain context (an actual linguistic expression) that brings out certain parts of the meaning, giving the category neutral lexeme a particular categorical (verbal, nominal, etc.) flavour.” (Hengeveld et al. 2004: 541)

However, Croft (2001: 71, 72) argues that problems with this vagueness approach arise in defining the meaning of a lexeme in such a way that it is neutral with respect to the various interpretations associated with it in each of its possible functions. He thus claims that it is impossible to come up with a definition that covers all and only the meanings of a flexible lexeme like Tongan si’i. A definition such as ‘some concept associated with smallness’ is too general because it may include all kinds of concepts that have the property of ‘smallness’. As mentioned above, the only alternative, according to Croft, is to specify by convention two or more different meanings of the lexeme in question, and to store these meanings together with the function to which they belong. This analysis undermines the very notion of flexibility.

\textsuperscript{27} The term ‘morphological markers’ refers to \textit{behavioural potential}, not to \textit{structural coding}; the presence of the latter would obviously block the analysis as a flexible lexeme. (Kees Hengeveld, p.c.)
Considering the Tongan data in (34) in light of Evans and Osada’s criterion of semantic compositionality, one may argue that the shift in the meaning of \textit{si’i} in (34a) (‘be small’) versus (34c) (‘small’) is indeed attributable to the different functional environments in which the lexeme appears, namely predication and attribution, respectively. In contrast, this type of analysis cannot easily be applied to (34b), where \textit{si’i} means ‘childhood’ (or ‘period in one’s life during which one is small’). This meaning does not follow predictably from the fact that \textit{si’i} functions as the head of referential phrase in this example. Meanings such as ‘smallness’ or ‘small one’ appear to be at least as plausible in this function. This would suggest that Tongan does not meet the compositionality criterion, at least not across the board.

However, Hengeveld and Rijkhoff (2005) also disagree with Evans and Osada’s assumption that a functional environment \textit{adds} semantic content to a flexible lexeme that it did not have before it was used. As mentioned, Hengeveld and his colleagues claim that flexible lexemes are semantically \textit{vague} (cf. the above quotation). This means that both the verbal and the nominal senses are already contained in the semantics, and that the functional slot “\textit{only highlights meaning components that are already there} (\ldots)” (Hengeveld & Rijkhoff 2005: 514). Thus, according to Hengeveld and Rijkhoff, non-compositional meaning shifts are not an argument \textit{against} flexibility. To the contrary, they say that semantic idiosyncrasies should be expected, since each lexeme has its own unique set of meaning components, which in combination with a particular functional environment yields the intended semantic interpretation.

This line of thought links up with work in the framework of Cognitive Grammar, as developed by Langacker (1987). As mentioned briefly in section 2.2, in Cognitive Grammar the meanings of lexemes are defined as image-schematic concepts. Verb meanings are conceptualized as temporal sequences and are therefore \textit{sequentially} scanned. Nouns, on the other hand, are conceptualized as bounded regions. If the meaning of a noun includes the notion of some temporal unfolding (like a verb), then it is scanned \textit{summarily}, as a whole, rather than \textit{sequentially}. Within this cognitive framework, flexible lexemes can be defined as having a single conceptual structure that is compatible with either a verbal or a nominal meaning. Depending on the functional frame in which the lexeme is used, a profiling of the concept is triggered that corresponds to the appropriate semantic interpretation of the word.
Farrell (2001) applies this cognitive approach to English lexemes. Even though English is probably most often thought of as a language with rigid PoS classes, it has long been noted that many lexemes in this language can indeed be used as nouns and verbs without any overt derivation (Clark & Clark 1979)\(^2\). Farrell argues that any English lexeme – irrespective of the behaviour of other lexemes in the language – that can occur in two different functions without derivational marking does not belong to a part of speech until it is inserted into a syntactic frame. In other words, he claims that at least some proportion of the English lexicon is flexible.

Consider for instance the lexeme *kiss*. Farrell analyzes this item as having a conceptual structure that defines contact between the lips of one entity, the agent, with another entity, the patient. This is shown in the upper box of Figure 2.14 (taken from Farrell (2001: 114, Figure 2)). When *kiss* is used as the head of a predicate phrase, its action-interpretation is profiled, as represented in the lower left box of Figure 2.14 by the bold lines of the temporal t-axis and the oval and circle designating the participants. If, on the other hand, *kiss* is used as the head of a referential phrase, then the abstract region defining the overall event is profiled, as represented in the lower right box by the bold print of the outermost oval.

According to Farrell, *kiss* is an example of a so-called *process-centred* word. There are also *thing-centred* words. An example of the latter would be *hammer*. The analysis of thing-centred words differs to the extent that using the lexeme in referential function triggers the profiling of one of the entities involved in the conceptual representation – the *hammer* in this case –, rather then the overall process, as in the case of *kiss*. As Farrell himself admits, however, there are many other possible types of semantic shifts, which do not correspond to the two types just discussed.

In fact, this seems to be exactly what also constitutes the flaw in Hengeveld and Rijkhoff’s analysis of flexibility: Even if both the lexeme’s particular semantic features and its functional environment are taken into account, it is not always predictable exactly which meaning elements are to be profiled in a particular function. On the other hand, an advantage of Hengeveld and Rijkhoff’s vagueness analysis and Farrell’s cognitive analysis is that neither of them needs to posit a process of *conversion*, that is,

\(^2\) See also Evans & Osada (2005) (and the discussion further on in this section), who analysed samples of 120 lexemes in Mundari (a presumably flexible language) and English, and found that around 75% of the Mundari lexemes could function as verbs and nouns, against around 65% of the English items.
zero-derivation from one lexical category to a new one, without any formal evidence for such a transition. In other words, under both analyses the lack of overt morphological marking of a category change is explained by the fact that there is no such change going on in the first place.

![Conceptual Structure of Kiss](image)

Figure 2.14: The conceptual structure of kiss.

In general, the above discussion makes clear that some types of meaning shifts accompanying the use of formally identical lexemes in different functional environments appear to be highly regular, while others seem to be less predictable. In addition, it has been shown that there are, broadly speaking, two views concerning the relation between lexical flexibility and semantic shift: One claiming that ‘true’ flexibility involves purely compositional semantics, and another claiming that idiosyncratic interpretations are expected on the basis of the specific meaning components that are part of the unique semantics of any flexible lexeme. In the next section, I will show that in fact both regular and irregular types of semantic interpretations occur in languages for which pervasive lexical flexibility has been claimed. It will be argued that the relevant data can be understood in terms of a mismatch between lexical and phrase-structural categorization in these languages.

2.5.2.2 An integrative approach to semantics in flexible languages

Theories about PoS often take for granted that there is a one-to-one correlation between lexical categories and syntactic categories. However, Sasse (1993a, b) explicitly advocates a clear distinction between these two types or levels of categorization. Following Sasse’s view, Himmelmann characterizes the levels as follows:
“One level is the level of terminal syntactic categories where lexical items are categorized according to their phrase-structural properties. (...). The second level is the level of lexical categories proper where lexical items are categorized according to those grammatical features which are not directly relevant for phrase structure [...] (i.e. phonological, morphological, and syntactic features).” (Himmelmann 2007: 263)

Making this distinction implies that there may or may not be a correlation between the categories distinguished at one level and those distinguished at the other level. The former situation, i.e. a one-to-one correlation between lexical and syntactic categorization, is typically attested in rigid languages, including many Indo-European languages. The latter option involves a mismatch between lexical and syntactic categories. A well-known example of a language for which this has been claimed is Tongan (see also the previous section). According to Broschart (1997), lexical categories in Tongan are defined in terms of morphological features, which have no relevance for syntactic distribution. Himmelmann (2007) makes a very similar claim for Tagalog: According to him, there are a number of morphologically defined lexical categories in Tagalog, which are not projected into the syntax. There is some kind of correlation between these so-called morpho-lexical categories and purely ontologically defined categories, but this correlation is far from perfect. This means that the lexical categories of Tagalog are definitely grammatical categories, even though they are not syntactic categories.

Don and Van Lier (forthcoming) relate the issue of compositional and non-compositional semantic interpretation in flexible languages to a mismatch between lexical and syntactic categorization in these languages. They assume that, at a sufficiently abstract level of representation, all languages have a set of un-categorized roots with certain ontological meanings. In rigid languages, the operation that assigns these roots to a lexical category also automatically determines their phrase-structural destination: Once it is categorized, a lexical item can be used in only one syntactic function (without further measures). Flexible languages, in contrast, are characterized by the fact that they can lexically categorize and re-categorize their roots without affecting the phrase-structural possibilities of the output forms. Such non-syntactic (re-)categorization may or may not involve overt formal marking. The crucial point is that, since this type of lexical categorization occurs prior to syntactic categorization, it may involve non-compositional semantic interpretations. In contrast, syntactic categorization, i.e. the
ultimate insertion of lexical material into a syntactic function, should yield fully compositional semantic interpretations. Typically, the latter type of categorization involves syntactic frames made up of tense-aspect-mood (TAM) operators (for verbal/predicative usage), or determiners and/or case markers (for nominal/referential usage).

Consider a language like Samoan. In this language, according to Mosel and Hovdhaugen (1992: 77), “the categorisation of words into nouns and verbs is not given a priori in the lexicon. It is only their actual occurrence in a particular syntactic environment which gives them the status of a verb or a noun”. When a lexical item is preceded by a TAM particle it functions syntactically as a verb, and when it combines with one of the following particles it functions syntactically as a noun: le (determiner), e (ergative; absolutive is zero-marked), o (possessive), or i (marker of non-core arguments and locative-directional adjuncts).

The semantic interpretation of lexemes in verbal versus nominal syntactic function is often completely predictable, as is illustrated in examples (35) and (36). In example (35b), the action-denoting root *alu* ‘go’ occurs in nominal function and means ‘(the action of) going’. In example (36a) the object-denoting root *uō* ‘friend’ occurs in verbal function and is interpreted compositionally as ‘be friends’.

**Samoan** (Mosel & Hovdhaugen 1992: 73, 77)

(35) a.  
\[ E \text{ alu le pasi i Apia } \]
\[ \text{genr go det bus dir Apia} \]
‘The bus goes to Apia.’

b.  
\[ \text{le alu o le pasi i Apia} \]
\[ \text{det go poss det bus dir Apia} \]
‘the going of the bus to Apia.’

(36) a.  
\[ E \text{ uō Tanielu ma Ionatana} \]
\[ \text{genr friend Daniel and Jonathan} \]
‘Daniel and Jonathan are friends.’

b.  
\[ E \text{ alofa Taniel i l=a=na uō} \]
\[ \text{genr love Daniel dir det=poss=3sg friend} \]
‘Daniel loves his friend.’
However, there are also many other possible semantic shifts, with less predictable interpretations. Some examples are given in (37). Note that the semantic relations between the paired forms make reference to actions and participants involved in those actions (actors, instruments, undergoers) or specific instances of those actions.

**Samoan** (Mosel & Hovdhaugen 1992: 82, 83)

(37)  a. *fana*: ‘gun’ (N) or ‘shoot’ (V)
    b. *lama*: ‘torch’ (N) or ‘fish by torch light’ (V)
    c. *gaoi*: ‘steal’ (V) or ‘thief’ (N)
    d. *solo*: ‘move forward’ (V) or ‘procession’ (N)
    e. *tusi*: ‘write’ (V) or ‘letter/book’ (N)

The semantic shifts that go with the conversions in the examples in (37) are not marked by any overt formal change. However, overt lexical derivation does occur in Samoan, in particular with the suffix -ga. Some examples are given in (38). Interestingly, the output forms of this process have irregular semantic patterns, which at least partly resemble the ones in (37). The derived forms may denote a participant, a location, or an institution associated with the action denoted by the base form, but also a specific instance of that action.

**Samoan** (Mosel & Hovdhaugen 1992: 195)

(38)  a. *amo* ‘carry’ - *āmo-ga* ‘person(s) carrying loads’
    b. *āo* ‘teach’ - *āo-ga* ‘school’
    c. *tipi* ‘cut’ - *tipi-ga* ‘surgical operation’
    d. *pule* ‘control’ - *pule-ga* ‘unit of church administration’

Since the semantic meanings of the derived forms in (38) are typically ‘nominal’, while those of the un-derived ones are ‘verbal’, -ga has traditionally been termed a nominalization suffix (Mosel & Hovdhaugen 1992: 194). However, according to Mosel (2004: 267) “words carrying the so-called nominalization suffix can occur in the head position of a verb complex.” In other words, -ga is *not* a syntactically categorizing derivational suffix: it creates a new lexical item with a particular semantic meaning, but does not determine the phrase-structural destiny of the item. This is illustrated in (39) with the form *to-ga* ‘plantation’, derived from the root *to* ‘plant’. The derived form has an unpredictable semantic interpretation, which may be described roughly as:
‘location associated with the action denoted by the base form’. After entering into a compound with niu ‘coconut’, the derived form togā is inserted into a verbal syntactic function by combining it with the perfective particle ‘ua’. This syntactic categorization, as expected, yields a fully compositional semantic interpretation, namely ‘was a (coconut) plantation’.

*Samoan* (Mosel 2004: 267):

(39) ‘*Ua* to-gā-niu ʻātoa le mea maupuʻepuʻe

\[\begin{array}{ll}
\text{TAM} & \text{plant-NMLZ-coconut whole ART place hill} \\
\end{array} \]

‘The whole hilly place was now a coconut plantation.’

A very similar process occurs in Tongan, which has the suffix -Canga that can be applied to action-denoting, property-denoting, and object-denoting roots. According to Broschart (1997: 146), this suffix, just like Samoan -ga, is not a nominalizer in the phrase-structural sense, but rather a derivational marker whose output forms are flexible and mean something like “domain where something is or takes place”. However, the exact semantic interpretation of Canga-derived forms is far from fully predictable, as the examples in (40a-c) make clear:

*Tongan* (Broschart 1997: 146)

(40) a. *pule* ‘govern’ – *pule-anga* ‘government’

b. *motu’a* ‘old’ – *motu’a-anga* ‘reason for having aged’

c. *api* ‘home’ – *api-tanga* ‘homestead’

Despite their ‘nominal’ semantics, Tongan -Canga derivations can be used as verbs by combining them with a TAM particle, as is illustrated in example (41):

*Tongan* (Broschart 1997: 145)

(41) ‘oku ʻikai ke puleʻanga

\[\begin{array}{ll}
\text{PRS NEG SUBJ government} \\
\end{array} \]

‘It does not belong to the government.’

(lit. ‘it is not that it government-s.’)

I argue that the semantic shifts in the Samoan lexemes listed in (37) above are due to a process of derivation that must be analyzed as the zero-marked counterpart of -ga derivation. The same analysis applies to Tongan -Canga derivation, as compared with Tongan zero-derivation of the type illustrated.
in (34b) above. Let us take the Samoan form *gaoi* in (37c) as a concrete example. The proposal is that there are two homophonous but structurally different forms *gaoi*: (i) an un-derived form with the meaning ‘steal’, and (ii) a zero-derived form with the meaning ‘thief’. Alternatively, the items in (37) can be analyzed as polysemous, i.e. as involving two meanings belonging to a single form. The crucial point, which is captured by both analyses, is that the cases in (37) are not instances of flexibility. As such they stand in contrast to the examples in (35) and (36), which involve syntactic rather than lexical categorization, and accordingly behave semantically fully regularly.

Notably, it is predicted that both forms of the examples in (37), just like *-ga* derivations and their base forms, are flexible, i.e. are open to usage in verbal and nominal syntactic functions, and that these usages will yield compositional semantic interpretations. Taking *gaoi* as an example again, the un-derived form with the meaning ‘steal’ would be interpreted as ‘to steal’ in verbal function, and as ‘(the act of) stealing’ in nominal function. The zero-derived form with the meaning ‘thief’ would be interpreted as ‘to be a thief’ in verbal function, and as ‘a/the thief’ in nominal function. Unfortunately, the available data are insufficient to confirm this prediction.

Interestingly, the Samoan *-ga* suffix can also be used as a syntactic marker, rather than as a marker of lexical derivation. As would be expected, in the former function its output receives fully compositional semantic interpretations, namely ‘the act of X-ing’, where X is the action denoted by the base. In addition, the two different levels at which *-ga* may operate correlate with a phonological difference: Lexical derivation involves vowel-lengthening in the base form, while syntactic derivation leaves the phonological structure of the base unaffected. The phonological and semantic differences between lexical and syntactic *-ga* formations are illustrated in (42), using the same examples as in (38) above.

*Samoan* (Mosel & Hovdhaugen 1992: 195)

(42)

<table>
<thead>
<tr>
<th>Base</th>
<th>Lexical derivation</th>
<th>Syntactic derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>amo ‘carry’</td>
<td>āmo-<em>ga</em> ‘person(s) carrying loads’</td>
<td>amo-<em>ga</em> ‘carrying’</td>
</tr>
<tr>
<td>a’o ‘teach’</td>
<td>ā’o-<em>ga</em> ‘school’</td>
<td>a’o-<em>ga</em> ‘education’</td>
</tr>
<tr>
<td>tipi ‘cut’</td>
<td>tipi-<em>ga</em> ‘surgical operation’</td>
<td>tipi-<em>ga</em> ‘cutting’</td>
</tr>
<tr>
<td>pule ‘control’</td>
<td>pule-<em>ga</em> ‘unit of church administration’</td>
<td>pule-<em>ga</em> ‘controlling’</td>
</tr>
</tbody>
</table>
A similar phenomenon is attested in Maori. This language also has a very high degree of lexical flexibility, and it has a suffix -tanga, which can be used as a marker of lexical and syntactic derivation (Bauer 1993: 254). Example (43) nicely shows both types of derivations from the same base form tangi (‘cry’) within a single clause. The lexical process yields the idiosyncratic interpretation ‘funeral’, while the syntactic formation is compositionally interpreted as ‘the crying (of the people)’.

Maori (Bauer 1993: 48)

(43) I rongo au i te [tangi-hanga botubotu-tanga
            T/A hear 1SG DIR.OBJ ART cry-NMLZ sob-NMLZ
O ngaa taangata i te tangi-hanga ki a
            POSS ART people at ART funeral LOC ART

Maui Poomare

‘I heard the people’s crying at Maui Poomare’s funeral.’

So far, we have seen that syntactic categorization of putatively flexible lexemes in languages like Samoan and Tongan involves fully compositional semantic interpretation. This is in accordance with Evans and Osada’s compositionality criterion for ‘true’ flexibility. However, the data discussed above also show that such flexibility does not exclude the presence of lexical (re-)categorization processes. When such lexical processes are not overtly marked, they may be confused with flexibility. However, on the basis of the irregular semantic characteristics of these lexical processes, I have argued that they do not involve flexibility, but should rather be analyzed as instances of conversion.

We may now compare the Tongan and Samoan data discussed above with similar patterns found in other supposedly flexible languages. Consider for instance Tagalog, a language in which, according to Himmelmann (2005: 361), “content words do not have to be sub-classified with regard to syntactic categories. They all have the same syntactic distribution, i.e. they all may occur as predicates [and] as (semantic) heads of noun phrases […]”. Himmelmann analyzes voice-marking in Tagalog as a derivational process, which produces “actor- or undergoer-oriented action expressions” that are still flexible in terms of their syntactic possibilities. This process is characterized by “pervasive formal and semantic idiosyncrasies” (Himmelmann 2007: 288). Some examples of idiosyncratic semantic interpretations of voice-marked forms are provided in (44):
**Tagalog** (Foley 1998)

(44) **Base** | **Voice-marked form**  
---|---  
a. *abogado* ‘lawyer’ | *mag-abogado* ‘become a lawyer’  
b. *bus* ‘bus’ | *mag-bus* ‘ride a bus’  
c. *kamay* ‘hand’ | *mag-may-an* ‘shake hands’  
d. *langgam* ‘ant’ | *langgam-in* ‘be infested with ants’  
e. *lubid* ‘rope’ | *lubir-in* ‘be made into rope’

Derivation with voice markers may also be irregular in terms of phonology: It may involve unpredictable deletion of root vowels, as illustrated in (45a), or sporadic insertion of */n/*, as in (45b). There are also completely irregular forms, such as in (45c):

**Tagalog** (Himmelmann 2007: 288)

(45) **Base** | **Voice-marked form**  
---|---  
a. *kain* | *kan-in*  
‘consumption of food’ | ‘to eat something’  
b. *tawa* | *tawa-n-an*  
‘laugh, laughter, laughing’ | ‘to laugh at someone’  
c. *kuha* | *kun-in*  
‘getting, a helping’ | ‘to get something’

As already mentioned, even though voice-marked forms denote actions, they can still be used in nominal syntactic function without further measures. As predicted, such syntactic categorization yields fully predictable semantic interpretations: An actor-voice form in nominal function denotes ‘the agent of the action’, a patient voice form in nominal function denotes ‘the patient of the action’, and so on (Himmelmann 2007)\(^{29}\). An example of a locative voice form in nominal syntactic function is given in (46):

\(^{29}\) These Tagalog data show that a compositional interpretation of flexible lexemes is not necessarily ontologically equal across languages: The interaction with the specific feature of voice-marking in Tagalog accounts for the fact that in this language action-denoting forms are interpreted as objects or individuals in referential function, while in Samoan they are interpreted as denoting the action as a whole. See also the examples from Kharia further on.
Like the other flexible languages discussed above, Tagalog also seems to display cases in which two or more unpredictably related meanings are associated with a single form. According to Himmelmann, when action-denoting roots are used in nominal function without voice-marking, they may have one of the following kinds of meanings: (i) a state that results from the performance of an action, (ii) the result of the performance of an action, (iii) an object typically involved in the performance of an action, or (iv) the name of an action. I argue that interpretation (iv) is the result of syntactic categorization, while the meanings (ii) and (iii) are the result of lexical (re-) categorization. Consider now example (47), in which the root lakad ‘walk’ has meaning type (ii):

Tagalog (Himmelmann 2007: 280)

(47)  Iyón  ay  ma-baba’-ng  lakad
      dist  pm  stat-length-LK  walk
‘That was/is a long walk’

In parallel with the analysis of Samoan gaoi, I propose that there is a simple root lakad₁, meaning ‘walk’, and a homophonous, zero-derived form lakad₂, meaning ‘the result of walking’ → ‘a walk’. It may be objected that a zero-derivation analysis is somewhat problematic, since, unlike in Samoan, there seems to be no overt counterpart of this particular process. This is why above I have described the phenomenon in terms of polysemy. The crucial point is that the meaning shift is not due to a syntactic, but to a lexical process, and as such does not involve flexibility.

It is predicted that both lakad₁ and lakad₂ are flexible, i.e. can be used in verbal and nominal syntactic functions, with compositional semantic interpretations. The base form lakad₁ would mean ‘to walk’ when used as a verb, and ‘(the act of) walking’ when used as a noun.

Meaning (i) is associated with attributive function, and I will not discuss it here.
is illustrated in example (48) below. In contrast, *lakad* would mean ‘a/the walk’ when used as a noun, and ‘be a walk’ when used as a verb (as is in fact the case in example (47) above).

Tagalog (Himmelmann 2007: 281)

(48) Ma-husay ang lakad ng mákiná
    stat-orderliness spec walk gen machine
    ‘The walking of the machine is good.’

These data show that, like the other languages discussed in this section, Tagalog displays ‘true flexibility’ in the sense that it freely uses simple and derived lexical material in all syntactic functions with compositional semantic interpretations. This flexibility is combined with either overtly or zero-marked lexical processes, which yield semantically non-compositional outcomes.

Finally, consider Kharia, a South Munda language that is unrelated to the Malayo-Polynesian languages discussed so far. Peterson (2005, 2006, forthcoming) claims that Kharia does not possess the lexical classes ‘noun’ and ‘verb’, but instead has two syntactically defined categories, which he calls ‘predicate’ and ‘complement’. According to Peterson (2006: 60), “lexemes in Kharia do not appear to be either inherently predicates or their complement but can generally appear in both functions, without any overt derivational morphology.” The categories ‘predicate’ and ‘complement’ are expressed through the combination of a so-called ‘content head’ with a functional head. Content heads may consist of any single lexical root, but also of multiple roots or complex phrases. A predicative functional head is expressed through enclitic markers for tense/voice and person, while a complement functional head is expressed through enclitic case marking (other than possessive/genitive).

The semantic interpretation of a content head in an actual utterance requires a combination of the intrinsic meaning of the root(s) that it consists of, the syntactic function in which it appears, and the morpho-syntactic distinctions belonging to that function (such as tense/voice in predicative function). According to Peterson (2006: 70), this combination yields “entirely predictable” semantic outcomes. In particular, if an object- or individual-denoting root is used as a predicate and marked for middle voice, it gets the meaning ‘to become X’ (where X is the denotation of the root). When the same type of lexeme in the same function is marked for active voice it means ‘to turn something into X’. Action-denoting roots that are used in
referential function get the meaning ‘(the act of) X-ing’ (where X is again the denotation of the root). Examples are given in (49) and (50):

Kharia (Peterson 2006: 60, 68)

(49) a. **lebu ḍel-ki**  
man come=m.pst  
‘The/a man came.’

b. **bhagwan lebu=ki**  
God man=m.pst  
‘God became a man.’ (= Jesus)

(50) a. **u kayom ondor=kon raṭa=ya? ayo=dom,**  
this talk hear=seq Rata=gen mother=3.poss  
‘Hearing this talking, Rata’s mother, Darhi’s wife says: …’

darṭi=ya? saw=ṇay=dom  
Darhi=gen spouse=woman=3.poss say=act.prst  
‘The people tell [this] story in the villages.’

b. **ni lebu=ki khoṛi=ki=te kayom=ta=ki**  
story people=pl village.section=pl=obl talk=prst=pl  
‘The people tell [this] story in the villages.’

The syntactic nature of Kharia categorization and the phrasal clitic-hood of the functional heads becomes even more evident when considering an example like (51), in which the complex content head **bidesa? ebukiya? rupra**  
‘the appearance of people from abroad’, is used in the syntactic function of predicate. This involves combining the content head with a functional head, which is expressed by clitic forms for voice, tense and person, and attaches to the phrase’s final element.

Kharia (Peterson 2006: 85)

(51) **bharat=ya? lebu=ki [bides=a? lebu=ki=ya?**  
India=gen person=pl abroad=gen person=pl=gen  
rupra]=ki=may.  
appearance=m.pst=3pl  
‘The Indians took on the appearance of people from abroad.’

Just like the other flexible languages discussed above, Kharia also has lexical processes that involve semantically and phonologically unpredictable outcomes. Consider the derivational process marked by **nasal-vowel-infixation** in Kharia. The output forms of this process may denote an object, an instrument, or a location typically involved in the action denoted by the
base form, as well as a specific instance or result of that action. With respect to phonology, the infixed vowel has the same quality as the vowel preceding the nasal, but the infixed nasal is of indeterminate quality: it is usually realized as /n/, but /m/ is also occasionally found (Peterson 2006: 79). Some examples of the process are given in (52):

\[
\begin{array}{ll}
\text{Kharia (Peterson 2006: 80)} & \\
(52) & \text{Base} & \text{Lexical derivation} \\
a. & \text{bel} & \text{‘spread out’} & \text{be-ne-l} & \text{‘bedding’} \\
b. & \text{bui} & \text{‘keep, raise’} & \text{bu-nu-i} & \text{‘pig’} \\
c. & \text{jo} & \text{‘sweep’} & \text{jo-no-} & \text{‘broom’} \\
d. & \text{jin} & \text{‘touch’} & \text{ji-ni-b} & \text{‘touch’} \\
e. & \text{rab} & \text{‘bury’} & \text{ra-na-b} & \text{‘burial ground’} \\
f. & \text{kuj} & \text{‘dance’} & \text{ku-nu-j} & \text{‘dance’} \\
g. & \text{muʔsiŋ} & \text{‘rise’} & \text{mu-nu-ʔsiŋ} & \text{‘east’} \\
\end{array}
\]

According to Peterson, and in line with the lexical (zero-)derivational processes discussed for other flexible languages, Kharia infixed forms can be used freely in nominal and verbal syntactic function, despite their ‘nominal’ meanings. As is the case for any other type of content head, this yields entirely compositional semantic interpretations. Thus, a derived form like bunui ‘pig’ (see (52b) above) will get the meaning ‘to become a pig’, when used as a predicate in middle voice, but ‘to turn something//someone into a pig’ when used as a predicate with active voice. In nominal syntactic function it will mean ‘a//the pig’. Obviously, the predicative uses of a form like bunui are pragmatically rather odd, and will therefore be seldom attested, if at all. However, examples like those in (49b) above make clear that, given the appropriate pragmatic context, the use of object- or individual-denoting lexemes in verbal syntactic function is grammatically totally unproblematic.

To summarize, I argued in this section that the patterns of semantic interpretation of lexical material in presumably flexible languages can be understood when taking into account the mismatch between lexical and syntactic categorization in these languages. In flexible languages, lexical categorization is not relevant for phrase structure. This stands in contrast to typical rigid languages, in which there is a one-to-one correlation between lexical categorization and syntactic categorization. Possibly, these two scenarios represent the extreme ends of a scale or cline. This would mean that there are also languages in which syntactic categorization takes place.
somewhere in between the level of simple lexical root and the level of phrase structure. In Chapter 8 I will discuss this idea and its repercussions for the present study in more detail.

2.5.3 The criterion of exhaustiveness: How flexible should a flexible language be?

The second criterion proposed by Evans and Osada (2005) concerns the amount or the proportion of flexible lexical items that is required in order for a language to qualify as flexible. This criterion is described as follows:

“It is not sufficient to find a few choice examples which suggest word class flexibility. Since word classes are partitionings of the entire lexicon, equivalent statements need to hold for all relevant words in the lexicon that are claimed to have the same class”. (Evans & Osada 2005: 378)

It seems, in fact, that there is considerable agreement between scholars on this point. For instance, Hengeveld et al. (2004: 538) remark that “in order for a language to qualify as flexible, it has to show systematic flexibility (…) for an entire class of lexemes.” And Baker (2003: 117) makes a similar point saying that “an important typological difference exists only if categorial ambiguity extends to an entire open class of inflectionally similar words”.

However, Evans and Osada (2005: 378) rightly observe that “in practice (…) it is difficult to check every one of tens and thousands of lexical items”. In spite of this rather obvious practical restriction, it seems that in many reference grammars and other descriptive sources quite unambiguous statements can be found about the pervasiveness of multi-functionality of (certain groups of) lexemes, be it not exactly in the quantified fashion that Evans and Osada aim at. This is clearly illustrated by the quotations in the previous subsection about such languages as Samoan, Tagalog, and Kharia. In Chapter 5 I will further discuss the criterion of exhaustiveness and its application to the relevant languages of the sample.

2.5.4 The criterion of equivalent combinatorics: Bi-directional flexibility

The third and final criterion for flexibility proposed by Evans and Osada is the criterion of equivalent combinatorics. It states that

“[M]embers of what are claimed to be merged classes [i.e. flexible classes, EvL] should have identical distributions […]. All members of both
putative classes should be equally acceptable in both primary functions as argument and predicate”. (Evans & Osada 2005: 366-367)

A corollary of this criterion is that flexibility should be bi-directional:

“To establish that there is just a single [i.e. flexible, EvL] word class, it is not enough for Xs to be usable as Ys without modification: It must also be the case that Ys are usable as Xs.” (Evans & Osada 2005: 375)

The part of this criterion that makes reference to the actual possibility to use a flexible lexeme in more than one function seems quite unproblematic. Probably anyone would agree that this is an essential part of any definition of flexibility. Rather less straightforward is the part of the bi-directionality criterion that refers to equal acceptability of every lexeme in all functions. In particular, the frequency with which flexible lexemes appear in predicative or referential function is bound to be influenced by semantic and pragmatic factors. Regarding this point, I follow Hengeveld and Rijkhoff, who argue as follows:

“There is no reason to assume that all our concepts are equally symmetrical with respect to predicating and referring functions in a particular language. We may expect certain flexible lexemes to occur more as predicates than as arguments, whereas other lexemes are used more often as arguments than as predicates.” (Hengeveld & Rijkhoff 2005: 412)

In other words, the semantic denotation of a root affects its relative markedness with respect to a particular propositional function. This is in fact convincingly shown by Croft’s prototype approach to PoS classes, as discussed in section 2.4.2. Also, recall in this respect the examples from Samoan, Tongan, and Kharia mentioned in the previous subsection, which involve semantically ‘nominal’ but syntactically flexible (derived) lexemes. Crucially, the fact that the use of a lexeme with a particular semantic meaning in a particular function is pragmatically marked and therefore infrequently attested does not mean that there is a syntactic restriction that excludes such usage. In Chapter 5 I will further illustrate this point for the relevant languages in the sample.

A final aspect of bi-directionality is specifically related to noun-adjective flexibility. There are languages in which property-denoting lexemes
(i.e. semantic adjectives) can be the sole element in a referential phrase. However, this does not necessarily warrant the conclusion that there is no lexical distinction between adjectives and nouns. Often, such cases of apparent noun-adjective flexibility in fact involve so-called *absolutely used adjectives*, i.e. adjectives which function as the modifier of a head that is understood from the discourse context and therefore remains unexpressed.

Moravcsik (2001) distinguishes two ways in which property-denoting lexemes can function as the sole element in a referential phrase, namely referentially *dependent* and referentially *independent*. In the former case, the head noun must be available from the discourse or situational context. In the latter case, the intended referent is invariant, such as ‘people’ in (53) below. This possibility of referentially *independent* use of property-denoting words seems to be a requirement for ‘true’ noun-adjective flexibility.

*Hungarian* (Moravscik 2001: 339)

(53) *Sok gazdag van itt*
many rich is here
‘There are many rich (people) here.’

However, it is often problematic to determine whether referentially independent usage of property-denoting lexemes as the head of a referential phrase is a real option in an individual language. This is especially so when the analysis is based on isolated examples from written sources. Sometimes however, evidence in favour of an independent usage analysis can be found in the domain of agreement. In particular, the understood head may trigger the expression of agreement morphology on the absolutely used adjective. This happens for instance in Spanish, as shown in example (54), where *moderna* takes a feminine singular agreement suffix because it modifies the un-expressed but understood head *la casa* ‘the house’.

*Spanish* (Hengeveld 1992: 61-62)

(54) *prefiero es-a modern-a*
prefer:1s.prs that=fem.sg modern=fem.sg
‘I prefer that modern (one).’ [e.g. the house]

---

31 Schachter & Shopen (2007: 17) use the term ‘adjectival-noun languages’ for this type of language.
Note furthermore that, if one takes seriously the criterion of bi-directionality, this implies that, in order to establish ‘true’ noun-adjective flexibility, it should be possible not only to use property-denoting lexemes as heads of referential phrases, but also to use object- or individual-denoting lexemes (i.e. semantic ‘nouns’) without further measures as modifiers in referential phrases. Again, this is not always easily determinable, since the distinction between nominal apposition or composition on the one hand, and attributive modification on the other hand may be unclear. In Chapter 5, where the PoS systems of the individual sample languages are categorized, I will discuss the relevant cases in more detail.

2.5.5 Lexical flexibility: A summary
In this section, I discussed the three criteria for lexical flexibility proposed by Evans and Osada (2005): Semantic compositionality, exhaustiveness, and equivalent combinatorics. Especially the first criterion has given rise to disagreement in the literature: While Evans and Osada contend that ‘true’ flexibility involves fully compositional semantic interpretation of lexemes in particular functional environments, other scholars argue that the occurrence of semantic idiosyncrasies is not an argument against flexibility. I have shown that a number of languages, which figure prominently in discussions on lexical flexibility, display both semantically compositional and non-compositional (zero-marked) derivational processes. The difference between these two types of processes lies in the fact that one pertains to syntactic, and the other to lexical (re-)categorization. The special characteristic of flexible languages, then, seems to be the lack of a one-to-one correspondence between these two levels of categorization.

As regards the remaining two criteria, the one concerning exhaustiveness seems theoretically unproblematic, but is hard to verify empirically. Finally, with respect to the criterion of equivalent combinatorics, it was argued that while bi-directionality is probably a central notion in flexibility, this does not mean that flexible lexemes are expected to appear equally frequently in all of their possible functions.

2.6 Summary
This chapter presented a number of functional-typological approaches to lexical categorization. I started out with a brief outline of some early studies, and the criteria that were used in them to define PoS classes. Subsequently, I provided a detailed discussion of Hengeveld’s PoS theory. In its most
recent version, this theory proposes an implicational map consisting of four propositional functions defined by two binary parameters: *predication-reference*, and *head-modifier*. Hengeveld’s definitions of PoS classes are based on the presence versus absence of structural coding accompanying the use of lexemes in each of these four possible functions. The dominance relations holding between the two parameters and between their respective values predict a number of PoS systems that are expected to occur in actual languages.

Hengeveld’s theory excludes two types of defining criteria for PoS, namely morphological criteria (or behavioral potential) and semantic criteria. The advantage of this approach is that it provides the possibility to compare PoS classes and PoS systems across languages. However, exactly this aspect of the Hengeveldian method has been criticized, most elaborately by Croft (2000, 2001, 2005). The latter objects that Hengeveld’s theory, since it ignores morphological and semantic data, amounts to methodological opportunism and misses out on cross-linguistic universals to be found in these two areas. On the other hand, the consequence of Croft’s method, namely taking into account the full set of distributional data, is that lexical categories, like any other type of language-specific formal category, can no longer be compared across languages. Croft’s solution to this problem is to define PoS as typological prototypes, i.e. as unmarked combinations of semantic classes and pragmatic functions, which correlate cross-linguistically with relative degrees of formal marking.

Another major point of discussion concerns the status of so-called flexible languages, and especially the role of semantics in defining what lexical flexibility really entails. Evans and Osada (2005, and cf. Croft 2005) propose that in truly flexible languages the interpretation of lexical material in different functional environments should be fully compositional. I discussed data from supposedly flexible languages, which show that this criterion is indeed met. At the same time, however, such languages display lexical processes of (zero-)derivation and polysemy, and these are semantically (as well as phonologically) non-compositional. I argued that the presence of both compositional and non-compositional semantic (re-)categorization in flexible languages can be understood in terms of the fact that these languages, unlike in the more familiar rigid languages, lack a one-to-one correlation between lexical and syntactic categorization.

In the next chapter, I will use Hengeveld’s method for lexical classification to develop a typology for a different set of constructions, namely dependent
clauses. Therefore, it seems appropriate at this point to emphasize that I fully acknowledge the limited scope of the Hengeveldian approach in terms of the type and amount of analytic primitives that it takes into account. However, unlike Croft, I do not regard this as methodological opportunism. Depending on the aim of one’s research, it seems perfectly defendable to focus on certain linguistic facts while ignoring others, as long as this is made explicit. As already mentioned, the lexical categories defined within Hengeveld’s framework are claimed to be cross-linguistically comparable only as far as his very specific set of criteria goes. And, as Cysouw (2007: 227) puts it: “Cross-linguistic identity is always just a matter of granularity of analysis”.
3.1 Introduction

In the previous chapter, parts of speech classes were defined as mappings onto a space consisting of four propositional functions, defined in terms of two hierarchically ordered implicational parameters: the predication-reference parameter and the head-modifier parameter. The aim of the present chapter is two-fold. First, it demonstrates that dependent clauses (henceforth DCs) can be defined in the same way as parts of speech (PoS), namely as mappings onto this functional space. Second, on the basis of many earlier functional-typological studies, it is shown how DCs can be classified according to their internal morpho-syntactic properties.

The central point of section 3.2 is that propositional functions constitute the common ground between PoS classes and DC constructions. More specifically, complement clauses have referential function, like lexical nouns; relative clauses function as modifiers in referential phrases, like adjectives; and adverbial clauses function as modifiers of predicates, predicate phrases or clauses, in the same way as lexical adverbs. Based on this parallelism, I show in section 3.2.1 that DCs can be defined, like PoS, in terms of the (set of) propositional function(s) that they express.

This implies that the notions of flexibility and rigidity, as introduced for PoS in Chapter 2, can also be applied to DCs. Recall that, according to

32 Apart from these three traditional types of DCs, there are so-called predicate clauses which function predicatively, like lexical verbs. I will come to these shortly.
the Hengeveldian approach, these notions depend crucially on the presence versus absence of structural coding accompanying the use of a particular construction in a particular propositional function. In order to arrive at parallel definitions for PoS and DCs, a definition is required of structural coding of DCs. This issue is addressed in section 3.2.2. Subsequently, in section 3.2.3, I define a number of rigid and flexible DC types, in parallel with the various types of PoS classes that figure in the PoS systems predicted in Chapter 2.

In section 3.3 I turn to the internal morpho-syntactic properties of DCs, i.e. their behavioural potential. In section 3.3.1 I explain that these properties are, at least in part, motivated by the fact that DCs involve a marked combination of semantic class and propositional function. In particular, DCs denote states of affairs (henceforth SoAs), or propositional contents\(^{33}\). As such, they are marked as expressions of reference and modification, since these propositional functions are prototypically expressed by object-denoting and property-denoting (lexical) constructions, such as nouns, adjectives, and manner adverbs (see Chapter 2, section 2.4.2). The marked combination of semantic class and propositional function displayed by DCs is reflected in their formal encoding, which in many cases shows a certain mixture of morpho-syntactic characteristics of independent clauses on the one hand, and features of lexical or phrasal constructions on the other hand.

There are cross-linguistic differences as regards the set of possible features that may be involved in the behavioural potential of DCs in the first place; not all languages display the same morpho-syntactic distinctions. In addition, language-specific DC constructions may differ in the amount and type of features that is selected from the available set. However, the variation is not random, and functional-typological research has come a long way in bringing to light universals in these areas. Among other factors, the possible combinations of morpho-syntactic features in the DCs of the world's languages are constrained in terms of universal hierarchical ordering patterns displayed by certain feature sets. In section 3.3.2 I give a general overview of the research on these patterns, which are usually referred to as verbal and nominal feature hierarchies (Bybee 1985; Dik 1997; Lehmann & Moravcsik 2000; Rijkhoff 2002). Subsequently, in section 3.3.2, I will discuss studies that focus particularly on the impact of these hierarchies on

\(^{33}\) There are also DCs that do denote objects or individuals (i.e. first order entities), such as headless relative clauses and actor/object nominalizations. However, the present study is not concerned with these.
mixing patterns of verbal and nominal features in DCs, and the functional motivations underlying these effects (Cristofaro 2003; Malchukov 2004).

At the end of this chapter, in section 3.4, I present a typological framework for DCs that combines the parameters of internal morphosyntactic properties (as discussed in section 3.3) with the classification system of rigid and flexible DCs (as developed in section 3.2). This integrated typology will be applied in Chapter 6 to the DCs of the sample languages. Finally, section 3.5 is a brief summary of the present chapter.

3.2 Dependent clauses and propositional functions

3.2.1 Slots and fillers
Traditionally, as mentioned in the introduction to this chapter, three types of DCs are distinguished: complement clauses, relative clauses, and adverbial clauses (Thompson et al. 2007: 238). They can be defined in terms of their respective pragmatic-syntactic functions: A complement clause functions as a referential unit, expressing an argument of the matrix clause predicate; a relative clause functions as a modifier of a matrix clause argument; and an adverbial clause functions as a modifier of the main predicate of the matrix clause. As such, the functions of these DC types correspond to the defining propositional functions of lexical nouns, adjectives, and manner adverbs, as presented in Chapter 2, section 2.2.

In addition to these three DC types, there are also clausal constructions that function predicatively. Such constructions are usually called predicate clauses. Recall from Chapter 2 that the use of non-verbal lexemes in predicative function may trigger one of three expression strategies: (i) a copula, (ii) a zero-2 strategy (i.e. no copula and no verbal behavioural potential), or (iii) a zero-1 strategy (i.e. no copula but full expression of verbal behavioural potential). Only when a particular PoS class triggers the third, zero-1 strategy, it is regarded as having the possibility to function predicatively, in addition to its defining, non-predicative use(s) (see also Chapter 5, section 5.5).

Note that, as far as adverbial clauses are concerned, this study is restricted to adverbial manner clauses (cf. lexical manner adverbs in Chapter 2, section 2.3.1), excluding other types of adverbial clauses, such as purposive, temporal, and conditional clauses. This means that, in terms of DCs, I will take into account only adverbial clauses that specify how the event expressed by the predicate of the matrix clause is performed (Kortmann 1998: 467). However, the borderline between adverbial manner clauses and other, semantically related types of adverbial clauses is not always easily drawn. This issue is taken up in Chapter 6, where the actual DCs of individual languages are discussed.
In the present chapter (and in Chapter 6) the same procedure will be followed for DCs\textsuperscript{35}.

In Chapter 2 an implicational map model was developed to formulate restrictions on the typology of PoS classes and PoS systems. The functional space pertaining to this model is re-presented in Figure 3.1a. In Figure 3.1b, the four functional slots of the space are filled with the four DC types mentioned above, instead of with lexeme classes, as was the case in Chapter 2. In other words: lexical units are substituted by \textit{configurational units}. The term ‘configurational’ refers to a complex linguistic construction, in which the elements stand in a nucleus-dependent (or predicate-argument) relationship (cf. Hengeveld & Van Lier 2008: 769-770).

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{Predicate phrase} & \textbf{Referential phrase} \\
\hline
\textbf{Head} & \textbf{Modifier} \\
\hline
\end{tabular}
\caption{The space of propositional functions}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Predicate phrase} & \textbf{Predicate clause} & \textbf{Adverbial manner clause} \\
\hline
\textbf{Referential phrase} & \textbf{Complement clause} & \textbf{Relative clause} \\
\hline
\end{tabular}
\caption{Propositional functions and dependent clause types}
\end{table}

The definitions of the four types of dependent clauses in Figure 3.1b can now be formulated as in (i):

(i)

(i) A predicate clause is a configurational unit that can be used as the head of a predicate phrase;
(ii) A complement clause is a configurational unit that is used as the head of a referential phrase;
(iii) A relative clause is a configurational unit that is used as a modifier in a referential phrase;
(iv) An adverbial manner clause is a configurational unit that is used as a modifier in a predicate phrase.

\textsuperscript{35} In Chapter 6 it will become clear that zero-1 strategies are hardly if ever attested in combination with clausal constructions. Nevertheless, in the present discussion the possibility of a predicate clause triggering a zero-1 strategy is included for reasons of theoretical consistency.
The parallel between these definitions and the definitions of nouns, adjectives, and manner adverbs presented in Chapter 2 will be obvious: DCs are regarded as the clausal equivalents of PoS classes, as far as their respective functions in the matrix clause are concerned.

Recall, however, that in the definition of the PoS classes in Chapter 2 the phrase ‘without further morpho-syntactic measures’ was included (see section 2.3.1). As was explained, these further measures make reference to structural coding rather than to behavioural potential. In order to complete the parallelism between PoS and DCs, the notion of structural coding needs to be defined with reference to the latter construction type. This is the topic of the next subsection.

3.2.2 Structural coding in dependent clauses:

Subordinating conjunctions and dependent verb forms

Two general types of structural coding of DCs can be distinguished:

(i) Subordinating conjunctions, further sub-divided into complementizers, relativizers, and adverbial conjunctions;
(ii) Special markers on the dependent predicate, further sub-divided into nominalizers and infinitive markers, participles, and converbs36.

These two groups of markers are regarded as structural coding because they function primarily as indicators of subordination, and, when they are specialized markers, of the function in which the DC appears. In the case of subordinating conjunctions, the structural coding occurs at the level of the DC construction as a whole, and the internal morpho-syntactic structure of the DC remains the same as in independent clauses (in the particular language under analysis). In contrast, special morphemes on the dependent predicate mark the construction at the level of its nucleus. This affects the categorial status of the predicate, which is reflected in a certain degree of deviation of the internal morpho-syntax of the DC from the structure of an independent clause. Structural coding by means of a subordinating conjunction is illustrated in (2a) below with the English that-construction. Example (2b) shows structural coding with a dependent predicate marker, namely the English -ing form. Note that in the DC in (2b) there is no tense expression, and that the subordinate subject is encoded as a possessor.

36 The term converb refers to “a non-finite verb form whose main function is to mark adverbial subordination” (Haspelmath 1995: 3).
(2) a. John regrets [that Paul will move to Finland].
   b. John regrets [Paul’s moving to Finland].

A number of issues concerning the two types of structural coding of DCs require some discussion. First, examples (2a-b) show that subordinating conjunctions are typically freestanding forms, which appear at the periphery (i.e. either the beginning or the end) of a DC, whereas dependent predicate markers are (obviously) bound forms that attach to the predicate. However, as Cristofaro (2003) explains, in some languages this distinction between the two types of structural coding is somewhat obscured, namely when subordinating conjunctions are clitics. These clitics may appear attached to the dependent predicate for one of the following independent reasons:

(i) They can be attached to any constituent in the clause and so may happen to be attached to verbs;
(ii) They are usually placed close to verbs because of word order rules;
(iii) There is a rule stating that they have to be attached to verbs.
(Cristofaro 2003: 58)

In such cases one might be tempted to regard the complex [dependent predicate + cliticized conjunction] as an instance of structural coding by means of a special marker on the dependent predicate. However, special markers on the dependent predicate, as mentioned above, affect the internal morpho-syntactic structure of the DC, while subordinating conjunctions do not. Therefore, Cristofaro argues, a cliticized subordinating conjunction can be recognized as such when the structure of the DC remains unaltered. Consider example (3), from Cofán. The relativizer =’cho cliticizes to the dependent predicate because (i) it must attach to the final element of the unit on which it operates, and (ii) DCs in Cofán are obligatorily predicate-final. Despite its position, =’cho can be identified as a conjunction because the internal structure of the DC is the same as in an independent clause.
Note that the DC in (3) is marked off by square brackets. This will be done in examples of DC constructions throughout the rest of the book

*Cofán* (Fischer & Van Lier, forthcoming)

(3) *Ingi=ta [avuña-tsb=e canse=’cho] a’i=’fa=gi.*
    *we=indef happy=stat=advr live=rel person=pl=fo.1*
    ‘As for us, we were people that lived happily.’
Second, it should be mentioned that not all DCs are overtly structurally coded by means of a subordinating conjunction or a special marker on the dependent predicate. Rather, DCs can also be zero-marked\(^3\). Consider for instance example (4) from Bukiyip:

**Bukiyip** (Conrad & Wogiga 1991: 179)

(4) Énan n-a-kli [yek i-wich umu énaniny moul]

he he-real-say I-I-IRR-enter ben his work

‘He said that I would have his job.’

Third, structural coding may be fused with the expression of a category that pertains to behavioural potential. For instance, determiners may function as complementizers. This occurs in Dhaasanac, as illustrated in (5):

**Dhaasanac** (Tosco 2001: 288)

(5) yú [kúun ʰi fuddir=a] fede

I you 3(pron) open.dep=det want.ipfv

‘I want you to open it.’

Another common case of accumulation of structural coding and behavioural potential is the fusion of a TAM marker with a special dependent predicate marker. This is attested in Imbabura Quechua, where different forms of the nominalization suffix express different relative tenses, as shown in (6):

---

\(^3\) Clauses marked with =’cho can also be used as object complements. In that function they are marked for accusative case, as illustrated in (1):

(1) Atesu-mbi ke ɲa=nga tеваen=’cho=ma

know-NEG you I=DAT write=comp=acc

‘I didn’t know that you had written to me.’

According to Cristofaro, the expression of case would be evidence of categorial change and would as such be a reason to analyze =’cho as a special dependent predicate marker. However, following Malchukov (2004; see section 3.3.3.3 of the present chapter and Fischer & Van Lier (forthcoming)), I regard case as an external category, which is in principle independent of the categorial status of the construction on which it operates.

\(^8\) DCs without morpho-syntactic marking may be marked through prosody. For practical reasons however, the latter coding strategy is not taken into account in the present study. Another possibility is of course structural coding by means of fixed constituent order.
Imbabura Quechua (Cole 1982: 33)

(6) \textit{niuka-ka} \quad [Juan kay-pi \quad ka-j/-shka/-na]-ta \quad ya-ni

\begin{itemize}
  \item I-TOP \quad Juan this-in be-NMZL.PRS/PST/FUT-ACC think-I
\end{itemize}

\textit{‘I think that Juan is/was/will be here.’}

Finally, as mentioned above, DCs marked by subordinating conjunctions typically have an internal structure that resembles the structure of independent clauses, while the internal structure of DCs marked by means of a special dependent predicate form deviates from the structure of independent clauses. However, this correlation between structural coding type and internal morpho-syntax is not absolute. In particular, the structure of DCs marked with subordinating conjunctions may differ in certain respects from that of independent clauses. Typically, this involves the use of dependent or subjunctive forms for the expression of verbal categories. Consider for instance example (7), which shows a Georgian complement clause marked with the complementizer \textit{rom} and with a predicate that bears a subjunctive aorist suffix.

\textit{Georgian} (Hewitt 1995: 628)

(7) \textit{Masp’injel-s u-nd-a, \quad [rom \quad male \ c’a-vid-e]}

\begin{itemize}
  \item host-DAT \quad ov-want-it \quad COMP \quad soon \quad PREV-go-AOR.SUBJ
\end{itemize}

\textit{‘The host wants me to go soon.’}

Subjunctive constructions tend to have fewer inflectional possibilities than their indicative counterparts, but this is not necessarily the case. As will become clear in section 3.4, I will regard subjunctive constructions as sentence-like DCs, i.e. as DCs with roughly the same syntactic form as independent clauses (cf. Hengeveld 1998; Malchukov 2006: 974; Noonan 2007: 59-63).

Taking these issues into account, I round off the discussion of structural coding of DCs and proceed to the next step: applying the notions of flexibility and rigidity to DC constructions.

3.2.3 Rigid and flexible dependent clauses

In Chapter 2, it was shown that PoS can either be \textit{rigid}, i.e. specialized for the expression of a single propositional function, or \textit{flexible}, i.e. usable in two or more functions. If DCs are assumed to be the configurational equivalents of PoS, then it follows that there are also rigid and flexible DCs. Like PoS, DCs are regarded as rigid when they are used in only one propositional
function, and as flexible when they can fulfil two or more functions without further measures, i.e. without any change in their structural coding. In this section I will define a number of possible flexible and rigid DC types. These DCs are the functional counterparts of the different types of PoS classes that figure in the systems predicted in Chapter 2, section 2.3.2.239.

The four predicted rigid DC types were in fact already presented in Figure 3.1b above, since they correspond to the three traditional DC types, plus predicate clauses. For the sake of completeness, these DCs are represented separately in (8)-(11) below. Notice that, as in Chapter 2, rigid constructions are presented with light shading and flexible ones (as in shown in (12) and further) with darker shading.

(8)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicate clause</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

(9)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Complement clause</td>
</tr>
</tbody>
</table>

(10)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Relative clause</td>
</tr>
</tbody>
</table>

39 As shown in Chapter 2, PoS classes enter into larger PoS systems. If DCs indeed mirror the functional possibilities of PoS classes, then it may be expected that the former are subject to the same restrictions as the latter in terms of the types of systems into which they may combine. However, DC systems are as such not the focus of this study. Rather, I will be more generally concerned with the question whether flexibility/rigidity in the realm of lexical classification is reflected in flexibility/rigidity in the realm of DC constructions. In addition, I will investigate whether particular PoS classes have DC counterparts with the same functional possibilities. Neither of these two research questions requires the definition of specific systems of DC constructions. Moreover, as I will argue shortly, languages often display a range of different DCs that can be used in (partly) the same propositional function(s).
Turning to flexible DCs, the predicted types are represented in (12)-(20).
The first flexible DC-type can be used in all four functions of the space.
This type will be called a contentive clause, in accordance with the term contentive used for maximally flexible lexemes. The functional potential of contentive clauses is represented in (12):

(12)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Contentive clause</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

The second flexible DC type can be used in all functions except the head of a predicate phrase. Thus, it has the same functional potential as lexical non-verbs. However, the term ‘non-verbal clause’ seems counter-intuitive, since dependent clauses typically do contain a verb. Therefore, I will use multifunctional clause instead. This type is represented in (13):

(13)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Multi-functional clause</td>
</tr>
</tbody>
</table>

Third, a flexible DC type is predicted that can be used to express both modifier functions, in the same way as lexical modifiers. This type will therefore be termed a modifier clause. It is represented in (14):
Fourth, a DC type is predicted that can be used as the head and modifier in a referential phrase, in parallel with lexical *nominals*. This type, termed *nominal clause*, is represented in (15):

The DC type in (16) represents the clausal counterpart of the PoS class of *predicatives*. This DC type, which will be called *predicative clause*, can express the functions of head and modifier in a predicate phrase:

The DC type represented in (17) parallels the distribution of lexical flexible *heads*. This *head clause* can function as the head of both predicative and referential phrases.
Finally, in (18)-(21) I represent four types of flexible clauses with the same distribution as four types of PoS classes predicted in Chapter 2 and indicated there as *Flex A, B, C,* and *D.* As the representations show, I have simply named the corresponding DC types *Flex clause A, B, C,* and *D,* respectively.

\[
\begin{array}{c|c}
(18) & \\
| Head & Modifier \\
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Flex clause A</td>
</tr>
<tr>
<td>Reference</td>
<td>-</td>
</tr>
</tbody>
</table>
\end{array}
\]

\[
\begin{array}{c|c}
(19) & \\
| Head & Modifier \\
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Flex clause B</td>
</tr>
<tr>
<td>Reference</td>
<td>-</td>
</tr>
</tbody>
</table>
\end{array}
\]

\[
\begin{array}{c|c}
(20) & \\
| Head & Modifier \\
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Flex clause C</td>
</tr>
<tr>
<td>Reference</td>
<td>Flex clause C</td>
</tr>
</tbody>
</table>
\end{array}
\]

\[
\begin{array}{c|c}
(21) & \\
| Head & Modifier \\
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Flex clause D</td>
</tr>
<tr>
<td>Reference</td>
<td>Flex clause D</td>
</tr>
</tbody>
</table>
\end{array}
\]

Note that there is one logically possible type of flexible DCs that is expected not to occur in actual languages. It is represented in (22)* below (the asterix indicates that that this type of DC is a ‘forbidden’ type). This system is the functional counterpart of a PoS class that figures only in systems that violate one or more of the constraints formulated in Chapter 2 (cf. Appendix II: (2)*/(3)*).
Having defined these DC types in terms of propositional functions, in parallel with the PoS classes in Chapter 2, a difference between DCs and PoS should be noted. In particular, subordination relations are semantically complex and this is reflected by structural variation in the DCs that express them. A consequence of this is that in many languages there are multiple DC constructions that can be used in the same or partly the same propositional function(s) (cf. note 39). For instance, Turkish has two types of nominal clause constructions: One in which the predicate is marked with \(-\text{AcAK}\), and another that takes the suffix \(-\text{DIK}\). Examples (23) and (24) show that both constructions can be used as complement clauses and as (object) relative clauses. The choice between the two forms is semantically motivated and concerns tense: \(-\text{DIK}\) indicates past and \(-(y)\text{AcAK}\) future tense.

**Turkish** (Göksel & Kerslake 2005: 423, 442)

(23) a. \[\text{Orhan-}m \text{bir şey yap-}m-a-yacağ-i] \nOrhan-gen anything do-NEG-NMLZ-3SG.POSS
\it{belliydi}.
\it{It was obvious that Orhan wouldn’t do/wasn’t going to do anything.}

b. \[\text{Fatma-‘nin yarın gör-eceğ-i}] \text{film} \nFatma-gen tomorrow see-PTC-3SG.POSS film
\it{‘the film that Fatma is going to/will be seeing tomorrow’}

**Turkish** (Kornfelt 1997: 50)

(24) a. \((\text{ben}) \text{[Ahmed-in öl-düğ-}i-}n]-ü duyu-du-m} \nI Ahmed-gen die-NMLZ-3SG-ACC hear-PST-1SG
\it{‘I heard that Ahmed died.’}
b. \[ adam-in \ git-tiğ-i \] \( okul \)
\[ man-gen \ go-\textit{ptc-3sg} \] \( school \)
‘the school that the man goes/went to’

It is also possible that two DC constructions in the same language show partial overlap in terms of the propositional functions that they can express. For instance, Turkish has, apart from relative clauses marked by \(-\textit{DIK}/-\textit{AcAK}\), another participial construction with the suffix \(-\textit{An}\). Unlike the \(-\textit{DIK}/-\textit{AcAK}\) construction, however, \(-\textit{An}\) clauses cannot function as complement clauses. Moreover, \(-\textit{An}\) clauses express subject relative clauses, whereas \(-\textit{DIK}/-\textit{AcAK}\) clauses express object relative clauses. An example of an \(-\textit{An}\) clause is given in (25):

*Turkish* (Göksel & Kerslake 2005: 440)

(25) \[ ögretmen \ ol-an \] \( haydar \)
\[ teacher \ be-\textit{ptc} \] \( Haydar \)
‘Haydar, who is a teacher’

These examples show that, when there are multiple DC constructions that can express (partly) the same propositional function(s), the choice between these constructions depends on the specific semantic and/or syntactic characteristics of the subordination relation that holds between the DC and the matrix clause.

### 3.2.4 Summary

In this section I have established the first part of a typology of DC constructions, namely the part that makes reference to the functional possibilities of DCs. These functional possibilities have been defined in parallel with the PoS classes in Chapter 2, i.e. in terms of the number and type of propositional function(s) that a particular DC can express without any difference in structural coding. In the next section I turn to the second part of the typology, namely the part that is concerned with the internal morpho-syntactic properties or the *behavioural potential* of DC constructions.
3.3 Behavioural potential in dependent clauses: De-categorization and re-categorization

3.3.1 Introduction

The internal structure of DC constructions has been investigated in a fair number of functional–typological studies. Many of these studies focus on the ‘mixed’ patterns of behavioural potential attested in DCs across languages, i.e. on the various combinations of morpho-syntactic properties typical of independent clauses on the one hand, and of lexical or phrasal constructions on the other hand.

These mixed patterns of morpho-syntactic properties can be understood as the outcome of two functionally motivated processes: de-categorization and re-categorization. The former process involves the non-expression of behavioural potential associated with the primary propositional function of a construction (Hopper & Thompson 1984). For a DC, i.e. a clausal construction denoting a SoA or a proposition, this primary function would be a predication. The categories that belong to this primary function, and which may be lost as a result of de-categorization, are often called verbal categories. The obvious reason for this is that the nucleus of a clause or predication – the predicate – is prototypically a verb, which expresses the relevant categories in the form of inflectional markers. Such verbal categories include tense-aspect-mood (TAM) distinctions and/or person marking (including agreement and cross-reference).40

The process of re-categorization, on the other hand, is reflected by the expression of behavioural potential associated with the secondary propositional function in which a construction appears (Bhat 1994). In the case of a DC, this secondary function is reference or modification. As we have seen, reference and modification are prototypically expressed by object-denoting and property-denoting constructions, i.e. nouns, adjectives, and manner adverbs (or phrases headed by these PoS). Cross-linguistically, the range of categories associated with nouns and NPs is more extensive than the range of adjectival or adverbial categories. Accordingly, the features that are expressed in a DC as a result of re-categorization are often termed nominal categories. They include gender, number, case, and definiteness. There are also some typical adjectival features, such as degree of comparison and head-modifier agreement. However, the former category is semantically marked

40 The process of de-categorization in DCs has alternatively been termed desententialization (Lehmann 1988: 193) and deverbalization (Croft 1991: 79).
with respect to event-denoting DCs, while the latter is cross-linguistically not very common (Haspelmath 2001: 16541). For manner adverbs there seem to be no prototypical features at all. This asymmetrical distribution of features over lexeme classes is also reflected in the literature on the internal structure of DCs, which is often concerned with clausal nominalizations.

Both verbal and nominal features have been shown to display universal patterns as regards the order in which they are expressed relative to their base unit, a verbal or nominal head. In what follows I first discuss functional-typological studies that provide evidence for these hierarchical ordering patterns (section 3.3.2). Subsequently, I turn to studies that investigate how these hierarchies motivate the particular mixtures of verbal and nominal features attested in DCs across languages (section 3.3.3).

3.3.2 Verbal and nominal feature hierarchies

3.3.2.1 Introduction
Different studies have taken into account somewhat varied sets of verbal and nominal features, depending on their theoretical orientation and on whether or not they distinguish between morphological and lexical expressions of categories. In this section I adopt the versions of the feature hierarchies proposed by Malchukov (2004), since they explicitly integrate various strands of functional-typological work.

3.3.2.2 The verbal feature hierarchy
Starting with (Malchukov’s version of) the verbal feature hierarchy, its most important ingredients are the work of Bybee (1985), combined with insights from research carried out in the framework of Functional Grammar (FG). Bybee’s well-known study shows that bound morphemes expressing verbal categories exhibit a universally preferred order in relation to their stem. This order is iconically motivated: it reflects the degree of relevance of each category for the interpretation of the verbal stem. In particular, those categories that occur more closely to the stem – the internal categories – affect the verb’s semantics, whereas those that are expressed further away from the stem – the external categories, in particular agreement phenomena – do not affect the meaning of the stem but are rather relevant to its syntactic

41 I will adopt the visual representations of the hierarchies that are used by Nikitina (2007) in her review of Malchukov’s work.
and/or *pragmatic* function. The verbal feature hierarchy as proposed by Bybee is represented in (26):

(26) $\text{STEM} \subset \text{Valency} \subset \text{Voice} \subset \text{Aspect} \subset \text{Tense} \subset \text{Mood} \subset \text{Agreement}$

Considering the hierarchy in some more detail, the most external category, agreement, is purely inflectional to the extent that its sole *raison d’être* is the existence of a syntactic relation between the verb and its argument(s) (Haspelmath 2002: 72). Agreement does not have any impact on the semantics of the verb. In contrast, the categories of tense, aspect and mood are not directly triggered by syntactic relations. These categories can be ordered with respect to each other in terms of their degree of semantic relevance to the verbal stem. Mood and tense have relatively little impact on the meaning of the verb, because the nature of the expressed action is in principle independent of the speaker’s evaluation of that action (mood), and of its temporal relation to the speech event (tense). Nevertheless, the categories of mood and tense do convey a certain amount of independent semantic content; a property that agreement typically does not have. The category of aspect is again of a different type: since aspectual distinctions modify the internal temporal constituency of an event, they directly affect the semantics of the verb. Finally, the two most internal categories, valency and voice, have an even higher amount of semantic impact on the verb, since they relate to the number and orientation of the participants involved in the event expressed by the verb.

Clearly, the functional difference between internal and external categories relates to the traditional distinction between derivation and inflection. However, this distinction is notoriously hard to make, since many of the criteria proposed to distinguish between the two types of morphological processes do not involve absolute properties. This is reflected in Bybee’s treatment of verbal categories as points on a scale ranging from typically inflectional (external) to increasingly derivational (internal).

The hierarchical structure of verbal categories also forms part of the theory of Functional Grammar (FG; Dik 1997), and its successor: Functional

---

42 In fact, the way in which verbal categories are expressed is also iconically motivated: More internal features tend to reveal more fusion with the stem. I will return to the functional principle of Iconicity in section 3.3.3.2.
Discourse Grammar (FDG; Hengeveld & Mackenzie 2008). According to F(D)G, utterances consist of several representational layers. The innermost layer hosts the predicate with its arguments. On each higher layer a new set of semantic distinctions is added, pertaining respectively to the domains of aspect, tense, modality, and illocutionary force.43

As Malchukov (2004: 16) remarks, the verbal feature hierarchy used in F(D)G’s layered model of the clause is quite similar to the hierarchy proposed by Bybee (1985), but there are also some differences. First, the F(D)G approach takes into account both morpho-syntactic and lexical expressions of TAM categories, whereas Bybee considers bound morphemes only. In F(D)G, morpho-syntactic expressions of categories are termed operators, whereas lexical expressions are called satellites. The inclusion of the latter implies that adverbs of manner, time, etcetera, count as expressions of verbal TAM categories.

A second difference between Bybee’s hierarchy and the framework of F(D)G is that agreement is included as a verbal category in the former but not in the latter. F(D)G represents agreement at a different grammatical level than TAM categories, because, as mentioned above, agreement reflects the syntactic relation between the predicate and its argument(s). As such, it has no semantic representation and belongs to the level of morpho-syntactic encoding. In contrast, TAM distinctions concern the semantics of their base and therefore belong to the representational level of grammar. As we will presently see, Malchukov includes agreement as a category in his version of the verbal feature hierarchy, and he applies the F(D)G distinction between operators and satellites to it. In particular, Malchukov regards an agreement marker as an operator expressing argument realization, and the actual encoding of the corresponding argument as the satellite-type expression of the same category.

A complicating factor in this respect is that the category of agreement can be decomposed into subject agreement and object agreement. According to Malchukov (2004: 19), these two sub-categories occupy different locations on the verbal feature hierarchy. Object agreement pertains to verbal valency, i.e. to the innermost level of the feature hierarchy. Subject agreement, on

43 The FDG model rather works top-down and outside-in. Thus, the structural building of an utterance starts from the speaker’s communicative intention, and works down to its actual morpho-syntactic and phonological encoding. In addition, the underlying representation of an utterance is build up starting with the most external categories, while FG representations start with the innermost categories followed by the more external ones.

82 | Parts of Speech and Dependent Clauses
the other hand, is more relevant to perspectivization, i.e. to the pragmatics of the base unit, and as such qualifies as an external category. Moreover, as just mentioned, the morphological expression of agreement is linked to the lexical expression of the corresponding argument(s). This means that lexical subject and object expression occupy the same respective locations on the verbal feature hierarchy as the concomitant agreement morphology.

A final, less important difference between Bybee’s feature hierarchy and FG’s layered clause structure involves illocutionary force. The fact that FG includes this category is a consequence of the fact that this framework takes the clause as its basic unit of analysis, whereas Bybee looks more narrowly at the verb phrase 44.

Integrating the points discussed above, Malchukov (2004: 20) proposes the verbal hierarchy represented in (27). This is the version of the hierarchy that will be adopted in the remainder of this chapter.

(27) VERB stem
   ⊑
   voice/valency, direct object, object agreement
   ⊑
   aspectual operators, adverbial satellites with aspectual value
   (manner adverbs)
   ⊑
   tense and mood operators and corresponding satellites
   (temporal/modal adverbs)
   ⊑
   subject agreement, clausal subject
   ⊑
   illocutionary force markers

3.3.2.3 The nominal feature hierarchy

As regards the nominal feature hierarchy, Malchukov (2004) again integrates several earlier proposals. One of them is the hierarchy of Lehmann and Moravcsik (2000: 753), which is represented in (28):

44 In FDG, the basic unit of analysis is a communicative Move, which may consist of several discourse Acts. These discourse acts do not necessarily correspond to clauses at the morphosyntactic level. For a full discussion, see Hengeveld & Mackenzie 2008.
As in the case of Bybee's verbal feature hierarchy, the ordering of the nominal features is presumably motivated in terms of iconicity: it reflects the relative degree of relevance of the morphological categories for the interpretation of the stem.

This becomes clearer when considering a second, more detailed proposal, namely Rijkhoff's (2002, 2008a, b) hierarchical model of the noun phrase. This model consists of five concentric layers at which nominal features may operate. Working in the F(D)G framework, Rijkhoff includes both morpho-syntactic and lexical expressions of nominal categories, i.e. both operators and satellites. Furthermore, he proposes a basic distinction between descriptive modification of the nominal base, which involves the four most internal levels of nominal features, and one external level of discourse-referential modification. In other words: the features belonging to the four internal levels influence the semantic meaning of the base, while the most external features are relevant to its pragmatics, i.e. “the status of the thing (or event) as a discourse entity” (Rijkhoff 2008a: 67). First, consider in more detail the four layers of descriptive modification:

(i) The kind layer, which has scope over the lexical head noun and accommodates classifying modifiers that further specify what kind of entity is being referred to by the speaker. In other words, they specify a particular subclass of the entity denoted by the base. They can be satellites, as in ‘a corporate lawyer’ or ‘an electric toothbrush’. Classifying operators are so-called ‘nominal aspect markers’, such as singulative/collective markers.

(ii) The quality layer, which has scope over the kind layer and accommodates lexical modifiers (satellites) that specify properties concerning size, color, weight, value, age, etcetera.\(^{45}\)

(iii) The quantity layer, which has scope over the quality layer and accommodates grammatical and lexical expressions of number.

(iv) The location layer, which has scope over the quantity layer and accommodates modifiers that specify properties concerning the

\(^{45}\) According to Rijkhoff (2008a: 794), qualifying operators, i.e. grammatical expressions of quality do not exist.
location of the referent. These can be satellites and take the form of case-marked or adpositional phrases, including possessive phrases. Operators that belong to the locational level are demonstratives.

The highest or outermost layer distinguished in Rijkhoff’s model is the layer of discourse-referential modification. Operators at this layer are markers of definiteness, i.e. articles or determiners. Examples of discourse-referential satellites in English would be *same* or *other*, i.e. lexical expressions that “provide the addressee with information about the referent as a discourse entity” (Rijkhoff 2008a: 789). The five layers of nominal features are represented in (29):

(29) \[N-STEM\]

- **Descriptive**
  - **Kind**: Nominal aspect, quantifying adjectives
  - **Quality**: Qualifying adjectives

- **Modification**
  - **Quantity**: Lexical/grammatical Number
  - **Location**: Demonstratives, locative/possessive phrases

- **Discourse-Referential Modification**
  - **Modification**: Determiners, lexical discourse-related modifiers

Note that Rijkhoff, unlike Lehmann and Moravcsik (2000), does not take into account case marking as a nominal feature. This is because (core) case marking reflects the syntactic function of the unit it operates on. Therefore, like verbal agreement, it is represented at the level of morpho-syntactic encoding in F(D)G.

The empirical justification for the iconic ordering of nominal features is somewhat fragmented. Early evidence for the hierarchy comes from Greenberg in the form of his Universals 20 and 39, which refer to the relative orderings of case markers, demonstratives, numerals and adjectives (Greenberg 1966: 87, 95). These findings are further supported by Hawkins’s research (1983) based on a sample of more than 300 languages, as well as by

Rijkhoff (2008a: 797) notes that (adnominal, restrictive) relative clauses can also be used for modification at the locational level, as in *the book, that’s lying on the table*. For the purpose of the present study, however, it is assumed that relative clauses are equivalent to lexical adjectives in terms of their propositional function.
Rijkhoff’s (2002) study of a balanced sample of 52 languages. Malchukov (2004: 23) also tests the full range of nominal features on his 50-language sample and reports confirmation of the hierarchical pattern.

Malchukov (2004: 47) integrates the findings from the studies discussed above into the nominal feature hierarchy represented in (30):^47^:

\[(30)\]

- NOUN stem
  - classifying/qualitative operators/satellites: singulative/collective markers, noun classifiers, nominal class markers, adjectives
  - quantitative operators/satellites: number markers, numerals
  - locative/possessive phrases
  - determiners
  - case markers/adpositions

3.3.2.4 Summary, outlook
Having introduced the hierarchies of verbal and nominal categories, the next question is how these hierarchies are reflected in the behavioural potential of DC constructions. In other words, what kinds of mixing patterns of verbal and nominal features are actually displayed by DC constructions of the world’s languages? This is the topic of the next subsection.

3.3.3 Feature mixing in DCs across languages: Cross-linguistic generalizations and functional motivations

3.3.3.1 Introduction
Since the 1970s, a considerable number of typological studies have been carried out on the behavioural potential of DCs (e.g. Comrie 1976; Noonan 1985/2007; Lehmann 1988; Koptjevskaja-Tamm 1993; Mackenzie 1996; Croft 1991, 2001; Dik 1997; Cristofaro 2003; Malchukov 2004, 2006; Dixon & Aikhenvald

^47^ Note that the most recent version of Rijkhoff’s model, as presented in Rijkhof 2008a/b, is not incorporated in Malchukov’s 2004 study. This is why the latter does not distinguish the innermost level of classifying modification. Since the distinction is not directly relevant to my research, I have not added it to Malchukov’s model as represented in (30).
Together, these studies provide a body of evidence showing that the cross-linguistic variation in DC constructions takes the form of a continuum, ranging from DCs that express the full set of available verbal categories and no nominal features, to DCs that have lost (almost) all verbal features and express (almost) all nominal features. In the literature, starting with Stassen (1985), DCs of the first type are called *balanced*; they can express the same set of verbal features as independent clauses in the particular language under analysis. All other DC types are called *deranked*, meaning that their behavioural potential deviates to a greater or lesser extent from the coding of an independent clause, in terms of de-categorization (i.e. elimination of verbal categories) and/or re-categorization (i.e. expression of nominal categories). As already anticipated by the discussion of the verbal and nominal feature hierarchies in the previous subsection, formal deviation from independent-clause-coding may affect different elements of the DC. The individual structural phenomena that may obtain are listed in (31a-f) (cf. Croft 2001):

(31)   a. (Partial) elimination of tense/mood/aspect marking (operators and/or satellites);
   b. (Partial) elimination of person marking;
   c. Expression of definiteness (determiners) and/or case/adpositions;
   d. Expression of number and/or class;
   e. Non-expression of argument(s) in the dependent clause, or expression of argument(s) in a different manner than in an independent clause;
   f. Adjectival rather than adverbial expression of lexical modifier(s)\(^{48,49}\).

Even though there are some minor differences, the typological studies mentioned above reveal broadly the same generalizations as regards the (co-)occurrence in DC constructions of the features listed in (31). In what follows I will not give an overview of this literature\(^{50}\). Rather, I will discuss two of its

\(^{48}\) For the time being I ignore the possibility that TAM and/or agreement are coded by means of dependent/subjunctive forms. This phenomenon is of minor importance to the present discussion, but I will return to it in section 3.4.

\(^{49}\) Usage of different word order than in simple main clauses is another possible form of structural deviation. However, since it is quite marginal and usually occurs in combination with some other strategy, I do not take it into account in the present study.

\(^{50}\) See Malchukov (2004: 9-12) for a brief overview of this kind.
most recent exponents in some detail, namely the work of Cristofaro (2003), and Malchukov (2004).

3.3.3.2 The typology of subordination (Cristofaro 2003)
First, Cristofaro’s work probably represents the broadest typological study of subordination to date: It takes into account complement clauses, relative clauses, and adverbial clauses in a sample of 80 languages. Cristofaro identifies a number of correlations between various deranking phenomena in DCs. These correlations are listed in (32a–h) (Cristofaro 2003, chapter 10):

(32) a. Agreement not expressed  ➔ TAM not expressed
   b. Case/adposition marking ➔ TAM not expressed
   c. Case/adposition marking ➔ Agreement not expressed
   d. Arguments expressed as POSS ➔ TAM not expressed
   e. Arguments expressed as POSS ➔ Agreement not expressed
   f. Arguments expressed as POSS ➔ Case/adposition marking
   g. Arguments not expressed ➔ TAM not expressed
   h. Arguments not expressed ➔ Agreement not expressed

Cristofaro proposes three types of functional factors underlying these correlations, two of which are the general functional principles of (Syntagmatic) Economy and Iconicity (Haiman 1983). These principles account for the non-expression of verbal features in DCs, including the non-expression of arguments.

First, consider the principle of Economy, which refers to “the tendency to reduce as much as possible the length or complexity of any utterance” (Cristofaro 2003: 9). As Cristofaro shows, in some subordination relations the semantics of the matrix clause predetermine certain TAM values and/or a participant of the SoA expressed by the dependent clause. There are three possibilities in this respect:

(i) The matrix clause predetermines both TAM value(s) and a participant of the dependent clause;

51 The generalizations in (32b) and (32c) are not supported by my data. In addition, as I will become clear shortly, they are not in accordance with the typology proposed by Malchukov (2004). See Lehmann (1988) for a proposal similar to Malchukov’s, and Mackenzie (1996) for a proposal that supports Cristofaro’s findings. See also Fischer & Van Lier (forthcoming) for more discussion.
(ii) The matrix clause predetermines TAM value(s) (but no participant) of the dependent clause;

(iii) The matrix clause predetermines neither TAM value(s) nor any participant of the dependent clause.

There are no subordination relations in which TAM values are predetermined, while participants are not, which means that predetermination of participants implies predetermination of TAM values. The Economy principle predicts that predetermined information may be left unexpressed (this is called the Principle of Information Recoverability by Cristofaro (2003: 258)). Thus, this principle explains the correlation mentioned under (32g) above, namely that non-expression of arguments implies non-expression of TAM distinctions. Moreover, the non-expression of arguments in turn implies the non-expression of concomitant agreement morphology, as is reflected in the correlation mentioned under (32h) above.

The principle of Iconicity refers to “the tendency to shape linguistic structure in conformity with the structure of conceptual experience” (Cristofaro 2003: 8). Iconicity explains the same structural phenomena as the principle of Economy, namely non-expression of verbal categories and arguments in DCs, but it does so in terms of semantic integration, rather than in terms of semantic predetermination. Increased semantic integration between the SoAs expressed by the two linked clauses arguably leads to reduced linguistic independence of the DC, which is in turn reflected in the non-expression of verbal features and/or arguments in the latter. This point was already made by Givón (1980 and later work), who proposed the so-called binding hierarchy or scale of event integration, stating that “the stronger the semantic bond between two events, the more extensive will be the syntactic integration of the two clauses into a single though complex clause.” (Givón 2001: 40)

However, neither Economy nor Iconicity can account for the expression of nominal features in DCs. Therefore, Cristofaro proposes a third explanatory principle, which has a basis in Cognitive Linguistics (Langacker 1987). According to this principle, DCs are conceptually not

---

52 This suggests that the principle of Iconicity is superfluous, to the extent that it does not explain any more than the Economy principle does. Indeed, Haspelmath (2008) argues that Iconicity is irrelevant for the explanation of cross-linguistic generalizations found in the coding of DCs. In his view, they should rather be explained in terms of frequency of use. Since frequency implies predictability, this amounts to an explanation in terms of the Economy principle.
construed as independent processes, and may instead be construed as things or properties. This means that they are interpreted ‘as wholes’, without any internal temporal development (cf. Chapter 2, section 2.2). Using the terminology of Cognitive Linguistics, DCs are said to be “summarily rather than sequentially scanned.” (Cristofaro 2003: 284)

In fact, according to Cristofaro, this principle of cognitive thing-construal explains both the non-expression of verbal features and the expression of nominal features in DCs. First, verbal categories such as TAM values express meanings that are semantically relevant to processes or SoAs, but not to things. Therefore, if a DC is construed as a thing, then TAM distinctions can be omitted. A similar argument can be made for verbal agreement: It marks the relationship between a predicate and its arguments, and as such it is relevant to an SoA, but not to a thing. The fact that non-expression of TAM and non-expression of agreement have a parallel motivation accounts for the correlation between these two phenomena, mentioned under (32a) above and repeated here for convenience as (33):

(33) Agreement not expressed → TAM not expressed

Second, Cristofaro claims that the conceptualization of DCs as things explains the expression of nominal properties such as case or adposition marking, and the coding of arguments as possessors. In particular, she argues that a DC may “display the same properties as the grammatical entities that prototypically code things, that is, nouns.” (Cristofaro 2003: 262-63)

Now, conceptualization as a thing presupposes non-sequential scanning. Thing-construal is formally reflected by the expression of nominal categories, while non-sequential scanning is reflected by non-expression of verbal categories. This explains the universals of co-occurrence, as listed under (32b-γ) and repeated here as (34a-e) (Cristofaro 2003: 284-285). Specifically, the expression of the nominal features of case/adposition and possessive argument coding implies the non-expression of the verbal features TAM and agreement. The correlation between coding of arguments as possessors and the expression of case/adpositions (see 34e) is explained by the fact that these phenomena both reflect conceptualization of the DC as a thing.

For ease of reference, in the remainder of the discussion I will use thing-construal, rather than to thing- or property-construal. Basically, however, the same claims are made for DCs that are cognitively construed as properties as for those that are construed as things. For further details on property-construal see Cristofaro (2003).
Clearly, this third explanatory principle – thing-construal of DCs – is most relevant to the present study, since it hinges on a functional parallel between DCs and PoS. Specifically, Cristofaro’s explanation makes reference to the notional similarity between nouns and DCs to explain the attested formal patterns: both are construed as ‘things’. However, recall from Chapter 2 that Croft defines prototypical nouns as lexical constructions that combine the semantic denotation of thing (or ‘object’ in Croft’s terms) with the pragmatic function of reference (see section 2.4.2). Recall further that, as argued in section 3.2, it is the pragmatic or propositional function, rather than the semantic denotation that DCs have in common with PoS. In fact, DCs do not denote things or properties; they denote higher order entities such as SoAs and propositions54,55. Therefore, it seems that the expression of nominal (or other non-verbal) features in DCs should be explained in terms of the fact that DCs share the propositional function of reference or modification with prototypical nouns, adjectives, and manner adverbs. Such a pragmatic explanation for patterns of nominal and verbal feature mixing in DCs is proposed in the second typological study of subordination that I want to discuss: Malchukov’s work on nominalizations, as outlined in the next subsection.

3.3.3.3 The typology of nominalization (Malchukov 2004)

Malchukov’s typological study of clausal nominalizations shows that verbal and nominal categories at the external end of their respective feature hierarchies (as discussed in section 3.3.2.3 above) are more readily affected by processes of de-categorization and re-categorization in DCs than

54 cf. Noonan, who defines complementation the as situation in which “a notional sentence or predication is an argument of a predicate” (Noonan 2007: 52; emphasis added, EvL).
55 It may be argued that cognitive thing-construal is not the same semantic thing-denotation. However, as Cristofaro (2003: 302-303) herself admits, there is no independent evidence for cognitive thing-construal. Cristofaro argues that the cross-linguistic generalizations attested in her study can be regarded as supporting her cognitive explanation. I would argue, however, that the linguistic facts to be explained in terms of cognitive construal cannot at the same time function as evidence for such cognitive construal.
internal features. According to Malchukov, the functional explanation for this generalization must be sought in the fact that external categories reflect the syntactic and/or pragmatic function of the linguistic unit on which they operate, while internal categories are relevant for the semantic interpretation of their base. Specifically, since “the ultimate task of nominalization is to produce a referential expression” (Malchukov 2004: 26), this explains why external nominal features, which are relevant to the pragmatic function of reference, are expressed in nominalizations before internal ones. In addition, it explains why external verbal features, which are relevant to pragmatic and syntactic aspects of (independent) predications, are the most likely ones to be lost in nominalization processes.

At this point, it should be mentioned that the idea of a pragmatic motivation for feature mixing in DCs ties in with Dik’s Principle of Formal Adjustment (PFA). This principle states that:

“Derived, secondary constructions of type X will be formally assimilated (adjusted) to non-derived, primary constructions of type X.”

(Dik 1997: 158)

Applying the PFA to the case at hand, we can say that DCs are derived, secondary referential or modifying constructions (rather than derived, secondary thing-denoting or property-denoting constructions). Their primary, non-derived referential or modifying counterparts are lexical nouns, adjectives, and adverbs. Thus, Dik’s PFA predicts that DCs, as secondary constructions, will adopt formal features of primary lexical constructions.

Turning back to Malchukov’s study, his finding that external features are affected in nominalizations before internal ones is further specified in terms of two implicational hierarchies. These implicational hierarchies are based directly on the feature hierarchies discussed in section 3.3.2 above. Regarding nominal categories, the generalization is that the expression of a particular feature in a nominalization implies the expression of any feature that occupies a more external position on the nominal feature hierarchy. Regarding verbal categories, in contrast, the loss (or non-expression) of a particular feature in a nominalization construction entails the loss of any feature that occupies a more external position on the verbal feature hierarchy.

In addition to these implicational effects of the verbal and nominal feature hierarchies, Malchukov claims that there are other factors that constrain cross-linguistic variation in verbal and nominal feature combinations. The most
important factor involves so-called blocking-effects. Basically, ‘blocking’ refers to the fact that certain combinations of nominal and verbal categories are incompatible as a result of their functional equivalence. In particular, arguments and modifiers can be expressed either verbally or nominally, but not in both ways at the same time. Therefore, the verbal expression of an argument or a modifier blocks the nominal expression of that element, and vice versa.

Malchukov (2004: 66–67) distinguishes three types of blocking effects: subject blocking, object blocking, and modifier blocking. In combination with the implicational effects of the verbal and nominal feature hierarchies, this yields a three-way typology of nominalizations, consisting of the following types:

(i) nominalizations with no blocking;
(ii) nominalizations with subject blocking;
(iii) nominalizations with object blocking or modifier blocking.

First consider nominalizations with no blocking. As far as nominal features are concerned, this type of nominalization may express only the two outermost nominal categories (in terms of the feature hierarchy), i.e. determiners and case markers. This is because these categories create no blocking effects, as they have no verbal functional counterparts. This means that all verbal features can be retained (with the exception of illocutionary force, which is unexpressed by definition in dependent clauses), including the expression of arguments and modifiers as in independent clauses \(^{56}\). The feature combination that thus characterizes nominalizations with no blocking is represented in (35) below, where the dashed line indicates the boundary between verbal features (above the line) and nominal features (below the line), i.e. the point where the two feature hierarchies intersect. The brackets around the nominal features indicate optionality: case/adposition and determiners can but need not be expressed in nominalizations of this type.

\(^{56}\) Note that this is contra Cristofaro’s generalizations in (34a) and (34b) above, according to which case or adposition marking implies the loss of verbal features (TAM and AGR).
Nominalizations with no blocking:

STEM

- voice/valency, direct object, object agreement
- aspect
- tense and mood
- subject agreement, clausal subject

The second type of nominalization involves subject blocking. This means that one argument – the subject or agent – is expressed as a possessor. This nominal subject encoding blocks verbal subject encoding and as such also disrupts concomitant verbal subject agreement. This in turn implies the non-expression of any verbal features external to subject agreement, i.e. the category of illocutionary force. All other verbal categories can be retained, and any arguments other than the subject/agent are expressed as in independent clauses. The feature combination of nominalizations with subject blocking is represented in (36), again with a dashed line marking the boundary between the verbal and the nominal feature hierarchies:
The third and final type concerns *nominalizations with object blocking or modifier blocking*. As its label makes clear, this type is in fact a combination of two types, which are merged because they look formally very much alike. In these nominalizations, the object or patient argument is coded as a possessor and/or modification is adjectival. These two features involve blocking effects that pertain to the innermost layers of both hierarchies: Object blocking affects the valency layer of the verbal hierarchy, while modifier blocking affects the qualitative layer of the nominal hierarchy. Therefore, they imply the loss of all other verbal categories and the expression of all other nominal categories, respectively. The morpho-syntactic effects of both types of blocking are thus quite similar. The only difference is that in the case of adjectival modification the valency layer of the verbal hierarchy need not be affected, as it always is in the case of possessive object expression. The feature combination of nominalizations with object/modifier blocking is represented in (37):
Malchukov’s framework thus predicts a three-way typology of nominalization constructions. However, not all nominalizations attested in the languages of his sample fit the above typology. According to Malchukov, this is because there are several structural factors that may complicate the combined effects of blocking and implicational hierarchies.

First, there is a difference between the behaviour of morpho-syntactic versus lexical instantiations of the same categories, i.e. between operators and satellites. In particular, the possibility to express a specific nominal or verbal category by means of an operator implies the possibility to express that category with a satellite, but not vice versa: the expression of a satellite does not mean that the corresponding operator is also available. Malchukov (2004: 37) calls this the **Operator–Satellite Asymmetry principle** (henceforth OSAP). Recall from section 3.3.2.2 that Malchukov regards lexical arguments and agreement morphology as, respectively, the satellite and the operator type expressions of the same category. Thus, since nominal argument expression blocks verbal argument expression (satellites), the OSAP predicts the loss of concomitant verbal agreement (operators). In contrast, verbal expression of arguments (satellites) can go together with either retention or loss of verbal agreement (operators). With respect to modifier expression, the OSAP explains why adverbial modification (satellites) does not entail the availability of morphologically expressed TAM distinctions (operators). This means that for nominalizations without modifier blocking no strong predictions can be made about TAM operators; they may be either lost or retained.
Second, the outcome of feature mixing in nominalization constructions can be complicated by cumulative expression of multiple categories through a single morpheme. If such categories belong to different levels of either the verbal or the nominal feature hierarchy, then a conflict arises that will result in either the retention or the loss of all categories expressed by the relevant morpheme. (See Malchukov (2004: 111) for more discussion and examples, and Malchukov (2007) for an Optimality Theory implementation of the interaction between functional and structural factors in determining the expression of nominalization constructions.)

It is noteworthy in the context of the present study that Malchukov explicitly claims his typology of nominalizations to be applicable to languages without a lexical distinction between nouns and verbs. He argues that the model works as long as a noun-verb distinction is present on the syntactic level, i.e. as long as verb phrases and noun phrases involve different categories (Malchukov 2004: 126). In other words, the expression of ‘nominal’ categories in a DC does not necessarily imply the existence of a class of rigid lexical nouns; it merely means that there is a certain phrase type – headed by a lexical element whose categorial value may be indeterminate in isolation – that is used to express the function of reference.

This is also the reason why in the present study the relevant set of DC constructions is initially determined on the basis of structural coding, rather than behavioural potential. The expression of behavioural potential in a DC reflects its non-prototypical combination of propositional function and semantic denotation; it does not have a direct connection with the lexical categories available in a particular language. However, once a particular DC construction is determined on the basis of the structural coding criterion, it may be further classified according to its behavioural potential. This allows one to address a second question, namely whether the internal structure of DC constructions is a factor of influence on the relation between DCs and PoS, in terms of their ability to express particular propositional functions.

3.3.4 Summary, outlook
In this section I discussed the typology of the internal morpho-syntax of DCs. Starting out with a presentation of verbal and nominal feature hierarchies, I subsequently discussed cross-linguistic patterns pertaining to the mixing of these two feature types in DC constructions, and the functional motivations that have been proposed to explain these patterns. This section thus provides the basis for the second part of the typology that I will present.
in the next section, and that will be used in Chapter 6 to classify the DCs of the languages in my sample.

3.4 A two-step typology of dependent clauses

The classification of DCs that I propose involves a two-step procedure, which integrates the two types of characteristics of DCs that have been at stake in the present chapter: structural coding and behavioural potential. First, DCs are identified purely on the basis of their structural coding, i.e. the form of either the subordinating conjunction or the special marker on the dependent predicate. For every DC construction identified in this way, it is determined which propositional function(s) it can express. This yields a classification of DCs in terms of the rigid and flexible types presented in section 3.2.2.

Secondly, every rigid or flexible DC construction can be classified according to its internal morpho-syntactic properties, i.e. its behavioural potential. For this part of the typology I will adopt a three-way distinction between DC types, to be presented shortly. This typology is quite similar to Malchukov’s classification of nominalizations discussed above, but also differs from it in some respects. A first difference is that Malchukov’s study does not include fully balanced DCs as a separate category. His nominalization type with no blocking must have verbal arguments, but it does not necessarily express all TAM/agreement categories. In contrast, I will distinguish between:

(i) DCs that express their arguments as in independent clauses (i.e. verbally) and retain all verbal (TAM/person) categories that can be expressed in independent clauses;

(ii) DCs that express their arguments as in independent clauses (i.e. verbally) but do not retain (all) the verbal (TAM/person) categories that can be expressed in independent clauses.

Note that both these DC types may or may not combine with case markers or determiners, an issue to which I return below.

Second, Malchukov uses the verbal versus nominal encoding of subject and object arguments as one of the decisive criteria for his classification of nominalizations. This means that his model does not provide an obvious locus for DCs in which one or more argument(s) remain un-expressed. The typology that will be used here does take this possibility into account, as I will explain below.
Third, while Malchukov makes a principled distinction between constructions with one possessive argument (nominalizations with subject blocking) and constructions with two possessive arguments (nominalizations with object blocking), I will regard these two types as belonging to a single DC type, in which at least one argument is expressed nominally.

In line with previous studies on subordination, the three-way typology of DCs that I will adopt is based on three formal parameters, as listed in (38):

(38)  a. The (partial) (non-)expression of TAM operators and/or person marking;  
      b. The (non-)expression of the nominal categories Determiner and/or Case/adposition;  
      c. The coding of the argument(s).

Some remarks are required concerning these three parameters. First, it may be noted that the verbal versus nominal expression of lexical modifiers (i.e. adverbial versus adjectival modifiers) is not taken into account. This is because this feature is not often well-documented in reference grammars.

Second, in terms of verbal operators (parameter (38a)), the following features will be taken into account: voice and valency marking; tense, aspect, and mood marking; person marking; and illocutionary force. Following Malchukov (2004; and see also Noonan 2007: 68), object marking will be regarded as pertaining to the innermost level of verbal valency.

Third, with regard to nominal features (parameter (38b)), I take into account only the expression of determiners and case markers or adpositions. Like Cristofaro (2003), but unlike Malchukov (2004), I do not systematically consider other nominal features, such as number and gender/class marking, since they are very infrequently attested in the DCs of the sample languages (cf. Chapter 6).

Finally, consider parameter (38c) in some more detail. Argument coding in DCs can take three possible forms, as listed in (39):

(39)  a. An argument is expressed as it would be in an independent clause;

57 As mentioned earlier, I use the term person marking to include both agreement and cross-referencing on the predicate.
b. An argument is expressed in a different way than it would be in an independent clause;

c. An argument is not expressed.

The possibilities in (39a) and (39c) are straightforward enough. Following the terminological convention of Koptjevskaja-Tamm (1993), I will call the type of expression in (39a) ‘sentential’, abbreviated as $\text{SENT}$, and the type in (39c) ‘zero’, indicated by $\varnothing$. The parameter value in (39b), however, may manifest itself in several different forms. Most typically, it involves coding of the argument as a possessor, as in example (40) from Turkish (note that the possessive subject triggers concomitant nominal agreement on the dependent predicate).

*Turkish* (Göksel & Kerslake 2005: 420)

(40) \[
\begin{array}{lll}
\text{Kerkes-in} & \text{birier} & \text{kikaye} & \text{anlat-ma-si}
\end{array}
\]
\[
\begin{array}{llllllll}
\text{everyone-gen} & \text{one.each} & \text{story} & \text{tell-NMLZ-3SG.POSS}
\end{array}
\]
\[
\begin{array}{llllllllllllllllllll}
\text{iste-n-iyor-mu}
\end{array}
\]
\[
\begin{array}{llllllllllllllllllll}
\text{want-PASS-IPFV-EV.COP}
\end{array}
\]

‘It seems that they want everyone to tell a story.’

In other cases, however, the argument may be coded as an oblique. This may involve an adposition, as in example (41) from Georgian, or a special case marker, such as the dative in example (42), also from Georgian.

*Georgian* (Hewitt 1995: 542, 609)

(41) \[
\begin{array}{llllllllllllllllllll}
\text{mokalake-ta} & \text{mier} & \text{upleb-eb-is}
\end{array}
\]
\[
\begin{array}{llllllllllllllllllll}
\text{citizen-PL(GEN )} & \text{by} & \text{rights-PL-GEN}
\end{array}
\]
\[
\begin{array}{llllllllllllllllllll}
\text{ga-mo-q'en-eb-a}
\end{array}
\]
\[
\begin{array}{llllllllllllllllllll}
\text{PREV-PREV-USE-THS-NMLZ}
\end{array}
\]

‘the making use of their rights by the citizens’

(42) \[
\begin{array}{llllllllllllllllllll}
\text{sa-upl-eb-ul-i}
\end{array}
\]
\[
\begin{array}{llllllllllllllllllll}
\text{literary language-DAT} & \text{PREV-master-THS-PST.PTC-NOM}
\end{array}
\]

‘(he) who has mastered the literary language’

A third possibility is the encoding of an argument in the form of an attributive or associative (nominal) modifier, as in example (43) from Lango.
Finally, an argument may be incorporated into the dependent predicate, as in the Ket example in (44):

**Ket** (Vajda 2004: 78)

(44) \[ \text{ə́ k-nà qús-bèt] bínùt} \]

1pl-anim.pl.gen tent-make.inf it.ended

‘We finished making the tent.’ (lit. ‘Our tent making ended.’)

In this study, all sub-types of non-sentential argument coding illustrated in examples (40)-(44) – possessor-coding, oblique-coding, attributive/associative coding, and incorporation – will be regarded as a single strategy. This strategy is characterized by the fact that it deviates from argument coding in independent clauses, and will be called *alternative coding*, abbreviated as ALT58.

A final issue, related to argument coding, concerns alignment patterns. In this study I will distinguish between first and second arguments only. In this way I generalize over constructions with nominative-accusative and ergative-absolutive alignment. This is unlike the approach taken by Koptjevskaja-Tamm (1993, 2003, 2005), who distinguishes between so-called *possessive-accusative* and *ergative-possessive* constructions. In the former type, the S and A arguments are expressed as possessors, while the P argument is expressed as in an independent clause (i.e. with accusative coding). In the latter case, the S and P arguments are expressed as possessors, in contrast to the A, which is coded as in an independent clause (i.e. with ergative coding). In my classification, these two types of alignment are collapsed into a single strategy, characterized by the fact that in a transitive DC construction one argument gets alternative coding (ALT), while the other one retains sentential coding (SENT).

58 cf. Koptjevskaja-Tamm (1993, 2003, 2005), who treats possessive coding, oblique coding, and incorporation all as separate types. Note that initially I have followed this more finegrained distinction, but abandoned it again in a later stage, as it did not influence the results in any interesting way.
Taking these remarks into account, I present below the three-way typology of DC constructions to be used in this study. Note that each type is indicated with both a number and a mnemonic label.

- **Type 1: Balanced constructions** (abbreviated as *B*)
  Characterized by the following three parameter values:
  a. All TAM distinctions and person marking are expressed as in independent clauses;
  b. All overt arguments are coded as in independent clauses (SENT-SENT);
  c. Determiners (DET) and/or case markers/adpositions (CASE)\(^{59}\) can but need not be expressed.

With respect to the second parameter value (argument expression), I regard relative clause constructions with a gap-strategy as balanced, even though one of the arguments (the relativized one) of the DC remains unexpressed. Relative clauses with a gap are defined by Comrie & Kuteva (2005: 494) as those involving “no overt case-marked reference to the head noun within the relative clause”.

The third parameter value, according to which balanced DCs may combine with determiners and/or case markers/adpositions, also requires some clarification. Recall that in Malchukov’s typology a balanced clause with a case marker or adposition would be a nominalization with no blocking, which allows for the retention of all verbal categories (see (35) above). This type of construction would constitute a counterexample to the universal claimed by Cristofaro (2003), namely that the expression of case or adposition marking implies the loss of TAM and agreement (see (32b-c)/(34a-b) above). Nonetheless, there are several instances of this type of construction in my sample. For instance, Guaraní has a complement clause construction that takes a determiner but is otherwise completely balanced, as can be seen in (45). Basque has balanced complement clauses with bound complementizers and determiners, as is illustrated in (46). A similar construction, with a bound complementizer and a case marker, is attested in Cofán, as shown in (47).

\(^{59}\) The abbreviation CASE is meant to include both case markers and adpositions.
Guaraní (Gregores & Suárez 1967:158)

(45) ai-kwaá [la [n o-ù mo?á i ha]]
    I-know ART NEG he-go MOD NEG COMP
    ‘I know that he does not intend to go.’

Basque (Hualde & Ortiz de Urbina 2003: 646)

(46) Entzun dut [[Amaiaren neba bil d-en]-a]
    hear AUX Amaia.gen brother die AUX-COMP-DET
    ‘I heard that Amaia’s brother died.’

Cofan (Fischer & Van Lier, forthcoming)

(47) Atesu-mbi [[ke ña=nga tevaen=’cho]=ma]
    know-NEG you I=DAT write=COMP=ACC
    ‘I didn’t know that you had written to me.’

The fact that these constructions are fully balanced, apart from the expression of a determiner or case marker, is in accordance with the nominal feature hierarchy and with Malchukov’s typology: Determiners and case-markers are external features that mark the syntactic-pragmatic function of their base unit, without affecting its internal morpho-syntactic structure.

Note that determiners and case markers may also be expressed on balanced DCs that are not marked by a complementizer. For determiners this was already illustrated with an example from Dhaasanac (see (5) in section 3.2.2). A similar construction is attested in Cofán, as is illustrated in (48):

Cofán (Fischer & Van Lier: forthcoming)

(48) [Doscientos ochenta dolar=ma gana-’je-’fa]=ma
    two.hundred eighty dollar=ACC earn-IPFV-PL=ACC
    in’jan kan-’se.
    think try-MIT
    ‘Imagine that they earn 280 dollars.’

Finally, it must be noted that I take balanced clauses (type 1) to include two sub-types that are strictly speaking not fully balanced. Both deviate slightly from independent clause coding, particularly in terms of the expression of verbal features. These two subtypes are:

In fact, case markers and adpositions are a common historical source for complementizers.
(i) DCs that are fully balanced except for the fact that they cannot express illocutionary force marking, while this is an obligatory category in independent clauses.

(ii) DCs in which the same or almost the same range of verbal categories can be expressed as independent clauses, but by means of dependent or subjunctive forms.

The first sub-type is attested for instance in Nama. In this language, the declarative marker *ke* occurs in every independent clause, while it cannot be expressed in DCs. In all other respects, however, DCs and independent clauses have the same structure. Abkhaz presents an example of the second subtype. In this language, certain DC constructions make use of the so-called ‘non-finite’ verbal paradigm, as illustrated in (49). This paradigm includes almost the same distinctions as the independent paradigm, but expresses them with different forms. As already mentioned in section 3.2.2 (see example (7)), I will regard DCs with dependent or subjunctive forms as belonging to the group of balanced DCs.

*Abkhaz* (Hewitt 1987: 238)

(49)  
\[
\left[ d\text{-}{{\text{şaa}}\text{-}wa} \right] \quad (\theta-)x\text{å}\text{-}s\text{-}c'\text{a}\text{-}w\text{a}\text{-}yt'.
\]

\[\text{he-come\-DY}(n.\text{FIN.PRS}) \quad \text{it-head.in-I-put\-DY\-FIN(PRS)}\]

‘I believe that he will come.’

We have now rounded off the discussion of the first DC type, and may consider the second one:

- **Type 2: Deranked DCs with SENT arguments** (Abbreviated as D-SENT)
  Characterized by the following parameter-values:
  d. (Partial) non-expression of TAM distinctions and/or Person marking;
  e. All overt arguments are expressed as in independent clauses (SENT-SENT);
  f. DET/CASE can but need not be expressed.

This D-SENT type also (like type 1) resembles Malchukov’s nominalizations with no blocking (see (35) above), but differs from it in two respects, namely:
(i) Malchukov’s Nominalizations with no blocking do not necessarily involve (partial) loss of TAM and/or person distinctions, whereas my Type 2 D-SENT DCs do;

(ii) Malchukov’s Nominalizations with no blocking do not explicitly include constructions in which one or more arguments remain unexpressed, whereas my Type 2 D-SENT DCs do.

An example of a D-SENT construction with two overt arguments is given in (50). This Tamil nominalization involves loss of agreement and partial loss of aspect and mood distinctions, but it retains tense.

*Tamil* (Asher 1982: 20)

(50) [tiruŋan nakeyellaam eguttukkiṭṭirukkar-ate]

\[
\begin{align*}
\text{thief} & \quad \text{jewellery.all} & \quad \text{take.progr.prs-nmlz:acc} \\
\text{naan} & \quad \text{patteen} & \\
\text{I} & \quad \text{see-pst-1sg}
\end{align*}
\]

‘I saw the thief taking the jewels.’

D-SENT constructions can but need not be marked for CASE or DET. Some D-SENT constructions are never marked for CASE/DET, such as the Polish infinitive construction illustrated in (51):

*Polish* (Bielec 1998: 19)

(51) 

\[
\begin{align*}
\text{Postaniwi} & \quad \text{em [kupić dom]}.
\end{align*}
\]

\[
\begin{align*}
\text{I.decided} & \quad \text{buy.inf} & \quad \text{house}
\end{align*}
\]

‘I decided to buy a house.’

Other D-SENT constructions are marked for case in some functions but not in others. For instance, Imbabura Quechua has a flexible D-SENT construction, which is case-marked in the function of complement clause (52a), but not in the function of relative clause (52b); except if the relative clause is extra-posed, then the head and the relative clause are both marked for case (52c).
Imbabura Quechua (Cole 1982: 14, 47, 51)

(52) a. Marya nin-n [Juzi jatun wasi-ta cari-shka]-ta
   María  say-3 José big house-ACC have-NMLZ:PST-ACC
   ‘Maria says that José had a big house.’

   b. [Marya riku-shka] runa
   María see-PTC:PST man
   ‘the man whom Maria saw’

   c. Kwitsa-ta juya-ni [Juan-wan tushu-shka
   girl-ACC love-1 Juan-with dance-PTC:PST
   ka-shka]-ta
   be-PTC:PST-ACC
   ‘I love the girl who had danced with Juan.’

These Polish and Imbabura Quechua examples also show that in D-SENT constructions one argument may remain unexpressed: Example (51) illustrates the zero-expression of a co-referential subject argument, and (52b-c) illustrate the use of a gap strategy in a D-SENT construction. Other D-SENT constructions with an unexpressed argument include converbal constructions with co-referential subjects, such as the one illustrated in (53) from Turkish:

Turkish (Kornfilt 1997: 73)

(53) Ben [etraf-ım-a bak-arak] yür-ür-üm
   I around-1SG-DAT look-CONV walk-AOR-1SG
   ‘I walk looking around (myself).’

Finally, D-SENT constructions may lack arguments altogether, even if they are derived from transitive clauses. Such cases typically involve past/passive or future participles, in which the agent argument is demoted and the patient is gapped (Haspelmath 1994). This is illustrated in (54) with a future participle construction from Georgian.
Having discussed DC constructions of type 2, I turn now to the third type:

- **Type 3: Deranked DCs with at least one alternatively coded argument**
  (Abbreviated as D-ALT)
  Characterized by the following parameter-values:
  - (Partial) non-expression of TAM distinctions and/or person marking;
  - At least one argument gets alternative (ALT) expression; the other one, if available/overt, gets either sentential (SENT) or alternative (ALT) expression;
  - DET/CASE are expressed (if relevant in the language).

The Turkish example in (40) above illustrates a DC of type 3 with two overt arguments, one of which has ALT coding, while the other remains SENT.

An example of a Type 3 DC with two ALT-coded arguments appears in (41) above; a Georgian nominalization construction, in which the (transitive) subject takes genitive case, and the object takes oblique coding, with the proposition *mier* ‘by’ (which also requires genitive case). Tense, mood, and person/number agreement are lost, but aspect is retained. This construction can be case-marked or combined with a postposition according to its function in the matrix clause.

Non-expression of an argument may also occur in a Type 3 DC, namely when it combines with a gap strategy. An example is given in (55); a participial construction (again from Georgian), in which the agent argument is in the genitive case and the relativized patient is gapped:
We have now defined three structural types of DCs. In Table 3.1, these three types and their respective parameter values are summarized:

<table>
<thead>
<tr>
<th>DC type</th>
<th>Argument expression</th>
<th>TAM/Person</th>
<th>DET/CASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>number</td>
<td>label</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Type 1</td>
<td>B</td>
<td>SENT/Ø</td>
<td>SENT/Ø</td>
</tr>
<tr>
<td>Type 2</td>
<td>D-SENT</td>
<td>SENT/Ø</td>
<td>SENT/Ø</td>
</tr>
<tr>
<td>Type 3</td>
<td>D-ALT</td>
<td>ALT</td>
<td>SENT/ALT/Ø</td>
</tr>
</tbody>
</table>

Table 3.1: DC types and their internal formal properties

3.5 Summary

In this chapter I have discussed two approaches to the typology of dependent clauses: One based on structural coding of DC constructions and their possibilities to express certain propositional functions; the other based on the internal morpho-syntactic properties of DCs, i.e. their behavioural potential.

Concerning the first approach, I showed that DCs can be defined as mappings onto a space of four propositional functions, in parallel with the definitions of PoS classes developed in Chapter 2. This involved the application of the notions of flexibility and rigidity to the domain of DC constructions, which in turn required a definition of structural coding of these constructions: subordinating conjunctions and dependent predicate markers. A number of expected rigid and flexible DC types were defined in this way.

Second, I discussed the internal morpho-syntax of DCs against the background of the functional-typological literature on verbal and nominal feature hierarchies, and the way they interact in DC constructions across languages. Based on this literature, I developed a classification system of three structural DC types, defined in terms of the (non-)expression of TAM distinctions and Person marking, the type of argument encoding, and the (non-)expression of determiners and/or case markers.

In Chapter 6 the integrated DC typology, based on structural coding and behavioural potential, will be used to classify the DC constructions attested in the languages of the sample.
4.1 Introduction
The aims of this chapter are as follows: First, in section 4.2, it presents the sample of languages investigated in this study. Second, in section 4.3, the concrete research questions that will be addressed in this study are formulated and operationalized. Third, section 4.4 explains the method that will be used to answer these questions. Finally, section 4.5 provides a brief outlook on the remaining part of the study.

4.2 The language sample
The sample used for this study consists of 50 languages and is composed by means of the Diversity Value Technique (henceforth DVT; Rijkhoff et al. 1993, Bakker forthcoming). This technique can be used with any language classification. A sample composed with the DVT consists minimally of one language per family, according to the classification chosen. Absolute isolates (as opposed to isolates within a genetic grouping) are always part of any sample. Once such a minimal sample is put together, it can be expanded – depending on the researcher’s wishes – on the basis of so-called diversity values. These values are assigned to all nodes in the language family tree. They express the complexity of the tree below the node in terms of the amount of daughter nodes and the way they are embedded. The diversity value determines the proportion of languages to be drawn from under the relevant node in relation to its sister nodes, given a certain desired sample
size. Once the number of languages to be selected from a particular family is determined, the linguist chooses those languages for which the best descriptive grammars (and/or other data sources) are available. Preferably, the languages should come from different subgroups within their respective families (Bakker, forthcoming).

In principle, the DVT composes so-called variety samples, which are used to explore the range of diversity in a relatively little-studied linguistic domain. Variety samples stand in contrast to probability samples, which are used to determine significant correlations between grammatical traits (Croft 2003: 23). In a sample of the latter type, the languages represent independently selected cases. Thus, this type of sample explicitly avoids the situation in which grammatical traits are shared between languages as a result of descent from a common ancestor or through language contact. Samples that are selected without controlling for these factors are said to be genetically and/or areally biased.

Although the DVT thus aims primarily at maximal genetic diversity, it can also be used for the composition of what Bakker (forthcoming) calls pseudo-probability samples. These are relatively small variety samples with a relatively high degree of independence between the selected cases. They can be composed by combining the DVT (applied to an up-to-date language classification) with an areal classification as a stratifying dimension. The rationale behind the use of small variety samples as pseudo-probability samples is that languages that belong to different families vary along certain structural parameters. Therefore, in the initial stage of the sampling process, the goal of covering maximal diversity (relevant for a variety sample) and the goal of avoiding genetic bias (relevant for a probability sample) can be reached by the same procedure: picking one language per family. Combining this method with an areal stratification allows one to avoid the second major source of bias for a probability sample: feature sharing as a result of language contact. Only when variety samples are expanded in order to cover a wider range of variation (which may be attested in specific genetic or areal groupings) do they become fundamentally ill-suited for probabilistic research, since the languages can no longer be regarded as independent cases.

Hengeveld et al. (2004) use a pseudo-probability sample for their study on the relation between PoS systems and word order constraints. This 50-language sample is composed by applying the DVT to Ruhlen’s (1991) classification and then choosing languages that are spoken in non-contiguous areas, at least whenever the bibliographical situation would permit it.
Notably, Ruhlen’s classification is controversial to the extent that it distinguishes a relatively small number of (large) language families: 19 families and a number of isolates. This number is much smaller than for instance in the Ethnologue classification, which distinguishes 120 families (Gordon 2005). As mentioned above however, the DVT requires a minimum of one language per family (see Rijkhoff & Bakker 1998). In view of this requirement, Ruhlen’s classification provides a suitable basis for composing a relatively small language sample, which is desirable in the case of a pseudo-probability sample, and also in practical terms of feasibility. The Ethnologue classification, in contrast, yields a sample of minimally 120 languages. Note also that the specific methodology of the DVT partly makes up for the more extensive ‘lumping’ in Ruhlen’s classification by dictating the selection of more languages from families with larger internal diversity.

The sample used in the present study is very similar to the sample used by Hengeveld et al. (2004). The first reason for this is a practical one: It allowed me to take as a point of departure for most languages the PoS system classifications as proposed by Hengeveld et al. (2004). Nevertheless, I have adapted my sample in some respects. First, I have replaced a few languages from the sample of Hengeveld et al. for bibliographical reasons. The relevant cases are listed below:

- Hdi instead of Gude (Chadic, Afro-Asiatic)
- Dhaasanac instead of Oromo (Cushitic, Afro-Asiatic)

Maslova and Nikitina (submitted) have recently challenged the assumption behind ‘classical’ probability sampling, namely that language family size is a ‘historical accident’, and that therefore every family should be represented in the sample with an equal number of languages, in order to avoid disproportional representation of certain language types. They argue that family size is determined by *transition probabilities*, which are in turn determined by *language constants*, i.e. universal properties of languages and their users. Therefore, they argue for the use of ‘truly’ random samples (so-called *R*-samples), rather than samples containing one randomly selected language per family (so-called *I*-samples). *R*-samples will contain a higher percentage of more stable language types. On the other hand, frequency distributions in *R*-samples may be due to an accidentally higher initial frequency of some source type, rather than to a systematically higher amount of transitions to this type. Therefore, *R*-samples and *I*-samples should be compared in order to find out whether the distribution attested in the modern language population is indeed the outcome of differences in transition probabilities. If yes, then it constitutes evidence for a statistical universal. If not, then the comparison can still indicate tendencies towards a linguistically meaningful equilibrium: If a more stable trait occurs less frequently synchronically, then there will be a gradual increase of languages with this trait. Conversely, if a less stable trait occurs more frequently, then the number of languages with this trait will decrease over time.

In the second edition of Ruhlen, Korean-Japanese-Ainu and Kartvelian are distinguished as separate families, while in the first edition they were considered sub-branches of Altaic and Caucasian, respectively (Hengeveld et al. 2004: 528, note 4).
• Lavukaleve instead of Nasioi (East Papuan, Indo-Pacific)
• Abun instead of Tidore (West Papuan, Indo-Pacific)
• Ma’di instead of Ngiti (Central Sudanic, Nilo-Saharan)
• Slave instead of Navaho (Na-Dene)
• Gooniyandi instead of Ngalakan (non-Pama-Nyungan (Bunaban), Australian)
• Kharia instead of Mundari (Munda, Austro-Asiatic, Austric)

The second difference between my sample and that of Hengeveld et al. (2004) is that I have expanded mine with two languages with a flexible PoS system, since the most interesting data were obtained for this type of languages. The added languages are:

• Santali, which comes from the same sub-family as Kharia: Munda, Austro-Asiatic, Austric;
• Kambera, which belongs to the same family as Samoan and Tagalog: Malayo-Polynesian, Austronesian.

Obviously, adding these languages results in areal and genetic biasing of the sample. For this reason, Santali and Kambera are not taken into account when calculating statistical dependencies between grammatical traits (see Chapter 7). Rather, they are used as a further back-up to the patterns found in the other flexible languages of the sample.

Finally, two further details should be noted. First, the present study does not include any data on the extinct languages that are part of the sample used by Hengeveld et al. (2004), namely Etruscan, Hurrian, Hittite, Nahali, and Sumerian. Furthermore, in two cases I use a different language name than Hengeveld et al. (2004): Bukiyip instead of Mountain Arapesh, and Hmong Njua instead of Miao.

As Hengeveld et al. (2004: 529) note, the inclusion of Tagalog in their sample (and also in mine) represents a violation of the genetic criterion, since it does not belong to the Formosan branch of the Austric family, which would have been the right choice. It was included in Hengeveld et al. in order to represent a language with a maximally flexible PoS system. Even though this choice is relevant in view of the present study as well, it undermines the value of the sample as a pseudo-probability sample, since it involves deliberately choosing a case (i.e. a language) on the basis of the grammatical phenomenon being studied. I thank Elena Maslova for emphasizing this point.

With the same goal in mind, some data from some other flexible languages have been and will be discussed, including Tongan, Maori, and Mundari (see Chapters 2 and 5).
The total number of languages in the sample is 50, as can be seen in Table 4.1 below. In Appendix 1, a more complete specification of the sample is provided, in which the languages are also classified for family and subfamily according to the Ethnologue, and for Genus according to the World Atlas of Language Structures (WALS)\textsuperscript{65}.

As this section makes clear, a language sample is composed with a specific typological research question in mind. Variety samples are in principle not appropriate to answer probabilistic research questions, since they are specifically designed to manipulate chances of occurrence: They intend to maximize the likelihood of capturing all the linguistic diversity for the phenomenon under study (Croft 2003: 21). However, small, genetically and areally balanced variety samples, of the type used in Hengeveld et al. (2004) and in the present study, can indeed be used as pseudo-probability samples, on which to test hypotheses about dependency relations between structural traits, in this case between various parameter values related to PoS and DCs. Bearing this in mind, in the following section I turn to the specific formulation and operationalization of the research questions of this study.

<table>
<thead>
<tr>
<th>Family (Ruhlen 1991)</th>
<th>Language(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afro-Asiatic</td>
<td>Hdi, Dhaasanac</td>
</tr>
<tr>
<td>Altaic</td>
<td>Turkish</td>
</tr>
<tr>
<td>Korean-Japanese</td>
<td>Japanese</td>
</tr>
<tr>
<td>Amerindian</td>
<td>Pipil, Hixkaryana, Tuscarora, Koasati, Guarani (Paraguayan), Warao, Imbabura Quechua</td>
</tr>
<tr>
<td>Australian</td>
<td>Kayardild, Nunggubuyu, Gooniyandi</td>
</tr>
<tr>
<td>Austric</td>
<td>Tagalog, Samoan, Kambera, Paiwan, Garo, Thai, Santali, Kharia, Hmong Njua</td>
</tr>
<tr>
<td>Caucasian</td>
<td>Abkhaz</td>
</tr>
<tr>
<td>Kartvelian</td>
<td>Georgian</td>
</tr>
<tr>
<td>Ckukchi-Kamchatkan</td>
<td>Itelmen</td>
</tr>
<tr>
<td>Elamo-Dravidian</td>
<td>Tamil</td>
</tr>
<tr>
<td>Eskimo-Aleut</td>
<td>West Greenlandic</td>
</tr>
<tr>
<td>Indo-Hittite</td>
<td>Polish</td>
</tr>
<tr>
<td>Indo-Pacific</td>
<td>Wambon, Alamblak, Lavukaleve, Abun, Bukiyp</td>
</tr>
<tr>
<td>Isolates</td>
<td>Ket, Burushaski, Basque, Nivkh</td>
</tr>
</tbody>
</table>

\textsuperscript{65} See \url{www.ethnologue.org} (or Gordon 2005) and \url{http://wals.info/} (or Haspelmath et al. 2005).
<table>
<thead>
<tr>
<th>Family (Ruhlen 1991)</th>
<th>Language(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khoisan</td>
<td>Nama</td>
</tr>
<tr>
<td>Na-Dene</td>
<td>Slave</td>
</tr>
<tr>
<td>Niger-Kordofanian</td>
<td>Babungo, Kisi, Bambara, Krongo</td>
</tr>
<tr>
<td>Nilo-Saharan</td>
<td>Lango, Ma’di</td>
</tr>
<tr>
<td>Pidgins and creoles</td>
<td>Berbice Dutch Creole</td>
</tr>
<tr>
<td>Sino-Tibetan</td>
<td>Nung, Mandarin Chinese</td>
</tr>
<tr>
<td>Uralic-Yukaghir</td>
<td>Hungarian</td>
</tr>
</tbody>
</table>

Table 4.1: The language sample

4.3 Research Questions

4.3.1 Introduction

In the previous chapters, it was shown that PoS classes and DCs can be defined as primary and secondary constructions, respectively, expressing the propositional functions of predication, reference, and/or modification. On the basis of this parallel, it can be hypothesized that there will be a match between the distributional patterns of PoS classes and DC constructions in any language. In particular, it is expected that the functional possibilities of the (secondary) DCs constructions in a language are dependent on those of the (primary) PoS categories available in that language. This hypothesis will be investigated from two converging perspectives:

(i) The perspective of global matches, i.e. in terms of the presence of functional flexibility versus rigidity in PoS systems as compared with systems of DCs constructions (section 4.3.2);

(ii) The perspective of specific matches between particular types of flexible and rigid PoS classes and DCs constructions, in terms of set of functions that can be expressed by them (section 4.3.4).

In addition, and cross-cutting these two perspectives, I will investigate the influence of another parameter, namely the internal morpho-syntactic characteristics of the DC construction(s) under scrutiny (see sections 4.3.3 and 4.3.5). I will take this parameter into account by testing each of the hypotheses to be formulated in the next subsections in two ways:

(i) Considering DCs as a single, undifferentiated construction type, and
(ii) Splitting up the hypothesis into several sub-hypotheses, each of which addresses a different structural DC type.

In what follows, a set of fully explicit hypotheses will be formulated and operationalized according to the general research design just outlined.

### 4.3.2 Global functional matching

First, it is predicted that a *global match* exists between flexibility or rigidity as attested in the PoS system of a particular language, and flexibility or rigidity in the set of DC constructions in that language. This hypothesis is operationalized in the form of the two-fold prediction A1/A2, as given in (1a, b):

Predictions A1/A2

1. **If a particular language has one or more flexible DCs, then it should also have one or more flexible PoS classes.**
2. **If a language has rigid DCs only, then it should also have rigid PoS classes only.**

The phrases in italics in the above predictions require some explanation. Recall from Chapter 2 that PoS systems are called *flexible* when they include *one or more* flexible lexeme class(es). This means that a language with a flexible PoS system does not need to have flexible PoS classes *only*. In fact, Tagalog and Kharia are the only two languages in the sample that do not have any rigid PoS classes. All other languages with flexible PoS systems display a mixture of flexible and rigid PoS classes. In contrast, rigid PoS systems were defined as consisting of rigid PoS classes *only* (even though not all rigid systems have a rigid PoS class for every propositional function). These asymmetric definitions also apply to the predictions in (1) above, as the phrases in italics are meant to indicate: (1a) makes reference to languages with flexible PoS systems, and (1b) to languages with rigid PoS systems.

### 4.3.3 Global matching including differentiation for structural DC type

The predictions in (1) can be made more specific by taking into account the internal morpho-syntactic characteristics of the DC construction(s) under study. In particular, this is done by adding a parameter that differentiates between the three structural types of DCs defined in Chapter 3, section 3.4. They are summarized once more in Table 4.2 (which is the same as Table 3.1):
<table>
<thead>
<tr>
<th>DC type</th>
<th>Argument expression</th>
<th>TAM/Person</th>
<th>DET/CASE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number</td>
<td>label</td>
<td>1</td>
</tr>
<tr>
<td>Type 1</td>
<td>B</td>
<td>SENT/Ø</td>
<td>SENT/Ø</td>
</tr>
<tr>
<td>Type 2</td>
<td>D-SENT</td>
<td>SENT/Ø</td>
<td>SENT/Ø</td>
</tr>
<tr>
<td>Type 3</td>
<td>D-ALT</td>
<td>ALT</td>
<td>SENT/ALT/Ø</td>
</tr>
</tbody>
</table>

Table 4.2: DC types and their internal formal properties

Considering the morpho-syntactic properties of the three DC types in Table 4.2, we can say that:

(i) Type 1 DCs (Balanced) do not show any formal reflection of de-categorization (i.e. TAM/Person is expressed as in independent clauses), nor of re-categorization (arguments, if overt, are SENT, i.e. expressed as in independent clauses).

(ii) Type 2 DCs (D-SENT) show formal reflections of de-categorization (i.e. TAM/Person is (partially) lost), but not of re-categorization (arguments, if overt, are SENT, i.e. expressed as in independent clauses).

(iii) Type 3 DCs (D-ALT) show formal reflections of both de-categorization (i.e. TAM/Person is (partially) lost) and re-categorization (at least one argument is expressed with an ALT strategy, i.e. different from independent clauses).

A larger amount of formal de-/re-categorization of a DC construction implies that it has less morpho-syntactic characteristics of a clause and more of a lexical construction. It is hypothesized is that the more formally similar a DC construction is to a lexical expression, the more functionally similar the DC will be to its lexical counterpart. This means that deranked DCs, i.e. type 2 (D-SENT) and 3 (D-ALT), are expected to show more functional similarity with lexical categories than balanced DCs, i.e. type 1 (B). Furthermore, within the group of deranked DCs (types 2 and 3) it is expected that the distributional patterns of more deranked DCs (type 3) will be functionally more similar to lexical expressions than those of less deranked DCs (type 2). Thus, we can formulate the two sub-predictions B1 and B2, as in (2a) and (2b):

66 Recall that determiners and/or case-markers are not regarded as features reflecting re-categorization, because they do not affect the internal structure of the construction on which they operate.
Prediction B1/B2:

(2)  

a. The functional possibilities of deranked DCs (type 2/3) are more similar to the functional possibilities of PoS than those of balanced DCs (type 1).

b. Within the group of deranked DCs, the functional possibilities of type 3 DCs are more similar to the functional possibilities of PoS than those of type 2 DCs.

Note that, even though it has never been tested empirically on a larger scale, the hypothesis of a functional connection between lexeme classes and deranked DC constructions is not new. In fact, deranked complement clauses (nominalizations/infinitives), relative clauses (participial clauses), and adverbal clauses (converbal constructions) are traditionally characterized as the syntactically derived clausal counterparts of nouns, adjectives and adverbs, respectively (Croft 1991, Koptjevskaja-Tamm 1993, Haspelmath 1994, 1995). These connections between PoS and deranked DCs can be summarized as in Table 4.3 (adapted from: Haspelmath 1995: 3-4)\textsuperscript{67}.

<table>
<thead>
<tr>
<th>Forms</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ref Head</td>
</tr>
<tr>
<td>PoS</td>
<td>noun</td>
</tr>
<tr>
<td>deranked DC</td>
<td>nominalization/infinitive</td>
</tr>
<tr>
<td>type</td>
<td>construction</td>
</tr>
<tr>
<td></td>
<td>Ref Mod</td>
</tr>
<tr>
<td></td>
<td>adjective</td>
</tr>
<tr>
<td></td>
<td>participle construction</td>
</tr>
<tr>
<td></td>
<td>(manner) adverb</td>
</tr>
<tr>
<td></td>
<td>Pred Mod</td>
</tr>
<tr>
<td></td>
<td>converb construction</td>
</tr>
</tbody>
</table>

Table 4.3: Functional connection between PoS and deranked DCs

We can now combine predictions B1 and B2 in (2) above with predictions A1 and A2 in (1) above. For flexible constructions, this yields the four sub-hypotheses given in (3a-d):

(3)  

a. If a language has one or more flexible balanced DC(s) of type 1, then it should also have one or more flexible PoS class(es).

b. If a language has one or more flexible deranked DC(s) of type 2/3, then it should also have one or more flexible PoS class(es).

\textsuperscript{67} Notably, the PoS classes and DCs defined in Chapters 2 and 3 make reference to 4 propositional functions: in Table 4.3 the function of head of a predicate phrase is not taken into account. In Chapter 6 it will become clear that this functional slot is irrelevant for the present study to the extent that there are no (or hardly any) DC constructions that express it.
c. If a language has one or more flexible deranked DC(s) of type 2, then it should also have one or more flexible PoS class(es).

d. If a language has one or more flexible deranked DC(s) of type 3, then it should also have one or more flexible PoS class(es).

On the basis of the structural differences between the DC types, it is expected that the prediction in (3a), involving balanced clauses, is less likely to receive empirical support than the one in (3b), concerning deranked clauses. Further, the prediction in (3c), involving less deranked clauses of type 2, is expected to be less probable than the one in (3d), with concerns more deranked clauses of type 3.

Note that, due to the asymmetrical definitions of languages with flexible versus rigid PoS systems, it does not make sense to formulate specific predictions about implicational relations between the presence of rigid PoS only and the presence of rigid DCs of a specific behavioural potential type only. This is because, on the basis of the available typological evidence discussed in Chapter 3, one would not predict that languages express all subordination relations with rigid DCs of just a single structural type. Rather, semantically different types of subordination relations are likely to be expressed by morpho-syntactically different DC constructions. Therefore, no separate sub-hypotheses are formulated that make reference to and rigid DCs of type 1/2/3 only.

4.3.4 Specific functional matching
I now turn to the second perspective on the relation between PoS and DCs, which involves the investigation of specific matches between the functional possibilities of particular PoS classes and DC constructions. We have seen in Chapters 2 and 3 that there are 10 predicted types of flexible constructions (lexical and clausal), depending on how many and which propositional functions they can express. These flexible construction types are listed once more under (4):

(4) a. Lexical and clausal constructions that can be used in all four propositional functions: contentives and contentive clauses;

b. Lexical and clausal constructions that can be used in all propositional functions except the head of a predicate phrase: non-verbs and multi-functional clauses;
c. Lexical and clausal constructions that can be used as the head and modifier in a referential phrase: *nominals* and *nominal clauses*;
d. Lexical and clausal constructions that can be used as modifiers in referential and predicate phrases: *modifiers* and *modifier clauses*;
e. Lexical and clausal constructions that can be used as the head and modifier in a predicate phrase: *predicatives* and *predicative clauses*;
f. Lexical and clausal constructions that can be used as heads of referential and predicate phrases: *heads* and *head clauses*;
g. Lexical and clausal constructions that can be used as heads of referential and predicate phrases and as modifiers in referential phrases: *Flex PoS A* and *Flex clause A*.
h. Lexical and clausal constructions that can be used as heads of referential and predicate phrases and as modifiers in predicate phrases: *Flex PoS B* and *Flex clause B*.
i. Lexical and clausal constructions that can be used as the head of a referential phrase and as a modifier in a predicate phrase: *Flex PoS C* and *Flex clause C*.
j. Lexical and clausal constructions that can be used as the head of a predicate phrase and as a modifier in a referential phrase: *Flex PoS D* and *Flex clause D*.

As regards rigid constructions, four construction types were predicted, each of them specialized for the expression of a single propositional function. They are listed in (5):

(5)  
a. Lexical and clausal constructions that are specialized for the function of head of a predicate phrase: *verbs* and *predicate clauses*.
b. Lexical and clausal constructions that are specialized for the function of head of a referential phrase: *nouns* and *complement clauses*;
c. Lexical and clausal constructions that are specialized for the function of modifier in a referential phrase: *adjectives* and *relative clauses*;
d. Lexical and clausal constructions that are specialized for the function of modifier in a predicate phrase: *manner adverbs* and *adverbial manner clauses*.
It is hypothesized that each DC of a specific flexible/rigid type, being a secondary construction, will have a PoS class of the same flexible/rigid type as its primary counterpart. This hypothesis is operationalized in the form of prediction C in (6):

Prediction C:

(6) If a language has a DC construction of a flexible or rigid type X, then it should also have a PoS class of type X.

Applying Prediction C to each of the specific flexible constructions listed in (4), we arrive at the set of predictions in (7):

(7) a. If a language has contentive clauses, then it should also have lexical contentives.
b. If a language has multi-functional clauses, then it should also have lexical non-verbs.
c. If a language has nominal clauses, then it should also have lexical nominals.
d. If a language has modifier clauses, then it should also have lexical modifiers.
e. If a language has predicative clauses, then it should also have lexical predicatives.
f. If a language has head clauses, then it should also have lexical heads.
g. If a language has Flex clause A, then it should also have Flex PoS A.
h. If a language has Flex clause B, then it should also have Flex PoS B.
i. If a language has Flex clause C, then it should also have Flex PoS C.
j. If a language has Flex clause D, then it should also have Flex PoS D.

A parallel list of predictions for the specific rigid constructions listed in (5) appears in (8):

(8) a. If a language has predicate clauses, then it should also have lexical verbs.
b. If a language has complement clauses, then it should also have lexical nouns.

c. If a language has relative clauses, then it should also have lexical adjectives.

d. If a language has adverbial manner clauses, then it should also have lexical manner adverbs.

The predictions in (7) and (8) can be further fine-tuned by taking into account the parameter of structural DC types. This is done in the next section.

4.3.5 Specific matching including differentiation for structural DC type

The predictions in (7) and (8) concerning specific flexible and rigid construction types are now combined with predictions concerning different structural DC types. Starting with maximally flexible constructions, i.e. lexical contentives and contentive clauses, this yields the set of predictions listed in (9a-d). The expected likelihood that each of the implicational relations in (9) will actually hold, is the same as for the set of implications in (3) above. The implication in (9a), concerning balanced DCs, is less likely to receive empirical support than the one in (9b), concerning deranked DCs. Further, the implication in (9c), which makes reference to less deranked DCs of type 2, is less likely to hold than the one in (9d), involving more deranked DCs of type 3.

(9)  a. If a language has balanced contentive clauses of type 1, then it should also have lexical contentives.

b. If a language has deranked contentive clauses of type 2/3, then it should also have lexical contentives.

c. If a language has deranked contentive clauses of type 2, then it should also have lexical contentives.

d. If a language has deranked contentive clauses of type 3, then it should also have lexical contentives.

Analogous sets of predictions can be set up for the remaining nine types of flexible constructions listed in (4)/(7) above, by means of substituting lexical contentives for non-verbs, nominals, modifiers, etcetera, and substituting contentive clauses for multi-functional clauses, nominal clauses, etcetera.
Similarly, the parameter of structural DC type is expected to interact with the predictions concerning different types of rigid constructions: For each of the rigid constructions listed in (5)/(8) a set of four sub-predictions can be set up along the lines of (9) above. Taking as an example lexical and clausal constructions specialized for the function of head of a referential phrase, i.e. nouns and complement clauses, we get the set of implicational relations in (10a-d). Again, the expectation is that the implication in (10a) is less likely to be born out than the one in (10b), and that the implication in (10c) is less probable than the one in (10d).

(10)  a. If a language has balanced complement clauses of type 1, then it should also have lexical nouns.

   b. If a language has deranked complement clauses of type 2/3, then it should also have lexical nouns.

   c. If a language has deranked complement clauses of type 2, then it should also have lexical nouns.

   d. If a language has deranked complement clauses of type 3, it should also have lexical nouns.

Analogous lists of testable implications can be set up for the other types of rigid constructions, by means of substituting nouns with verbs, adjectives, or manner adverbs, and by substituting complement clauses with predicate clauses, relative clauses, and adverbial manner clauses.

4.3.6 Summary

In sum, the general research question posed in this study is whether there is a match between the functional possibilities of lexical constructions (PoS) in a language and clausal constructions (DCs) in that language. This general question is approached from two converging perspectives: First, global matches are investigated between flexibility/rigidity as displayed by the PoS systems and the dependent clause constructions of particular languages. Second, specific matches are investigated between the functional patterns of particular types of flexible and rigid PoS classes and DC constructions. Across both perspectives cuts the additional parameter of structural DC type, defined in terms of behavioural potential. The specific predictions formulated above will be tested using a statistical method that is presented in the following section.
4.4 Method

The predictions formulated in the previous section take the form of classic Greenbergian implicational universals, namely: *(With more than chance frequency), if a language has structural characteristic A, then it has structural characteristic B* (Greenberg 1963; Cysouw 2005: 564). An implicational universal can thus be construed as a hypothesis about a particular type of dependency relation between A and B, such that one value of A (the positive one) constrains the value of B, whereas the other value of A (the negative one) does not (Maslova 2003: 103). This means that, in order to claim an implicational universal, two methodological steps are required, as described in (11a-b):

(11) a. Establishing a dependency relation between A and B, i.e. a statistically significant correlation between the co-occurrence in languages of characteristic A and characteristic B.

b. Establishing an asymmetrical dependency relation between A and B, such that the distribution of characteristic B among languages with characteristic A is more strongly skewed than the distribution of B among languages without A.

Consider first (11a): Establishing a dependency between A and B requires rejection of the hypothesis of independence, which states that the probability of any combination of values of A and B is just the product of the probabilities of these values taken in isolation. Statistical tests can be used to reject this hypothesis (Maslova 2003: 102).

In the present study, each of the predictions formulated in the previous section involves 2 binary parameters, one concerning PoS, the other concerning DCs. Each parameter can have two values: Either the particular type of PoS or DC is attested in the language ([+PoS], [+DC]), or it is not attested ([−PoS], [−DC]). For every prediction this yields a 2x2 contingency table with four cells. As is shown in Table 4.4 below, each of the cells in such a table represents one of the four possible combinations of the two binary parameters under investigation:

(i) [+PoS, + DC] = both the PoS and the DC feature are attested;
(ii) [+ PoS, − DC] = the PoS feature is attested, but the DC feature is not attested;
(iii) \([- \text{PoS}, + \text{DC}]\) = the PoS feature is not attested, but the DC feature is attested;
(iv) \([- \text{PoS}, - \text{DC}]\) = neither the PoS feature nor the DC feature is attested.

<table>
<thead>
<tr>
<th>PoS</th>
<th>DCs</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>(i)</td>
<td>(ii)</td>
</tr>
<tr>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>(iii)</td>
<td></td>
<td>(iv)</td>
</tr>
</tbody>
</table>

*Table 4.4: A 2x2 contingency table*

The observed frequencies, i.e. the numbers of languages in the sample that display a particular feature combination (as mentioned in (i)-(iv) above), can be compared with the frequencies that would be expected if the co-occurrence of the PoS and the DC features would be purely coincidental. If this is not the case, i.e. if there is indeed a dependency relation between a particular PoS feature and a particular DC feature, then the observed frequencies in cells (i) and (iv) will be higher than the frequencies that are expected on the basis of chance, whereas the observed frequencies in cells (ii) and (iii) will be lower than the expected frequencies. A Fischer’s Exact test can be used to calculate whether the deviation between observed and expected frequencies is statistically significant. This test produces a p-value between 0 and 1, which specifies how likely it is for the observed distribution to be the result of chance (Cysouw 2003: 91). The critical value used to identify a statistically significant correlation between two grammatical traits is p<0.05, meaning that there is a 0.5% chance that the observed frequencies are coincidental.

As Maslova (2003: 102) notes, if Fischer’s Exact fails to reject the hypothesis of independence, i.e. if no significant correlation between the parameters can be established, this does not automatically mean that they really are independent: it can also be the case that the sample is too small to reveal the dependency. Following the guidelines of Cohen (1995), a sample size of 26 is needed to have an 80% chance of detecting a large effect with a test like Fisher’s Exact. For the detection of a medium effect one needs a sample of 84, while small effects are detectable only with a sample as large as 785 cases. This means that the sample used in the present study suffices for the detection of medium to large effects. However, in the case of small effects, the sample will be too small for Fischer’s Exact to be able to reject the hypothesis of independence.
Since Fisher’s Exact is thus highly sensitive to sample size, I will report, in those cases where Fisher’s Exact is statistically significant, the contingency coefficient (CC). This CC is a value between 0 and 1, which estimates the effect size independent of the sample size. A CC value of approximately 0.10 indicates a small effect, around 0.30 is a medium effect, and from 0.45 upwards is a large effect (Everitt 1977).

As mentioned in (11b) above, if Fischer’s Exact yields a p-value below 0.05, then a second step must be taken, namely the identification of an asymmetrical dependency relation, such that the positive value of one parameter (in this case the PoS parameter) constrains the value of the other one (in this case the DC parameter). Maslova (2003) proposes a statistical method to establish such an asymmetrical dependency relation. The basic idea is to correlate each of the two parameters, in this case the PoS parameter and the DC parameter, to a third, derived parameter: PoS = DC, which refers to the event of the PoS parameter and the DC parameter having the same value. This third parameter thus contrasts languages that confirm the correlation, i.e. languages in which PoS and DC have the same value (PoS = DC: either [+ PoS, + DC] or [- PoS, -DC]), with languages that disconfirm the correlation, i.e. languages in which PoS and DC do not have the same value (PoS ≠ DC: either [+ PoS, - DC] or [- PoS, + DC]). This involves two extra contingency tables, as illustrated with Tables 4.5a and 4.5b:

<table>
<thead>
<tr>
<th></th>
<th>PoS +</th>
<th>PoS −</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 4.5a: Maslova test no. 1: Change DC parameter to PoS=DC*

<table>
<thead>
<tr>
<th></th>
<th>DC +</th>
<th>DC −</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 4.5b: Maslova test no. 2: Change PoS parameter to PoS=DC*

Submitting the observed frequencies in two extra tables of this kind to Fischer’s Exact tests may yield three possible results, listed in (12a–c)

---

See also Cysouw (2005) for a brief assessment of this method.
(12) a. A *symmetrical* dependency: None of the tests shows a significant interaction, i.e. in both cases $p > 0.05$.

b. A *one-sided asymmetrical* dependency: One of the two tests shows a significant interaction, but not the other, i.e. in one case $p < 0.05$, and in the other $p > 0.05$.

c. A *two-sided asymmetrical* dependency: Both tests show a significant interaction, i.e. in both cases $p < 0.05$.

An example of (12a), a *symmetrical dependency*, is given in Tables 4.6 and 4.7a-b below. The original distribution appears in Table 4.6. Applying Fischer’s Exact to these frequencies yields a significant correlation, i.e. there is a dependency relation between the PoS and DC parameters. The two additional tables are 4.7a and 4.7b, both of which yield a non-significant p-value. This means that neither the positive nor the negative value of the PoS or the DC parameter imposes a constraint on the event of the two parameters having the same value.

<table>
<thead>
<tr>
<th>Table 4.6: A symmetrical dependency; original distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \begin{array}{</td>
</tr>
<tr>
<td>$p &lt; 0.05$ (significant)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4.7a: Distribution for PoS and PoS = DC (Maslova-test no. 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \begin{array}{</td>
</tr>
<tr>
<td>$p &lt; 0.05$ (not significant)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4.7b: Distribution for DC and PoS = DC (Maslova-test no. 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \begin{array}{</td>
</tr>
<tr>
<td>$p &lt; 0.05$ (not significant)</td>
</tr>
</tbody>
</table>

69 These ‘observed’ frequencies are just illustrations. They are taken from examples in Maslova (2003:105–106).
Second, Tables 4.8 and 4.9a–b illustrate the situation in (12b) above, i.e. a one-sided asymmetrical dependency. Table 4.8 shows the original distribution, which again yields a significant p-value, so that a dependency relation is established. Tables 4.9a and 4.9b represent the two additional tests. Table 4.9a yields a significant p-value, while 4.9b does not. More specifically, Table 4.9a shows that the positive value of the PoS parameter poses a strong constraint on the event of PoS and DC having the same value. In contrast, Table 4.8 makes clear that the event of (non-)attestation of the DC characteristic and the event of the PoS and the DC parameters having the same value are independent of each other.

Table 4.8: A one-sided asymmetrical dependency; original distribution

<table>
<thead>
<tr>
<th></th>
<th>DC +</th>
<th>DC –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS +</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>PoS –</td>
<td>35</td>
<td>28</td>
</tr>
</tbody>
</table>

p < 0.05 (significant)

Table 4.9a: Distribution for PoS and PoS = DC (Maslova-test no. 1)

<table>
<thead>
<tr>
<th></th>
<th>Pos +</th>
<th>Pos –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>90</td>
<td>28</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>10</td>
<td>35</td>
</tr>
</tbody>
</table>

p < 0.05 (significant)

Table 4.9b: Distribution for DC and PoS = DC (Maslova-test no. 2)

<table>
<thead>
<tr>
<th></th>
<th>DC +</th>
<th>DC –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>90</td>
<td>28</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>35</td>
<td>10</td>
</tr>
</tbody>
</table>

p < 0.05 (not significant)

This type of result may be interpreted as a statistical basis to formulate a weak unidirectional implicational universal (Maslova 2003: 106; Dryer 2003: 111) of the following form:
Languages with a positive PoS value significantly more often have a positive DC value (regardless of whether languages with a negative PoS value also tend to have a positive DC value).

Finally, Tables 4.10 and 4.11a-b illustrate the third type of result, a two-sided asymmetrical dependency (see (12c) above). The original distribution appears in Table 4.10, which again reveals a significant dependency relation. Moreover, the two extra tests, applied to the figures in Tables 4.11a and 4.11b, both yield a significant p-value.

Table 4.10: A two-sided asymmetrical dependency; original distribution

<table>
<thead>
<tr>
<th></th>
<th>DC +</th>
<th>DC –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS +</td>
<td>40</td>
<td>6</td>
</tr>
<tr>
<td>PoS –</td>
<td>50</td>
<td>60</td>
</tr>
</tbody>
</table>

p < 0.05 (significant)

Table 4.11a: Distribution for PoS and PoS = DC (Maslova-test no. 1)

<table>
<thead>
<tr>
<th></th>
<th>Pos +</th>
<th>Pos –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>6</td>
<td>50</td>
</tr>
</tbody>
</table>

p < 0.05 (significant)

Table 4.11b: Distribution for DC and PoS = DC (Maslova-test no. 2)

<table>
<thead>
<tr>
<th></th>
<th>DC +</th>
<th>DC –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>50</td>
<td>6</td>
</tr>
</tbody>
</table>

p < 0.05 (significant)

As in the previous case, the positive value of the PoS parameter poses a constraint on the event of PoS and DC having the same value. The significant correlation in Table 4.11a reflects this. In addition, the original distribution in Table 4.10 shows that the distribution of PoS for the negative value of the DC parameter is clearly more skewed than for the positive value. In other words, it is not only the case that [+PoS] statistically implies [+ DC],

[+ PoS] ⇒ [+ DC]
but also that [-DC] implies [-PoS]. The latter dependency is shown by the significant correlation in Table 4.11b.

This type of result may be interpreted as a statistical basis to formulate a *strong* unidirectional implicational universal (Maslova 2003: 106; Dryer 2003: 111) of the following form:

\[
[+ \text{PoS}] \rightarrow [+ \text{DC}] \text{ and } [-\text{DC}] \rightarrow [-\text{PoS}]
\]

Languages with a positive PoS value significantly more often have a positive DC value, and the tendency for languages with a positive PoS value to have a positive DC value is significantly stronger than the tendency for languages with a negative PoS value to have a positive DC value.

In sum, the method used in this study involves observing frequencies of co-occurrence of formal linguistic traits. These frequencies are used to detect different types of dependency relations between various pairs of formal parameters, related to PoS classes and DC constructions, respectively, in terms of the hypotheses formulated in the previous section.

### 4.5 Outlook

This chapter rounds off the first, theoretical part of the book. In the second part, the data will be presented. Chapter 5 presents the classification of PoS classes and PoS systems attested in the languages of the sample. Chapter 6 presents the typology of DC constructions in the sample languages. In Chapter 7, these two typological data sets are combined to test the hypotheses formulated in the present chapter. Chapter 8 provides a more descriptive discussion of the results obtained in Chapter 7, explaining them from a functionalist perspective. Finally, Chapter 9 presents the study’s overall conclusions.
5.1 Introduction

In this chapter I present the classification of parts of speech (PoS) classes attested in the sample languages, following the definitions developed in Chapter 2. In particular, for every language in the sample I determine the set of PoS classes it displays, and define which propositional functions each PoS class can express without extra structural coding. These data are presented in section 5.2. Subsequently, in section 5.3, I evaluate the attested PoS classes and PoS systems in terms of the predictions of the implicational map model of parts of speech, as presented in Chapter 2.

The remainder of the chapter addresses a variety of issues regarding lexical classification in the sample languages. These issues have already been touched upon in the theoretical discussion on PoS typology in Chapter 2, but their practical implications are now further illustrated with actual language data. First, in section 5.4, I discuss the generality problem and the subclass problem. I show how the generality problem is circumvented by disregarding behavioural potential as a criterion for PoS classes. In addition, I discuss the problem of fuzzy boundaries between PoS classes. I first illustrate cases of lexeme classes with variable distributional patterns. Second, I consider different types of ‘restricted’ lexeme classes, i.e. small, closed classes and classes consisting of derived lexemes. Section 5.5 discusses the application of the typology of non-verbal predication to the sample languages. Section 5.6 focuses on the identification of flexible PoS classes. Finally, section 5.7 is a brief conclusion.
5.2 Overview of the language data

Table 5.1 below shows the PoS classes that are attested in the languages of the sample. The languages are listed in the leftmost column; they are ordered from maximally flexible, via less flexible, to fully rigid. Within the group of languages with rigid PoS systems, those with four large, open lexeme classes precede the ones with (an increasing amount of) restricted and/or lacking PoS classes. Every PoS class appears in a separate row. The terms for different types of PoS are abbreviated as follows: C for contentives, non-V for non-verbs, Pred for predicatives, Nom for nominals, V for verbs, N for nouns, Adj for adjectives and mAdv for manner adverbs\(^70\). These abbreviations appear in every column corresponding to a function that can be expressed without extra structural coding by the PoS class in question\(^71\). The names of the four propositional functions appear in the top row of the Table: Pred Head for head of a predicate phrase, Ref Head for head of a referential phrase, Ref Mod for modifier in a referential phrase, and Pred Mod for modifier in a predicate phrase. The two rightmost columns are used to add information about the size and status of a PoS class: When it is a small, closed class, an S appears in the penultimate column (with the heading Small). When a PoS class consists of derived members only, a D appear in the rightmost column (with the heading Derived). All other, unmarked cases involve large, open PoS classes.

<table>
<thead>
<tr>
<th>Language</th>
<th>Pred Head</th>
<th>Ref Head</th>
<th>Ref Mod</th>
<th>Pred Mod</th>
<th>Small</th>
<th>Derived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagalog</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kharia</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kambera</td>
<td>V</td>
<td>C</td>
<td>C</td>
<td>C mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Samoan</td>
<td>V</td>
<td>C</td>
<td>C</td>
<td>C mAdv</td>
<td>S</td>
<td>D</td>
</tr>
</tbody>
</table>

\(^70\) Some of the PoS classes predicted in Chapter 2 do not appear in Table 5.1. This is discussed further in section 5.2.

\(^71\) If a non-verbal PoS class has an additional predicative use (with a zero-1 expression strategy, see Chapter 2, section 2.3.1), then the abbreviation of the particular PoS class appears also in the column for Head of a Predicate Phrase.)
<table>
<thead>
<tr>
<th>Language</th>
<th>Pred Head</th>
<th>Ref Head</th>
<th>Ref Mod</th>
<th>Pred Mod</th>
<th>Small</th>
<th>Derived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guarani</td>
<td>C</td>
<td>V</td>
<td>C</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santali</td>
<td>C</td>
<td>V</td>
<td>C</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warao</td>
<td>V</td>
<td>non-V</td>
<td>non-V</td>
<td>non-V</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Santali</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td>V</td>
<td>non-V</td>
<td>non-V</td>
<td>non-V</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Kayardild</td>
<td>Pred</td>
<td>non-V</td>
<td>non-V</td>
<td>non-V</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Paiwan</td>
<td>Pred</td>
<td>N</td>
<td>N</td>
<td>Adj</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>I. Quechua</td>
<td>V</td>
<td>Nom</td>
<td>Nom</td>
<td>mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Ma'di</td>
<td>V</td>
<td>Nom</td>
<td>Nom</td>
<td>Mod</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Hungarian</td>
<td>V</td>
<td>Nom</td>
<td>Nom</td>
<td>mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>V</td>
<td>N</td>
<td>Nom</td>
<td>mAdv</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>V</td>
<td>N</td>
<td>Mod</td>
<td>Mod</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Pred Head</td>
<td>Ref Head</td>
<td>Ref Mod</td>
<td>Pred Mod</td>
<td>Small</td>
<td>Derived</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Lango</td>
<td>V</td>
<td>N</td>
<td>Mod</td>
<td>Mod mAdv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td>V</td>
<td>N</td>
<td>Mod Adj</td>
<td>Mod mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Koasati</td>
<td>V</td>
<td>N</td>
<td>Mod Adj</td>
<td>Mod mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Itelmen</td>
<td>V</td>
<td>N</td>
<td>Mod Adj</td>
<td>Mod mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Thai</td>
<td>V</td>
<td>Adj Mod</td>
<td>Adj Mod</td>
<td>Mod mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Basque</td>
<td>V</td>
<td>Adj Mod</td>
<td>Adj Mod</td>
<td>Mod mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Abun</td>
<td>V</td>
<td>Adj Mod</td>
<td>Adj Mod</td>
<td>Mod mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Bambara</td>
<td>V</td>
<td>Adj mAdv</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Pred Head</td>
<td>Ref Head</td>
<td>Ref Mod</td>
<td>Pred Mod</td>
<td>Small</td>
<td>Derived</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Georgian</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>mAdv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bukiyp</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Abkhaz</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Polish</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Burushaski</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>mAdv</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Alamblak</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Pipil</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wambon</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Pred Head</td>
<td>Ref Head</td>
<td>Ref Mod</td>
<td>Pred Mod</td>
<td>Small</td>
<td>Derived</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berbice D.C.</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babungo</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mAdv</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nama</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mAdv</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hdi</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandarin C.</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mAdv</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kisi</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nung</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garo</td>
<td>V</td>
<td>N</td>
<td>Adj</td>
<td>Adj</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.3 Results in relation to the implicational map of parts of speech

5.3.1 Introduction

In Chapter 2, three hierarchical relations were proposed as constraints on cross-linguistic variation in PoS classification: A primary relation *predication* ⊆ *reference*, a secondary relation *head* ⊆ *modifier*, and a ranking between these two: ((Predication/Reference) ⊆ (Head/Modifier)). These dominance relations were combined into an implicational map model of parts of speech. On the basis of this model, a number of possible PoS systems were predicted. In addition, it was hypothesized that, as long as no constraints are violated, in individual languages the predicted PoS systems can be mixed or supplemented with additional classes.

In this section, I evaluate the data from the sample languages, as presented in Table 5.1 above, in terms of the predictions of the implicational PoS map. I start out with a presentation of the languages that exhibit one of the predicted PoS systems in its ‘pure’ form. Second, I turn to languages with mixed or ‘intermediate’ PoS systems. Third, I discuss cases of PoS systems that were predicted in Chapter 2 but are not attested in the sample. There are also some potential examples of the reverse situation: PoS systems that
were excluded in Chapter 2 on the basis of constraint violation(s), but that are nonetheless attested in the sample.

5.3.2 Predicted and attested ‘pure’ systems
In this section I present the languages that exhibit one of the PoS systems predicted in Chapter 2, without any mixing or additional classes. I start with flexible systems, followed by rigid systems, in the same order as in Table 5.1.

Tagalog and Kharia both have contentives as their only PoS class. This type of system is represented in (1) (cf. (14) in Chapter 2):

(1)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Contentives</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

Hmong Njua has verbs, nouns, and a class of modifiers that can also be used predicatively. This system is represented in (2) (cf. (16) in Chapter 2):

(2)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
</tr>
<tr>
<td>Modifier</td>
<td></td>
</tr>
</tbody>
</table>

Gooniyandi and Hungarian have a PoS system with verbs, nominals, and manner adverbs, as in (3) (cf. (21) in Chapter 2). Note that Hungarian has both simple and derived manner adverbs.

(3)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
</tr>
<tr>
<td>Reference</td>
<td>Nominal</td>
</tr>
</tbody>
</table>

72 In section 5.6 the analysis of flexible PoS systems is discussed in full detail.
73 In Table 5.1 this is indicated by the abbreviation Mod in the column for Pred Head.
Georgian has four large, open classes, each of them specialized for a single function. This is represented in (4) (cf. (17) in Chapter 2):

(4)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>Manner adverb</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Noun</td>
<td>Adjective</td>
</tr>
</tbody>
</table>

The PoS system represented in (5) (cf. (18) in Chapter 2) has verbs, nouns, and adjectives, but no manner adverbs. It is attested in Pipil, Wambon, Dhaasanac, and Berbice Dutch Creole. Note that these languages differ in terms of the ability to use adjectives in predicative function with a zero-1 strategy (see Chapter 2, section 2.3.1 and the discussion in section 5.4 below):

In Pipil and Wambon adjectives cannot be used as verbs, while in Berbice Dutch they can. Dhaasanac has both types: a large class of non-predicative adjectives, and a small class of predicative ones. In Table 5.1, predicative adjectives are recognisable by the fact that the abbreviation Adj appears in the column for modifier in a referential phrase and in the column for head of a predicate phrase. In contrast, classes of non-predicative adjectives are indicated with Adj in the column for modifier in a referential phrase only (cf. notes 71 and 73).

(5)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>–</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Noun</td>
<td>Adjective</td>
</tr>
</tbody>
</table>

Finally, there are languages with the PoS system represented in (6) (cf. (19) in Chapter 2), with verbs and nouns, but no lexeme classes for the two modifier functions. This system is attested in Slave, Nivkh, West Greenlandic, Nunggubuyu, and Tuscarora.
The PoS system of Tuscarora may alternatively be analyzed as intermediate between the system in (6) above and the system represented in (7) below (cf. (20) in Chapter 2).

The reason for this is that Tuscarora (like other Iroquoian languages) often uses verbal constructions where other languages would use nouns. Mithun describes this phenomenon as follows:

“A striking feature of natural speech in Iroquoian languages […] is the relative rarity of nouns. It can be attributed in part to noun incorporation, as in ‘he suit-case-carried’. […] More important are the functions that verbs can serve. […] They can be used as descriptive labels for entities (objects, animals, people) and even proper names […].” (Mithun 2000: 412)

An example of such a descriptively used verb is given in (8):

*Tuscarora* (Mithun-Williams 1976: 30)

(8)  
\[Ra\_bren\_abs\]
\[MASC\_cut\_progr\]
\‘he cuts’ → ‘surgeon’

Apparently however, this “relative rarity of nouns” in Tuscarora does not extend to the extreme situation in which only verbs remain. I have tried to capture this in Table 5.1 by adding a bracketed S to Tuscarora’s noun class.
Many other languages, besides Tuscarora, also display a PoS system which combines features of more than one of the systems predicted in Chapter 2. To these ‘intermediate’ systems I turn in the next section.

5.3.3 Predicted and attested ‘intermediate’ systems
First, there are two languages in the sample that combine a class of contentives with a large, open class of verbs: Guaraní and Santali. Thus, the PoS system of these languages can be analyzed as a combination of the systems represented in (1) and (7) above. Other languages have contentives in combination with a class of derived verbs. This holds for Kambera and Samoan. In addition, these languages both display another extra class, namely a small, closed class of manner adverbs. This is an unexpected pattern to the extent it involves a head-modifier distinction (derived verbs versus adverbs) without a full-blown predication-reference distinction (see constraint (13) of Chapter 2). However, it seems that adverbs in Kambera and Samoan have many characteristics of function words, as opposed to content words (see also section 5.4.3.2).

Second, the system in (9) (cf. (15) in Chapter 2) is attested in Warao, but in combination with a class of derived nouns.

The PoS system of Warao may thus be represented as in Figure 5.1 below. This case is discussed in more detail in section 5.6.3.1.

There are two more languages with intermediate PoS systems involving the system in (9). First, Ma’di combines a PoS system of verbs and non-verbs with a system of verbs, nominals, and manner adverbs. The latter is represented in (10) (cf. (21) in Chapter 2):
Head Modifier

<table>
<thead>
<tr>
<th>Predication</th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb</td>
<td>Manner adverb</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Nominal</td>
<td></td>
</tr>
</tbody>
</table>

Turkish, as discussed in Chapter 2 (see Figure 2.5), also has a PoS system with verbs and non-verbs (see (9) above), but combines it with a class of derived modifiers, i.e. lexemes with the distributional possibilities of non-verbs minus the function of head of a referential phrase.

Imbabura Quechua has an intermediate PoS system that involves a combination of the system in (10) above, and the one in (11) below (cf. (22) in Chapter 2), since it has verbs, nominals, and a closed class of manner adverbs.

<table>
<thead>
<tr>
<th>Predication</th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Nominal</td>
<td></td>
</tr>
</tbody>
</table>

Japanese combines the system in (10) above with the system in (12), which has verbs, nouns, and manner adverbs (cf. (30) in Chapter 2)\(^{74}\).

<table>
<thead>
<tr>
<th>Predication</th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb</td>
<td>Manner adverb</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
<td>–</td>
</tr>
</tbody>
</table>

Another type of intermediate PoS system involves a system with verbs, nouns, and flexible modifiers, in combination with rigid classes of adjectives and/or manner adverbs. This is attested quite often, namely in Lango, Ket, Koasati, Itelmen, Thai, Abun, and Basque. There is variation in the size and status of the flexible and rigid PoS classes that can express the modifier functions in these PoS systems. Lango, for instance, has a large class of

\(^{74}\) Alternatively, the Japanese system can be described as the system in (10) supplemented with a class of rigid nouns. The PoS system of Japanese is discussed in more detail in section 5.4.2.1.
flexible modifiers and a large class of rigid manner adverbs, but no rigid adjectives. Thus, its PoS system can be analyzed as a combination of the system in (12) above and the one in (13) below:

(13)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
</tr>
</tbody>
</table>

Ket also has a large class of modifiers, but in combination with small classes of rigid adjectives and manner adverbs (cf. (4) above). In addition, Ket has derived adjectives.

Another possibility is that the rigid adjective and/or adverb class(es) are large, while the flexible modifier class is relatively small. In Abun, for instance, the classes of adjectives and manner adverbs are larger than the modifier class. Thai and Basque both have full-blown adjective classes, combined with (very) restricted classes of modifiers and manner adverbs. Finally, it may be the case that all three types of PoS classes, i.e. modifiers, adjectives and adverbs, are small and/or derived. This is attested in Koasati and Itelmen.

A number of other languages have a fully differentiated PoS system with four rigid PoS classes, in which either the adjective class or the manner adverb class or both are small and/or derived. The following languages have a small and/or a derived class of manner adverbs: Abkhaz, Polish, Burushaski, Lavukaleve, and Alamblak. This type of PoS system is represented in (14). It may be regarded as a system intermediate between the pure systems presented in (4) and (5) above.

(14)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
</tr>
</tbody>
</table>

Bukiyip has two types of manner adverbs. One type takes verbal inflection, but cannot appear independently as the head of a predicate phrase. Conrad and Wogiga (1991) do not give information about the size of this class,
nor about its open or closed status. An example of this type of adverb is given in (15) below. The second class is described as a small, closed class of uninflected manner adverbs, consisting of some 13 items (Conrad & Wogiga 1991: 41).

**Bukiyip** (Conrad & Wogiga 1991: 41)

(15) \[\text{Awou} \quad \text{w-a-gamu}\]

\[\text{3pl.fem} \quad \text{3pl.fem.sbj-real-well}\]

\[\text{w-a-dükemêch}\]

\[\text{3pl.fem.sbj-real-understand}\]

‘The women understand well.’

There are also languages in which both adjectives and manner adverbs constitute small and/or derived classes, as is represented in (16) below. This is the case in Babungo, Nama, Hdi, Mandarin Chinese, and Tamil.

(16)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb \quad S/D Manner adverb</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun \quad S/D Adjective</td>
</tr>
</tbody>
</table>

Finally, there are languages in which either the adjective class is small/derived and the manner adverb class is altogether missing, or in which the adjective class is missing and the adverb class is small/derived. The former situation is represented in (17). It is attested in Kisi and Nung. Kisi has both a small and a derived class of adjectives, neither of which can be used predicatively. Nung has only a small class of adjectives, which can also be used as verbs.

(17)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun \quad Adjective</td>
</tr>
</tbody>
</table>

Finally, the reverse situation – a restricted class of manner adverbs and no adjectives – is represented in (18). This type of system is attested in Garo, Krongo, and Hixkaryana.
5.3.4. Predicted but not attested systems  
(and attested but not predicted systems)

When considering all the predicted PoS systems in Chapter 2, we find that the following ones are not attested in any of the sample languages:

(i) PoS systems with flexible predicatives: see (19) and (20) below (and cf. (23) and (24) in Chapter 2);
(ii) PoS system with flexible heads: see (21) below (and cf. (25) in Chapter 2);
(iii) PoS systems with a class of flexible lexemes that can be used in all functions except one of the modifier functions (Flex A and Flex B): see (22) and (23) below (and cf. (26) and (27) in Chapter 2);
(iv) PoS systems that involve either one or both of the ‘cross-wise’ flexible classes Flex C and Flex D: see (24) and (25) below (and cf. (28) and (29) in Chapter 2).

(19)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicative</td>
</tr>
<tr>
<td>Reference</td>
<td>Nominal</td>
</tr>
</tbody>
</table>

(20)

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicative</td>
</tr>
<tr>
<td>Reference</td>
<td>–</td>
</tr>
</tbody>
</table>
I start out with the systems involving predicatives. First, the fact that the system in (20) is not attested does not appear to be very surprising, since this system has no lexical strategy to express any function within the referential domain. Even though constraint (11) in Chapter 2 determines that the predicative domain ranks higher than the reference domain, this constraint seems to have more bearing on functional specialization (as opposed to flexibility) of PoS classes, rather than predicting the presence versus
absence of a lexical strategy in the first place. Apparently, as the discussion of Tuscarora in the previous section also suggests, languages may make relatively extensive use of verbal as opposed to nominal constructions, but a complete lack of lexical strategies to express any function in the referential domain seems to be a marginal option, if it occurs at all.

In contrast, the system in (19) does have a lexical strategy for the expression of referential functions: it has predicatives and nominals. In fact, even though this system is not attested in its ‘pure’ form, there are two languages in the sample with PoS systems that resemble it to some extent. One of these languages is Paiwan. It seems that all lexemes in this language that usually occur as predicative heads (i.e. ‘verbs’) can also be used as predicate modifiers (‘manner adverbs’), and vice versa. There are only two items that possibly qualify as lexical adverbs, because they cannot be used predicatively (Egli 1990: 158, 313). Since only one of these items appears to be a manner adverb, I do not make reference to a (small) class of manner adverbs in Paiwan (see Table 5.1). Even though Paiwan thus has a class of predicatives, its PoS system differs from the system in (19): Paiwan has rigid nouns and a restricted class of rigid adjectives, whereas the system in (19) has nominals. Thus, the PoS system of Paiwan can be analyzed as intermediate between the systems represented in (26)* and (27)* below (cf. the systems in (15)* and (16)* of Appendix ii). As the asterixes are meant to indicate, these systems are both excluded by the implicational PoS map, in particular on the basis of constraint (11a) in Chapter 2, which states that the availability of a rigid class for heads of referential phrases (nouns) implies the availability of a rigid class for heads of predicate phrases (verbs).

(26)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicative</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
</tr>
</tbody>
</table>

(27)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicative</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
</tr>
</tbody>
</table>
There is, however, an alternative analysis, which hinges on the fact that it is not easy to distinguish in Paiwan between a predicative head modified by a simple lexical manner expressions and a serial verb construction, since the language has very little verbal morphology. Therefore, it is possible that Paiwan predicatives are in fact rigid verbs (cf. Egli 1990: 312). If this is the right analysis, then constraint (11a) of Chapter 2 is no longer violated.

The second language in the sample with predicatives is Kayardild. According to Evans, in this language

“[V]erbals primarily denote actions and processes, but may also provide adverbial type information about the manner in which these are carried out. […] A few lexemes only permit the modifier function.” (Evans 1995: 86)

In fact, Evans lists only four lexemes that are restricted to the function of modifier in a predicate phrase (i.e. rigid manner adverbs), and about these he remarks that “it is possible that a bigger corpus would see even these used as main verbs” (Evans 1995: 303–304). Thus, it seems that Kayardild indeed has predicatives. However, Kayardild’s PoS system also differs in some respects from the system in (19) above: First, Kayardild has a class of non-verbs, rather than nominals. In addition, there is a class of rigid manner adverbs, which are analyzed as a subclass of non-verbs, since, unlike predicatives, they take nominal inflection (Evans 1995: 227–229). There are thus two lexical strategies available to express the function of modifier in a predicate phrase: a ‘verbal’ and ‘nominal’ one. Moreover, Kayardild has a reasonably large class of rigid adjectives (about 100 items). This is unexpected, since it involves a specialized class for modification in the referential phrase, without a specialized class for the corresponding head function (i.e. without rigid nouns, cf. constraint (12a) of Chapter 2). In Figure 5.2 the PoS system of Kayardild is schematically represented:

```
Figure 5.2: The PoS system of Kayardild
```
I now turn to the predicted but unattested PoS system in (21) above. Even though this system does not occur in its pure form, there are some languages in the sample with a PoS system consisting of rigid nouns and verbs in combination with a restricted group of lexemes that can apparently be used as both. The relevant languages are Slave, Nivkh, West Greenlandic, and possibly Nunggubuyu. The question is, however, whether these apparent ‘heads’ are truly flexible lexemes, or whether they rather involve zero-derivation/conversion or polysemy. In section 5.6, where I discuss lexical flexibility in the sample languages, I will argue that the non-compositional semantic interpretation of the relevant lexemes in verbal and nominal function suggests that they do not represent cases of true flexibility.

The non-attestation of flexible heads may also be related to the fact that the propositional function(s) of modification appear not to be relevant in all languages. Rather, in some languages ‘modifiers’ take the form of either separate predications or of appositional referential expressions. In such cases the head-modifier parameter becomes irrelevant, which means that a hypothetical class of ‘heads’ would in fact be a class of ‘contentives’, in the sense of a lexeme class whose members can be used in all (relevant) functions 75.

The issue of irrelevance of propositional function(s) probably plays a similar role in explaining why the systems in (22) and (23) above are not attested. It is quite possible that some of the languages classified as having a single class of contentives, in fact have a PoS class with maximally flexible lexemes, which can be used in the complete set of relevant functions; a set that excludes one of the two modifier functions. In Tagalog, for instance, it seems that the slot for modification in a predicate phrase is irrelevant (see Hengeveld & Van Lier 2009).

Finally, the systems in (24) and (25) above seem intuitively rather improbable, even though they cannot be excluded on the basis of the constraints as formulated in Chapter 2. In particular, these systems involve one or two flexible lexeme class(es), which combine the opposite values of both the predication-reference and the head-modifier parameter. In other words, the relevant flexible classes combine the head function of one domain with the modifier function of the other domain. The markedness

75 Note that the possibility of irrelevant propositional function(s) in specific languages suggests that the functional space as defined in Chapter 2 is not universal. This is in line with the assumptions of the framework of Functional Discourse Grammar, which explicitly distinguishes between language-specific pragmatic and semantic functions, which belong to the grammar of a language, and universal conceptualisation, which belongs to general cognition (Hengeveld & Mackenzie 2008).
of this situation, i.e. the neutralization of maximally divergent functional distinctions, may explain the non-occurrence of the systems (24) and (25) in the sample languages.

Of course, in these and all other cases discussed above, it is possible that the non-attestation of a predicted PoS system is purely coincidental, and that the relevant system will turn up in an investigation of a larger sample. In addition, as I have suggested several times above, both the non-attestation of predicted systems and the attestation of excluded systems may be due to particular difficulties in the analyses of the available language data. Finally, it may very well be that the formulation of the restrictions on PoS systems in Chapter 2 is not yet optimal, and/or that the model should take into account more functional parameters.

5.3.5 Summary
To summarize, in this section I have related the PoS systems attested in the languages of the sample to the PoS systems predicted in Chapter 2. It was shown that most predicted systems are actually attested, either in their ‘pure’ form or in the form of an ‘intermediate’ system. I have suggested possible explanations for the cases of predicted but unattested PoS systems. Apart from the unexpected attestation of rigid adjective and noun classes in the two languages with predicatives (Paiwan and Kayardild), none of the systems that were excluded on the basis of the constraints in Chapter 2 (see Appendix II) have been found in the sample languages. In general, these results suggest that the implicational map model of PoS developed in Chapter 2 has a reasonable typological adequacy.

5.4 The generality and the subclass problems in practice

5.4.1 The generality problem
As explained in Chapter 2, the Hengeveldian method for identifying PoS classes crucially depends on whether or not extra structural coding (specific function-indicating morpho-syntax) is required to use a member of a particular lexeme class in a certain propositional function. In contrast, behavioural potential (the expression of morpho-syntactic categories belonging to a particular function) does not play a role in the definition of PoS classes.

In the present study, the practical application of this method has in some cases resulted in analyses of PoS systems that differ from the ones
proposed in the relevant descriptive sources. In all these cases, the difference in analysis is of the same nature: I propose a flexible PoS class where the reference grammar distinguishes two (or more) rigid word classes. The reverse situation – in which a reference grammar claims flexibility, while I propose multiple rigid PoS classes – does not occur. Obviously, this ‘lumping’ effect is due to the fact that structural coding is a rather restricted defining criterion. Most reference grammars identify PoS classes using the full set of language-specific distributional data, as advocated by Croft (2001). However, as discussed in Chapter 2, the Croftian method raises the generality problem: PoS classes identified in this way are not cross-linguistically comparable. In what follows, I will give examples to show how the Hengeveldian method gets around the generality problem. At the same time, these examples illustrate how the results obtained with the Hengeveldian method may differ from the results of a complete language-specific distributional analysis.

First, consider Kambera76. According to Klamer (1998: 91-95), Kambera has verbs, nouns, and a class of multifunctional lexemes that can be used as both. Nouns and verbs are distinguished by Klamer on the basis of the following criteria:

(i) Nouns can occur with articles and with the marker *bai/*bi, which expresses appreciation or derogation of (properties of) the noun. Noun phrases can be modified by emphatic or demonstrative pronouns. Nouns also have specific quantifying properties. They may occur, for instance, with a bare, un-derived numeral and may have a classifier.

(ii) Verbs can be derived and they can be modified by adverbs (adverbs do not modify nouns).

These criteria involve language-specific aspects of behavioural potential, rather than structural coding. Therefore, following the Hengeveldian approach, these facts are not regarded as evidence for a lexical distinction between nouns and verb.

Moreover, there is explicit evidence in favour of a flexibility analysis. First, Klamer (1998: 96) acknowledges the presence of “a considerable number

76 Recall that Kambera is one of the two languages added to the balanced sample (see Chapter 4), and will therefore not be included in the analyses of Chapter 7.
of multifunctional lexical items" in Kambera. In addition, ‘nouns’ and ‘verbs’ have identical distributional possibilities in terms of the ability to be used predicatively and referentially. This is illustrated in examples (28) and (29) below. In example (28a), an action-denoting lexeme is used as the head of a predicate phrase and takes the inflection belonging to this function, while in (28b) the same lexeme is used as the head of a referential phrase and combines with an article. In (29) two object-denoting lexemes are used as predicative heads77.

Kambera (Klamer 1998: 105, 107)

(28) a. fàka mănganga-ma-na-yna-i una,…
   if steal-emph-3sg.gen-3sg.dat-iter emph.3sg
   ‘If he does steal again,…’

   b. Na ma-kaloru-nya na mănganga…
   ART rel.subj-arrange-3sg.dat ART steal
   ‘Who is engaged in theft,…’

(29) Nina nda tustel-a-ya, senter-ya
   deic-3sg neg camera-mod-3sg.acc torch-3sg.acc
   ‘This is not a camera, it’s a torch.’

Apart from expressing the two head functions, ‘verbs’ and ‘nouns’ in Kambera can also be used, without structural coding, as modifiers in predicate phrases and referential phrases. The examples in (30) show an action-denoting lexeme (a) and an object-denoting lexeme (b) functioning as modifiers in a referential phrase. Examples (31a) and (31b) show an action-denoting and an object-denoting lexeme, respectively, in the function of modifier in a predicate phrase.

Kambera (Klamer 1998: 108, 109)

(30) a. iyang wàu b. meu rumba
   fish smell cat grass
   ‘smelly fish’ ‘wild cat’

77 These examples show that the semantic interpretation of the lexemes in either function is fully compositional: an action-denoting lexeme in nominal function denotes ‘the act of X-ing’, while an object-denoting lexeme in verbal function denotes ‘be X’. I return to this issue in section 5.6.
In short, these examples show that there is no difference between nouns and verbs in Kambera, at least not in terms of the definitions developed in Chapter 2.

Ma’di represents another case in which the Hengeveldian method results in a different classification of a language’s PoS system than the one proposed in the reference grammar. Blackings and Fabb (2003) distinguish between nouns and adjectives in Ma’di, on the basis of the following distributional facts:

(i) Most adjectives are inflected for number (except for colour terms), while most nouns are not;
(ii) Adjectives always combine with an article when they are used as the head of a referential phrase, whereas nouns can also appear without one.

However, adjectives and nouns do not differ in terms of their possibility to be used, without structural coding, as the head and the modifier of a referential phrase. This is illustrated in (32) and (33). In the first example a property-denoting lexeme is used as the head of a referential phrase; in the second example an individual-denoting lexeme is used as a modifier in a referential phrase. Therefore, I have classified Ma’di as a language with flexible nominals.

Ma’di (Blackings & Fabb 2003: 106, 304)

(32)  ḍalī  ṛi  pi  ḍ-tʃā  ḍdʒimī
short(pl)  del  pl.pron  (3)-ve-arrive yesterday
‘The short ones arrived yesterday.’
(33) Māɗí ədriópi  rí  sxá  rú  ilé-ní
   person  brother(INDEF)  DEF  act  REFL  that-like
   kʊ
   NEG(NON-PST)
   ‘A person who is a brother won’t behave like that.’

Note in passing that a determiner is also obligatory in Māɗi when both a modifier and a head are present in the referential phase. The ungrammatical example in (34) shows this:

Māɗi (Blackings & Fabb 2003: 302)
(34) *básí  ingwê  ē-tʃá  rá
    bus  white  (3)-ve-arrive  AFF
    ‘The/a white bus has certainly arrived.’

The case of Māɗi may be contrasted with a case like Hdi. Frajzyngier and Shay (2002: 71) use the following two criteria to define the class of adjectives in Hdi:

(i) Adjectives can modify a noun without any intervening marker (whereas nouns cannot);
(ii) Adjectives cannot be used as arguments (whereas nouns can).

Example (35a) and the ungrammatical (35b) show that these two criteria qualify the lexeme xɓùzá ‘big-bellied’ as an adjective:

Hdi (Frajzyngier & Shay 2002: 72)
(35) a. ndá  ngh-iyù  tá  gu  xɓùzá
    ASSOC  see-1SG  OBJ  goat  big-bellied
    ‘I saw a big-bellied goat.’

b. *ndá  ngh-iyù  tá  xɓùzá
   ASSOC  see-1SG  OBJ  big-bellied
   ‘I saw a big-bellied one.’

Thus, the distinction between nouns and adjectives, as proposed by Frajzyngier and Shay (2002) is based on functional possibilities and
structural coding only. Therefore, their PoS classification is the same as the one I arrive at using the Hengeveldian method.

In this section I have illustrated how the criterion of structural coding allows for the establishment of cross-linguistically comparable PoS classes and PoS systems. It should be clear that the down-side of this method is that it may result in lumping together groups of lexemes that would be regarded as separate classes when the full set of language-specific distributional facts would be taken into account. This brings us to the related subclass problem, which is discussed in the next subsection. Further on, in section 5.6, I return in more detail to the issue of defining flexible PoS classes.

5.4.2 The sub-class problem
In this section I discuss a number of issues related to the sub-class problem. First, as pointed out in Chapter 2, PoS classes in particular languages may have fuzzy boundaries in terms of their distributional behaviour. Such PoS classes display a combination of properties associated with two other groups of lexemes and/or they display varying behaviour in a single propositional function. These phenomena are illustrated in section 5.4.2.1. A second problem concerns the boundary between major, open PoS classes and various types of more restricted classes. In section 5.4.2.2 I consider two types of such restricted classes: (i) small, closed lexeme classes, and (ii) classes of derived lexemes.

5.4.2.1 Fuzzy boundaries
In Chapter 2, I discussed the distinction between nouns and adjectives in Japanese. This case was adduced by Croft (2001) to illustrate that the Hengeveldian approach to PoS definition cannot handle the phenomenon of fuzzy boundaries between language-specific lexeme classes. Since Japanese is one of the languages of my sample, I consider here once more the relevant data, which are presented in Figure 5.3. The different groups of lexemes identified in Croft’s analysis are indicated here with numerical codes (cf. Figure 2.13 in Chapter 2).
Following the Hengeveldian approach, the data in Figure 5.3 result in the distinction of three PoS classes in Japanese:

(i) A class of nouns (class number 1): lexemes that are used without structural coding as the head of a referential phrase. These lexemes need a copula when used in predicative function, and a genitive marker _no_ when used as modifier in a referential phrase;

(ii) A class of nominals (class number 2): lexemes that are used without structural coding as the head of a referential phrase, and with the linking element _na_ (rather than _no_) when used as a modifier in a referential phrase. In predicative function these lexemes need a copula.

(iii) A class of verbs (class number 3): lexemes that are used without structural coding as the head of a predicate phrase, and that cannot be used as the head of a referential phrase. When used as a modifier in a referential phrase, these lexemes remain verbal (i.e. are relativized): they inflect for tense and take neither the genitive nor the linking element, as do nouns and nominals, respectively.

The remaining classes, i.e. the ones with the codes 1/2, 2/3, 2/3a and 2/3b in Figure 5.2 above, are not taken into account as separate classes. This is reflected in Table 5.2, which repeats the relevant data for Japanese from Table 5.1.
In other languages of the sample I have come across comparable cases of fuzziness. For instance, in Mandarin Chinese the existence of a separate class of adjectives is still debated. In terms of the Hengeveldian approach the following facts are relevant: There is a large group of property-denoting lexemes (i.e. semantic ‘adjectives’) that can function as modifiers in referential phrases only when they are structurally coded with the marker de, whereas they can be the heads of predicate phrases without any structural coding. The modifying construction is illustrated in (36) below. This construction is exactly the same as the one used for action-denoting lexemes in modifier position, i.e. relative clauses, as illustrated in example (37). For this reason, semantic ‘adjectives’ are analyzed as (stative) verbs.

*Mandarin Chinese* (Li & Thompson 1981: 32, 581)

(36) \[ Hóng=hong \text{ de } hūa \]

red=red LK flower

‘flowers that are really red’

(37) \[ Jīntiān yíng \text{ de } qián \]

today win LK money

‘the money that we won today’

However, according to Paul (2005), there is also a relatively small group of so-called “non-predicative adjectives”. When these property-denoting lexemes modify a noun, they appear either with or without the linking element de, as is shown in (38) below. In addition, these lexical items cannot function as the head of a predicate phrase without structural coding; they need the copula shì in combination with de, as is illustrated in (39).

*Mandarin Chinese* (Paul 2005: 760, 759)

(38) \[ ta mai-le yi-ge fang (de) panzi \]

3sg buy-pfv 1-cl square LK plate

‘He bought a square plate.’

(39) \[ Zhei-ge panzi shì fāng de \]

this-cl plate COP square LK

‘This plate is square.’
On the basis of these distributional facts, I have analyzed Mandarin Chinese as having (i) a class of verbs: lexemes that can only be used predicatively, and (ii) a small class of adjectives: lexemes that can (at least sometimes) function without structural coding as modifiers in referential phrases, and that cannot be the heads of predicate phrases without a copula. The fact that the adjective class had ‘fuzzy edges’, in the sense that its members show variable behaviour in the modifier position, is not taken into account. This analysis is represented in Table 5.3, which repeats the data for Mandarin Chinese from Table 5.1:

<table>
<thead>
<tr>
<th>Language</th>
<th>Pred Head</th>
<th>Ref Head</th>
<th>Ref Mod</th>
<th>Pred Mod</th>
<th>Small</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandarine Ch</td>
<td>Verb</td>
<td>Adj</td>
<td>S</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.3: Verbs and Adjectives in Mandarin Chinese

In sum, this section further illustrated the practical repercussions of the restrictive structural coding criterion of the Hengeveldian method. In particular, fuzzy boundaries between language-specific lexeme groups and class-internal behavioural variability are ignored, in order to maintain the cross-linguistic comparability of PoS classes.

5.4.3.2 Restricted classes

Small, closed classes

Lexemes and lexeme classes in languages can be divided (broadly) into content words and function words. Generally, content word classes are open, which means that new members can easily be added to them. Also, content word classes are large; they comprise hundreds or thousands of items. Function word classes, on the other hand, tend to be closed and small.

It can be difficult, however, to determine the boundary between these two types of word classes in individual languages. Especially adjectives and/or (manner) adverbs (or flexible modifier classes, for that matter) are in some languages more like function words than like content words, in that they constitute relatively restricted, closed classes (Haspelmath 2001: 16539-42). The available descriptive sources do not always provide clear indications of the size and status of adjective, adverb, and/or modifier classes. And even if such information is available, definitions remain far from absolute.
I already briefly touched upon this point in Chapter 2, in connection with the adjective class in Alamblak. A comparable case is attested in Abun. According to Berry and Berry:

“The class of adjectives in Abun hovers on the borderline of being an open or a closed word class (...). In the corpus a total of 38 adjectives has been distinguished, but there may well be others, thus it is included as an open class.” (Berry & Berry 1999: 36)

The number of manner adverbs in Abun is somewhat more limited; Berry and Berry (1999: 25) identify 25 items. They state that:

“The subclass of manner adverbs is not as ‘open’ as a regular open class because many adverbial meanings are expressed in other ways, such as associative prepositional phrases or the use of adjectives without any special marking.” (Berry & Berry 1999: 37)

Even though manner adverbs are thus not very numerous, Berry and Berry include them as an open word class in Abun. In addition to adjectives and adverbs, there is a group of about 12 lexical items in Abun that can function as both (see also the second quotation above). This group is analyzed as a small (and probably closed) class of flexible modifiers.

In some other languages even smaller classes are distinguished. Hdi, for instance, has 8 lexical adjectives (Frajzyngier & Shay 2002: 71). Classes of manner adverbs may likewise consist of (very) few members. For example, Tamil is described as having only 3 manner adverbs: *mella* ‘softly’, *molla* ‘slowly’, and *jalti* ‘quickly’ (Asher 1985: 115).

In many other cases, the available descriptive sources do not give specific numbers or lists of items that belong to classes of adjectives, adverbs, or modifiers. Rather, such classes are characterized for instance as ‘restricted’ compared to other classes like nouns and verbs, which comprise much larger amounts of items. Clearly, the bottom line of this section is that at least some of the decisions to mark a PoS class as ‘small’ (S) in Table 5.1 involved a certain degree of arbitrariness.
Derived classes

PoS classes that are marked as ‘derived’ (D) in Table 5.1 involve items that are formed by means of productive derivational processes only\(^{78}\). Furthermore, derived PoS classes are mentioned separately only in cases where there is either no categorically equivalent class of simple items, or where the categorically equivalent simple items constitute a small, closed class.

For example, if a language has a large, open class of simple nouns and a productive process to derive nouns, then the latter is not mentioned in the data. In contrast, if a language has no class of simple adjectives, but does have a productive process to derive adjectives, then a derived class of adjectives is included in the data. Also, when a language has but a small, closed class of adjectives, then the possibility to productively derive adjectives is also mentioned separately. In the present section I give some examples of derived PoS classes from the sample languages.

Consider once more Kambera. We already saw that, as far as un-derived items are concerned, Kambera is analyzed as having contentives and adverbs. However, the language also has two productive verbalization processes: (i) \textit{pa}- prefixation, which derives verbs with causative, permissive, factitive, resultative, intensive, infinitive, or reciprocal interpretation; and (ii) \textit{–ng} suffixation, which forms applicative verbs. Some examples of these two processes are given in (40) and (41):

\begin{itemize}
\item Kambera (Klamer 1998: 179, 199)
\item (40) \textit{ànga} ‘foolish/useless’ – \textit{pa-ànga} ‘confuse, cheat on someone’
  \item \textit{ana} ‘child’ – \textit{pa-ana} ‘have children’
  \item \textit{dua} ‘two’ – \textit{pa-dua} ‘divide’
  \item \textit{ndia} ‘no’ (emphatic negator) – \textit{pa-ndia} ‘deny’
\item (41) \textit{bungga} ‘open X’ – \textit{bunggahu-ng} ‘open X for Y’
  \item \textit{riki} ‘laugh’ – \textit{riki-ng} ‘to laugh at/about X’
  \item \textit{angu} ‘friend’ – \textit{angu-ng} ‘have X as a friend’
  \item \textit{nàmu} ‘towards the speaker’ – \textit{nàmu-ng} ‘move towards the speaker’
\end{itemize}

\(^{78}\) The notion of ‘productivity’ is of course rather slippery. I have, as much as possible, relied upon unambiguous statements in reference grammars in deciding whether or not a certain derivational process is productive, i.e. applicable to all members of the input class(es).
In languages with rigid PoS systems, it is often the case that a class of derived items exists alongside a small, closed class of underived items with the same categorial value. Usually, the relevant lexeme classes are adjectives and/or manner adverbs. Nama, for instance, has a small class of simple adjectives and two types of productively derived adjectives, formed with the suffixes -xà (‘attributive’) and -o (‘privative’), respectively (Hagman 1979: 60). However, classes of derived lexemes also occur without the presence of any categorially equivalent class of simple items. Basque, for example, has derived manner adverbs only. These are productively formed from adjectival bases with the suffix -ki, as in sendo ‘strong’, sendo-ki ‘strongly’ (Hualde & Ortiz de Urbina 2003: 193).

A factor that may complicate the identification of classes of productively derived lexemes is the influence of diachronic processes. First, it may be the case that a certain proportion of the derived forms becomes lexicalized and that their base forms no longer occur independently. Consider for example Babungo: The large majority of the adjectives in this language is derived from verbs. However, there are also a few adjectives for which no corresponding verbal form exists. Historically, these forms probably did exist, but synchronically it seems appropriate to say that Babungo has a small class of simple adjectives, next to a class of derived adjectives.

The case of manner adverbs in Babungo is similar, except that the process of lexicalization is apparently more advanced: Many adverbs take the form of reduplications. Only a few of them, however, can still be identified as derivations from verbs. For instance, the manner adverb bwānɔnɔnɔ ‘gently’ is derived from the verb bwānɔnɔ ‘to be well, to be soft’. In contrast, for the lexeme wëe-wëe ‘slowly, gently’ no verbal base form is available (Schaub 1985: 246). In accordance with cases of the second type, Babungo is synchronically analyzed as having a restricted class of underived manner adverbs. Furthermore, since the derivational process is no longer productive, no class of derived manner adverbs is included for Babungo in Table 5.1.

A related potential problem concerns the distinction between lexical and clausal derivation. In Tamil, for example, it may seem as if there is a class of derived adjectives, formed from verbal and nominal bases by means

---

79 In one case, namely Hungarian, I have coded both a class of simple and a class of derived manner adverbs. In fact, however, the status of these classes is not entirely clear. According to Kenesei et al. (1998: 222), simple adverbs are ‘numerous’. However, according to Rounds (2001: 180) most manner adverbs are derived, which seems to suggests that the class of simple items is in fact relatively restricted.
of the suffixes -a and -aana/-ulla, respectively. However, these derivations are in fact all dependent clauses: The suffix -a operates on the phrase level, and the argument(s) of the dependent predicate retain the same expression as main clause arguments. This is shown in example (42):

*Tamil* (Asher 1985: 28)

(42) [Akkaa taykaccikki caata pooût-a] karantî
elder.sister younger.sister-DAT rice put: PST-PTC spoon
‘The spoon with which elder sister gave rice to younger sister.’

Derivations from nominal bases with -aana/-ulla are no different: These suffixes are in fact dependent, participial forms of the verbs aaku ‘to become’ and ullu ‘(existential) be’. This can be seen in example (43):

*Tamil* (Lehmann 2005: 13)

(43) [aZak-aan-a] kannati
beauty-become . PST-PTC mirror
‘beautiful mirror’ (lit. ‘mirror that has become beautiful’)

In sum, what may look like a derived adjective in Tamil is in fact a relative clause construction.

Before rounding off this section, it is worth mentioning that the notion of derived PoS classes clashes in some sense with the criterion of structural coding, since word-class changing derivation can be regarded as an instance of structural coding in itself. Therefore, it may seem strange to say that the members of a class of derived lexemes can be used without structural coding in a particular function. On the other hand, derivation is also a means by which new group of lexical stems is created, which is especially suited for the expression of a specific function. This group of stems may in turn form the input of further processes of morphological or syntactic derivation and structural coding. Moreover, in the context of the present study, the crucial question is whether derived PoS classes may serve as functional models for DC constructions, and there seems to be no principled reason why they may not.

### 5.5 Non-verbal predication

As explained in Chapter 2, the function of head of a predicate phrase has an exceptional status within Hengeveld’s PoS theory. This is because this function is the *only* possible function of the PoS class defined as ‘verbs’,
while it can be an *additional* function of PoS classes other than verbs. In Chapter 2 I discussed the possible strategies that languages may employ to express non-verbal predication. These strategies were divided into those with a copula and those without one. The strategies without a copula were further subdivided into zero-1 and zero-2 strategies. A zero-1 strategy means that the non-verbal predicate shows the same behavioural potential as a verbal predicate in the language under analysis. A zero-2 strategy means that the non-verbal predicate does not express any verbal categories and is simply juxtaposed to its argument.

The use of a copula strategy counts as structural coding. Therefore, a non-verbal PoS class that employs this strategy is analyzed as not including the function of head of a predicate phrase in its range of functional possibilities. The same holds for non-verbal PoS classes that use a zero-2 strategy: they are also analyzed as non-predicative, since they cannot express the same categories as verbal predicates. In addition, zero-2 strategies often occur in alternation with a copula strategy. In such cases, the copula does not appear when there is no functional need to express verbal features like TAM distinctions and/or Person. Typically this involves ‘default’ cases, such as present tense and third person singular. The Hungarian examples (44) (in present tense) and (45) (in past tense) illustrate the alternation of a zero-2 strategy and a copula strategy.

*Hungarian* (Kenesei et al. 1998: 62, 78)

(44) *A lány diákl*  
the girl student  
‘The girl is tall/a student.’

(45) *Péter diákvolt.*  
Peter student **cop(pst)**  
‘Peter was a student.’

In short, PoS classes that employ a copula or a zero-2 strategy when used as the head of a predicate phrase are regarded as non-predicative. In contrast, PoS classes that use the zero-1 strategy are analyzed as including the function of head of a predicate phrase in their range of distributional possibilities.

However, several difficulties may arise when classifying the strategies used for non-verbal predication in individual languages. First, it may be the case that only *part* of the behavioural potential of verbal predicates
can be carried over to non-verbal predicates. This happens for instance in Abkhaz and Hdi. Non-verbal predicates in Abkhaz can take verbal suffixes that express present tense, past tense, and imperative mood. However, other verbal suffixes can be added to dynamic stems only. A copula intervenes when one of the latter suffixes must be expressed on a non-dynamic, non-verbal predicate. In Hdi nouns do not need a copula when they are used in predicative function and tense can be expressed on them. However, unlike verbal predicates, nouns cannot be marked for aspect. The reason for this discrepancy is that tense is marked by means of independent morphemes, while aspect marking involves different stem forms of verbs, and these are not available for nouns. In cases like Abkhaz and Hdi I have opted for the strongest possible definition of the zero-1 strategy: a non-verbal predicate is analyzed as employing a zero-1 strategy only if the full set of verbal behavioural potential can be expressed on the non-verbal predicate.

Second, it can be problematic to apply this criterion for the zero-1 strategy to isolating languages. In these languages, which lack inflection by definition, it is not possible to decide in favour of a zero-1 strategy (instead of zero-2) on the basis of the expression of verbal inflection on the non-verbal predicate. In such cases, there are two possible ways to distinguish between a zero-1 and a zero-2 strategy:

(i) The possibility to add a copula;
(ii) The possibility to combine the non-verbal predicate with freestanding particles that express verbal categories such as TAM.

For instance, Paiwan does not have a copula, and non-present tense can be expressed with non-verbal predicates by means of the freestanding particles *na* (for past) and *urbi* (for future), as is shown in (46) below. This is interpreted as evidence that there is no difference between verbal and non-verbal predication in terms of behavioural potential, i.e. that a zero-1 strategy is used in Paiwan.

*Paiwan* (Egli 1990: 61)

(46) Na/*urbi* *sivitai* ti *kama*

\[\text{pst/fut} \text{ soldier foc father}\]

‘Father was/will be a soldier.’
In other isolating languages, such as Hmong Njua, Abun, Thai, and Nung, nominal predicates do take a copula. Under certain circumstances this copula can be omitted, which results in a zero-2 strategy. In contrast, adjectival predicates never take a copula. In addition, there appear to be no principled restrictions on the usage of grammatical particles with adjectival predicates. In (47) I give an example from Thai, which shows the expression of imperfective aspect with an adjectival predicate by means of the particle \( yiu \). On the basis of these data, Thai is analyzed as having predicative adjectives. Similar evidence is available for the other isolating languages mentioned above.

\[
\text{Thai (Iwasaki & Ingkaphirom 2005: 155)}
\]

(47) \( aab\,\text{man} \ kaml\,\text{ay} \ r\,\text{öm} \ yiu \)
\[
\begin{array}{c}
\text{food} \\
\text{ADV} \\
\text{hot} \\
\text{IPFV}
\end{array}
\]

The food is still hot.

It may be worthwhile at this point to stress the difference between (i) a situation in which adjectives can be used predicatively with a zero-1 strategy, and (ii) a situation in which there is no difference between adjectives and verbs. While the latter case involves bi-directional flexibility, in the former case there is only a unidirectional identity relation: all adjectives can be used as verbs, but not all verbs can be used as adjectives without structural coding\(^80\). Rather, when a verb is used as a modifier in a referential phrase, it must be relativized, as in example (48), again from Thai:

\[
\text{Thai (Iwasaki & Ingkaphirom 2005: 243)}
\]

(48) \( Khon \ [\text{thii} \ du\,\text{leee}] \ ni \ pen \ pen \ acaan \ l\,\text{is} \)
\[
\begin{array}{c}
\text{person} \\
\text{REL} \\
\text{take.care} \\
\text{PRT} \\
\text{COP} \\
\text{COP} \\
\text{teacher} \\
\text{Q}
\end{array}
\]

‘Is the person who takes care (of the students) a teacher?’

Notably, the use of a verb in the function of modifier in a referential phrase does not always require an overt relativizer in Thai. Subject relative clauses that ascribe a general property to their head can also appear without the relativizer, as shown in (49). Nonetheless, the availability of the relativizer as a structural coding mechanism, even though it is not always needed, is interpreted as evidence that there is no complete overlap between verbs and adjectives in Thai.

\(^{80}\) This amounts to what Stassen (1997) terms “verby adjectives”.

---

80 This amounts to what Stassen (1997) terms “verby adjectives”.
Thai (Iwasaki & Ingkaphirom 2005: 250)

(49) dek [rian kén]
child study well
‘a child who studies well.’

I will return to the issue of bi-directionality as a criterion for flexible PoS classes in the course of the following section, which is concerned more generally with the identification of lexical flexibility in the languages of the sample.

5.6 Identifying flexibility

5.6.1 Introduction
In Chapter 2, I discussed three criteria for lexical flexibility, as proposed by Evans and Osada (2005): (i) the criterion of semantic compositionality, (ii) the criterion of exhaustiveness, and (iii) the criterion of equivalent combinatorics. In this section, I present evidence for flexible PoS classes in the languages of the sample. As far as the data allow it, I organize the discussion along the lines of Evans and Osada’s criteria. In section 5.6.2 I apply these criteria to languages that appear to display the most radical type of flexibility: a class of contentives that can be used in all propositional functions. In section 5.6.3 I focus on languages that display less flexible lexeme classes, i.e. classes that can express the functions of head and modifier in referential phrases (nominals), modifier in predicate and referential phrases (modifiers), or all three functions (non-verbs).

5.6.2 Languages with contentives
In some languages contentives are the only available lexeme class. This is the case is Tagalog and Kharia. Other languages combine a large class of contentives with one or more rigid PoS classes, which can but need not be small or derived. This is attested in Samoan, Kambera, Santali, and Guaraní.

5.6.2.1 Semantic compositionality
I argued in Chapter 2 (section 2.5.2.2) that flexible languages are characterized by a mismatch between lexical categorization and syntactic categorization. Derivational processes pertaining to the former type of categorization typically involve semantically (and phonologically) unpredictable interpretations. Syntactic categorization of flexible lexemes, in contrast, yields compositional semantic meanings. The presumably flexible
languages discussed in Chapter 2 were shown to exhibit the latter type of categorization, and can thus be analyzed as behaving in accordance with Evans and Osada’s compositionality criterion. The relevant languages are Samoan, Tagalog, and Kharia. At the same time, however, these languages were shown to also display (zero-marked) processes of lexical derivation, with unpredictable semantic outcomes. In this section, I will further discuss this situation. In addition, I will show that similar situations are attested in Kambera and Santali.

For Kambera I have given examples in (28) and (29) above, showing that the insertion of presumably flexible lexemes into verbal versus nominal syntactic functions yields compositional semantic interpretations. However, there are also other types of processes in this language that are accompanied by more idiosyncratic meaning shifts. As illustrated in (50) below, the types of shifts are comparable to the ones discussed in Chapter 2. In particular, they involve meaning such as: ‘the typical cognate object of the action denoted by the root’, ‘a location associated with the action denoted by the root’, ‘the result of an action’, etcetera. As in Chapter 2, I argue that the examples in (50) are instances of lexical zero-derivation, whereas examples (28) and (29) illustrate ‘true’ flexibility.

*Kambera* (Klamer 1998: 110)

<table>
<thead>
<tr>
<th>Form</th>
<th>Nominal meaning</th>
<th>Verbal meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hukung</td>
<td>‘law/penalty’</td>
<td>‘punish’</td>
</tr>
<tr>
<td>hindu</td>
<td>‘attic’</td>
<td>‘dry/smoke’</td>
</tr>
<tr>
<td>bandil</td>
<td>‘rifle’</td>
<td>‘shoot’</td>
</tr>
<tr>
<td>dadu</td>
<td>‘dice’</td>
<td>‘play dice’</td>
</tr>
<tr>
<td>lindi</td>
<td>‘bridge’</td>
<td>‘cross a bridge’</td>
</tr>
<tr>
<td>ramuk</td>
<td>‘pool/swamp’</td>
<td>‘be in a pool/swamp’</td>
</tr>
<tr>
<td>tanda</td>
<td>‘sign’</td>
<td>‘know’</td>
</tr>
<tr>
<td>ludu</td>
<td>‘song’</td>
<td>‘emit sound/sing’</td>
</tr>
</tbody>
</table>

Like the other flexible languages discussed in Chapter 2, Kambera also has overtly marked lexical derivations. In fact, these forms are only formally derived; the process itself is no longer productive. As would be predicted, formally derived forms can be used in nominal as well as in verbal syntactic environments. This is illustrated in (51a-b):
Kambera (Klamer 1998: 113)

(51) a. Ku-manàhal-nya na n-jala-nggu
   1sg.nom-regret-3sg.dat art mistake-1sg.gen
   ‘I regret my sins.’

b. Eha! N-jala-mbu-nggunja-i-ka nú kawài
   excl be.wrong-also-1sg.ca-iter-pfv deic just.now
   ‘Darn! I was mistaken here too.’

On the other hand, Kambera also has derivational processes that produce verbal output forms, i.e. forms that are no longer flexible, but rather categorized in terms of phrase structure. I have illustrated these verbalizing processes in (40) and (41) above.

Recall from Chapter 2 that the flexible language Kharia has the possibility to insert complex phrases consisting of multiple roots or stems into verbal as well as nominal syntactic functions (see example (51) in Chapter 2). Interestingly, complex phrases in Kambera are flexible as well. Consider for instance example (52), in which the translational equivalent of an English NP functions as the head of a predicate phrase. The semantic interpretation is compositional: ‘be X’.

Kambera (Klamer 1998: 107)

(52) [Tau mayila]-mbu-kai nyimi ná
   person poor-also-2pl.a you(pl) deic
   ‘Moreover, you are also poor people.’

Similar examples can be found in Samoan, as is shown in (53): The complex phrase *atu mata tasi* ‘one-eyed bonito’ appears as an object argument with the possessive form *la’u* in (53a), and as a predicate with the future tense particle *ole’a* in (53b). Again, the semantic interpretation of the phrase in these two functional environments is entirely predictable.

Samoan (Mosel 2004: 286, 287)

(53) a. fia maua la’u [atu mata tasi]
    want find my bonito eye one
    ‘I want to find me a one-eyed bonito.’
b. *Ia, ‘ole’ā [atu mata tasi]*
   well FUT bonito eye one
   ‘Was that going to be the one-eyed bonito?’

Consider now Santali, which resembles the other flexible languages in that it combines very regular meaning shifts with more idiosyncratic ones. The former type of shift, which characterizes ‘true’ flexibility, is illustrated in (54) and (55) below. First, examples (54a-b) show the compositional interpretation of action-denoting roots in the nominal functions of direct object (zero-marked for accusative), and dative-marked locative adjunct, respectively:

*Santali* (Neukom 2001: 17; Rau, forthcoming)

(54)  a. *Uni-then odi ruhet'-in jom-akat'-ma*
   that-DAT much scold-1SG.SBJ eat-PFV:AFT:ACT:IND
   ‘I got scolded badly by him.’ (lit: ‘I ate much scolding from him’)

b. *Gapa-do am-geg si-ok'-then*
   tomorrow-TOP you-FOC plough-MID-DAT
   *dangra-do laga-ogu-kin-me*
   bullock-TOP drive-bring-3DAT.OBJ:2SG.SBJ
   ‘Tomorrow you shall drive the bullocks to where I am ploughing.’

The examples in (55) illustrate the compositional interpretation of semantically nominal lexemes in predicative function. In (55a) an individual-denoting lexeme is used as the head of a predicative phrase. Examples (55b-d) show that the same possibility exists for proper names, onomatopoeic forms, and complex NPs, respectively81.

*Santali* (Rau, forthcoming; Neukom 2001: 15)

(55)  a. *ad-e raj-en-a*
   then-3SG.SBJ king-PST:M:IND
   ‘So he became king.’

81 See Neukom (2001: 15) for comparable examples with pronouns, numerals, adverbs, demonstratives, and quantifiers. Similar examples are provided by Peterson (2006: 62ff) for Kharia.
Santali (Neukom 2001: 60)

(58) orak’-ip etxɔp’-akat’a
    house-1SG.SUBJ beginning-PFV:ACT-IND
    ‘I have begun (to build) the house.’

Thus, the process exemplified in (56) is interpreted as the zero-marked counterpart of overt lexical derivation, as illustrated in (57). If this is the right analysis, then it is predicted that both the ‘nominal’ and the ‘verbal’ instantiations of the examples in (58) can be used nominal as well as verbal syntactic environments. Unfortunately, however, the available data are not conclusive on this point 82.

Finally, as mentioned in section 5.3.4 above, there are some languages in the sample that seem to combine rigid classes of verbs and nouns with a class of flexible heads that can be used as both. The languages in which this is attested are Slave, Nivkh, West Greenlandic, and possibly Nunggubuyu. However, it appears that in all cases the meaning shifts displayed by the relevant lexemes are unpredictable, which disqualifies them as instances of flexibility. In (59) I give some examples of presumably flexible items in Nivkh.

Nivkh (Matissen & Drossard 1998: 62)

(59) Form | Nominal meaning | Verbal meaning
---|---|---
k’u | ‘arrow’ | ‘shoot’
γa | ‘beast’ | ‘pursue’
kelma | ‘step’ | ‘step’
k’ɔr | ‘hunger’ | ‘starve’
mos | ‘porridge’ | ‘grind’
ɔlv | ‘roof’ | ‘thatch’

Similarly, Sadock (2003: 79) notes for West Greenlandic that “there is usually an incompletely predictable meaning relation between homophonous noun/verb pairs of stems.” Some examples are provided in (60):

---

82 For Guaraní, there are unfortunately no clear data about the semantic relationship between nominal and verbal uses of flexible items.
West Greenlandic (Sadock 2003: 4)

(60)  Form       Nominal meaning       Verbal meaning
  iga/igavoq  ‘cooking pot’      ‘cook’
  imeg/imerpog  ‘water’      ‘drink’
  kalleg/kallerpoq  ‘thunder’       ‘be thundering’
  kuuk/kuuppoq  ‘river’       ‘flow’
  niu/niuvoq  ‘leg’          ‘get out of a vehicle’
  sianeq/sianerpoq  ‘bell’         ‘ring up, telephone’
  siku/sikuvo  ‘ice’      ‘be frozen over’

Finally, also in Slave there appear to be different types of unpredictable semantic shifts associated with the nominal and the verbal use of the same root. Some examples appear in (61):

Slave (Rice 1989: 161)

(61)  Form       Nominal meaning       Verbal meaning
  t'éh  ‘charcoal’       ‘cook’
  tê  ‘ice’      ‘freeze’
  seeh  ‘saliva’         ‘spit’
  dzéh  ‘gum’         ‘be sticky’
  tthêb  ‘axe’      ‘chop with an axe’

On the basis of the unpredictable semantics of the meaning shifts illustrated in (59)-(61), I argue that these are not to be analyzed as instances of flexibility. In addition, there is a crucial difference between the data in (61)-(63) from Nivkh, West Greenlandic, and Slave on the one hand, and the lexical zero-derivations discussed for the flexible languages on the other hand (see examples (50) and (56)). This difference is due to the separation in flexible languages between lexical and syntactic categorization. In flexible languages, such as Samoan, the two meanings pertaining to a particular form reflect lexical categorization only. Syntactically, i.e. in terms of phrase-structural possibilities, both zero-derived forms are flexible, i.e. can be inserted in verbal as well as in a nominal syntactic frames, with compositional semantic interpretations. In contrast, in rigid languages, such as Nivkh, lexical and syntactic categorization go hand in hand. Therefore, the lexical processes

83 The case of Nunggubuyu is less clear. As I have no data for this language of the type discussed for Nivkh, West Greenlandic and Slave in (59)-(61), I postpone the discussion of Nunggubuyu to the next subsection.
which zero-derive the two unpredictably related meanings of a single root form also determine the phrase-structural possibilities of their output. In other words, any lexically categorized item is also syntactically categorized: it is either a noun or a verb, and can no longer be used in another function than its defining function, at least not without further measures (see also Don & Van Lier, forthcoming).

In sum, truly flexible languages can, in accordance with Evans and Osada’s compositionality criterion, use simple roots, (zero-)derived stems, and complex phrases in more than one syntactic function without extra structural coding and with predictable semantic interpretations. The possibility of semantic compositionality is interpreted as a direct consequence of the fact that in flexible languages lexical and syntactic categorization do not coincide as they typically do in rigid languages.

In the next sub-section I address Evans and Osada’s second criterion, which is concerned with the quantitative measurement of lexical flexibility.

### 5.6.2.2 Exhaustiveness

According to the criterion of exhaustiveness, ‘true’ flexibility requires that all members of a specific PoS class be flexible (see Chapter 2, section 2.5.3). Thus, if a language is claimed to have a class of simple contentives, then all content items belonging to that class should be usable in all relevant propositional functions.

I already indicated in Chapter 2 that my descriptive sources for Tagalog, Kharia, and Samoan are quite unambiguous about the fact that flexibility in these languages concerns all or almost all content words (depending on whether the class of contentives is combined with some (restricted) additional rigid PoS class). First, Himmelmann claims for Tagalog that

“(...)all Tagalog content words (both roots and derived words) are categorically indistinct, i.e. they may all occur in essentially the same basic syntactic positions.” (Himmelmann, 2007: 249; emphasis added, EvL)

For Kharia, Peterson writes the following:

“Almost all underived lexical morphemes in Kharia may be considered precategorial in the sense that they may appear in referential, attributive and predicative function, and [...] there is no reason to assume that
this is any different for derived stems.” (Peterson 2006: 83; emphasis added, EvL)

Finally, for Samoan it is claimed that:

“The categorisation of full words into nouns and verbs is not given a-priori in the lexicon. […] Derivations are more restricted in their usage; some derivations can occur only in verbal functions.” (Mosel & Hovdhaugen 1992: 73)

For the other languages in the sample that I have analyzed as displaying a class of contentives the data are somewhat less straightforward. Regarding Kambera, I explained earlier (see section 5.4.1) that Klamer (1998) distinguishes between nouns and verbs in this languages on the basis of differences in behavioural potential. However, if structural coding is the only criterion, then all lexemes in Kambera that are not adverbs or derived verbs apparently belong to a single class of contentives.

Regarding Guaraní and Santali, I have already mentioned that both combine a class of contentives with a class of rigid, un-derived verbs. (Nordhoff 2004: 56, 59 and p.c.; Neukom 2001: 13, 17). According to Neukom, about one third of all content lexemes in Santali are verbs; the rest are contentives. For Guarani, no such quantitative information is available.

Finally, there are some indications of the proportion of roots that serve as the basis of both nominal and verbal zero-derivations in Slave, Nivkh, West Greenlandic and Nunggubuyu, as discussed above. Even though these cases are not regarded as cases of true flexibility, their degree of pervasiveness seems interesting to report. For Nivkh, Matissen and Drossard (1998: 61) remark that “most root types are fixed with respect to lexical categories”. This suggests that the class of un-categorized roots is at least restricted as compared to the classes of rigid nouns and verbs.

For West Greenlandic, Sadock points out that:

“The neat division of stems into two distinct part-of-speech classes is somewhat complicated by the fact that there are quite a few stems that occur both as a noun and a verb. Bergsland (1955) estimated that about 200 out of 1500 roots are noun-verb homophones, the rest being roughly equally divided between the two classes.” (Sadock 2003: 4-5)
In Slave, category-less roots are converted into verbal or nominal stems by means of stem formation rules. There are three stem formatives that create nouns, the most common of which is a null suffix. Verb stems are composed of an uncategorized root and a suffix that indicates mode and aspect. This process may involve overt suffixes but can also be zero-marked (Rice 1989).

Finally, for Nunggubuyu the story is rather less clear. Heath (1984: 152) distinguishes “nominal adjectives” as a subclass of nouns. However, the difference between the two classes is that nouns cannot be used as predicates using a zero-1 strategy, while nominal adjectives can. In predicative function, nominal adjectives take a pronominal prefix, like verbs, while in referential function they take a nominal class prefix, like ‘regular’ nouns. In addition, it appears that the propositional function of modification is irrelevant in Nunggubuyu (see also section 5.3.4); modifiers always take the form of appositional constructions. In my view, this means that the so-called nominal adjectives in fact constitute a (relatively restricted) class of roots that can receive either a nominal or a verbal instantiation.

In sum, as far my descriptive sources allow it, I have shown that most relevant languages, i.e. those that I have classified as displaying a class of contentives, satisfy the criterion of exhaustiveness.

5.6.2.3 Equivalent combinatorics
The criterion of equivalent combinatorics states that flexibility must be bi-directional. This means that the members of a flexible word class should be combinable in an equal way and to an equal extent with multiple functions. Notably, Evans and Osada (2005) claim that flexibility in Mundari, the language to which they apply their criteria, is in fact not bi-directional. Their argument for this claim is that most action-denoting lexemes in Mundari must be turned into headless relative clauses in order to be usable in a nominal functional environment. An example is given in (62):

\[\text{Mundari (Evans & Osada 2005: 377)}\]

\[(62) \quad \text{susun-}\text{ta-}n=\text{iq} \quad \text{landa-}j\text{a-}n-a\]

\[\text{dance-Progr. Or-Intr-3Sg.Subj} \quad \text{laugh-Incep-Intr-Ind}\]

‘The one who is dancing has laughed.’

However, it seems that Evans and Osada confuse the semantic denotation of action-denoting lexemes in nominal function with their distributional
possibilities. A lexeme like *susun* ‘dance’ in (62) needs to be relativized only if it is to receive the interpretation of ‘the one who is dancing’. As Evans and Osada remark themselves, all verbs can indeed be used directly in what they call “clausal argument positions”, and this yields a compositional semantic interpretation, in which the relevant lexeme still denotes an action rather than an object or an individual. Example (63) illustrates this:

Central, Mundari (Evans & Osada 2005: 377)

(63) **dub** *= ko la:ga-ja-n-a**

sit=3PL be.tired.of-incep-intr-ind

‘They are tired of sitting.’

However, in some cases other semantic interpretations are found. An example would be the root *jom*, which means ‘to eat’ in predicative function, and ‘food’ (rather than ‘eating’) in referential function. In fact, this pattern is very reminiscent of the situation in other flexible languages. In particular, I would argue that *jom* is a case of conversion: lexical, zero-marked derivation with an unpredictable semantic outcome. This is in contrast to the case in (63), which illustrates flexibility. In short, by adhering to the criterion of semantic compositionality as proposed in the previous subsection, we may also solve this potentially problematic aspect of the criterion of equivalent combinatorics.

A second aspect of equivalent combinatorics involves the claim that truly flexible items should be usable to an equal extent in multiple functions. As already pointed out in Chapter 2 (section 2.5.4), I disagree with Evans and Osada that each flexible lexeme should appear equally frequently in all propositional functions. In some of my descriptive sources, this issue of usage frequency is explicitly addressed. For instance Peterson, in his comment on Evans and Osada’s (2005) article, argues for Kharia that:

“[I]n fact, the only real restriction on this type of flexibility is what might be termed semantic compatibility. For example, the loan *tebal/tebul* ‘table’ can function as the complement of a predicate with the meaning ‘table’, in the middle voice with the meaning ‘become a table’ or in the active voice with the meaning ‘turn (something/someone) into a table’. Needless to say, objects seldom turn into tables, hence the predicative use of this lexeme is virtually never found in actual conversations.”

(Peterson 2005: 396)
Crucially however, such predicative usage is not excluded for grammatical reasons. Peterson argues that a similar situation is likely to hold for Mundari, since speakers of this language accept constructions like the one in (64):

*Mundari* (Peterson 2005: 400)

\[(64) \text{Siŋbonga \textit{am-e tebal-ke-d-me-a}.} \]

\begin{tabular}{l}
God & 2SG-3SG \\
\text{table-act.pst-2SG-ind} & \\
\end{tabular}

‘God turned me into a table.’

In the same vein, Mosel and Hovdhaugen say the following about Samoan:

“Although certain full words seem to be used more as a verb or more as a noun phrase nucleus for semantic reasons, there are no lexical or grammatical constraints on why a particular word cannot be used in one or the other function.” (Mosel & Hovdhaugen 1992: 73)

Regarding Tagalog, Himmelmann makes clear claims about the frequency with which action-denoting roots are used in nominal function (he terms this “unaffixed uses”, i.e. without voice marking):

“It should be clearly understood that in general unaffixed uses of action roots are not in any way exceptional. Instead, they are reasonably common both in terms of types as well as in terms of tokens. […] Although there are differences with regard to how frequent and natural it is for a given action root to occur without affixes, it is clearly the case that Tagalog action roots quite generally allow for unaffixed uses.” (Himmelmann 2007: 284)

In sum, the available data strongly suggest that the languages I have analyzed as flexible indeed satisfy the criterion of equivalent combinatorics, as far as the possibility to use lexemes in multiple functions is concerned. As regards usage frequency, there are indications of asymmetry, but this only to be expected on the basis of semantic-pragmatic markedness.

5.6.3 Non-verbs, nominals, and modifiers

In this sub-section I focus on languages that are characterized by a certain degree of lexical flexibility regarding the following three functions: (i) head of a referential phrase, (ii) modifier in a referential phrase, and (iii)
modifier in a predicate phrase. As we have seen in section 5.2, there are some languages, in which these three functions are expressed by a single PoS class termed *non-verbs*. This is the case in Warao, Turkish, Kayardild, and (if small classes are included) Ma’di. Other languages have a class of *nominals* that can fulfil the two functions of head and modifier in a referential phrase. This is attested in Imbabura Quechua, Ma’di, and Hungarian. Finally, a number of languages have flexible modifiers, i.e. lexemes that can function as *modifiers* in both referential and predicate phrases. Below I discuss data from languages with non-verbs, nominals, and modifiers in turn.

### 5.6.3.1 Non-verbs

In Warao, according to Romero-Figeroa (1997: 49) “there is no clear-cut distinction between nouns and adjectives; Warao nouns may function as *attributives* modifying other nouns.” Romero-Figeroa illustrates this claim with examples involving lexemes that denote abstract properties, such as yak ‘beauty’ in (65):

**Warao (Romero-Figeroa 1997: 50)**

(65) *Hiaka* yak-*era* auka saba tainisa-n-a-e

*garment* beauty-*AUGM* for she *buy-SG-PUNCT-PST*

‘She bought a beautiful dress for her daughter.’

Note however, that it is not clear in how far object- and individual-denoting lexemes can also function without structural coding as modifiers in referential phrases. Moreover, there is no explicit evidence for bi-directional flexibility. That is to say, there are no examples in which items like yak are used without structural coding as heads of referential phrases (even though this possibility is expected on the basis of the ‘nominal’ translation of the lexeme).

In addition to the fuzzy noun-adjective distinction, the lexical flexibility in Warao presumably also includes the manner function. Example (66) below shows that the item yak ‘beauty’ can be used without structural coding as a modifier in a predicate phrase:

**Warao (Romero-Figeroa 1997: 70)**

(66) *Tai* bi saba yak-*era* ana tan-a-e

*it you for* goodness-*AUGM* neg happen-SG-PUNCT-PST

‘It happened not so well for you.’
On the basis of these data, I have classified Warao as having a class of non-verbs, but admittedly the evidence is rather patchy.

Turkish has a class of non-verbs as well (Göksel & Kerslake 2005: 49-50). In Chapter 2 this was illustrated with the lexeme güzel ‘beauty’. Note that, like yak in Warao, Turkish güzel denotes a property concept. Therefore, it may be questioned in how far object-denoting lexemes have the same distributional possibilities. According to Göksel and Kerslake however, indeed many lexemes display noun-adjective flexibility, even though each lexeme has a ‘primary’ function in which it occurs most frequently for semantic reasons.

Finally, Kayardild has a large class of lexemes that can be used flexibly as heads and modifiers in referential phrases. Some examples are given in (67):

Kayardild (Evans 1995: 85, 238)

(67) maku: ‘woman/female’
    balumbanda: ‘westerner/western’
    wurkura: ‘boy/male’
    jambanda: ‘hollow log/hollow’

Such items are also productively used as what Evans (1995) calls ‘secondary predicates’. I regard secondary predicates as participant-oriented manner expressions (see Himmelmann & Schultze-Berndt 2005). This means that in Kayardild the functions of nouns, adjectives, and manner adverbs can be expressed by a single class of lexemes: non-verbs. Notably, there is also a small class of rigid manner adverbs in Kayardild (see figure 5.1 above). Evans interprets these as a subclass of nominals, because they take nominal morphology, as is illustrated in (68) below. However, he explicitly mentions the difference in distributional freedom between these ‘manner nominals’ on the one hand and non-verbs on the other: “The distinguishing feature of this minor subclass is that it is restricted to one function, whereas the other nominal classes are versatile.” (Evans 1995: 227; emphasis added, EvL).

Kayardild (Evans 1995: 259)

(68) Ngawarri bukawa-th
    thirsty.nom die-act
    ‘(He) died thirsty.’
5.6.3.2 Nominals

Languages with non-verbs and nominals have in common that they lack a distinction between nouns and adjectives. As explained in section 2.5.4 of Chapter 2, this type of flexibility can be difficult to ascertain. In particular, it is often problematic to establish whether property-denoting lexemes can be referentially independent, or whether they rather function as modifiers of understood heads.

Hengeveld (1992: 63) argues for Imbabura Quechua that property-denoting lexemes can indeed be referentially independent. Example (69) shows the use of the property-denoting word *hatun* ‘big’ as the sole element of a referential phrase.

*Imbabura Quechua* (Schachter & Shopen 2007: 17)

(69) Rikaška hatun-ta
     see.pst-1sg big(pl)-acc
     ‘I saw the big one(s).’

In contrast, Beck (2002: 146-149; see also Cerrón-Palomino 1987) claims that Imbabura Quechua does have a lexical noun-adjective distinction. The evidence Beck advances for this claim is that modifying lexemes with nominal semantics, such as *rumi* ‘stone’ in *rumi pan* ‘stone road’, can occur only once in a referential phrase, whereas modifying lexemes with ‘adjectival’ semantics are recursive. According to Beck, this makes constructions with ‘nominal’ modifiers look very much like compounds. Support for this analysis is found in the fact that when such ‘compounds’ function themselves as modifiers in a referential phrase, the only possible reading is the one where the combination of the two nouns in the ‘compound’ modifies the head. Thus, in *hara čakra rumi* ‘stone from a cornfield’ the combination *hara čakra* ‘corn field’ modifies the head *rumi* ‘stone’. On the other hand, as Beck acknowledges, there is no phonological evidence for a ‘compound’ analysis. On a more general note, Beck observes that, on the basis of isolated sentences in a descriptive grammar, it is quite hard to decide upon the issue of noun-adjective flexibility in any language. This is obviously true, and in the case of Imbabura Quechua it poses a serious problem. Nonetheless, for lack of conclusive evidence to the contrary, I maintain that Imbabura Quechua has a class of flexible nominals.

The second language with nominals is Hungarian. It was already briefly mentioned in Chapter 2, where I pointed out that Hungarian property-denoting lexemes can be referentially independent (Moravcsik 2001).
Examples (70a-b) and (71a-b) show that object-denoting and property-denoting lexemes can both function as heads and modifiers in referential phrases. Moreover, these usages can be distinguished from appositional constructions, which are also open to both semantic types of lexemes. This is illustrated in (72a-b)

**Hungarian** (Moravcsik 2001: 339)

(70)  
\[ \text{a.} \quad A \ \text{ház-ak-at} \ \text{látom} \]
\[ \quad \text{the house-PL-ACC} \ \text{I.see} \]
\[ \quad \text{‘I see the houses.’} \]

\[ \text{b.} \quad A \ \text{nagy-ok-at} \ \text{látom.} \]
\[ \quad \text{The big-PL-ACC} \ \text{I.see} \]
\[ \quad \text{‘I see the big ones.’} \]

(71)  
\[ \text{a.} \quad \text{Ezr} \quad a \ \text{bestia nő-t} \ \text{utálom.} \]
\[ \quad \text{this:ACC} \ \text{the beast woman-ACC} \ \text{I.hate} \]
\[ \quad \text{‘I hate this beast of a woman.’} \]

\[ \text{b.} \quad A \ \text{kél ház-ak-at} \ \text{látom.} \]
\[ \quad \text{the blue house-PL-ACC} \ \text{I.see} \]
\[ \quad \text{‘I see the blue houses.’} \]

(72)  
\[ \text{a.} \quad A \ \text{ház-at}, \ \text{a szülöhely-em-et}, \ \text{látom} \]
\[ \quad \text{The house-ACC} \ \text{the birth.place-sg1-ACC} \ \text{I.see} \]
\[ \quad \text{‘I see the house, my birthplace.’} \]

\[ \text{b.} \quad A \ \text{ház-at}, \ \text{a kék-et}, \ \text{látom} \]
\[ \quad \text{The house-ACC} \ \text{the blue-ACC} \ \text{I.see} \]
\[ \quad \text{‘I see the house, the blue one.’} \]

The third language in the sample with a class of nominals is Gooniyandi. McGregor states that in this language:

“[N]early all words which can realise the Entity role can also realise the role of Qualifier: in other words, words which can be referential can also be qualifying, and may indicate qualities or properties of things.”
(McGregor 1990: 142).
It should be noted that the above quotation makes reference to unidirectional flexibility only: the use of entity-denoting lexemes (semantic ‘nouns’) in modifying function. With respect to the use of semantic ‘adjectives’ in referential function, things are slightly more complicated. As McGregor notes, for a number of property-denoting lexemes such usage has not yet been attested. However, according to the author, this is due to pragmatic markedness and “limitations of the data”, since many semantically similar (i.e. other property-denoting) lexemes do occur as heads of referential phrases. Therefore, I conclude with McGregor that there is insufficient evidence to justify a noun-adjective distinction in Gooniyandi.

Noun-adjective flexibility in Ma’di has already been addressed in section 5.4.1. It seems that in this language property-denoting lexemes can indeed be referentially independent. Note that, in addition to nominals, Ma’di has a small, closed class of what appear to be flexible non-verbs. These are property-denoting lexemes, which can be used as modifiers in both referential and predicative phrases (and also as heads of referential phrases if one assumes that there is no lexical noun-adjective distinction). Some examples are given in (73):

<table>
<thead>
<tr>
<th>Ma’di (Blackings &amp; Fabb 2003: 125)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Form</td>
<td>Adjectival meaning</td>
<td>Adverbial meaning</td>
</tr>
<tr>
<td>pélére</td>
<td>clean</td>
<td>properly</td>
</tr>
<tr>
<td>lɔsɔ</td>
<td>good</td>
<td>well</td>
</tr>
<tr>
<td>riri</td>
<td>quick</td>
<td>quickly</td>
</tr>
<tr>
<td>tʃe tʃe</td>
<td>slow</td>
<td>slowly</td>
</tr>
</tbody>
</table>

At this point, it may be worthwhile to consider, for the sake of contrast, some examples of languages that may seem to display noun-adjective flexibility, while in fact the absolutive use of adjectives requires the presence of an understood head in the discourse context. In these languages, evidence for such an understood head comes from the fact that it triggers agreement on the absolutive modifier. In Chapter 2, I already gave an example of this phenomenon from Spanish. Abkhaz is a similar case. Chirikba (2003: 29) remarks that in this language it may be difficult to distinguish adjectives from nouns, when the latter are used as the sole element of a referential phrase.

84 It is not clear whether these lexemes would remain property-denoting (i.e. become abstract nouns such as ‘cleanliness’) or would rather be interpreted as an object/individual with the property denoted by the base form (‘clean one’) when used as the head of a referential phrase.
phrase and combine with determiners and number markers. According to Hewitt (1979: 226) however, the morphological form of an absolutive adjective is determined by the unexpressed noun to which it functions as an attribute. Also in Basque, absolutive use of adjectives is “exceedingly common”. However, the animacy marking on the adjective is dependent on the animacy status of its understood head. For instance beltza ‘the black one’ takes animate morphology when the understood head is a horse, but inanimate morphology when it is a skirt (Hualde & Ortiz de Urbina 2003: 149–150).

5.6.3.2 Modifiers
I will now discuss what can be regarded as the most marginal form of lexical flexibility, involving the two modifier functions. As already mentioned, there are some languages in which flexible modifiers constitute quite large, open classes, while other languages rather have small or derived modifier classes.

As a first example, consider Lango, a language in which modification in predicate and referential phrases is expressed in the same way: with the linking particle à and a member of the class of flexible modifiers (Noonan 1992: 181). Notably, this class of modifiers exists alongside a class of rigid manner adverbs, which take the particle nǐ, as illustrated in (74) below. Noonan (1992: 181) explicitly mentions the difference between flexible modifiers and rigid manner adverbs: “Unlike type 1 [i.e. modifiers, EvL], these forms [i.e. manner adverbs, EvL] can never have another grammatical function”.

Lango (Noonan 1992: 181)

(74)  ngámô        cèm nì mrwɔk–mrwɔk
   3SG.SUBJ.chew.HAB   food   PRT   noisily
   ‘He chews (his) food noisily.’

85 The derived class of modifiers in Turkish had been mentioned several times already. It is interesting to note that, although -CA derivations can be used as modifiers in both predicative and referential phrases, they are always event-oriented (also when used attributively). Attributive use is only possible with nominal bases (notional nouns or nominalizations) that describe a result, or can at least take a resultative interpretation. The attributively used -CA form expresses the manner in which the action has been performed, and stands in opposition with an non-derived attribute, which simply predicates a property of the resultative noun, without specifying the orientation of this property:

Akıllı-ca birplan     Akıllı birplan
smart–MOD a.plan     smart a.plan
‘A smartly made plan.’  ‘a smart plan.’ (Schroeder 2004: 201)
Also in Ket, most modifying lexemes belong to the same flexible class. However, there are a few lexemes that can be used as adjectives only, and some others which are used exclusively as manner adverbs (Werner 1997: 119-120; Vaida 2004: 40).

In Itelmen, modifying lexemes can be divided into a morphologically regular class and a morphologically irregular class (Georg & Volodin 1999: 106). The members of the former group have different forms when they are used as modifiers in referential versus predicate phrases: In the former function they take the suffix -lah, and in the latter the suffix -q. This is illustrated in (75a) and (75b), respectively:

**Itelmen** (Georg & Volodin 1999: 111, 112)

(75)  

(a. K-k'ol-knen newen as-lah nejg-anke.  
inf.iii>come<inf.iii dem high-adj mountain-dat  
‘He came to a high mountain.’

(b. A t'salaj iseŋ-sal te-s-kinen  
but fox quiet-adv follow-prs-3.sg:3.sg.i  
‘But the fox was following him quietly.’

However, Itelmen also has a small, closed group of lexemes that are not morphologically recognizable as adjective or adverb: They have the same form in both functions. This is illustrated in (76a-b) with the lexeme miça ‘beautiful’:

**Itelmen** (Georg & Volodin 1999: 119)

(76)  

(a. Knino-kit miça mimsx t-onnkso-çen.  
2.sg-caus beautiful woman 1.sg-let.go-3.sg.pat  
‘Because of you I have let go of a beautiful woman.’

(b. E, çasit miça sun-s-ç  
interj now beautifully live-prs-2.sg  
‘Oh well, now he is doing well.’

Finally, Thai also has a small class of flexible modifiers, existing alongside rigid adjectives and rigid (derived) manner adverbs. An example of a flexible modifier is dii ‘good, well’, which can be used as an modifier in a referential phrase, as in khon dii ‘good person’, and as a modifier in a predicate phrase,
as in *tham dii* ‘do well’. (Iwasaki & Ingkaphirom 2005: 92). A similar system attested in Abun has already been discussed in section 5.2.2.

### 5.6.4 Summary

In this section I presented data in favour of lexical flexibility in the relevant languages of the sample. First, I discussed languages with contentives in terms of Evans and Osada’s three criteria for flexibility. In line with the discussion in Chapter 2 (section 2.5), I have proposed that the crucial characteristic of a ‘truly’ flexible language lies in the separation of lexical and syntactic categorization. This allows for the use of basically all content material in all syntactic functions, with semantically compositional interpretations. The only restriction on this flexibility concerns differences in frequency of use, which can be explained in terms of relative semantic-pragmatic markedness rather than absolute grammatical constraints.

Second, I discussed types of lexical flexibility with respect to the functions of head and modifier in a referential phrase, and modifier in a predicate phrase. In general, it seems that there has been less debate in the literature – both theoretical and typological – regarding these types of lexical flexibility. As a result, the kind of argumentation that should support the identification of non-verbs, nominals, and modifiers in actual languages is less developed, and sources provide less data that are relevant to the issue. This also has repercussions for the thoroughness of the discussion of these issues in the above section.

### 5.7 Conclusion, outlook

In this chapter I have presented the data on PoS classification in the languages of the sample, and evaluated them in light of the predictions and theoretical issues of Chapter 2.

In the next chapter I turn to the classification of the dependent clause constructions in the sample languages, within the functional-typological framework developed in Chapter 3.
6.1 Introduction
This chapter presents a classification of the DC constructions attested in the languages of the sample, in terms of the typology developed in Chapter 3. First, in section 6.2, the DC constructions of each language are identified on the basis of their structural coding, i.e. the form of either the subordinating conjunction or the special marker on the dependent predicate. For every DC defined in this way I determine which propositional function(s) it can express. This yields a classification of DCs in terms of several rigid and flexible types. These attested types are compared with the DC types predicted in Chapter 3 (section 3.2.3). Second, in section 6.3, the DCs identified in section 6.2 are classified according to their internal morpho-syntactic properties, i.e. in terms of the (non-)expression of verbal and nominal features and argument encoding. On the basis of these data, every DC construction is assigned to one of the three structural DC types defined in Chapter 3 (section 3.4). Finally, in section 6.4, the data presented in 6.2 and 6.3 are integrated into a single typology. Section 6.5 rounds off with a brief summery.

6.2 DC constructions and propositional functions

6.2.1 Introduction
This section is organized as follows: in 6.2.2 I start out with a preliminary discussion, in order to demarcate the type of data that are (and are not)
included in the analysis. Following this discussion, all relevant DC types are presented, together with the propositional functions that they can express. Subsequently, Section 6.2.3 compares the attested DC types with those predicted in Chapter 3. Section 6.2.4 is a summary.

6.2.2 Rigid and Flexible DCs: Attested types

6.2.2.1 Preliminary considerations
Before presenting the DCs attested in the languages of the sample and their distributional patterns, a few remarks are required. They concern three issues:

(i) The treatment of DCs that are normally used as heads or modifiers in referential phrases, but also occur within larger constructions that function in turn as predicate modifiers;
(ii) The definition of adverbial manner clauses, as opposed to other types of adverbial clauses;
(iii) The distinction between clausal and lexical derivation (see also Chapter 5, section 5.4.3.2).

I will discuss these issues in turn.

Nominal/adjectival DCs as predicate modifiers

In a number of languages, DCs that are normally used as heads and/or modifiers in referential phrases (i.e. as complement clauses, relative clauses, or nominal clauses) also appear within adverbial clauses, specifically in one of the following construction types:

(i) Marked for oblique case;
(ii) Dependent upon an adposition;
(iii) As a modifier of a nominal head that means something like ‘manner’ or ‘way’.

Such cases are not regarded as instances of flexibility, since they involve DCs that still function as a referential heads or modifiers, only within larger constructions that function in turn as a predicate modifiers.
Consider for instance examples (1) and (2) from Tamil. In (1), the clausal nominalization enka ammaa kaapi pooṭar-atu ‘our mother making coffee’ functions as the head of a referential phrase within the postpositional phrase headed by poola ‘like’. This postpositional phrase has the function of a predicate modifier. In (2) the same type of nominalization construction biicile maṇalviṭṭu kattar-atu ‘building a sandhouse on the beach’ functions referentially within a case-marked locative phrase.

Tamil (Asher 1982: 48, 21)

(1) \[\begin{array}{ll}
\text{enka ammaa kaapi pooṭar-atu} & \text{poola-vee} \\
\text{our} & \text{mother} \\
\text{coffee} & \text{put:PRS-NMLZ like-EMPH} \\
en & \text{my} \\
\text{mandevi} & \text{put-PRS-3SG.FEM} \\
\text{pooṭaraa} & \\
\text{wife} & \text{make coffee just as our mother makes it.} \\
\end{array}\]

(2) Kozantenkā \[\begin{array}{ll}
\text{biic-ile maṇalviṭṭu kattar-atu} & \text{ile} \\
\text{child-PL} & \text{beach-LOC} \\
\text{sand-house} & \text{build:PRS-NMLZ-LOC} \\
curucuruppaa iruntaanka & \\
\text{busy} & \text{be:PST.3PL} \\
\text{The children were busy building a sand-house on the beach.’} \\
\end{array}\]

It is not always obvious how to distinguish between, on the one hand, a nominalization with an oblique case marker or an adposition, and, on the other hand, a dedicated adverbial clause construction\(^6\). The same holds for the difference between relative clauses modifying a nominal head and dedicated adverbial clause constructions.

As regards the former case, in order to decide whether something is an oblique case marker or a (bound) adverbial marker, it should be determined whether or not the morpheme in question is part of a larger case marking paradigm. The locative marker in the Tamil example in (2) above, for instance, can be identified as a case marker because it is part of a paradigm of eight other case forms (Asher 1982: 103). For adpositional constructions, the relevant question is whether the marker is restricted to coding adverbial manner clauses, or whether it can also be used productively with other construction types, in particular with lexical nouns or NPs. The postposition

\(^6\) Adpositional or oblique case-forms of nominalizations can become independent from their original paradigm. This diachronic development is a common scenario for the genesis of rigid converb constructions (Haspelmath 1995: 17).
poola in (1) above does occur with ‘regular’ NPs and is therefore not analyzed as (part of) a dedicated adverbial conjunction (Asher 1982: 104).

In cases involving DC constructions that look like relative clauses, it must be determined whether or not the extra coding element is a lexical (nominal) head. Consider for instance Abun. In this language, adverbial manner clauses are marked by *sa gato*, as illustrated in (3). The second part of this complex marker (*gato*) has the same form as the relativizer, as can be seen in example (4). However, the first part (*sa*) is a conjunction rather than a lexical noun (Berry & Berry 1999: 158). Therefore, *sa gato* is analyzed as a specialized complex marker of rigid adverbial manner clauses, rather than as an instance of a relative clause construction modifying a nominal head *sa*.

*Abun* (Berry & Berry 1999: 158, 146):

(3)  
\[ An\ da\ ben\ mó\ sarewo\ an\ yo\ ben\  
\[ 3S\ sg\ actual\ do\ exist\ however\ 3S\ sg\ neg\ do\  
\[ kete\ bado\ yo\ teker\ [sa\ gato\ nyim\  
\[ too.much\ maybe\ neg\ too.much\ ADV\ earlier\  
\[ ne\ nde\ re].\  
\]  
\[ DET\ neg\ PFV\  
\]  
‘Although she does [these things] she does not do [them] very much, I mean, not like [she did them] before.’

(4)  
\[ Men\ mu\ gu\ ye\ [gato\ man\ siri\ su\ men\  
\[ 1Pl\ go\ kill\ person\ REL\ do\ wrong\ with\ 1Pl\  
\[ bi\ nggon].\  
\[ poss\ woman\  
\]  
‘We will go and kill the person who committed adultery with our (clans) woman.’

Basque displays a similar construction, but here the choice between a relative and an adverbial construction is somewhat less straightforward. Consider examples (5a-b); they show a DC construction marked with the subordinating conjunction *(e)n, which can be used as a complement clause (5a) and as a relative clause (5b). In example (6) the same DC appears within a larger construction, which involves the extra element *bezala* ‘like’ and functions as a predicate modifier:
Basque (Hualde & Ortiz de Urbina 2003: 646, 764, 722):

(5)  a. Entzun dut [Amaiaren neba hil d-en]-a
    hear aux Amaia.gen brother die aux-comp-det
    ‘I heard that Amaia’s brother died.’

    b. [Pellok ekarri du-en] dirua galdut dut
    Peter.erg bring aux-rel money.det lose aux
    ‘I lost the money that Peter brought.’

(6)  [Zeuk esan didaz-en bezala] egin dut lana
    you.emph say aux-adv do aux job
    ‘I did my job the way you told me.’

Adverbial constructions of the type illustrated in (6) also occur with other forms than *bezala*, to express different meanings such as ‘until’ and ‘though’. According to Hualde and Ortiz de Urbina (2003: 721), these other forms are either “true postpositions” or “lexicalized inflected nouns”. These analyses would characterize the construction in (6) as either a complement clause depending on a postposition, or a relative clause modifying a nominal head *bezala*, but not as a specialized adverbial clause construction. However, it seems that *bezala* occurs exclusively with clausal complements. For comparative constructions with simple nouns different (although related) forms are used, such as *bezain*, as shown in (7):

Basque (Hualde & Ortiz de Urbina 2003: 830)

(7)  Koldo bezain azkarra zara
    Koldo as.much.as fast.det are
    ‘You are as fast as Koldo.’

Moreover, when *bezala* does occur with a simple noun or pronoun, it must take additional structural coding in the form of the attributive suffix *ko*:

Basque (Hualde & Ortiz de Urbina 2003: 833)

(8)  Zü bezala-ko pollitararik
    you like-attr beautiful.abl
    ‘from [someone] as beautiful as you’
These distributional facts argue in favour of analyzing the construction [clause + -en bezala] as a dedicated adverbial manner construction.

Another Basque DC is marked by -ela, again a bound conjunction. This DC can be used as a complement clause, as in example (9), and without further measures as an adverbial manner clause, as illustrated in (10). Thus, this case, unlike the previous ones, is indeed an instance of flexibility. Note however, that there is a dialectal variant of (10), which involves the partitive case marker -rik, as shown in (11), and as such does not involve flexibility. Rather, it is a rigid complement clause construction, which takes oblique case when expressing an adjunct.

**Basque** (Hualde & Ortiz de Urbina 2003: 635, 712, 713)

(9) *batzuek uste dute [bauk oro kazeten eta some.erg think aux these all journals.gen and kazeta-egileen egitekoak dir-ela] journal-makers.gen duties are-comp*  
Some think that all these are duties of journals and journalists.’

(10) *[[Zer egin ez neki-ela] geratu nintzen what do not knew-adv stay aux*  
‘I stood there not knowing what to do.’

(11) *[[Jaikitzen da, jauzi egiten du-ela]-rik]*  
rise.ipfv aux, jump do.ipf aux-comp-part*  
‘(S)he gets up, jumping.’

In short, there are languages that express adverbial clauses by means of a larger construction, in which a DC functions as a referential head or modifier. Since these cases are not analyzed as instances of flexibility, there is no reflection of the ‘adverbial use’ of these DCs in Table 6.1 below.

Finally, some languages express adverbial manner clauses through reduplication of an infinitival construction. In my sample this strategy is attested in Kharia and Imbabura Quechua. Example (12a) illustrates the use of the Kharia infinitival construction as a complement clause (without reduplication), and (12b) shows the same construction, with reduplication,

---

87 Nevertheless, I have included this information in Appendix III, where I mention, for instance, that the Tamil *atu-*nominalization construction (see (1)-(2) above) can function adverbially in combination with a postposition or an oblique case marker.
used as an adverbial clause. A similar pair of examples from Imbabura Quechua appears in (13a-b). The reduplication is regarded as an iconically motivated formal reflection of the fact that the SoA denoted by the DCs stretches out temporally over the SoA denoted by the matrix clause, rather than as additional structural coding. This means that the examples in (12a-b) and (13a-b) are analyzed as instances of flexibility.

Kharia (Peterson 2006: 259, 249)
(12)  a.  [ip u ikuʔq sundar kontheʔq=ki=te
1sg this very beautiful bird=pl=obli
satay=na] um=ip lam=te
torment=inf neg=1sg want-act.prs
'I don’t want to torment these beautiful birds.'

b.  …lekin [lam=na lamna] souʔb=te ikuq jughay
but search=inf rdp all=obli very much
daʔ-piyas laʔ=ki.
water-thirst emot=m.pst
‘But searching and searching, [they] all became very thirsty.’

Imbabura Quechua (Cole 1982: 40, 62)
(13)  a.  ñuka-ka [shuj ali wagra-ta-mi randiʔ]-ta muna-ni
1-top one good cow-acc-va infl buy-inf-acc want-1sg
‘I want to buy a good cow.’

b.  [Kandaʔ y kanday] shamu-rka-ni
sing-inf rdp come-pst-1
‘I came singing.’

Delineating manner clauses

A number of the DCs that are listed in Table 6.1 as adverbial manner clauses also have ‘non-manner’ adverbial uses: They are used for instance as purpose clauses or as cause clauses. Since this study is restricted to adverbial manner clauses, such non-manner adverbial functions are not taken into account. For the same reason, rigid DCs that express non-manner adverbial functions only are not included in Table 6.1.
However, in many cases it is not straightforward to determine what constitutes a ‘true’ manner clause, as opposed to other adverbial clause types with related semantics. In a number of typological studies sets of adverbial meanings have been proposed, sometimes in the form of semantic maps. In such studies, the following types of adverbial meanings are typically mentioned in close relation to manner: simultaneity, instrumentality, accompanying circumstance, and comparison (Konig 1995, Kortmann 1998, Hengeveld 1998, Himmelmann & Schulze-Berndt 2005, Thompson et al. 2007). In the present study, I make an (admittedly fairly rough) distinction between two major types of adverbial clauses that are semantically close to manner clauses. I will call these similative clauses and simultaneity clauses, respectively. The former type has comparative meaning; it expresses similarity between the SoA of the dependent clause with some other action or psychological state expressed by the matrix clause. An example of a similative construction is given in (14):

\[
(14) \quad Ab, ni-pa:xa:lua \quad [ke:n-aken \quad taba \quad ti-ki:s-tuk \quad ti-pa:xa:lua]
\]

Oh I-walk just.as you you-leave-pfv you-walk

‘Oh, I’m taking a walk just like you have come out to take a walk.’

Simultaneity clauses are strictly speaking temporal clauses; they describe the two linked SoAs as temporally overlapping. As such, ‘simultaneity’ is meant here to include ‘accompanying circumstance’. In many languages, simultaneity clauses are used by semantic extension to express that the dependent SoA further specifies the manner in which the action expressed by the matrix clause is carried out. An example of a simultaneity clause in Babungo is given in (15):  

\[
(15) \quad gwöö \quad bëwëy \quad [ki \quad shöö \quad gwööa]
\]

he sleep-impf sim rest-ipfv body

‘He was asleep, resting his body.’

88 This is not meant to be an exhaustive list. The terms largely speak for themselves; the exact definitions proposed for each of them differ from one author to the next, but they are not crucial for the present discussion.

89 Instrumental clauses will be regarded as a sub-type of adverbial manner clauses.
Similative and simultaneity clauses are included in the data whenever I could not find evidence for a DC construction with ‘true’ manner semantics, and whenever the boundary between manner clauses and manner-like clauses was not clear-cut\(^\text{90}\).

**Clausal versus lexical derivation**

It can be problematic to decide whether a construction is truly an instance of clausal derivation, as opposed to lexical derivation. Consider for instance the Ket infinitive and the Basque perfective participle constructions, illustrated in (16) and (17), respectively. Since in these constructions no overt arguments are expressed, it is not clear whether they should be analyzed as relative clauses or rather as lexical adjectives derived from verbs.

**Ket** (Vajda 2004: 79)

(16) \(b\,i\,y\quad s\,\ddot{a}g\,\ddot{d}i\)

\(*\text{find.INF} \quad \text{boot}*

‘A boot that is found’

**Basque** (Hualde & Ortiz de Urbina 2003: 199)

(17) \(g\,i\,z\,o\,n\quad i\,k\,a\,s-i-a\)

\(*\text{man} \quad \text{learn-PFV.PTC-DET} \quad \text{a learned man.}*

In the cases of Ket infinitives and Basque perfective participles, the distinction between lexical versus clausal derivation is directly relevant to the question whether these DCs are flexible or rigid, since both constructions can also be used as complement clauses, in which function they do take overt argument(s). This is shown in examples (18) and (19)\(^\text{90}\):

**Ket** (Vajda 2004: 78):

(18) \([s\,k-n\,\ddot{a} \quad q\,\ddot{u}\,s-b\,\ddot{e}t] \quad b\,\ddot{i}n\,\ddot{u}\,\ddot{t} \quad \text{2PL-ANIM.PL.GEN} \quad \text{tent-making.INF} \quad \text{it.ended} \)

‘We finished making the tent.’ (Lit: ‘Our tent-making ended.’)

\(^{90}\) In Appendix III, a remark on semantics is added to similative and simultaneity adverbial clauses.
Basque (Hualde & Ortiz de Urbina 2003: 668):

(19) Damu dut [zu irain-du]-a
    regret aux you offend-PFV.NMLZ-DAT
    ‘I regret having offended you.’

Of course, it is quite possible that the same suffix is used for both a lexical and a clausal derivational process. This is even to be expected from a diachronic point of view. However, this does not solve the problem of determining the synchronic status of the relevant suffix. For the Basque case there is some evidence in favour of a lexical analysis for (17): In this language, relative clauses precede the head noun, while adjectives follow it (Hualde & Ortiz de Urbina 2003: 791). Since the ‘participle’ in (17) is postnominal, this can be taken as evidence that it is a derived adjective, rather than a relative clause. Consequently, the perfective nominalization in (19) is analyzed as a rigid DC, which can only be used in the function of head of a referential phrase. In Ket, however, there is no evidence against analyzing the construction in (16) as a relative clause. Therefore, the Ket infinitive construction is analyzed as a flexible DC.

6.2.2.2 Language data

Table 6.1 displays the set of relevant DC constructions in the language of the sample, and their possibilities to express each of the four propositional functions. The languages are listed in the leftmost column, in the same order as in Table 5.1 of the previous chapter. The DC constructions appear in separate rows in the second column and are identified by their structural coding. The four remaining columns represent the four propositional functions of DCs: Head of a predicate phrase (Pred Head), head of a referential phrase (Ref Head), modifier in a referential phrase (Ref Mod), and modifier in a predicate phrase (Pred Mod). A plus sign in any of these four columns indicates that the relevant DC can express that particular propositional function. An empty cell indicates that the DC cannot express

91 The fact that in the Ket example both arguments have ALT coding still makes it difficult to decide whether this is a deranked DC or a regular derived NP. In Malchukov (2006: 989), the Ket infinitive construction is characterized as exemplifying the final stage of de-verbalization, since it not only loses the inflectional slots of polysynthetic independent verb forms, but cannot take derivational categories pertaining to valency and aspect (causative, resultative etc.) either.

92 In Appendix III, all DC constructions are listed and illustrated in the same order in which they appear in Table 6.1. In this Appendix, the DCs are also identifiable through their structural coding.
the propositional function in question. In a few cases the evidence was not conclusive, which is indicated with question marks in the relevant cells.

Table 6.1: DCs in the sample languages and their possibilities concerning the expression of propositional functions

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Propositional Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Tagalog</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pag-</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>na/-ng</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kharia</td>
<td>RDP/Ø</td>
<td>+</td>
</tr>
<tr>
<td>-na</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-na-wala</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-al</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-ker etc.</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-ga + RDP</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-ta + RDP</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-na-wala</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>gam-kon</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>je etc.</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>a/i etc.</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Ø</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Kambera</td>
<td>pa-</td>
<td>+</td>
</tr>
<tr>
<td>ma-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Ø (NMLZ)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>wà</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>=ga</td>
<td>+</td>
</tr>
<tr>
<td>Ø (NMLZ)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>ona/ina</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-e</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Ø</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Guarani</td>
<td>hä(gwe)/Ø</td>
<td>+</td>
</tr>
<tr>
<td>va</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>vo</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Santali</td>
<td>Ø (no INDIC)</td>
<td>+</td>
</tr>
<tr>
<td>-kate</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>CORREL</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>mente</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Propositional Functions</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Warao</td>
<td>-kitane</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>kotai</td>
<td>+</td>
</tr>
<tr>
<td>Turkish</td>
<td>-DIK/-{(y)AcAK</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-mAK</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-mA</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-An</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(y)ArAk</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(y)A...-(y)A</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ki</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>diye</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ø</td>
<td>+</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-n-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-Thirri-n</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-n-garrba</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ntha</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ø</td>
<td>+</td>
</tr>
<tr>
<td>Paiwan</td>
<td>tu(a)/tjai</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a parhu</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(j)i...-(j)an + a</td>
<td>+</td>
</tr>
<tr>
<td>l. Quechua</td>
<td>-j/-shka/-na</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ngapaj</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-chun</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-y</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-shpa</td>
<td>+</td>
</tr>
<tr>
<td>Ma’di</td>
<td>-lɛ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-dʒo</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ka</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-rɛ/-bá</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-zi+sí</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Propositional Functions</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>-woo</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-wadda</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-mawoo</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-bari</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ya/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-gowaaya</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (+RSP)</td>
<td>+</td>
</tr>
<tr>
<td>Hungarian</td>
<td>-nî</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ds/-és</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ó</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ótt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-andó/-endó</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-vá/-vé</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-vén</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(azt) hogy (etc.)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>úgy (a-)hogy</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td>Japanese</td>
<td>no/mono etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (RSP)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-te/-de/-ite</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-i/-Ø</td>
<td></td>
</tr>
<tr>
<td>Hmong</td>
<td>qhov</td>
<td>+</td>
</tr>
<tr>
<td>Njua</td>
<td>kuam/(has)tas</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>kws</td>
<td></td>
</tr>
<tr>
<td>Lango</td>
<td>-(kk)ɔ̀</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>nî</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>à-mè</td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td>Ø (bare INF)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ëta qor’a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ásqà</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-PROSEC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-s/-bes</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Propositional Functions</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Itelmen</td>
<td>INF (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>quätz</td>
<td>+</td>
</tr>
<tr>
<td>Koasati</td>
<td>NMLZ (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-:sáya</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-:yólli</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-:ka</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-:kíttta</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-laho:li:sáya</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-n</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-k</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-t</td>
<td>+</td>
</tr>
<tr>
<td>Thai</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>thíii</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>wáa</td>
<td>+</td>
</tr>
<tr>
<td>Basque</td>
<td>-t(z)e</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-tu/-du/-i/-Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-en</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-en bezala</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(e)la</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>bait</td>
<td>+</td>
</tr>
<tr>
<td>Abun</td>
<td>do/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>gato</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>sa gato</td>
<td>+</td>
</tr>
<tr>
<td>Bambara</td>
<td>ka</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-le/-ne</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-min(u)/ -mun(u)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-tò</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Propositional Functions</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Georgian</td>
<td>-a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>m- (-a-)(-el/ -al))</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ul/-il/m--ar/ -al</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>sa- (-el/-al/r)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rom</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ra</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rogorc</td>
<td></td>
</tr>
<tr>
<td>Bukiyp</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(ú)ji</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bwidouk -(u)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mu</td>
<td></td>
</tr>
<tr>
<td>Abkhaz</td>
<td>-ra</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-NFIN</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-NFIN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ş- + -NFIN</td>
<td></td>
</tr>
<tr>
<td>Polish</td>
<td>-nie</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>INF (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-c</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-any etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(PST.PASS) PTC (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>PRS.PL-c-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>že</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>jak (gyd)by</td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td>-(á)as</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-im/-um/-am</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>n-STEM-(a)n</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ke/ki</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>sén-/ét-</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Propositional Functions</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pred Head</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>-e/-i</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>PERSON.SUFF</td>
<td>+</td>
</tr>
<tr>
<td>Alamblak</td>
<td>-nef</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(kfë)t</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ind</td>
<td>+</td>
</tr>
<tr>
<td>Pipil</td>
<td>ka(h)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ne</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ke</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ke:n-aken</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ADV</td>
<td>+</td>
</tr>
<tr>
<td>Wambon</td>
<td>-e</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-a + o</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-mo/-o</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ka</td>
<td>+</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>DET(+DEM)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-n/-an</td>
<td>+</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>fu/fi/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>bifi/dati/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Wh/Ø</td>
<td></td>
</tr>
<tr>
<td>Babungo</td>
<td>lāa</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>fāŋ/yūu</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ki(i)/Ø</td>
<td></td>
</tr>
<tr>
<td>Nama</td>
<td>!xāis-ā</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(no INDIC)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (no INDIC)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(RSP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-se/l’aa/tsii/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(no INDIC)</td>
<td></td>
</tr>
<tr>
<td>Hdi</td>
<td>ká</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>tá</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>tà</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>tà + NMLZ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Propositional Functions</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Mandarin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ø</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>de</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-zhe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-atu</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-(kk)+a</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kisi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(m)àà/Ø</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Ø</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>CL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nung</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ø</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>(tì-va) (+DEM) (+FOC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bät</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in-e</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-a</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-a-ni</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-na</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-kan/-kan-a</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-gip-a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-e/-e-min/-e-r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krongo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ânî tînj</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-(t)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>m-/n-/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hixkaryana</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ni/-thi/-</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-hito + ri</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slave</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ni/Ø</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>gha/gú</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>i/sji/lli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gháré</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nivkh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-vut/-vur</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Ø</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>NFIN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r/-t/-i/-n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Propositional Functions</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Greenlandic</td>
<td>-niq</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ta/-sa</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>PTC mood</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>CONT mood</td>
<td>+</td>
</tr>
</tbody>
</table>

6.2.3 Discussion

6.2.3.1 Attested and predicted types

The data in Table 6.1 show that most of the attested DCs represent predicted types. First, as far as rigid DCs are concerned, there are complement clauses, relative clauses, and adverbial manner clauses. The functional distribution of these DC types is represented in (19), (20), and (21), respectively (cf. (9), (10) and (11) in Chapter 3):

(19)

<table>
<thead>
<tr>
<th>Predication</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Complement clause</td>
</tr>
</tbody>
</table>

(20)

<table>
<thead>
<tr>
<th>Predication</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Relative clause</td>
</tr>
</tbody>
</table>

(21)

<table>
<thead>
<tr>
<th>Predication</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Manner adverbial clause</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

I will illustrate each of these rigid DC types with examples from the sample languages. Example (22) is a rigid complement clause from Kisi; this construction can be used in the function of head of a referential phrase only.
Kisi (Childs 1995: 280)
(22)  đò  cà  kò  bèê  [màà  té  đà รง  pé  đù  là  len]  
   he  see  only  indeed  COMP  if  he  cry  if  he  go.out  inside  
   ‘He really thinks that if he cries he will be free.’

Note that the behavioural potential of DCs does not play a role in the 
typology yet. Thus, rigid deranked nominalization constructions, such as in 
(23) from Turkish, also count as complement clauses.

Turkish (Göksel & Kerslake 2005: 420)
(23)  [Kerkes-in  birier  kikaye  anlat-ma-sî]  
   everyone-gen  one.each  story  tell-NMLZ-3SG.POS  
   iste-n-iyor-mus  
   want-PASS-IPFV-EV.COP  
   ‘It seems they want everyone to tell a story.’

Example (24) is a rigid relative clause construction from Hmong Njua:

Hmong Njua (Harriehausen 1990: 141)
(24)  Tug  txiv  neej  [kws  kuv  saib]  sab  
   CL  man  REL  1SG  see  big  
   ‘The man that I saw was tall.’

Relative clauses may also involve deranked participle constructions, as in 
example (25) from Hungarian:

Hungarian (Kenesei et al. 1998: 45)
(25)  [A  könyv-et  a  fiù-nak  gyorsan  olvas-b]  
   the  book-ACC  the  boy-DAT  fast  read-PRP.PRS.PTC  
   lány  itt  van  
   girl  here  is  
   ‘Here is the girl who reads the book to the boy fast.’

A rigid adverbial manner clause construction appears in (26), from Guaraní.

Guaraní (Gregores & Suárez 1967: 180)
(26)  H-asê  [o-sê  vo]  
   she-cry  she-go.out  ADV  
   ‘She goes out crying.’
Rigid adverbial manner clauses may also be deranked converbal constructions, such as the one in (27) from Polish:

*Polish* (Bielec 1998: 71)

(27)  
\[ \text{Słuchaj-conv music dress.I (my)self} \]

‘Listening to music, I got dressed.’

Turning to flexible DC types, the following predicted types are attested: multi-functional clauses, modifier clauses, nominal clauses, and flexible clauses of type C (which can be used as complement clauses and adverbial manner clauses). The functional distributions of these DC types are represented in (28), (29), (30), and (31), respectively (cf. (13), (14), (15) and (20) in Chapter 3):

(28)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Predication</th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Multi-functional clause</td>
<td></td>
</tr>
</tbody>
</table>

(29)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Predication</th>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Modifier clause</td>
</tr>
</tbody>
</table>
Below I will again illustrate each of these DC types with examples from the sample languages. Multifunctional clauses are attested in the isolating languages Mandarin Chinese and Nung. In both cases the DC is zero-marked. Examples (32a-c) are from Mandarin, and show that the DC can be used in the functions of head of a referential phrase (32a), modifier in a referential phrase (32b), and modifier in a predicate phrase (32c)\(^9\).

**Mandarin Chinese** (Li & Thompson 1981: 599, 611, 597)

(32) a. ˇ\text{Wǒ pàn wàng } [ˇ\text{nǐ kuài yìdiǎn biyè}]
   
   I hope you soon a.little graduate
   
   ‘I hope you’ll graduate a bit sooner.’

b. ˇ\text{tā yǒu yì-ge móimei } [ˇ\text{běn xǐhuān kàn diànyìng}]
   
   3sg exist one-cl younger.sister very like see movie
   
   ‘S/he has a younger sister who like to see movies.’

c. ˇ\text{Tāmén } [ˇ\text{yòng shǒu } chī-fàn]
   
   they use hand eat-food
   
   ‘They eat food using their hands.’

Deranked multi-functional clauses are attested in Kayardild. So-called ‘nominalizations’ in this language can be used as the head and modifier of a referential phrase, and as a modifier in a predicate phrase, as is illustrated in (33a-c) below. Note that the structural coding of the multi-functional clause,

\(^9\) Burushaski appears to have an overtly marked, balanced multi-functional DC, borrowed from Persian: This *ke/ki* construction is claimed to be usable as a complement, relative and adverbial clause, but my sources provide no clear examples (see Tikkanen 1995: 498).
the -n- marker, is glossed according to the function of the DC in the relevant example. Thus, in (33a) -n- is glossed as a nominalizer, in (33b) as a participle, and in (33c) as a converb. A similar procedure is followed for the glossing of structural coding in other types of flexible clauses, including balanced clauses. In the latter, the subordinating conjunction will be glossed as a complementizer (COMP), a relativizer (REL), or an adverbial subordinator (ADV), according to the function performed by the DC in the sentence (cf. for instance examples (9) and (10) above). This is to avoid tedious terminology such as ‘dependent predicate marker of a multi-functional balanced clause construction’, and opaque corresponding glosses. Note however, that the disadvantage of this procedure is that the gloss of an isolated example does not reveal whether the DC is a rigid or a flexible construction.

Kayardild (Evans 1995: 476, 474,
(33) a. *Ngada kurri-ja [niwan-ji budii-n-marri]*
     1sg.nom see-act 3sg.poss-mloc fun-nmlz-priv
     ‘I saw that he was not running.’

     b. *Nga-ku-l-da [wirr-n-ku] dangka-wu kurri-ju*
     1-inc-pl-nom dance-ptc-mprop man-mprop see-pot
     ‘We will watch the dancing man.’

     c. *[Bilaangka-nurru kari-i-n-da] ngada warra-j*
     blanket-assoc cover-m-conv-nom 1sg.nom go-act
     ‘I went along, covering myself in a blanket.’

Flexible modifier clauses are attested in Babungo, Mandarin Chinese, and Krongo. The balanced modifier construction in Babungo is illustrated below. In (34a) the DC functions as a modifier in a referential phrase, and in (34b) as a modifier in a predicate phrase.

Babungo (Schaub: 1985: 32, 39)
(34) a. *mɔ ye wɔ ntiɔ [fʌŋ (ŋwɔ) shɔw ngu ye]*
     I see.pfv person that rel he steal.pfv fowl your
     ‘I have seen that man who has stolen your fowl.’
b. ŋwọ sọ sọ́y [fáŋ nshù wi nàyś tī ŋwọ́]
she plant.PFV corn ADV mother her tell.PFV to her

‘She has planted the corn like her mother told her.’

The Krongo modifier clause construction is illustrated in (35a-b). This is a deranked construction to the extent that it does not express agreement and takes a possessive subject. Tense, aspect, and valency/voice marking are retained, however.

Krongo (Reh 1985: 256, 333)
(35)  a.  n-úllà àrìàñ kí-ùt-àndìñ
1/2-IPFV:love I LOC-SG-clothes
    [n-úufò-ñ kò-nìimò kàti]
    PTC:NEUT-IPFV:SEW-TR poss-mother my

    ‘I love the dress that my mother is sewing.’

b.  n-àa t-ànkwà-ànì
    CONN-NEUT-COP INK-gO.ROUND-DTR
    [n-ùrùná-ñ úuní kàndày]
    CONV:NEUT-IPFV:PAY.ATTENTION.TO footprints poss:3.PL

    ‘She goes round, paying attention to their footprints.’

Nominal clauses, i.e. DCs that can be used as the head and the modifier in a referential phrase, occur quite frequently. An example of a balanced nominal clause construction in Basque appears in (36a-b):

Basque (Hualde & Ortiz de Urbina 2003: 646, 764)
(36)  a.  Banekien [Mikel berandu etoriko z-en]-a
    ba-knew Mikel late arrive.FUT AUX-COMP-DET

    ‘I knew that Mikel would arrive late.’

b.  [Pellók ekarri du-en] dirua galdu dut
    Peter.ERG bring AUX-REL money.DET lose AUX

    ‘I lost the money that Peter brought.’

Deranked nominal clauses are attested for instance in Turkish, as shown in example (37a-b). Note that the difference in form of the special marker on the predicate in the two clauses is due to vowel harmony.
Finally, a few cases are attested of flexible DC clauses of type C, which can be used as complement clauses and as adverbial manner clauses (see (32) above). One case is the Basque –ela construction, which was already illustrated in (9) and (10) above. Recall that this flexible construction has a dialectal variant in which the adverbial use is accompanied by an extra case-marker –rik (see (11) above).94

Another case of a flexible clause of type C is attested in West Greenlandic. It is called the contemporative or conjunctive mood construction and it is formed with the mood-marker –lu/-llu, followed by a person marker. This is illustrated in (38a-b).

94 In addition, Hualde & Ortiz de Urbina (2003: 712) quote sources claiming that the marker –ela used for adverbial manner clauses is not exactly homophonous with the complementizer –ela, since they have a different accentual pattern.
that this DC type is flexible between the head function of one domain (reference) with the modifier function of the other domain (predication). As such, it merges two functions that are maximally different in terms of both the predication-reference parameter and the head-modifier parameter\(^{95}\). This marked situation is apparently reflected by the infrequent attestation of the relevant DC type.

On the other hand, a possible connection between complement clauses and adverbial clauses lies in the fact that they function as arguments and adjuncts, respectively. Complementizers often develop out of adpositions or case markers, which may already have been used for the marking of both direct object arguments and (certain types of) adjuncts (Cristofaro 1998). This flexibility may carry over to the coding of clausal constructions expressing these two functions, i.e. to complement clauses and (some types of) adverbial clauses.

In this subsection I presented all DC types attested in the sample, according to their functional distribution. These attested types do not cover all the types predicted in Chapter 3, section 3.2.3. In the next sub-section I discuss the predicted but unattested DC types.

6.2.3.2 Predicted but not attested DC-types: The problem of predicate clauses
The following predicted DC types were not attested in any of the sample languages: predicate clauses, contentive clauses, predicative clauses, head clauses, and flexible clauses of types A, B, and D\(^{96}\). These DC types are presented in (39)-(45) below (cf. (8), (12), (16), (17), (18), (19), and (21) in Chapter 3):

\[(39)\]

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicate clause</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

\(^{95}\) cf. Chapter 5, section 4.3.4, on the non-attestation of PoS classes with this distributional pattern.
In fact, there appears to be one instance of a contentive clause construction, attested in Kharia, which I will discuss shortly.
Notably, all these DC-types involve the functional slot for head of a predicate phrase. As explained in Chapter 2 (section 2.3.1) and Chapter 5 (section 5.5), this function has a special status in Hengeveld’s theory of PoS classification. In particular, the function of head of a predicate phrase is the only function of verbs, while it is a possible additional function of lexeme classes that are not verbs. Such a non-verbal lexeme class is regarded as having an additional predicative function only when its members display the same behavioural potential as verbal predicates (i.e. when a zero-1 strategy is employed).

In order to preserve the parallel between PoS classes and DC constructions, the latter are subjected to the same criterion as the former: DC constructions (which are normally used in one or more of the functions of head and modifier in a referential phrase and modifier in a predicate phrase) are regarded as having an additional predicative function only when they employ a zero-1 strategy in that function.

Hypothetically, to use a zero-1 strategy with a DC in predicative function would mean that verbal features take scope over the clausal construction as a whole. In the case of a balanced, finite DC (with verbal morphology already present on the dependent predicate), this would involve double expression of verbal categories, i.e. a kind of ‘stacking’ of verbal inflection. In the case of a deranked clause it would mean that constructions with typically non-finite predicates would express verbal features. In both cases one could imagine that verbal features somehow appear at the periphery of the DC construction. However, none of these scenarios is actually attested in the sample languages. Rather, the use of a DC construction in predicative function triggers either a copula strategy or a zero-2 strategy. Consider for example the English infinitival construction in (46), which in predicative function needs a copula:

\[(46) \text{ My plan is [to take you on a trip to the beach]} \]

In short, DC constructions with non-predicative defining uses do not employ the zero-1 strategy when used predicatively. One reason for this may be the structural problems involved in inflecting a DC in the same way as a lexical verb. This may explain why flexible DC constructions involving the function of head of a predicate phrase (as represented in (40)-(45) above) are not attested. It may also explain the non-occurrence of DCs that are specialized for the function of head of a predicate phrase, i.e. rigid predicate clause constructions (cf. (39) above).
In addition to this formal restriction, DCs (in particular deranked ones) appear to be conceptually marked with respect to the function of predicative head: It has been claimed that he SoAs expressed by DCs lack an autonomous cognitive profile, and this arguably makes them unlikely candidates for the expression of independent predications (Langacker 1987; Cristofaro 2003, cf. Chapter 3).

In view of this last point, it is interesting to consider in more detail the contentive clause type (represented in (40) above). As noted above, this clause type, although predicted in Chapter 3, is practically non-existent in the sample languages. However, there is in fact one DC construction that comes close to this type, since it can be used in all four functions, and apparently employs a zero-1 strategy in predicative function. This is the Kharia DC construction with so-called freestanding forms. These are bare, non-finite dependent predicate forms, sometimes with stem reduplication. When a freestanding form construction is used in predicative function, it takes clitics for voice/tense and person, just like any lexical predicate. As such, this DC construction seems to meet the criterion for a zero-1 strategy. However, the use of voice marking on freestanding form constructions is restricted, compared to lexical predicates: Freestanding form clauses can only take the middle voice, and they always have a habitual interpretation. Compare the example in (47a), showing a freestanding form construction in the function of predicative head, with the example in (47b), its counterpart with a lexical predicate.

Kharia (Peterson 2006: 74)

(47) a. 1sg | qa? | bi?d-bi?d=ki=m
       water pour.out-rdp=m.pst=1sg
       ‘I used to pour water out.’ (i.e. that was my job)

        1sg | qa? | bi?d=oj
       water pour.out:act=pst.1sg
       ‘I poured water out.’

It seems that the special behaviour of freestanding form clauses in this respect is motivated in cognitive-semantic terms. As Peterson puts it:

97 They are called masdars in Peterson (forthcoming).
“I would argue that the habitual interpretation of such forms in Kharia results from the fact that [they are] intimately connected to a depiction of the event without explicit reference to its internal temporal structure. If such a form is nevertheless marked as a finite predicate, what results is a habitual situation, not an activity or event in the usual sense […].” (Peterson 2006: 74)

The case of Kharia freestanding froms may be compared to so-called gerund constructions in Tagalog. Like Kharia, Tagalog has a class of lexical contentives. Unlike Kharia, however, it has no phrasal clitics. The Tagalog gerund construction is a deranked DC that can be flexibly used in all propositional functions. However, in the function of predicative head it does not employ a zero-1 strategy: Lexical predicates in Tagalog appear in sentence-initial position and are marked for voice (Himmelmann 2007). A predicatively used gerund construction also stands in first position, but it is not marked for voice, as can be seen in example (48).

Tagalog (Himmelmann 2005: 372)

(48) [pag-\textit{lu-luto?} \textit{ng pagkain} \textit{ang trabaho niyá}]
\textit{GER-RDP-cook GEN food SPEC work 3SG.POSS}

‘His/her job is cooking food.’

Thus, the Tagalog gerund construction should be analyzed as employing a zero-2 strategy in predicative function. However, it seems that there is a functional reason for the lack of voice-marking on gerund constructions. In particular, Tagalog voice-markers orient the action denoted by the predicate towards one of the participants of that action. Gerunds, in contrast, “refer to actions or states without orienting them towards one of the participants” (Himmelmann 2005: 372). In other words, gerund marking and voice-marking are mutually exclusive in functional terms. In fact, there is a systematic co-variation between the allomorphs of voice markers and the allomorphs of the gerund marker. For instance a form taking -\textit{um}- as the active voice marker, takes \textit{pag}- as a gerund marker, whereas a form taking \textit{mag}- as the active voice marker, takes \textit{pag-} plus reduplication for gerund formation (see Schachter & Otanes 1972: 160-161 for the full paradigms).

In general, it seems that there are both structural and cognitive-semantic factors that inhibit the use of DCs as predicative heads. This explains the non-attestation of the predicted flexible DC types that involve the function
of predicative head, as well as the non-attestation of rigid predicate clauses. In addition, it means that the function of head of a predicate phrase is in practice largely irrelevant to the purpose of the present study, namely the comparison of the distributional patterns of PoS and DCs with respect to propositional functions. Whereas the function of head of a predicate phrase appears to be the most central function in terms of PoS classification, it is the most marginal one for DCs.

6.2.4 Summary
To summarize, in this section I have shown that all DC types predicted in Chapter 3 are attested in the languages of the sample, except those involving the functional slot for head of a predicate phrase. This latter finding was explained in terms of the structural incompatibility of DC constructions with the zero–1 strategy, and their cognitive-semantic markedness with respect to the function of independent predication.

6.3 DC constructions and behavioural potential

6.3.1 Introduction
This section describes the second step towards the DC typology as developed in Chapter 3, namely the categorization of every DC construction (as listed in Table 6.1 of the previous section) in terms of its internal morpho-syntactic properties. I start out with a brief discussion of the formal parameters that are taken into account. In section 6.3.2 I then present the basic data as regards the (non-)expression of verbal categories and nominal categories in the DCs, and the realization of their arguments. Subsequently, in section 6.3.3, these data are used to assign every DC to one of the three structural DC types defined in Chapter 3 (section 3.4). In section 6.3.4 the two datasets of 6.3.2 and 6.3.3 are combined: The internal formal properties of all DCs are presented in the form of three separate tables corresponding to the three structural DC types.
6.3.2 Basic Data

6.3.2.1 Formal parameters

*Verbal and nominal categories*

Data on the (non)expression of following verbal categories are included in Table 6.2:

- voice and valency marking
- tense marking
- aspect marking
- mood marking
- person marking (including agreement and cross-reference)
- illocutionary force (indicative or declarative marking)

The nominal categories included in Table 6.2 are:

- determiner expression
- case/adposition marking

Number and gender/class marking are not systematically taken into account, because these categories are very infrequently attested in the DCs of the sample languages.

For each DC it is determined which verbal and nominal categories are expressed (indicated with a ‘+’ in Table 6.2) or not expressed (indicated with a ‘–’ in Table 6.2). An empty cell means either that the category is not relevant in the language in question, or that my sources provided no information about the (non-)expression of the category. The appearance of ‘+/–’ indicates one of the following situations:

(i) The expression of the relevant category is possible but not obligatory;

---

98 Recall from Chapter 3 that agreement refers to subject agreement, whereas object agreement is documented in the column for voice-valency marking.
99 In Appendix III I do include information about these categories, whenever relevant and available. This holds also for number/gender/class agreement on relative clause constructions.
(ii) Some distinctions of the relevant category are retained while others are lost (as compared with the range of distinctions available in independent clauses);

(iii) The category is expressed in some but not all of the propositional function(s) and/or specific subordination relations that can be expressed by the DC in question.

Below I will provide examples of each of these three situations. The first one is applicable to complement clauses in Guaraní. Examples (49a-b) show that this DC construction may appear either with or without a definite article.

**Guaraní (Gregores & Suárez 1967:158)**

(49) a.  *Rey-anú* [šé še-rasɨ̀ há]
    you-hear I I-be.sick **comp**
    ‘You heard that I was sick.’

b.  *ai-kwaá la* [n o-ù moñaɨ i há]
    I-know **art** neg he-go **mod** neg **comp**
    ‘I know that he does not intend to go.’

The second situation occurs in Hixkaryana. Nominalization constructions in this language retain three out of the seven tense distinctions that can be expressed in independent clauses (Derbyshire 1979: 25).

A clear example of the situation mentioned under (iii) above is found in Santalı. In this language DCs without the indicative marker –a exhibit variation in the expression of TAM and person, depending on the propositional function in which they are used, as well as the particular semantic and syntactic subordination relation that they express. When the construction functions as a complement clause in subject function, it shows no subject marking, while middle voice markers and TAM can be expressed. In the function of object complement clauses, it lacks subject and tense marking, but object marking is retained. Only complements of perception predicates can express subject marking as well as all TAM distinctions. Finally, when the DC functions as a relative clause, subject marking is lost, while all TAM distinctions may be expressed.

Note further that in some cells of Table 6.2 a plus sign appears between brackets: ‘(+).’ When referring to a verbal feature, this indicates that the relevant category is expressed by means of special dependent forms, either
because the structural coding of the construction fuses with the expression of a TAM value, or because there is a special dependent/subjunctive paradigm for the expression of TAM and/or person marking. In Chapter 3 I gave some examples of these situations, as attested in Imbabura Quechua, Georgian, and Abkhaz (see examples (6), (7), and (47) of Chapter 3).

As regards nominal categories, a ‘(+)' appears only in the column for case. This occurs when a DC is marked for oblique case when it functions as an adjunct, while in core argument function (i.e. as a subject or object complement) and/or in the function of referential modifier (i.e. as a relative clause), the case-marker is a zero-morpheme. Of course, the presence of a zero-morpheme is assumed only when it fits into a larger paradigm of overt case markers.

*The coding of the arguments*

As explained in Chapter 3, the parameter of argument coding has three possible values:

(i) An argument is expressed as it would be in an independent clause (*SENT*);
(ii) An argument is expressed in a different way than it would be in an independent clause (*ALT*);
(iii) An argument is not expressed (*Ø*).\(^{100}\)

A number of DC constructions display some variation in the realization of their argument(s). Despite this variation, such DCs are viewed as single constructions, as long as their structural coding remains unchanged. The selection of an argument coding strategy may depend on the propositional function in which the relevant DC appears (when it is a flexible construction) and/or on the specific semantic/syntactic characteristics of the subordination relation that it expresses. In section 6.3.3 I will discuss the repercussions of variable argument coding for the classification of DC constructions in terms of structural types.

---

\(^{100}\) For examples of each of the parameter values, see Chapter 3, section 3.4.
6.3.2.2 Language data

Keeping the remarks of the previous subsection in mind, consider the data in Table 6.2 below. As in Table 6.1 above, the languages are listed in the leftmost column. The DCs appear in the second column and are named after their structural coding. The other columns display the behavioural potential of every DC in terms of the three formal parameters:

(i) (non-)expression of verbal categories;
(ii) (non-)expression of nominal categories;
(iii) coding of the argument(s).

Verbal categories are abbreviated as follows: \(VV\) for voice and valency markers; \(T, A,\) and \(M\) for tense, aspect and mood marking, respectively; \(P\) for person marking, and \(IF\) for illocutionary force. Nominal categories are abbreviated as \(DET\) for determiners and \(CASE\) for case markers and adpositions. Arguments are simply labelled 1 and 2 for first and second argument.

Table 6.2: Basic data on behavioural potential of all DC constructions

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural coding</th>
<th>Behavioural potential</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Verbal categories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W T A M P IF DET CASE 1 2</td>
</tr>
<tr>
<td>Tagalog</td>
<td>(pag-)</td>
<td>– – + + ALT SENT</td>
</tr>
<tr>
<td></td>
<td>(na/-ng)</td>
<td>+ + + + + + + + + + + +</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SENT SENT</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>Verbal categories</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W T A M P</td>
<td>IF DET CASE</td>
</tr>
<tr>
<td>Kharia</td>
<td>RDP/Ø</td>
<td>+/- – – –</td>
</tr>
<tr>
<td>-na</td>
<td>– – – –</td>
<td>(+) ALT/Ø</td>
</tr>
<tr>
<td>-na-wala</td>
<td>– – – –</td>
<td>Ø SENT</td>
</tr>
<tr>
<td>-al</td>
<td>– – – –</td>
<td>ALT SENT</td>
</tr>
<tr>
<td>-ker etc.</td>
<td>– – – –</td>
<td>Ø ALT</td>
</tr>
<tr>
<td>-ga</td>
<td>– – – –</td>
<td>Ø SENT</td>
</tr>
<tr>
<td>-ta + RDP</td>
<td>– – – –</td>
<td>Ø SENT</td>
</tr>
<tr>
<td>no</td>
<td>+ + + –</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>gam-kon</td>
<td>+ + + –</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>je etc.</td>
<td>+ + + –</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>a/i etc.</td>
<td>+ + + –</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>Ø</td>
<td>+ + +/-</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td>Kambera</td>
<td>pa-</td>
<td>– – – –</td>
</tr>
<tr>
<td>ma-</td>
<td>– – – –</td>
<td>Ø SENT</td>
</tr>
<tr>
<td>Ø (NMLZ)</td>
<td>+/- +/-</td>
<td>ALT SENT</td>
</tr>
<tr>
<td>wà</td>
<td>+ + +</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>Samoan</td>
<td>=ga</td>
<td>– – – –</td>
</tr>
<tr>
<td>DET (NMLZ)</td>
<td>– – – –</td>
<td>+ + ALT</td>
</tr>
<tr>
<td>ona/ina</td>
<td>– – – –</td>
<td>– – SENT SENT</td>
</tr>
<tr>
<td>-e</td>
<td>+ + +</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td>Ø</td>
<td>+ + +</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>Guaraní</td>
<td>hágwe)/Ø</td>
<td>+ + +/-</td>
</tr>
<tr>
<td>va</td>
<td>+ +</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td>vo</td>
<td>+ +</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>Santali</td>
<td>Ø (no INDIC)</td>
<td>+/- +/- +/- +/- +/-</td>
</tr>
<tr>
<td>-kate</td>
<td>+ – – – –</td>
<td>Ø SENT</td>
</tr>
<tr>
<td>CORREL</td>
<td>+ + + +/- +/-</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>mente</td>
<td>+ + + + +</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Warao</td>
<td>-kitane</td>
<td>-det</td>
</tr>
<tr>
<td></td>
<td>kotai</td>
<td>?</td>
</tr>
<tr>
<td>Quechua</td>
<td>-j/-shka/-na</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>-ngapaj</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-chun</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-y</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-shpa</td>
<td>-</td>
</tr>
<tr>
<td>Turkish</td>
<td>-DIK/-(y)AcAK</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>-mAK</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-mA</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-An</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-(y)ArAk</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-(y)A...-(y)A</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ki</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>diye</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-n-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-Thirri-n</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-n-garrba</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-ntha</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>tu(a)/tjai</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a parhu</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(-)in- [...]-an + a</td>
<td>-</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>Verbal categories</td>
<td>Nominal categories</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>T</td>
</tr>
<tr>
<td>Ma`di</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-lɛ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-dɔɔ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ka</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ɛ̄/-bá</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ø</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-zi+sí</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gooniyandi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-woo</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-wadda</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-mawoo</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-bári</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-ya/</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-gowaya</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Ø (+RSP)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Hungarian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ni</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-ás/és</td>
<td></td>
<td>−</td>
</tr>
<tr>
<td>-ó</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ótt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-andó/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-endó</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-vá/-vé</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-vén</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hogy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RELPRON</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no/momo</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ø (RSP)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>-te/-de/</td>
<td></td>
<td>−</td>
</tr>
<tr>
<td>-ite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-i/-Ø</td>
<td></td>
<td>−</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Verbal categories</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>qhov</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>kuam/ (has)tas</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>kws</td>
<td>-</td>
</tr>
<tr>
<td>Lango</td>
<td>-(kk)b</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ni</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>à-mè</td>
<td>+</td>
</tr>
<tr>
<td>Ket</td>
<td>Ø (bare INF)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>età qor’a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ásqà</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-PROSEC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-s/-bes</td>
<td>+</td>
</tr>
<tr>
<td>Itelmen</td>
<td>INF (various forms)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>quatz</td>
<td>+</td>
</tr>
<tr>
<td>Koasati</td>
<td>NMLZ (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>:-sáya</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>:-yólli</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>:-ka</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>:-kítta</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>:-laho: li:sáya</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>:-n</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>:-k</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>:-t</td>
<td>-</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Thai</td>
<td>Ø</td>
<td>w t a m p if det case</td>
</tr>
<tr>
<td></td>
<td>thiï</td>
<td></td>
</tr>
<tr>
<td></td>
<td>wâa</td>
<td></td>
</tr>
<tr>
<td>Basque</td>
<td>-t(z)e</td>
<td>-tu/-du/-i/-Ø</td>
</tr>
<tr>
<td></td>
<td>-en</td>
<td>-en bezala - (e)la</td>
</tr>
<tr>
<td></td>
<td>bait</td>
<td></td>
</tr>
<tr>
<td>Abun</td>
<td>do/Ø</td>
<td>gato</td>
</tr>
<tr>
<td></td>
<td>sa gato</td>
<td></td>
</tr>
<tr>
<td>Bambara</td>
<td>ka</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-le/-ne</td>
<td>-min(u)/-mun(u)</td>
</tr>
<tr>
<td></td>
<td>-tò</td>
<td></td>
</tr>
<tr>
<td>Georgian</td>
<td>-a</td>
<td>m- (-a-) (-el/-all)</td>
</tr>
<tr>
<td></td>
<td>-ul/-il/m- -ar/-al</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sa- (-el/-al/r)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rom</td>
<td></td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>ra</td>
</tr>
<tr>
<td></td>
<td>rogore</td>
<td></td>
</tr>
<tr>
<td>Bukiyip</td>
<td>Ø</td>
<td>(ú)li</td>
</tr>
<tr>
<td></td>
<td>bwidouk</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-(u)mu</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>Verbal categories</td>
<td>Nominal categories</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>T</td>
</tr>
<tr>
<td>Abkhaz</td>
<td>-ra</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-NFIN</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>REL.PRON + -NFIN</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>-š-+ -NFIN</td>
<td>(+)</td>
</tr>
<tr>
<td>Polish</td>
<td>-nie</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>INF (various forms)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-c</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>-any etc.</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>(PST. PASS) PTC (various forms)</td>
<td>(+)</td>
</tr>
<tr>
<td></td>
<td>PRS.PL-c-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>že</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>jak (gdy) by</td>
<td>+</td>
</tr>
<tr>
<td>Burushaski</td>
<td>-(á)as</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-im/-um/-am</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>n-STEM -(a)n</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ke/ki</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>sén/-ét-</td>
<td>+</td>
</tr>
<tr>
<td>Lavu-kaleve</td>
<td>-e/-i</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>PERSON. SUFF</td>
<td>+</td>
</tr>
<tr>
<td>Alamblak</td>
<td>-nef</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-(kfé)t</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ind/Ø</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>Verbal categories</td>
<td>Nominal categories</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>T</td>
</tr>
<tr>
<td>Pipil</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ka(h)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ne</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ke</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ke:n-aken</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ADV</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Wambon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-e</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-a (+o)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ø</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-mo/-o</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>ka</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Dhaasansac</td>
<td>DET (+DEM)</td>
<td>+</td>
</tr>
<tr>
<td>ḷ/-an</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>fu/fi/Ø</td>
<td>-</td>
</tr>
<tr>
<td>bif/o</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Dati/o</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Wh/Ø</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Babungo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lāa</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>fān/yūu</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>kī(i)/Ø</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Nama</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lxáis-a</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>(no INDIC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ø (no INDIC)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-se/l’aa/ tsii/Ø (no INDIC)</td>
<td>?</td>
<td>-</td>
</tr>
<tr>
<td>Hdi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kā</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>tā</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>tā (+)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tā + NMLZ</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>-a</td>
<td>(+)</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal categories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W T A M P IF DET CASE</td>
</tr>
<tr>
<td>Mandarin</td>
<td>Ø</td>
<td>Ø/SENT SENT</td>
</tr>
<tr>
<td></td>
<td>de</td>
<td>Ø/SENT SENT</td>
</tr>
<tr>
<td></td>
<td>-zhe</td>
<td>Ø/SENT SENT</td>
</tr>
<tr>
<td>Tamil</td>
<td>-atu</td>
<td>Ø/SENT SENT</td>
</tr>
<tr>
<td></td>
<td>-(kk)a</td>
<td>Ø/SENT SENT</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td>Kisi</td>
<td>(m)âà/Ø</td>
<td>SENT SENT</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td>Nung</td>
<td>Ø</td>
<td>Ø/SENT SENT</td>
</tr>
<tr>
<td></td>
<td>(ti-va) (+DEM) (+FOC)</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td></td>
<td>bêt</td>
<td>SENT SENT</td>
</tr>
<tr>
<td>Garo</td>
<td>in-e</td>
<td>SENT SENT</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>ALT SENT</td>
</tr>
<tr>
<td></td>
<td>-a-ni</td>
<td>+ Ø Ø</td>
</tr>
<tr>
<td></td>
<td>-na</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td></td>
<td>-kan/-kan-a</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td></td>
<td>-gip-a</td>
<td>ALT/Ø SENT/Ø</td>
</tr>
<tr>
<td></td>
<td>-e/-e-min/-e-r</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td>Krongo</td>
<td>ânî tiŋ</td>
<td>+ + + + + + - SENT SENT</td>
</tr>
<tr>
<td></td>
<td>-(t)</td>
<td>+ + + - - +/- ALT/Ø SENT</td>
</tr>
<tr>
<td></td>
<td>m-/n</td>
<td>+ + +/- - - ALT/Ø SENT</td>
</tr>
<tr>
<td>Hixkar-</td>
<td>ni/-thi/-hito + ri</td>
<td>+/- - - - ALT ALT</td>
</tr>
<tr>
<td>yana</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slave</td>
<td>nî/Ø</td>
<td>SENT SENT</td>
</tr>
<tr>
<td></td>
<td>gha/gú</td>
<td>SENT SENT</td>
</tr>
<tr>
<td></td>
<td>i/-sí/-lìi</td>
<td>Ø/SENT Ø/SENT</td>
</tr>
<tr>
<td></td>
<td>gharé</td>
<td>SENT SENT</td>
</tr>
</tbody>
</table>
### 6.3.3 A three-way typology of DCs according to behavioural potential

#### 6.3.3.1 Introduction

Based on the behavioural potential of all DCs presented in Table 6.2 above, these constructions can now be assigned to one of the three structural types defined in Chapter 3. The properties of the three types are briefly repeated below:

- **Type 1: Balanced constructions** (abbreviated as B)
  - All TAM distinctions and Person are expressed as in independent clauses;
  - All overt arguments are coded as in independent clauses (SENT-SENT);
  - Determiners (DET) and/or case markers/adpositions (CASE) can but need not be expressed.

- **Type 2: Deranked DCs with SENT arguments** (Abbreviated as D-SENT)
  - (Partial) non-expression of TAM distinctions and/or Person marking;
  - All overt arguments are expressed as in independent clauses (SENT-SENT);
  - DET/CASE can but need not be expressed.
- **Type 3: Deranked DCs with at least one alternatively coded argument**  
  (Abbreviated as D-ALT)
  - (Partial) non-expression of TAM distinctions and/or Person marking;
  - One argument is expressed with an ALT strategy; the other one, if available/overt, is either SENT or ALT as well.
  - DET/CASE are expressed (if the language has these categories).

Table 6.3 summarizes the formal characteristics of the three DC types (this table is the same as Table 3.1 and Table 4.2):

<table>
<thead>
<tr>
<th>DC type</th>
<th>Argument expression</th>
<th>TAM / Person</th>
<th>DET / CASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>number</td>
<td>label</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Type 1</td>
<td>B</td>
<td>SENT / Ø</td>
<td>SENT / Ø</td>
</tr>
<tr>
<td>Type 2</td>
<td>D-SENT</td>
<td>SENT / Ø</td>
<td>SENT / Ø</td>
</tr>
<tr>
<td>Type 3</td>
<td>D-ALT</td>
<td>ALT</td>
<td>SENT / ALT / Ø</td>
</tr>
</tbody>
</table>

*Table 6.3: DC types and their internal formal properties*

Before presenting the classification of the DCs in the sample languages (see Table 6.5, in section 6.3.3.3) some preliminary considerations are required.

### 6.3.3.2 Preliminary considerations

*Variable argument coding*

As mentioned in the previous section, and as can be seen in Table 6.2, many DC constructions display some variation in the coding of their arguments, depending on their propositional function and/or on the specific semantic/syntactic subordination relation that they express. Such cases of variable argument coding come in several types.

First, there are complement constructions that can take the form of either an infinitive or a nominalization. That is to say, the first argument of such DCs is either unexpressed under co-referentiality or, when overt, it is expressed with an ALT strategy. The second argument (if present) remains SENT in both cases. This means that, in terms of the typology in Table 6.3 above, such a DC would classify sometimes as a type 2 construction (when the first argument is Ø) and sometimes as a type 3 construction (when the first argument is ALT). However, since the first argument is ALT whenever
it is overtly expressed, these DCs will be assigned to the group of type 3 (D-ALT) constructions. The examples in (50) illustrate this for the -ka construction in Ma’di, which has an un-expressed first argument in (50a); a possessive first argument in (50b); and an object argument that is expressed as in an independent clause in (50c):

Ma’di (Blackings & Fabb 2003: 22, 21, 213)
(50)  a. má ɓi [mū-ka] kɔ̀rò
     1SG try (n)-go-NMLZ NEG(pst)
     ‘I have not tried/did not try to go/going.’

     b. Má  ndrē [áriággwá rì ɬà ṑ-dé-ka] rá
     1SG see bird DEF POSS (n)-VE-fall-NMLZ AFF
     ‘I saw the bird’s falling/fall.’

     c. s-lè [tòbà sè-ka] dìà  kò
     IND-want cigarette (n)-smoke-NMLZ here NEG
     ‘Smoking is not permitted here.’

Note that this type of DC stands in contrast to constructions that are invariably infinitival. In the latter construction type the first argument always remains unexpressed, and therefore it is classified as type 2 (D-SENT).

A second type of DC with variable argument coding involves deranked relative clause constructions with a gap strategy: When the first argument of such a DC is gapped, the second argument (if present) remains SENT. However, when the second argument is gapped, the first argument is expressed with an ALT strategy. As in the previous case, since the first argument gets ALT coding whenever it is expressed overtly, this type of DCs is classified as type 3 (D-ALT). An example from Garo is given in (51), which shows the alternation of ALT expression (51a) and zero-expression (gapping) of the subject argument (51b):

Garo (Burling 2004: 301, 299)
(51)  a. [mèchik-ni skang-o den’gip-a] a-bol
     women-GEN previously-LOC cut-PTC firewood
     ‘firewood that the women chopped previously’
b.  [nok-o  pi'-sa-ko  nik-gip-a]  me-tra
   house-LOC  child-ACC  see-PTC  woman
   ‘the woman who saw the child at the house’

Note that there are also deranked relative clauses with gapping that do not have the possibility of ALT coding of the first argument. Rather, whenever an argument is overt (i.e. not gapped) in such DCs, it gets SENT expression. In addition, when the relativized item is another argument than the subject or object, both core arguments are sententially expressed. Therefore, these DCs are classified as type 2: D-SENT\textsuperscript{101}. This construction type is attested in Burushaski and Tamil. Examples from Tamil appear in (52), which show a relative clause with a gapped subject (52a), a gapped object (52b), and a gapped instrumental adjunct (52c):

\textit{Tamil} (Asher 1982: 28)

(52) a.  [vanqaan  aticc-a]  taccan
   carpenter-ACC  beat.PST-PTC  washerman
   ‘the washerman who beat the carpenter’

b.  [taccan  aticc-a]  vanqaan
   carpenter  beat.PST-PTC  washerman
   ‘the washerman whom the carpenter beat’

c.  [akkaa  tagkaccikki  caata  poott-a]  karanți
   elder.sister  younger.sister-DAT  rice  put.PST-PTC  spoon
   ‘The spoon with which elder sister gave rice to younger sister’

A third case of variable argument coding is found with certain types of nominal clauses, i.e. DCs that can be used as complement clauses and relative clauses. In the former function, the first argument of the DC must remain unexpressed (because of co-referentiality with the matrix clause subject), while the second argument is SENT. In the latter function, however, this type of construction expresses object relative clauses and employs a gap strategy. The first argument then gets ALT coding and the second remains unexpressed. Thus, there are two possible argument realizations depending

\textsuperscript{101} Of course, the same classification applies to relative clause construction in which the subject is obligatorily the relativized and gapped argument (and in which the second argument is SENT).
on the propositional function in which the DC is used: Ø-SENT in referential function and ALT-Ø in modifier function. Since the zero-coding of the subject argument is obligatory in referential function, this construction type receives a double classification in terms of behavioural potential: type 2 (D-SENT)/type 3 (D-ALT). There are only three such cases attested, namely the pa-construction in Kambera, which is illustrated in (53a-b), and the le- and dƷɔ́-constructions in Ma’di, illustrated in (54a-b) and (55a-b), respectively. In all pairs of examples the first one illustrates the complement clause function and the second the relative clause function of the DC. Note that these DC constructions, even though they receive a double classification in terms of behavioural potential, are nonetheless regarded as flexible constructions, since their structural coding is the same in both functions.

**Kambera** (Klamer 1998: 338, 326)

(53) a. *Tà-pakiring [pa-tinu-nya na lau]*
   1PL.NOM-start COMP-weave-3SG.DAT ART sarong haromu
   tomorrow
   We will start to weave the sarong tomorrow.’

b. *na kalæmbi na [pa-kei wà-nggu-nya]*
   ART shirt ART REL-buy use-1SG.GEN-3SG.DAT
   ‘the shirt that I bought’

**Ma’di** (Blackings & Fabb 2003: 202, 22, 207, 206)

(54) a. *Má lè-à [èbì ‘pà-le] rà*
   1SG (N)want-OBJ fish N-eat-NMLZ AFF
   ‘I certainly want to eat fish.’

   car Opi Poss (N)-take-PTC DEF leg AFR deflate AFF
   ‘The car which Opi took certainly has a flat tyre.’

(55) a. *spì édò [sì-dƷɔ́] rà*
   Opi start N-build-NMLZ AFF
   ‘Opi has certainly started to build it.’
b. òòòu [mà-à òòòu] rì ri ògwà dì
   spear 1sg-poss n-build-ptc def foc lose com
   ‘The spear with which I killed it is lost.’

There is one more flexible DC construction, in Krongo, with a double classification Type 2/3, but it is a flexible modifier clause construction, rather than a nominal clause construction. It has already been illustrated in (35a-b) above. These examples show that the construction has an ALT-coded subject when it is used as a relative clause (35a), while in adverbial function (35b) the subject always remains unexpressed, because it must be co-referential with the subject of the matrix clause.

An isolated case of double classification concerns Japanese nominalizations. These constructions are balanced to the extent that they retain tense marking, but deranked in the sense that they can, albeit optionally, take a possessive subject argument, as is shown in (56). Because of this contradictory evidence, the construction is classified as type 1/3 (B/D-ALT).


(56) [Àno hito ga/no hon o kai-ta koto]
   that person nom/gen book acc write-past nmlz
   ga yoku sira-re-te iru
   nom well know-pass-ger be
   ‘It is well known that that person wrote a book.’

Finally, consider the so-called ‘participial mood’ construction in West Greenlandic. As Fortescue (1984: 49) notes, this construction is “fully inflectible” when used as a complement clause, while it lacks person marking when used as a relative clause, as is illustrated in (57a) and (57b), respectively. Although this DC is thus not fully finite in both functions, it is nevertheless classified as a type 1 construction.

West-Greenlandic (Fortescue 1984: 36, 49)

(57) a. Ilisima-vaa [urni-ssa-giga]
   know-3sg.3sg.ind come.to-fut-1sg.3sg.ptc
   ‘He1 knew I would come to him2.’
b. **Niviarsiaq** [kalaallisut ilinnia-lir-suq]
   
girl Greenlandic learn-begin-{INTR.PTC}
   ‘the/a girl who has begun learning Greenlandic.’

_The balancing/deranking distinction in languages with little or no verbal inflection._

Not surprisingly, there are some cases in which the decision to classify a specific DC as a balanced or a deranked construction is empirically rather vacuous, as a result of the (near) absence of verbal inflectional categories in certain languages. This holds most obviously for isolating languages. For instance in Thai, the construction in (58) is classified as balanced by default; since there are no verbal inflectional categories that can be lost (and since no nominal categories are acquired), there is no empirical basis on which to classify the construction as deranked.

_Thai_ (Iwasaki & Ingkaphirom 2005: 243)

(58) **Khon** [tbi dulee] ni pen pen acaan lɔ
   
   person rel take.care prt cop cop teacher {INTERR}
   ‘Is the person who takes care [of the students] a teacher?’

A similar situation obtains in other languages with little verbal morphology. Consider for instance Ma’di: As regards inflectional morphology, verbs in this language can only take a low-tone prefix, which in independent clauses expresses non-past tense. Dependent predicates, which are marked with one of the subordinating suffixes kə, lɛ, rɛ, ᵇa, or dʒɔ̄, take a homophonous prefix. However, according to Blackings and Fabb (2003: 192) it is not really clear whether this is the same prefix as in independent clauses. An argument against identity of the two forms is that in dependent clauses the prefix is compatible with any tense interpretation. Moreover, when the subject in DCs with the subordinating suffixes lɛ, dʒɔ̄, and kə is overt, then it is coded as a possessor. On the basis of this evidence I classify these Ma’di constructions as deranked.

### 6.3.3.3 Language data

Keeping the above considerations in mind, the three-way DC classification on the basis of behavioural potential is represented in Table 6.4:
Table 6.4: Three-way classification of DCs

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural coding</th>
<th>DC type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagalog</td>
<td>pag-</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>na/-ng</td>
<td>1</td>
</tr>
<tr>
<td>Kharia</td>
<td>RDP/Ø</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-na</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-na-wala</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-al</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-ker etc.</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-ga</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-ta + RDP</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>gam-kon</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>je etc.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>a/i etc.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1</td>
</tr>
<tr>
<td>Kambera</td>
<td>pa-</td>
<td>2/3</td>
</tr>
<tr>
<td></td>
<td>ma-</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Ø (NMLZ)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>wà</td>
<td>1</td>
</tr>
<tr>
<td>Samoan</td>
<td>=ga</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Ø (NMLZ)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>ona/ina</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-e</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1</td>
</tr>
<tr>
<td>Guarani</td>
<td>há(gwe)/Ø</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>va</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>vo</td>
<td>1</td>
</tr>
<tr>
<td>Santali</td>
<td>Ø (no INDIC)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-Kate</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>CORREL</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>mente</td>
<td>1</td>
</tr>
<tr>
<td>Warao</td>
<td>-kitane</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>kotai</td>
<td>1</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>DC type</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Imbabura Quechua</td>
<td>-j/-shka/-na</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-ngapaj</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-chun</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-y</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-shpa</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Turkish</td>
<td>-DIK/-(y)AcAK</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>-mA</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-mA</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>-An</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-(y)ArAk</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-(y)A...-(y)A</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>ki</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>diyé</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1 B</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-n-</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>-Thiri-n</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>-n-garrba</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-ntha</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1 B</td>
</tr>
<tr>
<td>Paiwan</td>
<td>tu(a)/tjai</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>a parhu</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>(-)in- -an + a</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Ma’di</td>
<td>-ɛ</td>
<td>2/3 D-SENT/D-ALT</td>
</tr>
<tr>
<td></td>
<td>-dʒɔ</td>
<td>2/3 D-SENT/D-ALT</td>
</tr>
<tr>
<td></td>
<td>-ka</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>-ɛɛ/-bá</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>-zi+ si</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>-woo</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-wadda</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-mawoo</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-bari</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-ya/-gowaaya</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>Ø (+RSP)</td>
<td>1 B</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>DC type</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Hungarian</td>
<td>-ni</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-ás/-és</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-ó</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-ött</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-andó/-endó</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-vá/-vé</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-vén</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(azt) hogy (etc.)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ugy (a-)hogy</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>1</td>
</tr>
<tr>
<td>Japanese</td>
<td>no/mono etc.</td>
<td>1/3</td>
</tr>
<tr>
<td></td>
<td>Ø (RSP)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-te/-de/-ite</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-i/-Ø</td>
<td>2</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>qhov</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>kuan/(has)tas</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>kws</td>
<td>1</td>
</tr>
<tr>
<td>Lango</td>
<td>-(kk)ɔ</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>nî</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>à-më</td>
<td>1</td>
</tr>
<tr>
<td>Ket</td>
<td>Ø (bare INF)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>eta qor’a</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ásqà</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-PROSEC</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-s/-bes</td>
<td>1</td>
</tr>
<tr>
<td>Itelmen</td>
<td>INF (various forms)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>quatz</td>
<td>1</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>DC type</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Koasati</td>
<td>NMLZ (various forms)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-:sáya</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-:yólli</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-:ka</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-:kittta</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-:laho:li:sáya</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-n</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-k</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-t</td>
<td>D-SENT</td>
</tr>
<tr>
<td>Thai</td>
<td>ø</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>thii</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>wáa</td>
<td>B</td>
</tr>
<tr>
<td>Basque</td>
<td>-t(z)e</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-tu/-du/-i/-Ø</td>
<td>D-SENT</td>
</tr>
<tr>
<td></td>
<td>-en</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-en bezala</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-(e)la</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>bait</td>
<td>1</td>
</tr>
<tr>
<td>Abun</td>
<td>do/Ø</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>gato</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>sa gato</td>
<td>B</td>
</tr>
<tr>
<td>Bambara</td>
<td>ka</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-le/-ne</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-min(u)/-mun(u)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-tò</td>
<td>2</td>
</tr>
<tr>
<td>Georgian</td>
<td>-a</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>m-(-a-)(-el/-al))</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-ul/-il/m--ar/-al</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>sa-(-el/-al/r)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>rom</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>ra</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>rogorc</td>
<td>1</td>
</tr>
<tr>
<td>Bukiyip</td>
<td>ø</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(ú)li</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>bwidouk -(u)mu</td>
<td>B</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>DC type</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Abkhaz</td>
<td>-ra</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-NFIN</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>REL.PRON + -NFIN</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-ş+ -NFIN</td>
<td>1</td>
</tr>
<tr>
<td>Polish</td>
<td>-nie</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>INF (various forms)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-c</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-any etc.</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(PST.PASS) PTC (various forms)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>PRS.PL-c-</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Ŝe</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>jak (gdy)by</td>
<td>1</td>
</tr>
<tr>
<td>Burushaski</td>
<td>-(á)as</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-im/-um/-am</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>n-STEM-(a)n</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>ke/ki</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>sén-/ét-</td>
<td>1</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>-e/-i</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>PERSON.SUFF</td>
<td>1</td>
</tr>
<tr>
<td>Alamblak</td>
<td>-nef</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>-(kfe)t</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>ind/Ø</td>
<td>2</td>
</tr>
<tr>
<td>Pipil</td>
<td>ka(h)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ne</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ke</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ke:n-aken</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ADV</td>
<td>1</td>
</tr>
<tr>
<td>Wambon</td>
<td>-e</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-a (+ o)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>-mo/-o</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>ka</td>
<td>1</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>DET(+DEM)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>-n/-an</td>
<td>3</td>
</tr>
<tr>
<td>Language</td>
<td>Structural coding</td>
<td>DC type</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Berbice Dutch Creole</td>
<td>fu/fi/Ø</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>bifi/dati/Ø</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>Wh/Ø</td>
<td>1 B</td>
</tr>
<tr>
<td>Babungo</td>
<td>làa</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>fāñ/yüu</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>ki(i)/Ø</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Nama</td>
<td>lxáis-ā (no INDIC)</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>Ø (no INDIC)</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>-se/l’aa/tsii (no INDIC)</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Hdi</td>
<td>ká</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>tá</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>tà</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>tà + NMLZ</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>1 B</td>
</tr>
<tr>
<td>Mandarin Chinese</td>
<td>Ø</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>de</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>-zhe</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Tamil</td>
<td>-atu</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-(kk)a</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Kisi</td>
<td>(m)àà/Ø</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>1 B</td>
</tr>
<tr>
<td>Nung</td>
<td>Ø</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>(tì-và) (+DEM) (+FOC)</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>bàt</td>
<td>1 B</td>
</tr>
<tr>
<td>Garo</td>
<td>in-e</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>-a-ni</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-na</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-kan/-kan-a</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td></td>
<td>-gip-a</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>-e/-e-min/-e-r</td>
<td>2 D-SENT</td>
</tr>
<tr>
<td>Krongo</td>
<td>ànì tíŋ</td>
<td>1 B</td>
</tr>
<tr>
<td></td>
<td>-(t)</td>
<td>3 D-ALT</td>
</tr>
<tr>
<td></td>
<td>m- /n-</td>
<td>2/3 D-SENT /D-ALT</td>
</tr>
</tbody>
</table>
In this section, the data in Table 6.2 on the behavioural potential of the DCs in the sample languages are re-ordered in the form of three separate tables, each of which corresponds to one of the three structural DC types. In particular, Table 6.5 presents the behavioural potential of all type 1 balanced DCs; Table 6.6 presents the behavioural potential of all type 2 D-SENT DCs; and Table 6.7 presents the behavioural potential of all type 3 D-ALT DCs. Note that those DCs that received double classifications in Table 6.4 above, appear in two of the three tables below.

**Table 6.5: Behavioural potential of Type 1 balanced DCs**

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Behavioural potential</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Verbal categories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W</td>
</tr>
<tr>
<td>Tagalog</td>
<td>na/-ng</td>
<td>+</td>
</tr>
<tr>
<td>Kharia</td>
<td>no</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>gam-kon</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>je etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a/i etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Kambera</td>
<td>wà</td>
<td>+</td>
</tr>
</tbody>
</table>

6.3.4 Behavioural potential of different structural DC types

In this section, the data in Table 6.2 on the behavioural potential of the DCs in the sample languages are re-ordered in the form of three separate tables, each of which corresponds to one of the three structural DC types. In particular, Table 6.5 presents the behavioural potential of all type 1 balanced DCs; Table 6.6 presents the behavioural potential of all type 2 D-SENT DCs; and Table 6.7 presents the behavioural potential of all type 3 D-ALT DCs. Note that those DCs that received double classifications in Table 6.4 above, appear in two of the three tables below.
<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Behavioural potential</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Verbal categories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W</td>
</tr>
<tr>
<td>Samoan</td>
<td>-e</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Guaraní</td>
<td>hágwé/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>va</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>vo</td>
<td>+</td>
</tr>
<tr>
<td>Santali</td>
<td>CORREL</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>mente</td>
<td>+</td>
</tr>
<tr>
<td>Warao</td>
<td>kotai</td>
<td>+</td>
</tr>
<tr>
<td>Turkish</td>
<td>ki</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>diye</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-ntha</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Paiwan</td>
<td>tu(a)/tjai</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a parhu</td>
<td>+</td>
</tr>
<tr>
<td>Ma’di</td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>Ø (+RSP)</td>
<td>+</td>
</tr>
<tr>
<td>Hungarian</td>
<td>hogy (etc.)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>RELPRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ugy (a-)</td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>no/mono etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (RSP)</td>
<td>+</td>
</tr>
<tr>
<td>Hmong</td>
<td>qhov</td>
<td></td>
</tr>
<tr>
<td>Njua</td>
<td>kuam/ (has)tas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kws</td>
<td></td>
</tr>
<tr>
<td>Lango</td>
<td>ní</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>à-mê</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>Verbal categories</td>
<td>Nominal categories</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>T</td>
</tr>
<tr>
<td>Ket</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>eta qo r'a</em></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>dåsgà</em></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-PROSEC</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-s/-bes</td>
<td>+</td>
</tr>
<tr>
<td>Itelmen</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>min</em></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>quartz</em></td>
<td>+</td>
</tr>
<tr>
<td>Thai</td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>thii</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>wàa</em></td>
<td></td>
</tr>
<tr>
<td>Basque</td>
<td>-en</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-en bezala</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(*e)la</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>bait</em></td>
<td>+</td>
</tr>
<tr>
<td>Abun</td>
<td>do/Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>gato</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>sa gato</em></td>
<td></td>
</tr>
<tr>
<td>Bambara</td>
<td><em>ka</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-min(u)/ -mun(u)</td>
<td></td>
</tr>
<tr>
<td>Georgian</td>
<td>rom</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>ra</em></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rogorc</td>
<td>+</td>
</tr>
<tr>
<td>Bukiyip</td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ú)li</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bwidouk</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-(*u)mu</td>
<td></td>
</tr>
<tr>
<td>Abkhaz</td>
<td>-NFIN</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-s- + NFIN</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Behavioural potential</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>verbal categories</td>
<td>nominal categories</td>
</tr>
<tr>
<td></td>
<td>W  T  A  M  P  IF</td>
<td>DET  CASE  1  2</td>
</tr>
<tr>
<td>Polish</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>245</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ + + + -</td>
<td>SENT  SENT</td>
</tr>
<tr>
<td>Polish</td>
<td>RELPRON</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ + + + +</td>
<td>Ø/SENT  Ø/SENT</td>
</tr>
<tr>
<td>Polish</td>
<td>jak (gdy) by</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ + + +</td>
<td>SENT  SENT</td>
</tr>
<tr>
<td>Burushaski</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ke/ki</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ + + +</td>
<td>SENT  SENT</td>
</tr>
<tr>
<td>Burushaski</td>
<td>sén-/ét-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ + + +</td>
<td>SENT  SENT</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PERSON. SUFF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ + + + +</td>
<td>+  SENT  SENT</td>
</tr>
<tr>
<td>Pipil</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ka(h)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ +</td>
<td>SENT  SENT</td>
</tr>
<tr>
<td>Pipil</td>
<td>ne</td>
<td>+</td>
</tr>
<tr>
<td>Pipil</td>
<td>ke</td>
<td>+</td>
</tr>
<tr>
<td>Pipil</td>
<td>ke:n-aken</td>
<td>+</td>
</tr>
<tr>
<td>Pipil</td>
<td>ADV</td>
<td>+</td>
</tr>
<tr>
<td>Wambon</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-e</td>
<td>+ + + +</td>
</tr>
<tr>
<td>Wambon</td>
<td>-a (+ o)</td>
<td>+ + + +</td>
</tr>
<tr>
<td>Wambon</td>
<td>ka</td>
<td>+ + +</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>DET (+DEM)</td>
<td>+ + +</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>bifi/dati/ Ø</td>
<td>+ + +</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>Wh/Ø</td>
<td>+ + +</td>
</tr>
<tr>
<td>Babungo</td>
<td>lâa</td>
<td>+</td>
</tr>
<tr>
<td>Babungo</td>
<td>fâŋ/yúu</td>
<td>+</td>
</tr>
<tr>
<td>Nama</td>
<td>lxás-à (no INDIC)</td>
<td>+ + –</td>
</tr>
<tr>
<td>Nama</td>
<td>Ø (no INDIC)</td>
<td>+ + –</td>
</tr>
<tr>
<td>Hdi</td>
<td>ká</td>
<td>(+)</td>
</tr>
<tr>
<td>Hdi</td>
<td>tà</td>
<td>(+)</td>
</tr>
<tr>
<td>Hdi</td>
<td>-a</td>
<td>(+)</td>
</tr>
<tr>
<td>Mandarin Chinese</td>
<td>Ø</td>
<td></td>
</tr>
<tr>
<td>Mandarin Chinese</td>
<td>de</td>
<td></td>
</tr>
<tr>
<td>Kisi</td>
<td>(m)àà/Ø</td>
<td>+ + +</td>
</tr>
<tr>
<td>Kisi</td>
<td>CL</td>
<td>+ + +</td>
</tr>
</tbody>
</table>
### Table 6.6: Behavioural potential of type 2 D–SENT DCs

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Verbal categories</th>
<th>Nominal categories</th>
<th>Argument expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>Kharia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-na-wala</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ga</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ta + RDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kambera</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>ona/ina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santali</td>
<td>Ø (no INDIC)</td>
<td>+/-</td>
<td>+/</td>
<td>+/</td>
</tr>
<tr>
<td>-kate</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warao</td>
<td>-kitane</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Verbal categories</th>
<th>Nominal categories</th>
<th>Argument expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>Nung</td>
<td>Ø</td>
<td>Ø/SENT</td>
<td>SENT</td>
<td></td>
</tr>
<tr>
<td>(ti-va)</td>
<td>(+DEM)</td>
<td>Ø/SENT</td>
<td>Ø/SEN</td>
<td></td>
</tr>
<tr>
<td>(FOC)</td>
<td></td>
<td>Ø/SENT</td>
<td>Ø/SEN</td>
<td></td>
</tr>
<tr>
<td>bat</td>
<td></td>
<td>SENT</td>
<td>SENT</td>
<td></td>
</tr>
<tr>
<td>Garo</td>
<td>in-e</td>
<td>SENT</td>
<td>SENT</td>
<td></td>
</tr>
<tr>
<td>Kromo</td>
<td>anà tiŋ</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Slave</td>
<td>nj/Ø</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>gha/qú</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>i/šji/ií</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>gháré</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Nivkh</td>
<td>-vut/-vur</td>
<td>Ø/SENT</td>
<td>Ø/SENT</td>
<td></td>
</tr>
<tr>
<td>Ø</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>NFIN</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>Ø/SENT</td>
</tr>
<tr>
<td>West</td>
<td>PTC mood</td>
<td>(+)</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Greenlandic</td>
<td>CONT mood</td>
<td>(+)</td>
<td>(+)</td>
<td>(+)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Verbal categories</th>
<th>Nominal categories</th>
<th>Argument expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>Kharia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kambera</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>ona/ina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santali</td>
<td>Ø (no INDIC)</td>
<td>+/-</td>
<td>+/</td>
<td>+/</td>
</tr>
<tr>
<td>Warao</td>
<td>-kitane</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Behavioural potential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nominal categories</td>
<td>Argument expression</td>
<td>Verbal categories</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>Quechua</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-j/-shka/na</td>
<td>(+) +</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-chun</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-ngapaj</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-y</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-shpa</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-n-garba</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Paiwan</td>
<td>(-)in- -an</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-mAK</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-An</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-(y)ArAk</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-(y)A</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ma’di</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-lɛ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-dɔo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-rɛ/-bá</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-zɪ+sí</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>-woo</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>-wadda</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-mawoo</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-bari</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-ya/ -gowaaya</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Hungarian</td>
<td>-ó</td>
<td>(+)</td>
<td>(+)</td>
<td>–</td>
</tr>
<tr>
<td>-andó/ -endó</td>
<td>–</td>
<td>(+)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-vá/-vé</td>
<td>–</td>
<td>(+)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-vén</td>
<td>–</td>
<td>(+)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Japanese</td>
<td>-te/-de/ -ite</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-i/-Ø</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Itelmen</td>
<td>INF (various forms)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Behavioural potential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verbal categories</td>
<td>Nominal categories</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
<td>M</td>
</tr>
<tr>
<td>Koasati</td>
<td>NMLZ (various forms)</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-sáya</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>-yólli</td>
<td>-</td>
<td>(+)</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>-ka</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>-kítta</td>
<td>-</td>
<td>(+)</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>-laho:li:sáya</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>-n</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-k</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-t</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Basque</td>
<td>-t(że)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-tu/-du/-i/-Ø</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-(+)</td>
</tr>
<tr>
<td>Bambara</td>
<td>-le/-ne</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-tò</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>SENT</td>
</tr>
<tr>
<td>Georgian</td>
<td>sa- (-el/-al/r)</td>
<td>+</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>Polish</td>
<td>INF (various forms)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-c</td>
<td>(+)</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-any etc.</td>
<td>(+)</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(PST. PASS) PTC (various forms)</td>
<td>(+)</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PRS.PL-c-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Buru-shaski</td>
<td>-ásas</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>-im/-um/-am</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>n-STEM</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>-(a)n</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-(+)</td>
</tr>
<tr>
<td>Lavu-kaleve</td>
<td>-e/-i</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Behavioural potential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verbal categories</td>
<td>Nominal categories</td>
<td>Argument expression</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
<td>M</td>
</tr>
<tr>
<td>Alamblak</td>
<td>ind/Ø</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Wambon</td>
<td>Ø</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>-mo/-o</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>fu/fi/Ø</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Babungo</td>
<td>k[i]/Ø</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Nama</td>
<td>-se/l'aa/tsii/</td>
<td>+/-</td>
<td>+/-</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>tà + NMLZ</td>
<td>(+)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Mandarin C.</td>
<td>-zhe</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Tamil</td>
<td>-atu</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>-(kk)a</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kisi</td>
<td>Ø</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Garo</td>
<td>-a-ni</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>-na</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>-kan/kan-a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>-/-e-min/-e-r</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Krongo</td>
<td>m/-n-</td>
<td>+</td>
<td>+/-</td>
<td>–</td>
</tr>
<tr>
<td>Nivkh</td>
<td>r/-t/-i/-n</td>
<td>–</td>
<td>+</td>
<td>+/–</td>
</tr>
</tbody>
</table>
Table 6.7: Behavioural potential of type 3 D-ALT DCs

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Verbal categories</th>
<th>Nominal categories</th>
<th>Argument expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>Tagalog</td>
<td>pag-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kharia</td>
<td>RDP/Ø</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-na</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-al</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-ker etc.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kambera</td>
<td>pa-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ø (NMLZ)</td>
<td>+/-</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>Samoan</td>
<td>=ga</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Ø (NMLZ)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Turkish</td>
<td>-DIK/-(y)AcAK</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-mA</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-n-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-Thirri-n</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ma’di</td>
<td>-lɛ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-dʒɔ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ka</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td>-ás/-és</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-ni</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-ótt</td>
<td>(+)</td>
<td>(+)</td>
<td>-</td>
</tr>
<tr>
<td>Japanese</td>
<td>no/mono etc.</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ø (RSP)</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lango</td>
<td>-(kk)ɔ</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ket</td>
<td>Ø (bare INF)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### Language Structural Coding

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Behavioural potential</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Verbal categories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W</td>
</tr>
<tr>
<td>Georgian</td>
<td>-a</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>m- (-a-) (-el/-all)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-ul/-il/m- -ar/-al</td>
<td>-</td>
</tr>
<tr>
<td>Abkhaz</td>
<td>-ra</td>
<td>-</td>
</tr>
<tr>
<td>Polish</td>
<td>-nie</td>
<td>-</td>
</tr>
<tr>
<td>Alamblak</td>
<td>-nef</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-(kfrē)t</td>
<td>-</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>-n/-an</td>
<td>-</td>
</tr>
<tr>
<td>Hdi</td>
<td>tā</td>
<td>-</td>
</tr>
<tr>
<td>Garo</td>
<td>-a</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-gip-a</td>
<td>-</td>
</tr>
<tr>
<td>Krongo</td>
<td>-(t)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>m/-m</td>
<td>+</td>
</tr>
<tr>
<td>Hixkan-ryana</td>
<td>ni/-thi/-hito + rî</td>
<td>+/-</td>
</tr>
<tr>
<td>W. Greenlandic</td>
<td>-niq</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ta/-sa</td>
<td>(+)</td>
</tr>
</tbody>
</table>

### 6.4 An integrated DC typology: Functional possibilities of three structural DC types

In this section, the typologies developed in sections 6.2 and 6.3 are integrated. Below, the functional possibilities of the DCs in the sample languages are presented in three separate tables, one for each of the structural DC types. The functional distribution of type 1 balanced clauses is presented in Table 6.8; the functional distribution of type 2 D-SENT DCs in Table 6.9; and the functional distribution of type 3 D-ALT DCs in Table 6.10.
<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Functional possibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagalog</td>
<td>na/-ng</td>
<td>+</td>
</tr>
<tr>
<td>Kharia</td>
<td>no</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>gam-kon</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>je etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a/i etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Kambera</td>
<td>wà</td>
<td>+</td>
</tr>
<tr>
<td>Samoan</td>
<td>-e</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Guarani</td>
<td>hâ(gew)/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>va</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>vo</td>
<td>+</td>
</tr>
<tr>
<td>Santali</td>
<td>CORREL</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>mente</td>
<td>+</td>
</tr>
<tr>
<td>Warao</td>
<td>kotai</td>
<td>+</td>
</tr>
<tr>
<td>Turkish</td>
<td>ki</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>diye</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-ntha</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Paiwan</td>
<td>tu(a)/tjai</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>a parhu</td>
<td>+</td>
</tr>
<tr>
<td>Ma’di</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>Ø (+RSP)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hungarian</td>
<td>hogy (etc.)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>REL.PRON</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>úgy (-a-hogy)</td>
<td>+</td>
</tr>
<tr>
<td>Japanese</td>
<td>no/mono etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (RSP)</td>
<td>+</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>qhov</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>kuam/(has)tas</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>kws</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Functional possibilities</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Lango</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Itelmen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Itelmen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Itelmen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basque</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basque</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basque</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bambara</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bambara</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georgian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georgian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georgian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bukiyip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bukiyip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bukiyip</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Functional possibilities</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Abkhaz</td>
<td>-NFIN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>REL.PRON + -NFIN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ş- + -NFIN</td>
<td></td>
</tr>
<tr>
<td>Burushaski</td>
<td>ke/ki</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>sén-/ét-</td>
<td>+</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>PERSON.SUFF</td>
<td></td>
</tr>
<tr>
<td>Pipil</td>
<td>ka(h)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ne</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ke</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ke:n-aken</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ADV</td>
<td></td>
</tr>
<tr>
<td>Wambon</td>
<td>-e</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-a (+ o)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ka</td>
<td></td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>DET(+DEM)</td>
<td>+</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>bifi/dati/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Wh/Ø</td>
<td></td>
</tr>
<tr>
<td>Babungo</td>
<td>lāa</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>fāñ/yūu</td>
<td></td>
</tr>
<tr>
<td>Nama</td>
<td>!xāis-à (no INDIC)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (no INDIC) (RSP)</td>
<td>+</td>
</tr>
<tr>
<td>Mandarin Chinese</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>de</td>
<td></td>
</tr>
<tr>
<td>Kisi</td>
<td>(m)àà/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td></td>
</tr>
<tr>
<td>Nung</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(ti-va) (+DEM) (+FOC)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>bɑt</td>
<td></td>
</tr>
<tr>
<td>Garo</td>
<td>in-e</td>
<td>+</td>
</tr>
<tr>
<td>Krongo</td>
<td>ànî tìŋ</td>
<td>+</td>
</tr>
<tr>
<td>Slave</td>
<td>nj/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>gha/gú</td>
<td></td>
</tr>
<tr>
<td></td>
<td>i/sji/lii</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gháré</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Functional possibilities</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Nivkh</td>
<td>-vut/-vur</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>NFIN</td>
<td>+</td>
</tr>
<tr>
<td>West Greenlandic</td>
<td>PTC mood</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>CONT mood</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 6.9: Functional possibilities of type 2 D-SENT DCs

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Functional possibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pred Head</td>
<td>Ref Head</td>
</tr>
<tr>
<td>Kharia</td>
<td>-na-wala</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ga</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ta + RDP</td>
<td>+</td>
</tr>
<tr>
<td>Kambera</td>
<td>pa-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ma-</td>
<td>+</td>
</tr>
<tr>
<td>Samoan</td>
<td>ona/ina</td>
<td>+</td>
</tr>
<tr>
<td>Santali</td>
<td>Ø (no INDIC)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-kate</td>
<td>+</td>
</tr>
<tr>
<td>Warao</td>
<td>-kitane</td>
<td>+</td>
</tr>
<tr>
<td>Quechua</td>
<td>-ji/-shka/-na</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-chun</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ngapaj</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-y</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-shpa</td>
<td>+</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-n-garrba</td>
<td>+</td>
</tr>
<tr>
<td>Paiwan</td>
<td>(-)in- -an + a</td>
<td>+</td>
</tr>
<tr>
<td>Turkish</td>
<td>-mAK</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-An</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(y)ArAk</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(y)A…-(y)A</td>
<td>+</td>
</tr>
<tr>
<td>Ma’di</td>
<td>-lɛ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-dʒɔ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-reʔ/-bəd</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-zɛɾ+ si</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Functional possibilities</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>-woo</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-wadda</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-mawoo</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-bari</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ya/-gowaaya</td>
<td>+</td>
</tr>
<tr>
<td>Hungarian</td>
<td>-ó</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-andó/-endó</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-vá/-vé</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-vén</td>
<td>+</td>
</tr>
<tr>
<td>Japanese</td>
<td>-te/-de/-ite</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-i/-Ø</td>
<td>+</td>
</tr>
<tr>
<td>Itelmen</td>
<td>INF (various forms)</td>
<td>+</td>
</tr>
<tr>
<td>Koasati</td>
<td>NMLZ (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-sáya</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-yólli</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ka</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-kíutta</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-laho:lisáya</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-n</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-k</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-t</td>
<td>+</td>
</tr>
<tr>
<td>Basque</td>
<td>-t(z)e</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-tu/-du/-i/-Ø</td>
<td>+</td>
</tr>
<tr>
<td>Bambara</td>
<td>-le/-ne</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-tò</td>
<td>+</td>
</tr>
<tr>
<td>Georgian</td>
<td>sa- (-el/-al/r)</td>
<td>+</td>
</tr>
<tr>
<td>Polish</td>
<td>INF (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-c</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-any etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(PST.PASS) PTC (various forms)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>PRS.PL-c-</td>
<td>+</td>
</tr>
<tr>
<td>Burushaski</td>
<td>-á(as)</td>
<td>+ +</td>
</tr>
<tr>
<td></td>
<td>-im/-um/-am</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>n-STEM-(a)n</td>
<td>+</td>
</tr>
</tbody>
</table>
## Chapter 6 – Dependent Clauses in the Languages of the Sample

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Functional possibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pred</td>
<td>Ref</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>-e/-i</td>
<td>+</td>
</tr>
<tr>
<td>Alamblak</td>
<td>ind/Ø</td>
<td>+</td>
</tr>
<tr>
<td>Wambon</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-mo/-o</td>
<td>+</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>fu/fi/Ø</td>
<td>+</td>
</tr>
<tr>
<td>Babungo</td>
<td>ki(i)/Ø</td>
<td>+</td>
</tr>
<tr>
<td>Nama</td>
<td>-se/l’aa/tsii no INDIC</td>
<td>+</td>
</tr>
<tr>
<td>Hdi</td>
<td>tà + NMLZ</td>
<td>+</td>
</tr>
<tr>
<td>Mandarin C.</td>
<td>-zhe</td>
<td>+</td>
</tr>
<tr>
<td>Tamil</td>
<td>-atu</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(kk)a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-a</td>
<td>+</td>
</tr>
<tr>
<td>Kisi</td>
<td>Ø</td>
<td>+</td>
</tr>
<tr>
<td>Garo</td>
<td>-a-ni</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-na</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-kan/-kan-a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-e/-e-min/-e-r</td>
<td>+</td>
</tr>
<tr>
<td>Krongo</td>
<td>-(t)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>m-/n-</td>
<td>+</td>
</tr>
<tr>
<td>Nivkh</td>
<td>r/-t-f/-n</td>
<td>+</td>
</tr>
</tbody>
</table>

**Table 6.10: Functional possibilities of type 3 D-ALT DCs**

<table>
<thead>
<tr>
<th>Language</th>
<th>Structural Coding</th>
<th>Functional possibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pred</td>
<td>Ref</td>
</tr>
<tr>
<td>Tagalog</td>
<td>pag-</td>
<td>?</td>
</tr>
<tr>
<td>Kharia</td>
<td>RDP/Ø</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-na</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-al</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ker etc.</td>
<td>+</td>
</tr>
<tr>
<td>Kambera</td>
<td>pa-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (NMLZ)</td>
<td>+</td>
</tr>
<tr>
<td>Samoan</td>
<td>=ga</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (NMLZ)</td>
<td>+</td>
</tr>
<tr>
<td>Language</td>
<td>Structural Coding</td>
<td>Functional possibilities</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pred Head</td>
</tr>
<tr>
<td>Turkish</td>
<td>-DİK/-(y)AcAK</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-mA</td>
<td>+</td>
</tr>
<tr>
<td>Kayardild</td>
<td>-Thirri-n</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-n-</td>
<td>+</td>
</tr>
<tr>
<td>Ma’di</td>
<td>-lɛ</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-dʒo</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ka</td>
<td>+</td>
</tr>
<tr>
<td>Hungarian</td>
<td>-ás/-és</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ni</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-őtt</td>
<td>+</td>
</tr>
<tr>
<td>Japanese</td>
<td>no/mono etc.</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Ø (RSP)</td>
<td>+</td>
</tr>
<tr>
<td>Lango</td>
<td>-(kk)ọ</td>
<td>+</td>
</tr>
<tr>
<td>Ket</td>
<td>Ø (bare INF)</td>
<td>+</td>
</tr>
<tr>
<td>Georgian</td>
<td>-a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>m- (-a-)(-el/-al))</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ul/-il/m- -ar/-al</td>
<td>+</td>
</tr>
<tr>
<td>Abkhaz</td>
<td>-ra</td>
<td>+</td>
</tr>
<tr>
<td>Polish</td>
<td>-nie</td>
<td>+</td>
</tr>
<tr>
<td>Alamblak</td>
<td>-nef</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-(kfë)t</td>
<td>+</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>-n/-an</td>
<td>+</td>
</tr>
<tr>
<td>Hdi</td>
<td>tá</td>
<td>+</td>
</tr>
<tr>
<td>Garo</td>
<td>-a</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-gip-a</td>
<td>+</td>
</tr>
<tr>
<td>Krongo</td>
<td>-(t)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>m/-m-</td>
<td>+</td>
</tr>
<tr>
<td>West Greenlandic</td>
<td>-niq</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-ta/-sa</td>
<td>+</td>
</tr>
</tbody>
</table>
6.5 Summary, outlook

In this chapter, the dependent clause constructions of the sample languages were identified on the basis of their structural coding and classified in two ways: according to their functional possibilities (section 6.2) and according to their behavioural potential (section 6.3). These two typologies were then presented in an integrated fashion in section 6.4.

In Chapter 7, the DC typology presented in the present chapter and the PoS typology presented in Chapter 5 are combined in order to investigate dependency relations between the functional possibilities of PoS classes and of different structural types of DC constructions in the sample languages.
Results

7

7.1 Introduction
This chapter presents the results for the hypotheses formulated in Chapter 4. First, in section 7.2, I discuss the results for the most general hypothesis, namely that there is a global match between flexibility versus rigidity in the domains of PoS and DCs. This prediction is tested first considering DCs as a single group of constructions. Subsequently, the same prediction is assessed while distinguishing between the three structural DC types (see Chapter 3 section 3.4 and Chapter 6, section 6.3.3).

Second, in section 7.3, the functional relation between PoS and DCs is investigated in terms of specific matches, i.e. one-to-one matches between particular types of PoS classes and DC constructions, as regards the set of propositional functions that they can express. As in the case of global matches, the predictions concerning specific functional matches will be tested first considering DCs as an undifferentiated group, and then looking at the three structural DC types separately.

As explained in Chapter 4, section 4.4, the results are presented as frequency counts in 2x2 contingency tables, which are submitted to Fischer’s Exact tests in order to identify dependency relations between pairs of parameters. One parameter of such a pair relates to PoS, and the other to DCs. Whenever a significant dependency is found, its nature will be further specified using the method proposed by Maslova (2003).
It is shown that significant dependency relations exist between flexible rather than rigid PoS and DCs; between very flexible rather than less flexible PoS and DCs; and between PoS and deranked DCs, as opposed to balanced DCs. These dependencies are asymmetrical in nature, and can as such be interpreted as statistical bases for implicational universals.

Section 7.4 provides a brief summary of the overall results. The language data that pertain to the analyses presented in the present chapter are summarized in a set of tables in the final section (7.5).

7.2 Global functional matches: Flexibility versus rigidity in the PoS and DC domains

7.2.1 Global matching without differentiating for structural DC-types

7.2.1.1 Introduction
In this section I test the hypothesis that there is a global match between the functional characteristics of the PoS system of a language and its DC constructions, in terms of flexibility versus rigidity. In Chapter 4 this hypothesis was operationalized in the form of a two-fold prediction A1/A2, repeated here in (1a-1b):

Predictions A1/A2:

(i) a. If a particular language has one or more flexible DCs, then it should also have one or more flexible PoS classes.
   b. If a language has rigid DCs only, then it should also have rigid PoS classes only.

7.2.1.2 Global match for flexible constructions
Starting with the results for flexible constructions, Table 7.1 below presents the frequency counts for Prediction A1. This table (and similar ones in the remainder of this chapter) should be read as follows: The two grammatical parameters under investigation appear in the leftmost column and in the top row, respectively. In this case, i.e. in Table 7.1, these parameters concern the presence/absence of one or more flexible PoS class(es) and the presence/absence of one or more flexible DC construction(s). Thus, each of these parameters represents a variable with two possible values: Either the relevant trait is attested in the language (+), or it is not attested (-). Together, these two binary parameters yield four possible value combinations: +/+ , +/− ,
The numbers of languages in the sample that exhibit each of these four value combinations are presented in the four shaded cells in Table 7.1. Each of these shaded cells belongs to a set of two cells, the upper (shaded) one of which represents the observed frequency, and the lower (transparent) one the frequency that would be expected if the pattern of co-occurrence of the two grammatical traits in question were purely coincidental. If, however, the distribution is not co- incidental, then the observed frequencies in the two darker shaded cells are higher than the expected frequencies, whereas the observed frequencies in the lighter shaded cells are lower than the expected ones. As explained in Chapter 4, Fischer’s Exact yields a p-value that specifies how likely it is for the observed distribution to be the result of chance; the critical value is p<0.05 (see Cysouw 2003: 91). In Table 7.1, the observed and expected counts deviate in the predicted way, but not very much; p = 0.225 (2-sided), which means that there is no significant correlation.

Table 7.1: Frequencies for languages with/without (one or more) flexible PoS class(es) and DC(s)

<table>
<thead>
<tr>
<th>Flex PoS</th>
<th>Flex DCs (total)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>–</td>
<td>total</td>
<td></td>
</tr>
<tr>
<td>Flex PoS</td>
<td>observed freq</td>
<td>10</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Flex PoS</td>
<td>expected freq</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>observed freq</td>
<td>11</td>
<td>18</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>21</td>
<td>25</td>
<td>46</td>
<td></td>
</tr>
</tbody>
</table>

It should be noted that the total number of 17 languages with (one or more) flexible PoS class(es) in Table 7.1 includes those languages that have a large, open and/or a productively derived class of flexible lexemes. When looking back at the data in Chapter 5 (Table 5.1), we see that this means that all languages with a small class of modifiers (Itelmen, Thai, Basque, and Abun) are counted as languages without flexible PoS, while all languages that are listed above Itelmen in Table 5.1 are counted as languages with one or more flexible PoS class(es). Note further that Nunggubuyu and Tuscarora are excluded from all analyses of the present chapter, since they don’t seem to have any DC constructions. Rather, these languages express all subordination relations by means of separate, independent clauses. Therefore, no comparison can be made between the distributional patterns of the PoS classes in these languages and any DC construction. The language
data on which Table 7.1 above is based can be found in Tables 7.33 and 7.34 in section 7.5 at the end of this chapter.

7.2.1.3 Global match for rigid constructions
Consider now Table 7.2 below, which shows the results for prediction A2 (see (1b) above), concerning the correlation between rigid PoS only and rigid DCs only.

As in the case of flexible constructions, no significant correlation was found: Fisher’s exact yields \( p = 0.552 \) (2-sided). The relevant language data can again be found in Tables 7.33 and 7.34 of section 7.5.

<table>
<thead>
<tr>
<th>Rig PoS</th>
<th>Rig DCs (total)</th>
<th>+</th>
<th>−</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>15</td>
<td>14</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>14</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>observed freq</td>
<td>7</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>22</td>
<td>24</td>
<td>46</td>
</tr>
</tbody>
</table>

7.2.1.4 Summary
In its most general form, the hypothesis of a global match between flexibility versus rigidity in the domains of PoS and DCs is rejected. In the following section this hypothesis is refined by taking into account the parameter of structural DC type.

7.2.2 The parameter of structural DC type

7.2.2.1 Introduction
In this section I introduce the parameter of the internal structural properties of DCs. The prediction, as formulated in Chapter 4, is that the more formally similar a DC construction is to a lexical expression, the more functionally similar it will be to this lexical counterpart. In other words, it is expected that functional connections will be stronger between PoS and deranked DCs (type 2,3) than between PoS and balanced DCs (type 1). In addition, within the group of deranked DCs, it is expected that more deranked DCs (type 3, D-ALT) will show more functional similarity with PoS than less deranked
DCs (type 2, D-SENT). In Chapter 4 this was operationalized in the form of the two predictions B1 and B2, which are repeated in (2a-b) below:

Prediction B1/B2:

(2)  a. The functional possibilities of deranked DCs (type 2/3) are more similar to the functional possibilities of PoS than those of balanced DCs (type 1).
   b. Within the group of deranked DCs, the functional possibilities of type 3 DCs are more similar to the functional possibilities of PoS than those of type 2 DCs

These predictions will now be combined with the predictions concerning global matches between flexible versus rigid PoS systems and DC constructions.

7.2.2.2 Global match of flexible constructions, differentiating for structural DC type

Applying predictions B1 and B2 to flexible PoS and DC systems yields a set of four hypotheses, which are given in (3a-d) below (cf. (3) of Chapter 4). It is expected that the prediction in (3a) is less likely to hold true than the one in (3b). Furthermore, the prediction in (3c) is less likely to be confirmed than the one in (3d).

(3)  a. If a language has one or more flexible balanced DC(s) of type 1, then it should also have one or more flexible PoS class(es).
   b. If a language has one or more flexible deranked DC(s) of type 2/3, then it should also have one or more flexible PoS class(es).
   c. If a language has one or more flexible deranked DC(s) of type 2, then it should also have one or more flexible PoS class(es).
   d. If a language has one or more flexible deranked DC(s) of type 3, then it should also have one or more flexible PoS class(es).

The distributions pertaining to the predictions in (3a-d) are presented in Tables 7.3, 7.4, 7.6, and 7.7 below. The language data can be found in Tables 7.33 and 7.34 of section 7.5.

First, the frequencies relevant for the prediction in (3a), concerning balanced DCs, appear in Table 7.3. This table shows that there is hardly any difference between the observed and the expected distribution; Fischer’s exact yields $p = 0.755$ (2-sided).
Table 7.3: Frequencies for languages with/without flexible PoS and flexible DCs of type 1

<table>
<thead>
<tr>
<th>Flex PoS</th>
<th>observed freq</th>
<th>expected freq</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>7</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>-</td>
<td>10</td>
<td>11</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 7.4: Frequencies for languages with/without flexible PoS and flexible DCs of type 2/3

<table>
<thead>
<tr>
<th>Flex PoS</th>
<th>observed freq</th>
<th>expected freq</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>7</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>-</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

Second, the frequencies pertaining to the prediction in (3b), concerning deranked DCs, are presented in Table 7.4. As expected, this distribution does yield a significant correlation: \( p = 0.025 \) (2-sided). The contingency coefficient (CC) is 0.339, which suggests a moderately strong effect (see Chapter 4, section 4.4).

Table 7.5a: Maslova test no. 1: change DC parameter to PoS=DC

<table>
<thead>
<tr>
<th></th>
<th>Flex Pos +</th>
<th>Flex Pos –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>

Following Maslova (2003; and see Chapter 4, section 4.4), the PoS parameter and the DC parameters are each separately correlated to the derived parameter PoS = DC (i.e. the parameter of PoS and DC having the same or different values), in order to assess the nature of the dependency relation identified in Table 7.4. This yields Tables 7.5a and 7.5b:
Tables 7.5a and 7.5b indicate a one-sided asymmetrical dependency between flexible PoS and deranked flexible DCs. In particular, Table 7.5a shows that the negative value of the PoS parameter constrains the event of the PoS and DC parameters having the same value\(^{102}\). In contrast, Table 7.5b shows that the value for the DC parameter and the event of the PoS and DC parameters having the same value are completely independent of each other.

This outcome is interpreted as the statistical counterpart of the typical situation encountered with implicational universals, namely a 2x2 contingency table with exactly one (nearly) empty cell. In the case at hand, this is the cell with the value combination [-Flex PoS, + Flex DC type 2/3]. The observed frequency in this cell (3) is both low and significantly lower than the expected frequency (6). This constitutes statistical evidence for the universal in (4):

\[
(4) \quad [- \text{Flex PoS}] \rightarrow [- \text{Flex DC 2/3}]
\]

Languages without any flexible PoS class(es) also lack flexible deranked DC(s) of type 2/3 (independent of whether languages with flexible PoS class(es) have flexible deranked DC(s) of type 2/3).

Relating this result to the prediction in (3b), we can say that indeed the availability of flexible DC constructions of type 2/3 in a language is dependent on the availability of at least one flexible PoS class(es) in that language.

I turn now to the predictions in (3c) and (3d) above, differentiating within the group of deranked flexible DC between DCs of type 2 (D-SENT) and DCs of type 3 (D-ALT). First, the frequencies pertaining to the prediction in (3c), concerning type 2 DCs, appear in Table 7.6. They do not show any correlation: the observed and expected counts are identical, so that Fischer’s Exact yields \( p = 1 \).

\(^{102}\) Note that the pattern in Table 7.5a pattern differs from the one in Table 4.9a of Chapter 4 (the hypothetical illustration of a one-sided asymmetrical dependency relation), to the extent that in the latter case it is the positive rather than the negative value of the PoS parameter for which the distribution of flexible DCs is skewed.
At first sight, this result is unexpected, especially in view of the universal in (4). However, this finding can be explained by the fact that there are hardly any flexible DCs of type 2 attested in the sample. That is to say, when comparing Table 7.6 with Table 7.4, it can be seen that only 2 out of the total 10 flexible deranked DCs involve a construction of type 2; all others are of type 3. Therefore, it is not surprising that we do not find a significant correlation when testing the prediction in (3c).

On the other hand, and as expected, we do find a significant correlation when testing the prediction in (3d), which is concerned with flexible DCs of type 3. The relevant frequencies are presented in Table 7.7 below, which differs only very slightly from Table 7.4 above. They yield a significant p-value of 0.038. The CC value is 0.340, which indicates a moderately strong effect.

When applying Maslova’s method to these data, again a one-sided asymmetrical dependency is revealed, as the data in Table 7.8a and 7.8b show:

### Table 7.6: Frequencies for languages with/without flex PoS and flex DCs of type 2

<table>
<thead>
<tr>
<th>Flex PoS</th>
<th>Flex DCs type 2 (D-SENT)</th>
<th>+</th>
<th>−</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>1</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>1</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>observed freq</td>
<td>1</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>1</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>2</td>
<td>44</td>
<td>46</td>
</tr>
</tbody>
</table>

### Table 7.7: Frequencies for languages with/without flex PoS and flex DCs of type 3

<table>
<thead>
<tr>
<th>Flex PoS</th>
<th>Flex DCs type 3 (D-ALT)</th>
<th>+</th>
<th>−</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>6</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>3</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>observed freq</td>
<td>2</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>5</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>8</td>
<td>38</td>
<td>46</td>
</tr>
</tbody>
</table>

When applying Maslova’s method to these data, again a one-sided asymmetrical dependency is revealed, as the data in Table 7.8a and 7.8b show:
Table 7.8a: Maslova test no. 1: change DC parameter to PoS=DC

<table>
<thead>
<tr>
<th></th>
<th>Flex Pos +</th>
<th>Flex Pos −</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>

*p = 0.000* (significant)

Table 7.8b: Maslova test no. 2: Change PoS parameter to PoS=DC

<table>
<thead>
<tr>
<th></th>
<th>Flex DC3 +</th>
<th>Flex DC3 −</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>2</td>
<td>11</td>
</tr>
</tbody>
</table>

*p = 1* (not significant)

Table 7.8a shows that the negative value of the PoS parameter constrains the event of the PoS and DC parameters having the same value, while Table 7.8b shows that the DC parameter and the PoS=DC parameter do not interact. This is interpreted as evidence for the universal in (5), which is in fact a more precise version of (4) above.

(5)  
\[-\text{Flex PoS}] \xrightarrow{\text{a}} [-\text{Flex DC 3}]

Languages without any flexible PoS class(es) also lack flexible deranked DC(s) of type 3 (independent of whether languages with flexible PoS class(es) have flexible DC(s) of type 3).

In sum, these data show that taking into account structural DC type reveals a dependency relation between flexible PoS and flexible deranked DCs. In particular, flexible deranked DCs hardly ever occur in languages without any flexible PoS class(es). However, this does not mean that all languages with a flexible PoS class also have a flexible deranked DC construction.

7.2.2.3 Global match for rigid constructions, differentiating for structural DC type

In this section I turn back to prediction A2 about rigid constructions (see (1b) above), and combine it with the parameter of structural DC type (see B1/B2 in (2a,b) above). Recall the definitions of languages with flexible versus rigid constructions: Whereas flexible PoS and DC systems are defined as those that have at least one flexible construction, rigid PoS and DC systems consist of rigid constructions only. It was explained that, due to this asymmetry,
it does not make sense to formulate specific predictions about correlations between languages with rigid PoS only and rigid DCs of a specific structural type only. This is because there is no reasonable basis on which to expect languages to have only a single structural type of rigid DCs. On the contrary, it is expected that semantically different types of subordination relations will be expressed by DC constructions with different internal morpho-syntactic properties. Indeed, as can be seen in Table 7.2 in section 7.5, there are only a few languages in the sample that have both rigid PoS only and rigid DCs only, and in which all DCs are of a single structural type. The languages that display such a pattern are Abun, Bukiyip, Tamil, and Hixkaryana. All other languages with rigid PoS only and rigid DCs only (11 in total) show a mixture of two or three structurally different DC types.

Nevertheless, we may investigate which structural DC types occur in groups of languages with different values for the parameters of rigid PoS only and rigid DCs only. Specifically interesting are languages with unexpected value combinations, i.e. languages that do not show a match between the functional possibilities of PoS and DCs. Such cases may involve (i) languages with one or more flexible PoS class(es) and with rigid DCs only, and (ii) languages with rigid PoS only and with one or more flexible DC construction(s).

Consider once more the data in Table 7.2, repeated below as Table 7.2’ for convenience, which presents the frequencies for languages with and without rigid PoS and DCs only (without differentiating for structural DC type).

<table>
<thead>
<tr>
<th>Rig PoS</th>
<th>Rig DCs (total)</th>
<th>+</th>
<th>–</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>15</td>
<td>14</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>14</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>observed freq</td>
<td>7</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>22</td>
<td>24</td>
<td>46</td>
</tr>
</tbody>
</table>

Languages with flexible PoS systems and rigid DCs only

I start with a discussion of those languages that display the unexpected feature combination mentioned under (i) above: one or more flexible PoS
class(es) and rigid DCs only ([- Rig PoS only; + Rig DCs only]). Table 7.2’ shows that there are 7 languages of this type. They are: Guarani, Warao, Hungarian, Japanese, Hmong Njua, Lango, and Koasati. Considering Table 7.33 in section 7.5, we see that in some of these cases the rigid DCs in question are balanced (type 1). Even though the attestation of these constructions contradicts the prediction of a global mapping of rigid PoS and DCs, this result is relatively un-surprising, since balanced DCs were predicted to be least likely to show functional similarity to PoS.

What is apparently more surprising is the attestation of a number of deranked (type 2 or 3) rigid DC constructions in the group of languages with flexible PoS systems. However, in almost all of these cases the rigid deranked DC can be regarded as the clausal counterpart of a rigid PoS class (possibly a derived or a small class) that forms part of a system that also includes one or more flexible PoS class(es). Therefore, these cases in fact confirm the prediction that PoS and (deranked) DCs exhibit parallel functional patterns. In what follows, I will discuss in more detail the cases of languages with flexible PoS systems and rigid (balanced or deranked) DCs only.

Guaraní and Warao are both languages with a class of very flexible lexemes: they have contentives and non-verbs, respectively. Guaraní does not have any relevant extra rigid PoS classes to which its rigid DCs can be linked. However, all rigid DCs in this language are balanced (type 1), and as such not very likely to behave functionally like PoS. Warao, in contrast, has a balanced rigid relative clause construction and a deranked rigid complement construction of type 2. While the attestation of the former may again be explained in terms of lack of formal (and therefore functional) similarity with PoS, the presence of the latter DC can be related to the a class of derived nouns in Warao. This class of nouns was not taken into the account in the original analysis, since – in order to avoid double values – classes of simple, un-derived items are given prominence over classes of derived items in cases where both are available to express a particular propositional function (in this case the function of head of a referential phrase).

The other languages with flexible PoS systems and rigid DCs only (Hungarian, Japanese, Hmong Nua, Lango, and Koasati) have either a class of nominals or a class of modifiers. Hungarian, for instance, has lexical nominals and, contrary to the predictions, rigid nominalizations and participle constructions of types 2 and 3. A possible explanation for this counterexample may again be sought in the realm of derivational morphology. For one thing, the form -ás/-és, which is used to mark Hungarian rigid nominalizations of
type 3, can also be used for the lexical derivation of action nouns, as in: olt ‘extinguish’ ➔ olt-as ‘extinguishing’. There is thus a functional match between the deranked DC construction and a derived PoS class. Interestingly, one of the rigid participial constructions in Hungarian, namely the active (present) participle form in -ő/-ő, is also productively used to derive lexical nouns, as for example in: olvás ‘read’ ➔ olvás-ő ‘reader’, and ebédel ‘have dinner’ ➔ ebédl-ő ‘dining room’. Thus, even though the participial construction is not flexible, its marker also appears in the function of head of a referential phrase, as a marker of lexically derived nouns. In addition to rigid deranked nominalizations and participles, Hungarian has three balanced rigid DCs: complement clauses, relative clauses, and adverbial manner clauses. The functional rigidity of the former two constructions does not match with the flexibility of lexical nominals, but their balanced form makes them unlikely candidates to exhibit such a match in the first place.

Japanese also has nominals, but unlike Hungarian it has a class of simple rigid nouns as well. Therefore, the rigid complement clauses attested in Japanese indeed have a lexical counterpart, and as such do not constitute a counterexample. Lango is a language with two rigid complement clause constructions: a deranked one (of type 3) and a balanced one. These DCs are again not to be regarded counterexamples, since Lango has a class of rigid simple nouns, which serves as the lexical counterpart for the rigid complement clauses. In addition to these constructions, Lango has rigid balanced relative clauses, which do not match with its flexible lexical modifiers. Hmong Njua, similarly, has flexible lexical modifiers and rigid balanced relative clauses. In both cases, the fact that the DC construction is balanced may explain the lack of functional similarity with the relevant PoS class.

Finally, there is one language, Koasati, the PoS system of which is classified as flexible because it contains a derived class of modifiers, but which is otherwise fully rigid, i.e. displays open classes of nouns and verbs, a small class of adjectives, and small and derived classes of manner adverbs. This means that the nominalizations, participial and converbal constructions attested in Koasati, all of which are deranked (type 2), have the expected lexical counterparts.

In sum, there is only one truly problematic case, i.e. one language with a rigid, deranked DC for which no rigid (derived) lexical counterpart is

103 Recall from Chapter 6 that the classification of Japanese nominalizations in terms of structural type is not evident: They were classified as 1/3, because they retain tense, but also have the possibility of expressing the subject as a possessor.
available: Hungarian has rigid deranked participle constructions of type 2 and 3, without (simple or derived) rigid adjectives.

Languages with flexible PoS systems and a mixture of flexible and rigid DCs

Table 7.2 shows that, apart from the group of 7 languages with at least one flexible PoS class and with rigid DCs only, there is a group of 10 languages with at least one flexible PoS that do not have rigid DCs only ([- Rig PoS only; - Rig DCs only]). In principle, this latter value combination is expected. Notably however, 9 out of the 10 relevant languages have one or more rigid (deranked) DCs alongside the expected flexible DCs. In other words, these 9 languages with flexible PoS systems have a mixed system of flexible and rigid DCs.

Consider for instance a language like Ma’di: It has lexical nominals and, as predicted, flexible deranked nominal clauses. The latter construction is illustrated in (6a-b):

Ma’di (Blackings & Fabb 2003: 202, 22)
(6) a. Mâ lè-à [èbî ‘nà-łè] rà
   1sg (N)want-obj fish N-eat-NMLZ AFF
   ‘I certainly want to eat fish.’

   b. àràbìà [sɔ̀ ɔ̀ ɗì-łè] rì pá nà àdî.
   car Opi poss (N)-take-PTC DEF leg AFR deflate
   ‘The car which Opi took has a flat tyre.’

However, in modifier function the lè-construction is used for object relative clauses only. For subject relative clauses, a different rigid deranked relative clause construction (of type 2) is used, as is illustrated in (7):

Ma’di (Blackings & Fabb 2003: 193)
(7) ágà àm-à dʒɔ̀ nì sì-ɓà rì
   man [1PL-POSS house SPEC PRON N-build-PTC(PL)] DEF
   ‘One of the men who built / are building our house.’

Similarly, Turkish has flexible lexical non-verbs and flexible deranked nominal clauses of type 3. The latter are illustrated in (8a-8b):
**Turkish** (Kornfilt 1997: 50)

(8) a. *(ben) [Ahmed-in öl-düğ-ün]-ü duy-du-m*
   
   I Ahmed-GEN die-NMLZ-3SG-ACC hear-PST-1SG
   
   ‘I heard that Ahmed died.’

   b. *[adam-in gi-tığ-i] okul*
   
   man-GEN go-PTC-3SG school
   
   ‘the school that the man goes/went to’

However, Turkish also has a number of rigid deranked DCs. Examples (9) and (10) show the infinitival construction in –mAk (type 2), and the nominalization in –mA (type 3), respectively. Both constructions can be used in the function of head of a referential phrase only. Example (11) illustrates the rigid participle construction marked by –Aн (type 2), which is used exclusively as a modifier in a referential phrase, expressing subject and possessive relative clauses. Finally, example (12) shows the rigid converb construction in –(y)ArAk (also of type 2), which can express only the function of modifier in a predicate phrase.

**Turkish** (Kornfilt 1997: 51, 73; Göksel & Kerslake 2005: 420, 440)

(9) *[Lütfen pencere-yi aç-mağ]-ı unut-ma*

   please window-ACC open-INF-ACC forget-NEG

   ‘Please, don’t forget to open the window!’

(10) *[Kerkes-in birier kikaye anlat-ma-sı]*

   everyone-GEN one.each story tell-NMLZ-3SG.poss

   iste-n-iyor-muş want-PASS-IPFV-EV.COP

   ‘It seems they want [everyone to tell a story].’

(11) *[öğretmen ol-an] Haydar*

   teacher be-PTC Haydar

   ‘Haydar, who is a teacher’

(12) *Ben [etraf-im-a bak-arak] yür-ür-üm*

   I around-1SG-DAT look-CONV walk-AOR-1SG

   ‘I walk looking around (myself).’
In short, almost all languages with a flexible PoS system (i.e. with at least one flexible PoS class) have at least one rigid deranked DC construction; sometimes in combination with flexible deranked DC(s), sometimes not.

In addition, the examples from Turkish point to yet another interesting issue, namely the fact that flexible deranked DCs in languages with flexible PoS sometimes do not show exactly the same distributional pattern as their lexical counterparts, but are rather somewhat less flexible than the relevant PoS class. In particular, Turkish nominal clauses can express two out of the three functions that can be expressed by the lexical non-verbs in this language\(^{104}\). In Chapter 8 I will return to this issue, and suggest that there are good functional reasons for deranked DCs to be less flexible than PoS classes in the same language.

Languages with rigid PoS systems and with flexible DCs

Having discussed languages with flexible PoS systems and rigid (as well as flexible) DCs, I will now consider the second group of languages with an unexpected combination of parameter values: rigid PoS only and one or more flexible DC construction(s) ([+ Rig PoS only; - Rig DCs only]). Table 7.2’ above shows that half of the total amount of languages with rigid PoS only (14 out of 29) do not have rigid DCs only. However, when looking at Table 7.33 in section 7.5, we see that the majority of these cases (11 out of 14) involve balanced DCs\(^{105}\). This suggests again that a lack of formal similarity ties in with a lack of functional similarity between DCs and PoS. In addition, there may be other factors at work. First, the relevant flexible balanced DCs often have the possibility to use some extra morpho-syntactic means (i.e. in addition to their ‘regular’ structural coding) in order to disambiguate the multiple functions in which they can appear (Hengeveld & Van Lier 2008). Second, diachronic developments also seem to play a role. These issues will be discussed further in Chapter 8.

\(^{104}\) In combination with a postposition *gibi* (‘like’), nominal clauses in -DIK can be used as simulative clauses (cf. Chapter 6, section 6.2.2), as illustrated in (i). Note however, that this is not a case of flexibility since -DIK clauses can only function as predicate modifiers in combination with a postposition (see Chapter 6, section 6.2.2.1).

\(^{105}\) There is one language with rigid PoS, Burushaski, which has both balanced (type 1) and deranked (type 2) flexible DCs.
Apart from flexible balanced DCs, there are also three unexpected cases of flexible deranked DC constructions in languages with rigid PoS only (cf. the observed frequency of the ‘nearly empty cell’ in Table 7.4). The relevant cases are attested in Burushaski, Hdi, and Krongo, and I will discuss them one by one.

First, Burushaski has rigid PoS classes to express the functions of head and modifier in a referential phrase, i.e. nouns and adjectives, in combination with a flexible deranked nominal clause of type 2 that can express both these functions. The flexible DC is illustrated in (13a-b).

_Burushaski_ (Anderson 2002: 545, Berger 1998a: 171)

(13) a. buṭ muškil bilá [gőō-liṭir-as]
very difficult be.IV 2-show-INF
'It is very difficult to show (it) to you.'

b. [Chá-aṭe oōo-rvγ-as] bōk
post-superess neg-sit-PTC dog
'a dog which doesn’t sit at its post.'

I have not been able to find a plausible explanation for this counterexample, except that there are some indications that the lexical distinction between nouns and adjectives in Burushaski is also not particularly clear-cut. Lorimer (1935: 102), for instance, characterizes the distinction as “messy” and observes that “nouns borrowed from other languages are in many cases used as adjectives, which seems to show a slowness to appreciate the distinction between noun and adjective”. Some examples of such flexible items would be zo’r ‘power/powerful’, šorom ‘shame/ashamed’, and xstor ‘danger/dangerous’ (cf. Berger 1998a: 78). The Burushaski dictionary (Berger 1998b) also provides some indications of flexible items, such as aasāan ‘easy/easiness’, and ajōono ‘strange/stranger’. Although these data obviously do not suffice as a full explanation of the flexibility displayed by Burushaski -as clauses, at least they put their unexpected distributional pattern into some perspective.

Second, Hdi has a deranked (type 3) nominal clause construction, which is marked by tā. Notably, this marker is not exclusively used for structural coding of DCs; it also combines with non-clausal constituents and as such can have two functions: Either it marks a lexical object, or it is a so-called ‘comment marker’ in a focus construction. When introducing a complement clause, tā is glossed as having the former function, i.e. that of an object marker. This is shown in example (14a). In contrast, as can be seen in (14b),
tá is glossed as a comment marker when marking a relative clause. Thus, the flexibility of this DC construction seems to be due to the general multifunctionality of tá, the morpheme used for its structural coding.

*Hdi* (Frajzyngier & Shay 2002: 480, 406)

(14) a. *Sí tà ɗv-ά-y-xόn [tá ɓliy-ά-mú]*

`PST IPFV want-PO-3PL OBJ leave-GEN-1PL.INCL`

‘They wanted us to leave.’ (lit. They wanted our leaving.)

b. *ghúrùm [tá là-gbúw-i ndá mà xàdík]*

`hole COMM go-D:SO-REF ASSOC in ground`

‘a hole that went deep into the ground’

In Krongo, finally, the flexible deranked DC is a modifier clause; it can be used as a modifier of either a referential or a predicative head. This construction has been exemplified in Chapter 6 (see example (35a-b)). Crucially, Krongo does not have any (flexible or rigid) lexical strategy to express the two modifier functions. This means that it is not possible to compare the distributional pattern of the Krongo modifier clause to that of any PoS class(es). Therefore, this counterexample will be excluded in the analyses of section 7.3, where we will be concerned with specific matches between individual PoS classes and DC constructions.

In short, there are only very few cases of deranked flexible DCs in languages with rigid PoS systems; the large majority of flexible DCs in these languages is balanced.

### 7.2.2.4 Summary

In this section it has been shown that, when considering DCs as an undifferentiated group, there is no global match between flexibility versus rigidity in the domains of PoS and DCs. However, when taking into account the parameter of structural DC type, the following dependency relation is revealed: Deranked flexible DCs are almost completely absent in languages without any flexibility in their PoS system, but these languages quite often do have balanced flexible DCs.

On the other hand, languages with flexible PoS systems exhibit both flexible and rigid deranked DCs. Some of these languages display a combination of flexible and rigid DCs, while others have rigid DCs only. In fact, however, almost all languages with a flexible PoS system have at least one rigid DC construction. These rigid DCs can be deranked or balanced.
7.3 Specific functional matches: Types and amounts of flexibility/rigidity in the PoS and DC domains

7.3.1 Introduction
In this section I investigate whether there are correlations between the availability in a language of specific types of flexible and rigid PoS classes, which can express a particular (set of) propositional function(s), and the availability of DC constructions with the same functional possibilities. It is expected that each secondary DC construction of a specific flexible/rigid type will have a primary lexical counterpart of the same flexible/rigid type. This hypothesis was operationalized in Chapter 4 in the form of Prediction C, repeated in (16) below (cf. (6) in Chapter 4):

Prediction C:
(16) If a language has a DC construction of a specific flexible or rigid type X, then it should also have a PoS class of type X.

In what follows, Prediction C is tested first for specific flexible constructions, and then for specific rigid constructions. In every case, I will first consider DCs as a single group of constructions, and then differentiate them according to internal structural type.

7.3.2 Specific matches for flexible constructions

7.3.2.1 Introduction
Starting with flexible PoS and DC constructions, Prediction C in (16) above can be broken down into three sub-predictions, one for each functional type, as in (17a-c):

(17) a. If a language has contentive and/or multi-functional clauses, then it should also have lexical contentives and/or non-verbs.
   b. If a language has nominal clauses, then it should also have lexical nominals.
   c. If a language has modifier clauses, then it should also have lexical modifiers.

Note that, unlike in Chapter 4 (see (7) in that Chapter), the prediction in (17a) treats two flexible construction types as one: (i) constructions
that can be used in all four functions (lexical contentives and contentive clauses), and (ii) constructions that can be used in all functions except the head of a predicate phrase (lexical non-verbs and multifunctional clauses). This is because, as shown in Chapter 6, contentive clauses are very rare.

No predictions are formulated either that make reference to flexible DC types involving the functional slot for head of a predicate phrase (*predicative clauses, head clauses, and Flex Clauses A, B, and D*, see Chapter 3, section 3.2.3), since these constructions are not attested in the sample languages (see Chapter 6, section 6.2.3.2). Finally, while flexible DCs of type C (that can be used as the head of a referential phrase and the modifier in a predicate phrase) are indeed attested, their lexical equivalent (*Flex PoS C*) is not (see Chapter 5, section 5.3.4). Therefore, no predictions are formulated about constructions with this particular functional pattern.

Each of the predictions in (17) can be combined with the parameter of structural DC type. For example, the prediction in (17a) can be broken down into four sub-predictions, as in (18a-d) (cf. (9a-d) of Chapter 4). As before, it is expected that the prediction in (18a) is less likely to hold than the one in (18b), and that the prediction in (18c) is less likely to hold than the one in (18d).

(18)  a. If a language has balanced contentive and/or multifunctional clauses of type 1, then it should also have lexical contentives and/or non-verbs.

b. If a language has deranked contentive and/or multifunctional clauses of type 2/3, then it should also have lexical contentives and/or non-verbs.

c. If a language has deranked contentive and/or multifunctional clauses of type 2, then it should also have lexical contentives and/or non-verbs.

d. If a language has deranked contentive and/or multifunctional clauses of type 3, then it should also have lexical contentives and/or non-verbs.

In the remainder of this section I present the results for the three predictions in (17a-c) one by one. Each of these three predictions will first be tested

---

106 It is noteworthy that there are thus DC types without a functional lexical counterpart. However, flexible DCs of type C are infrequently attested (N=3), and involve either balanced clauses (in Basque and West Greenlandic) or an infinitive construction that is reduplicated in the function of modifier in a predicate phrase (Imbabura Quechua).
considering DCs as a single group, and then differentiating for structural DC types, along the lines of (18a–d).

7.3.2.2 Flexible match 1: Lexical contentives/non-verbs and contentive/multi-functional clauses

Consider first the results for the prediction in (17a). This prediction involves the most pervasively flexible construction types, namely lexical contentives/non-verbs and contentive/multi-functional clauses. The relevant frequencies are given in Table 7.9. As expected, they reveal a significant correlation: $p = 0.037$ (2-sided). The CC value is 0.351, indicating a moderately strong effect. The relevant language data are summarized in Tables 7.35 and 7.36 of section 7.5.

<table>
<thead>
<tr>
<th>PoS contentives/non-verbs</th>
<th>DCs: contentive/multi-functional clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq 3</td>
</tr>
<tr>
<td></td>
<td>expected freq 1</td>
</tr>
<tr>
<td>−</td>
<td>observed freq 3</td>
</tr>
<tr>
<td></td>
<td>expected freq 5</td>
</tr>
<tr>
<td>total</td>
<td>6</td>
</tr>
</tbody>
</table>

Applying Maslova’s method to the data in Table 7.9 reveals a two-sided asymmetrical dependency, as Tables 7.10a and 7.10b make clear:

<table>
<thead>
<tr>
<th>PoS = DC</th>
<th>Pos contentive/non-verb +</th>
<th>Pos contentive/non-verb –</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

$p = 0.006$ (significant)
Table 7.10b: Maslov test no. 2: change PoS parameter to PoS=DC

<table>
<thead>
<tr>
<th></th>
<th>DC contentive/multi-functional +</th>
<th>DC contentive/multi-functional -</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS = DC</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>PoS ≠ DC</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

p = 0.037 (significant)

The significant correlations in Tables 7.10a and 7.10b show that the negative values of both the PoS and the DC parameters constrain the event of the PoS and DC parameters having the same value 107. These results are interpreted as evidence for the universal in (19), which is in fact a pair of two implicational universals:

(19)   [- PoS contentive/non-verb] ⇔ [- DC contentive/multi-functional]

Languages without lexical contentives or non-verbs also lack contentive or multi-functional clauses, and vice versa (languages without contentive or multi-functional clauses also lack lexical contentives and non-verbs).

Notably however, the original distribution in Table 7.9 makes clear that the PoS and the DC parameters under investigation, involving the availability of maximal flexibility, are both strongly skewed towards the negative value. In other words, lexical contentives/non-verbs and contentive/multi-functional clauses are rare phenomena, independently of each other. This is why there are three nearly empty cells in Table 7.9, rather than exactly one nearly empty cell as would be expected in the case of a ‘classical’ implicational universal. This means that the implicational universals in (19), which have negative values in the consequent parts, are not particularly informative.

A much more interesting generalization that can be made on the basis of Table 7.9 is that the occurrence of a positive value on the PoS parameter strongly increases the likelihood that the DC parameter will also have a positive value, even though this likelihood is still only about 50%. In other words, while maximally flexible PoS classes are rare, when they are attested in a language, this dramatically increases the chances that the same type of flexibility will be attested in the DC system, resulting in a (nearly) even

107 Note that this pattern differs again from the hypothetical one for a two-sided asymmetrical dependency relation as illustrated in Chapter 4, Tables 4.11a-b. In the latter case, the positive value of the PoS parameter (Table 4.11a) and the negative value of the DC parameter (Table 4.11b) constrain the PoS = DC parameter.
distribution. This finding supports the general hypothesis that there is a strong tendency for the PoS and DC parameters to have the same value.

I will now investigate the effect of the parameter of structural DC type on the correlation between maximally flexible PoS and DCs, as operationalized in the set of predictions listed in (18a-d) above. First, consider the results for the prediction in (18a), concerning balanced clauses. The relevant frequencies are given in Table 7.11. Like in all the previous cases involving balanced clauses, these counts do not reveal a significant correlation: \( p = 0.160 \) (2-sided).

<table>
<thead>
<tr>
<th>PoS contentives/ non-verbs</th>
<th>DCs: contentive/multi-functional clauses type 1 (balanced)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq 2 5 7</td>
</tr>
<tr>
<td></td>
<td>expected freq 1 6</td>
</tr>
<tr>
<td>-</td>
<td>observed freq 3 36 39</td>
</tr>
<tr>
<td></td>
<td>expected freq 4 35</td>
</tr>
<tr>
<td>total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 41 46</td>
</tr>
</tbody>
</table>

Turning to deranked contentive and multi-functional clauses (see (18b-d) above), we find that all such constructions attested in the sample are of type 3 rather than type 2 (cf. Table 7.35 in section 7.5). Therefore, it is not possible to differentiate between clauses of type 3 and clauses of type 2. The frequencies for deranked contentive and multi-functional clauses of type 3 are presented in Table 7.12. As expected, these frequencies reveal a significant correlation: \( p = 0.000 \) (2-sided). Moreover, the CC value is 0.589, which indicates a strong effect.

\[108\] I am indebted to Elena Maslova for her helpful comments on the interpretation of these data.
Table 7.12: Frequencies for languages with/without lexical contentives/non-verbs and contentive clauses/multi-functional clauses of type 3

<table>
<thead>
<tr>
<th>PoS contentives /non-verbs</th>
<th>DCs: contentive/multi-functional clauses type 3 (D-ALT)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>observed freq</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>expected freq</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>−</td>
<td>0</td>
<td>39</td>
</tr>
<tr>
<td>observed freq</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>expected freq</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>total</td>
<td>4</td>
<td>42</td>
</tr>
</tbody>
</table>

Applying Maslova’s method makes clear that this correlation involves a one-sided asymmetrical dependency. The relevant data are presented in Tables 7.13a and 7.13b:

Table 7.13a: Maslova test no. 1: change DC parameter to PoS=DC

<table>
<thead>
<tr>
<th>PoS = DC</th>
<th>PoS contentive/non-verb +</th>
<th>PoS contentive/non-verb −</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>39</td>
<td>0</td>
</tr>
</tbody>
</table>

p = 0.002 (significant)

Table 7.13b: Maslova test no. 2: change PoS parameter to PoS=DC

<table>
<thead>
<tr>
<th>PoS = DC</th>
<th>DC contentive/multi-functional +</th>
<th>DC contentive/multi-functional type 3 −</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>39</td>
<td>0</td>
</tr>
</tbody>
</table>

p = 1 (not significant)

Table 7.13a shows that the negative value of the PoS parameter constrains the event of PoS and DCs having the same value. Table 7.13b, in contrast, shows that the DC parameter and the PoS=DC parameter do not interact. On the basis of these results, the universal in (20) can be formulated:
Languages without lexical contentives or non-verbs also lack deranked contentive/multi-functional clauses of type 3.

Notably, the distribution in Table 7.12 is similar to the one in Table 7.9, to the extent that it displays three (nearly) empty cells, as a result of the fact that both the PoS and the DC parameter are, independently of each other, strongly skewed towards the negative value. In view of this, the implicational universal in (20), which has a negative value in the consequent part, does not have much explanatory power. Again, it is much more telling that the likelihood of the DC parameter having the rare positive value increases to over 50% under the influence of the PoS parameter having this rare positive value.

7.3.2.3 Flexible match 2: Nominals and nominal clauses
Continuing with less pervasive flexibility, I will now investigate the relationship between lexical nominals and nominal clauses. First consider the data in Table 7.14, in which nominal clauses are considered as one undifferentiated group. The relevant language data can be found in Tables 7.37 and 7.38 of section 7.5. The observed frequencies in Table 7.14 do not deviate at all from the expected ones, so that $p = 1$ (2-sided). This means that there is no dependency relation between the presence of lexical nominals and nominal clauses.

<table>
<thead>
<tr>
<th>PoS: nominals</th>
<th>DCs: nominal clauses</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>nominals</td>
<td>observed freq 2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>expected freq 2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>observed freq 14</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>expected freq 14</td>
<td>27</td>
</tr>
<tr>
<td>total</td>
<td>16</td>
<td>30</td>
</tr>
</tbody>
</table>

Taking into account the parameter of structural DC types, we can formulate the more specific predictions in (21a-d), in parallel with the ones listed in (18a-d) for contentives/non-verbs and contentive/multi-functional clauses.
a. If a language has balanced nominal clauses of type 1, then it should also have lexical nominals.
b. If a language has deranked nominal clauses of type 2/3, then it should also have lexical nominals.
c. If a language has deranked nominal clauses of type 2, then it should also have lexical nominals.
d. If a language has deranked nominal clauses of type 3, then it should also have lexical nominals.

First consider the prediction in (21a). In line with the other results found so far, there is no correlation between lexical nominals and balanced nominal clauses. The relevant frequencies appear in Table 7.15; Fisher’s Exact yields $p = 0.301$ (2-sided). In fact, as this table shows, none of the 13 languages with nominal clauses of type 1 has lexical nominals. (For the language data pertaining to Table 7.15, see again Tables 7.37 and 7.38, in section 7.5.)

<table>
<thead>
<tr>
<th>PoS: nominals</th>
<th>DCs: nominal clauses type 1 (balanced)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>1</td>
</tr>
<tr>
<td>-</td>
<td>observed freq</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>12</td>
</tr>
<tr>
<td>total</td>
<td>13</td>
<td>33</td>
</tr>
</tbody>
</table>

Regarding the prediction in (21b), the frequencies for deranked nominal clauses of type 2 and 3 appear in Table 7.16. Contrary to the expectations, the observed frequencies do not yield a significant correlation ($p = 0.120$).
Table 7.16: Frequencies for languages with/without nominals and nominal clauses of type 2/3

<table>
<thead>
<tr>
<th>PoS: nominals</th>
<th></th>
<th>DCs: nominal clauses type 2/3 (deranked)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>observed freq</td>
<td>2, 3</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>observed freq</td>
<td>4, 37</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>5</td>
<td>36</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>6</td>
<td>40</td>
</tr>
</tbody>
</table>

Notably, the total number of languages with deranked nominal clause constructions attested in the whole sample is quite low (6). Therefore, splitting up the group of deranked nominal clauses into constructions of type 2 versus type 3 does not influence the results in any interesting way. For the sake of completeness, the relevant frequencies are presented in Table 7.17 (for nominal clauses of type 2) and Table 7.18 (for nominal clauses of type 3). Not surprisingly, no significant correlations are found: The relevant p-values are \( p = 0.208 \) (2-sided) for Table 7.17, and \( p = 0.379 \) (2-sided) for Table 7.18.

Table 7.17: Frequencies for languages with/without nominals and nominal clauses of type 2

<table>
<thead>
<tr>
<th>PoS: nominals</th>
<th></th>
<th>DCs: nominal clauses type 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>observed freq</td>
<td>1, 4</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>observed freq</td>
<td>1, 41</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>2</td>
<td>41</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>2</td>
<td>44</td>
</tr>
</tbody>
</table>

Table 7.18: Frequencies for languages with/without nominals and nominal clauses of type 3

<table>
<thead>
<tr>
<th>PoS: nominals</th>
<th></th>
<th>DCs: nominal clauses type 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>observed freq</td>
<td>1, 4</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>observed freq</td>
<td>3, 38</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>4</td>
<td>39</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>4</td>
<td>42</td>
</tr>
</tbody>
</table>
7.3.2.4 Flexible match 3: Modifiers and modifier clauses

Finally, consider the results for the third flexible construction type: lexical modifiers and modifier clauses, presented in Table 7.19 below. It must be noted that the total number of languages relevant for this prediction is only 38. This is because in 10 languages neither of the two modifier slots can be expressed by means of a large, open PoS class\(^{109}\). Some of these 10 languages do have a small, rigid class of adjectives and/or adverbs, while others have no lexical strategy at all. Both groups are interpreted as lacking a value for the PoS parameter. These cases are excluded from the analysis, since it is impossible to compare the distributional pattern of the DC construction with the pattern of any large, open PoS class. The language data pertaining to lexical modifiers and modifier clauses can be found in Tables 7.39 and 7.40 in section 7.5.

<table>
<thead>
<tr>
<th>PoS: modifiers</th>
<th>DCs: modifier clauses</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>−</td>
<td>observed freq</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>1</td>
<td>37</td>
</tr>
</tbody>
</table>

The data in Table 7.19 make clear that there is no correlation at all between the availability of lexical and clausal constructions that can express both modifier functions: \( p = 1 \). In fact, there is only one language with modifier clauses, and this language, Babungo, does not have lexical modifiers, but rather rigid derived adjectives (and a small set of simple ones) and a small class of rigid manner adverbs\(^{110,111}\). Moreover, there are only four languages with lexical modifiers attested in the sample, and none of these have modifier clauses.

\(^{109}\) These languages are Mandarin Chinese, Tamil, Nung, Krongo, Hixkaryana, Slave, Nivkh, West Greenlandic, Nunggubuyu, and Tuscarora.

\(^{110}\) Note that the modifier clause construction in Babungo is balanced. This means that it does not show any formal similarity to a lexical expression, which links up with the lack of functional similarity.

\(^{111}\) There are two more languages in the sample that display a flexible modifier clause construction: Mandarin Chinese and Krongo (see Table 7.40 in section 7.5). However, neither of these languages is taken into account in Table 7.19, since both lack (large, open) lexical classes to express the functions of modification, and thus do not allow for a functional comparison between PoS and DCs.
7.3.2.5 **Summary**

To sum up the results for specific matches between flexible constructions, it was shown that the most pervasively flexible (deranked) type of DC is hardly ever attested in languages without a PoS class with the same type of flexibility. However, this generalization is rather uninformative to the extent that maximal flexibility in both the PoS and the DCs domains are rare phenomena in the first place. The interesting observation is that, even though the presence of maximally flexible PoS classes in a language does not imply the presence of maximally flexible deranked DCs, the availability of the former strongly increases the chances of also having the latter.

In contrast, no dependency relations were found between the (non-)attestation of less flexible lexical and clausal constructions, i.e. between nominals and nominal clauses, and between modifiers and modifier clauses (neither balanced nor deranked). The number of cases with a positive value on either the PoS or the DC parameter was again found to be low, but unlike in the case of contentives/non-verbs and contentive/multi-functional clauses, there was no tendency for the two parameters to have the same value.

### 7.3.3 Specific matches for rigid constructions

#### 7.3.3.1 **Introduction**

I now turn back to Prediction C in (16) above and apply this prediction to the various types of rigid PoS classes and DC constructions. This allows for the formulation of the three sub-predictions listed in (22):

(22) a. If a language has complement clauses, then it should also have lexical nouns.

b. If a language has relative clauses, then it should also have lexical adjectives.

c. If a language has adverbial manner clauses, then it should also have lexical manner adverbs.

Note that, unlike in Chapter 4 (see (8) in that chapter), constructions specialized for the function of head of a predicate phrase (verbs and predicate clauses) are not taken into account, since, as shown in Chapter 6, rigid predicate clauses are not attested in any of the sample languages.

Each of the three predictions in (22) can be further differentiated when taking into account the parameter of structural DC type. In particular, the
prediction in (22a) can be split up into the sub-predictions listed in (23a-d) (cf. (10) of Chapter 4). As in the previous cases, the prediction in (23a) is less likely to be confirmed than the one in (23b), and the one in (23c) is less likely to hold than the one in (23d).

(23)  a. If a language has balanced complement clauses of type 1, then it should also have lexical nouns.
   b. If a language has deranked complement clauses of type 2/3, then it should also have lexical nouns.
   c. If a language has deranked complement clauses of type 2, then it should also have lexical nouns.
   d. If a language has deranked complement clauses of type 3, it should also have lexical nouns.

Parallel lists of sub-hypotheses can be set up for the predictions in (22b) and (22c) above, concerning lexical and clausal constructions that are specialized for the function of modifier in a referential phrase, and for modifier in a predicate phrase, respectively. In what follows, I present the results for every rigid construction type, first considering DCs as a single group, and then differentiating according to structural DC type.

7.3.3.2 Rigid match 1: Nouns and complement clauses
I start out with the results for the prediction in (22a), involving rigid nouns and rigid complement clauses, without differentiating for structural DC type. The observed frequencies are presented in Table 7.20 below. The relevant language data can be found in Tables 41 and 42, section 7.5. They do not deviate from chance frequency, so that p = 1 (2-sided).

112 Note that, as mentioned earlier, there is one language, Warao, which has two lexical strategies available to express the function of head of a referential phrase: a class of simple non-verbs and a class of derived nouns. To avoid double values, prominence is given to the class of simple lexemes. This means that Warao is counted as a language without a class of rigid nouns.
Table 7.20: Frequencies for languages with/without nouns and complement clauses

<table>
<thead>
<tr>
<th>PoS: nouns</th>
<th>DCs: complement clauses</th>
<th>+</th>
<th>–</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>30</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>30</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>observed freq</td>
<td>9</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>39</td>
<td>7</td>
<td>46</td>
</tr>
</tbody>
</table>

Turning to the sub-predictions regarding the different structural types of complement clauses, consider the results for the prediction in (23a), which appear in Table 7.21. In accordance with the findings for all balanced clause types considered so far, there is no correlation between rigid lexical nouns and rigid complement clauses of type 1. The counts in Table 7.21 hardly deviate from chance frequency: \( p = 0.730 \) (2-sided).

Table 7.21: Frequencies for language with/without nouns and complement clauses of type 1

<table>
<thead>
<tr>
<th>PoS: nouns</th>
<th>DCs: complement clauses type 1 (balanced)</th>
<th>+</th>
<th>–</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>20</td>
<td>15</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>19</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>observed freq</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>25</td>
<td>21</td>
<td>46</td>
</tr>
</tbody>
</table>

Unexpectedly, considering deranked DCs separately does not reveal a correlation either. The frequencies for complement clauses of type 2/3, and for types 2 and 3 separately, are presented in Tables 7.22, 7.23 and 7.24, respectively. None of these yields a significant result. For Table 7.22, \( p = 0.497 \) (2-sided); for Table 7.23, \( p = 0.702 \) (2-sided); and for Table 7.24, \( p = 1 \) (2-sided).
Table 7.22: Frequencies for language with/without nouns and complement clauses of type 2/3

<table>
<thead>
<tr>
<th>PoS: nouns</th>
<th>DCs: complement clauses type 2/3 (deranked)</th>
<th>+</th>
<th>–</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>17</td>
<td>18</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>18</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>observed freq</td>
<td>7</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>24</td>
<td>22</td>
<td>46</td>
</tr>
</tbody>
</table>

Table 7.23: Frequencies for language with/without nouns and complement clauses of type 2

<table>
<thead>
<tr>
<th>PoS: nouns</th>
<th>DCs: complement clauses type 2</th>
<th>+</th>
<th>–</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>9</td>
<td>26</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>10</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>observed freq</td>
<td>4</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>3</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>13</td>
<td>33</td>
<td>46</td>
</tr>
</tbody>
</table>

Table 7.24: Frequencies for language with/without nouns and complement clauses of type 3

<table>
<thead>
<tr>
<th>PoS: nouns</th>
<th>DCs: complement clauses type 3</th>
<th>+</th>
<th>–</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq</td>
<td>11</td>
<td>24</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>11</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>observed freq</td>
<td>4</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>expected freq</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>25</td>
<td>31</td>
<td>46</td>
</tr>
</tbody>
</table>

In sum, no significant correlation exists between rigid nouns and rigid complement clauses, neither for balanced nor for (different types of) deranked clauses.

7.3.3.3 Rigid match 2: Adjectives and relative clauses
Consider now lexical and clausal constructions that are specialized for the function of modifier in a referential phrase: adjectives and relative clauses. The distribution pertaining to the prediction in (22b) (without differentiation
for structural DC type) is presented in Table 7.25. This table shows that the observed frequencies are again equal to the expected frequencies, so that the p-value is 1 (2-sided). For the relevant language data, see Tables 7.43-7.47 in section 7.5.

Note that the total number of analyzed cases (N) in Table 7.25 is 34. This is because there are 14 languages that either lack a class of adjectives altogether (8 languages) or have only a small set of them (6 languages). These 14 cases are not taken into account because they are interpreted as lacking a construction in the PoS domain that can serve as the functional comparative standard\(^{113}\). This means that the 15 cases with a negative value on the PoS parameter in Table 7.25 involve languages with a flexible lexeme class that can express the function of modifier in a referential phrase\(^{114}\).

<table>
<thead>
<tr>
<th>PoS: adjectives</th>
<th>DCs: relative clauses</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ observed freq</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>expected freq</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>- observed freq</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>expected freq</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>total</td>
<td>27</td>
<td>7</td>
</tr>
</tbody>
</table>

Taking into account the parameter of structural DC type, we can formulate the set of predictions listed in (24), in parallel to those in (23) above. The same expectations as above hold regarding the likelihood of each of the predictions to be confirmed: (24a) is less likely than (24b), and (24c) less likely then (24d):

(24) a. If a language has balanced relative clauses of type 1, then it should also have lexical adjectives.

\(^{113}\) One of the languages without adjectives, namely Hixkaryana, also lacks a relative clause construction (Derbyshire 1979: 26).

\(^{114}\) Note that there is one language, namely Ket, which has a large, open class of flexible modifiers, as well as a class of derived rigid adjectives, and small classes of simple adjectives and manner adverbs. In order to avoid double values, prominence is given to the large, open lexeme class, i.e. the modifiers, and the derived and small classes are not taken into account. This means that Ket is counted as a language without adjectives. In the case of a choice between a derived and a small lexeme class, the former is considered primary and the latter is disregarded. This occurs in one language, namely Koasati, which has a class of derived modifiers and a small class of adjectives, but is counted as a language without rigid adjectives.
b. If a language has deranked relative clauses of type 2/3, then it should also have lexical adjectives.

c. If a language has deranked relative clauses of type 2, then it should also have lexical adjectives.

d. If a language has deranked relative clauses of type 3, then it should also have lexical adjectives.

The results for the prediction in (24a) are presented in Table 7.26. The observed and expected frequencies deviate only very slightly from each other, so that there is no significant correlation: $p = 0.724$ (2-sided).

<table>
<thead>
<tr>
<th>PoS: adjectives</th>
<th>DCs: relative clauses of type 1</th>
<th>(\text{balanced})</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq 13</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>expected freq 12</td>
<td>7</td>
</tr>
<tr>
<td>-</td>
<td>observed freq 9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>expected freq 10</td>
<td>5</td>
</tr>
<tr>
<td>total</td>
<td>22</td>
<td>12</td>
</tr>
</tbody>
</table>

The results for the prediction in (24b), concerning deranked relative clauses (type 2 and 3) are given in Table 7.27. As can be seen, there is again only a minimal deviation between the observed and the expected frequencies: $p = 0.718$ (two-sided).

<table>
<thead>
<tr>
<th>PoS: adjectives</th>
<th>DCs: relative clauses of type 2/3</th>
<th>(\text{deranked})</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>observed freq 5</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>expected freq 6</td>
<td>13</td>
</tr>
<tr>
<td>-</td>
<td>observed freq 5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>expected freq 4</td>
<td>11</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>249</td>
</tr>
</tbody>
</table>
In Tables 7.28 and 7.29 below the data are split up for deranked relative clauses of type 2 and of type 3, respectively. Neither the results in Table 7.28 nor those in Table 7.29 are in accordance with the predictions: There is no correlation between the availability of lexical adjectives and deranked relative clauses of type 2 and 3, respectively. The p-values are 0.718 (2-sided) for Table 7.28, and p = 1 (2-sided) for Table 7.29.

It should be mentioned that there are some languages (in particular: Kharia, Hungarian, Kayardild, and Georgian) that have both relative clauses of type 2 and of type 3, which is why the totals for relative clauses of Tables 7.28 and Table 7.29 add up to 14, i.e. more than the total of 10 languages with deranked relative clause constructions in Table 7.27. Note further that there are no languages with a relative clause construction of type 3 that do not also have a relative clause construction of type 2. This explains why the frequencies in Tables 7.27 and 7.28 are identical.

Table 7.28: frequencies for languages with/without rigid adjectives and relative clauses of type 2

<table>
<thead>
<tr>
<th>PoS: adjectives</th>
<th>DCs: relative clauses of type 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
</tr>
<tr>
<td>observed freq</td>
<td>5</td>
</tr>
<tr>
<td>expected freq</td>
<td>6</td>
</tr>
<tr>
<td>observed freq</td>
<td>5</td>
</tr>
<tr>
<td>expected freq</td>
<td>4</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 7.29: frequencies for languages with/without rigid adjectives and relative clauses of type 3

<table>
<thead>
<tr>
<th>PoS: adjectives</th>
<th>DCs: relative clauses of type 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
</tr>
<tr>
<td>observed freq</td>
<td>2</td>
</tr>
<tr>
<td>expected freq</td>
<td>2</td>
</tr>
<tr>
<td>observed freq</td>
<td>2</td>
</tr>
<tr>
<td>expected freq</td>
<td>2</td>
</tr>
<tr>
<td>total</td>
<td>4</td>
</tr>
</tbody>
</table>

In sum, the same conclusion is reached for adjectives and relative clauses as for nouns and complement clauses: There is no correlation between the
availability of lexical and (different structural types of) clausal constructions specialized for the function of modifier in a referential phrase.

7.3.3.4 Rigid match 3: Manner adverbs and adverbial manner clauses
Finally, consider the data for lexical and clausal constructions that are specialized for the function of modifier in a predicate phrase: manner adverbs and adverbial manner clauses. The distribution pertaining to the prediction in (22c) appears in Table 7.30. The relevant language data can be found in Tables 7.48–7.52 in section 7.5.

Note that, as in the case of adjectives, only those languages are taken into consideration that have a large, open class of simple or derived lexemes to express the function of modifier in a predicate phrase. Languages with a small, closed lexical class (8 cases), or no lexical option at all (11 cases) are excluded from the analysis. In addition, there are 3 languages with a large simple or derived class of manner adverbs, but without a DC strategy for the function of modifier in a predicate phrase. There are another 3 languages that do have a class of flexible lexemes that can be used in this function, but which are excluded from the analysis as well, since they lack a DC strategy to express this function. This leaves us with a rather small number of relevant languages, namely $N = 23$.

The observed frequencies in Table 7.30 deviate from the expected ones in the predicted way, but not significantly: $p = 0.083$ (2-sided).

Table 7.30: frequencies for languages with/without a (derived) manner adverbs and adverbial manner clauses

<table>
<thead>
<tr>
<th>PoS: manner adverbs</th>
<th>DCs: adverbial manner clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
</tr>
<tr>
<td>observed freq</td>
<td>16</td>
</tr>
<tr>
<td>expected freq</td>
<td>15</td>
</tr>
<tr>
<td>observed freq</td>
<td>5</td>
</tr>
<tr>
<td>expected freq</td>
<td>6</td>
</tr>
<tr>
<td>total</td>
<td>21</td>
</tr>
</tbody>
</table>

We may now formulate the specific sub-predictions for manner adverbs and adverbial manner clauses, taking into account the different structural DC types. These sub-predictions are listed in (25), in parallel with (24) and (23) above, and the by now familiar expectations regarding the likelihood of their confirmation apply.
(25) a. If a language has balanced adverbial manner clauses of type 1, then it should also have lexical manner adverbs.
b. If a language has deranked adverbial manner clauses of type 2/3, then it should also have lexical manner adverbs.
c. If a language has deranked adverbial manner clauses of type 2, then it should also have lexical manner adverbs.
d. If a language has deranked adverbial manner clauses of type 3, then it should also have lexical manner adverbs.

The results for the prediction in (25a) appear in Table 7.31. The observed frequencies equal the expected frequencies, so that p = 1 (2-sided).

Table 7.31: Frequencies for languages with/without (derived) manner adverbs and adverbial manner clauses of type 1

<table>
<thead>
<tr>
<th>DCs: adverbial manner clauses of type 1 (balanced)</th>
<th>+</th>
<th>–</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoS: manner adverbs (incl. derived)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ observed freq</td>
<td>8</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>expected freq</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>– observed freq</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>expected freq</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>11</td>
<td>12</td>
<td>23</td>
</tr>
</tbody>
</table>

The results for the prediction in (25b), concerning deranked adverbial manner clauses, appear in Table 7.32. The observed frequencies deviate from the expected frequencies in the predicted direction, but not enough to yield a significant correlation: p = 0.193 (2-sided).
Table 7.32: Frequencies for languages with/without (derived) manner adverbs and adverbial manner clauses of type 2/3

<table>
<thead>
<tr>
<th>PoS: manner adverbs (incl. derived)</th>
<th>DCs: adverbial manner clauses of type 2/3 (deranked)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>+ observed freq</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>expected freq</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>– observed freq</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>expected freq</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>total</td>
<td>12</td>
<td>11</td>
</tr>
</tbody>
</table>

Since there is only one language in the sample with an adverbial manner clause construction of type 3, the data are insufficient to further differentiate within the group of deranked adverbial manner clauses.

In sum, the findings for rigid constructions specialized for the function of modifier in a predicate phrase parallel those for the other two rigid construction types: there is no correlation between the presence of lexical manner adverbs and the presence of (different structural types of) adverbial manner clauses.

7.3.3.5 Summary
No dependency relations are identified between the availability of rigid lexical and clausal constructions that are specialized for the expression of the same propositional function. Differentiating between balanced and (different types of) deranked DCs does not influence this result.

7.4 Summary, conclusion
To conclude, the results presented in this chapter make clear that significant correlations between the functional possibilities of PoS and DCs are found along the following lines:

(i) For flexible rather than for rigid constructions;
(ii) For pervasively flexible constructions rather than for less flexible constructions;
(iii) For deranked DCs rather than for balanced ones.

I have presented the following evidence for these findings: First, Tables 7.4 and 7.5a-b reveal a one-sided asymmetrical dependency between flexible
PoS and flexible deranked DCs. In particular, it was shown that flexible deranked DCs occur almost exclusively in languages with a flexible PoS system. Second, Tables 7.7 and 7.8a-b show the same type of dependency relation, but specify it for deranked flexible DCs of type 3 (D-ALT). Third, Tables 7.9 and 7.10a-b contain the data that allow for a further fine-tuning of these findings in terms of specific flexible construction types. In particular, these data shows that maximal flexibility at the level of DC constructions (contentive/multi-functional clauses) is possible only when a maximally flexible PoS class (contentives/non-verbs) is also available. More interestingly, in view of the fact that both maximally flexible PoS classes and maximally flexible DC constructions are rare phenomena independently of each other, the data show that the presence of a maximally flexible PoS class in a language strongly increases the chances of also finding a maximally flexible DC construction in that language. On the basis of the data in Tables 7.12 and 7.13 a-b this generalization could be further specified as pertaining to deranked contentive and multi-functional clauses of type 3 (D-ALT).

These results suggest, in sum, that the presence of (pervasive) flexibility in the domain of deranked DC constructions is dependent on the presence of (pervasive) flexibility in the domain of PoS classes. On the other hand, flexibility in the PoS system of a language does not imply flexibility in the domain of deranked dependent clauses. In Chapter 8 I will further explore the nature of the relationship between distributional patterns of PoS and DCs, as well as the functional factors that may motivate this relationship.

### 7.5 Language data

<table>
<thead>
<tr>
<th>Language</th>
<th>Flex DCs</th>
<th>Structural type</th>
<th>Functional Type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Tagalog</td>
<td>+</td>
<td>3, 1</td>
<td>contentive/ multi-functional</td>
<td>pag-, Ø</td>
</tr>
<tr>
<td>2 Kharia</td>
<td>+</td>
<td>3</td>
<td>contentive/ multi-functional</td>
<td>RDP/Ø, -na</td>
</tr>
<tr>
<td>3 Samoan</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
<td>Ø</td>
</tr>
<tr>
<td>4 Guaraní</td>
<td>–</td>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Warao</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Flex DCs</td>
<td>Structural type</td>
<td>Functional Type</td>
<td>Structural coding</td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>6</td>
<td>Turkish</td>
<td>+</td>
<td>3, 1</td>
<td>nominal</td>
</tr>
<tr>
<td>7</td>
<td>Kayardild</td>
<td>+</td>
<td>3, 1</td>
<td>multi-functional, nominal</td>
</tr>
<tr>
<td>8</td>
<td>Paiwan</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
</tr>
<tr>
<td>9</td>
<td>Quechua</td>
<td>+</td>
<td>2</td>
<td>nominal</td>
</tr>
<tr>
<td>10</td>
<td>Ma’di</td>
<td>+</td>
<td>2/3</td>
<td>nominal</td>
</tr>
<tr>
<td>11</td>
<td>Gooniyandi</td>
<td>+</td>
<td>1</td>
<td>multi-functional</td>
</tr>
<tr>
<td>12</td>
<td>Hungarian</td>
<td>–</td>
<td>(1, 2, 3)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Japanese</td>
<td>–</td>
<td>(1, 2, 1/3)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Hmong</td>
<td>–</td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Lango</td>
<td>–</td>
<td>(1, 3)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Ket</td>
<td>+</td>
<td>3, 1</td>
<td>nominal</td>
</tr>
<tr>
<td>17</td>
<td>Koasati</td>
<td>–</td>
<td>(2)</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.34: Availability and type of flexible DCs in languages with rigid PoS classes only

<table>
<thead>
<tr>
<th>Language</th>
<th>Flex DCs</th>
<th>Structural type</th>
<th>Functional type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Itelmen</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Thai</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
</tr>
<tr>
<td>3</td>
<td>Basque</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
</tr>
<tr>
<td>4</td>
<td>Abun</td>
<td>–</td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Bambara</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Georgian</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
</tr>
<tr>
<td>7</td>
<td>Bukiyp</td>
<td>–</td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Abkhaz</td>
<td>–</td>
<td>(1, 3)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Polish</td>
<td>–</td>
<td>(1, 2, 3)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Burushaski</td>
<td>+</td>
<td>2, 1</td>
<td>nominal, multi-functional</td>
</tr>
<tr>
<td>11</td>
<td>Lavukaleve</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Alamblak</td>
<td>–</td>
<td>(2, 3)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Pipil</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
</tr>
<tr>
<td>14</td>
<td>Wambon</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Dhaasanac</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
</tr>
<tr>
<td>Language</td>
<td>Flex DCs</td>
<td>Structural type</td>
<td>Functional type</td>
<td>Structural coding</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------</td>
<td>-----------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>16 Berbice Dutch</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Babungo</td>
<td>+</td>
<td>1</td>
<td>modifier</td>
<td>fan/yúu</td>
</tr>
<tr>
<td>18 Nama</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 Hdi</td>
<td>+</td>
<td>3</td>
<td>nominal</td>
<td>tá</td>
</tr>
<tr>
<td>20 Mandarin Ch.</td>
<td>+</td>
<td>1</td>
<td>multi-functional, modifier</td>
<td>Ø, de</td>
</tr>
<tr>
<td>21 Tamil</td>
<td>–</td>
<td>(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 Kisi</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Nung</td>
<td>+</td>
<td>1</td>
<td>multi-functional</td>
<td>Ø</td>
</tr>
<tr>
<td>24 Garo</td>
<td>–</td>
<td>(1, 2, 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 Krongo</td>
<td>+</td>
<td>2/3</td>
<td>modifier</td>
<td>m-/n-</td>
</tr>
<tr>
<td>26 Hixkaryana</td>
<td>–</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 Slave</td>
<td>+</td>
<td>1</td>
<td>nominal</td>
<td>Ø</td>
</tr>
<tr>
<td>28 Nivkh</td>
<td>–</td>
<td>(1, 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29 Greenlandic</td>
<td>+</td>
<td>1</td>
<td>nominal, comp/adv</td>
<td>PTCmood/CONT mood</td>
</tr>
<tr>
<td>30 Nunggubuyu</td>
<td>none</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 Tuscarora</td>
<td>none</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.35: Availability and type of contentive/multi-functional clauses in languages with lexical contentives/non-verbs

<table>
<thead>
<tr>
<th>Language</th>
<th>Contentive/multifunctional clause</th>
<th>Type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Tagalog</td>
<td>+</td>
<td>3,1</td>
<td>pag-,Ø</td>
</tr>
<tr>
<td>2 Kharia</td>
<td>+</td>
<td>3</td>
<td>RDP,-na</td>
</tr>
<tr>
<td>3 Samoan</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Guarani</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Warao</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Turkish</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Kayardild</td>
<td>+</td>
<td>3</td>
<td>-n-,</td>
</tr>
<tr>
<td>Language</td>
<td>Contentive / multi-functional clause</td>
<td>Type</td>
<td>Structural coding</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------</td>
<td>------</td>
<td>-------------------</td>
</tr>
<tr>
<td>1 Paiwan</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Quechua</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Ma’di</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Gooniyandi</td>
<td>+</td>
<td>1</td>
<td>Ø</td>
</tr>
<tr>
<td>5 Hungarian</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Japanese</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Hmong Njua</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Lango</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Ket</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Koasati</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Itelmen</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Thai</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Basque</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Abun</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 Bambara</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 Georgian</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Bukiyip</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Abkhaz</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 Polish</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 Burushaski</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 Lavukaleve</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 Alamblak</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Pipil</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 Wambon</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 Dhaasanac</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 Berbice Dutch</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 Babungo</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28 Nama</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29 Hdi</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 Mandarin Ch.</td>
<td>+</td>
<td>1</td>
<td>Ø</td>
</tr>
<tr>
<td>31 Tamil</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32 Kisi</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33 Nung</td>
<td>+</td>
<td>1</td>
<td>Ø</td>
</tr>
</tbody>
</table>
Table 7.37: Availability and type of nominal clauses in languages with lexical nominals

<table>
<thead>
<tr>
<th>Language</th>
<th>Nominal clauses</th>
<th>Type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quechua</td>
<td>+</td>
<td>2</td>
<td>-j/-shka/-na</td>
</tr>
<tr>
<td>Ma’di</td>
<td>+</td>
<td>2/3</td>
<td>-lɛ/-dɔ</td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.38: Availability and type of and nominal clauses in languages without lexical nominals

<table>
<thead>
<tr>
<th>Language</th>
<th>Nominal clause</th>
<th>Type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagalog</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kharia</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>+</td>
<td>1</td>
<td>Ø</td>
</tr>
<tr>
<td>Guarani</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warao</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td>+</td>
<td>3,1</td>
<td>-DIK/-AcAK, ki (borrowed)</td>
</tr>
<tr>
<td>Kayardild</td>
<td>+</td>
<td>1</td>
<td>-ntha</td>
</tr>
<tr>
<td>Paiwan</td>
<td>+</td>
<td>1</td>
<td>a</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lango</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket</td>
<td>+</td>
<td>3,1</td>
<td>Ø, Ø (bare INF)</td>
</tr>
<tr>
<td>Koasati</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Nominal clause</td>
<td>Type</td>
<td>Structural coding</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------</td>
<td>------</td>
<td>-------------------</td>
</tr>
<tr>
<td>13 Itelmen</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Thai</td>
<td>+</td>
<td>1</td>
<td>Ø, thiï</td>
</tr>
<tr>
<td>15 Basque</td>
<td>+</td>
<td>1</td>
<td>-en</td>
</tr>
<tr>
<td>16 Abun</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Bambara</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Georgian</td>
<td>+</td>
<td>1</td>
<td>rom</td>
</tr>
<tr>
<td>19 Bukiyip</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 Abkhaz</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 Polish</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 Burushaski</td>
<td>+</td>
<td>2,1</td>
<td>-(á)as, ke/ki</td>
</tr>
<tr>
<td>23 Lavukaleve</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 Alamblak</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 Pipil</td>
<td>+</td>
<td>1</td>
<td>ka(h)</td>
</tr>
<tr>
<td>26 Wambon</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 Dhaasanac</td>
<td>+</td>
<td>1</td>
<td>DET(+DEM)</td>
</tr>
<tr>
<td>28 Berbice Dutch</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29 Babungo</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 Nama</td>
<td>-</td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>31 Hdi</td>
<td>+</td>
<td>3</td>
<td>tā</td>
</tr>
<tr>
<td>32 Mandarin Ch.</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33 Tamil</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34 Kisi</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35 Nung</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36 Garo</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37 Krongo</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38 Hixkaryana</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39 Slave</td>
<td>+</td>
<td>1</td>
<td>nj</td>
</tr>
<tr>
<td>40 Nivkh</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41 Greenlandic</td>
<td>+</td>
<td>1</td>
<td>PTCmood</td>
</tr>
<tr>
<td>42 Nunggubuyu</td>
<td>none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43 Tuscarora</td>
<td>none</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 7.39: Availability of modifier clauses in languages with (derived) lexical modifiers**

<table>
<thead>
<tr>
<th>Language</th>
<th>Modifier clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hmong Njua</td>
<td>-</td>
</tr>
<tr>
<td>2 Lango</td>
<td>-</td>
</tr>
</tbody>
</table>
### Table 7.40: Languages without (derived) lexical modifiers, but with modifier clauses

<table>
<thead>
<tr>
<th>Language</th>
<th>Modifier clause</th>
<th>Type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babungo</td>
<td>+</td>
<td>1</td>
<td>fan/yúu</td>
</tr>
<tr>
<td>Mandarin Ch.</td>
<td>+</td>
<td>1</td>
<td>de</td>
</tr>
<tr>
<td>Krongo</td>
<td>+</td>
<td>2/3</td>
<td>m-/n-</td>
</tr>
</tbody>
</table>

### Table 7.41: Availability and type of complement clauses in languages without nouns

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid complement clauses</th>
<th>Type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagalog</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Kharia</td>
<td>+</td>
<td>1</td>
<td>no, gam-kon</td>
</tr>
<tr>
<td>Samoan</td>
<td>+</td>
<td>3</td>
<td>-ga, unmarked NMLZ</td>
</tr>
<tr>
<td>Guarani</td>
<td>+</td>
<td>1</td>
<td>há(gwe)</td>
</tr>
<tr>
<td>Warao</td>
<td>+</td>
<td>2</td>
<td>-kitane</td>
</tr>
<tr>
<td>Turkish</td>
<td>+</td>
<td>1, 2, 3</td>
<td>-ma, -mak, diye, Ø</td>
</tr>
<tr>
<td>Kayardild</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Quechua</td>
<td>+</td>
<td>2</td>
<td>-y, ngapay</td>
</tr>
<tr>
<td>Ma’di</td>
<td>+</td>
<td>1, 2/3</td>
<td>ka, Ø</td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>+</td>
<td>2</td>
<td>-woo</td>
</tr>
<tr>
<td>Hungarian</td>
<td>+</td>
<td>3, 1</td>
<td>-ni, -ás/-és, hogy</td>
</tr>
</tbody>
</table>

### Table 7.42: Availability and type of complement clauses in languages with nouns

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid complement clauses</th>
<th>Type</th>
<th>Structural coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paiwan</td>
<td>+</td>
<td>1</td>
<td>tu(a)/tjai</td>
</tr>
<tr>
<td>Ket</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>+</td>
<td>1/3</td>
<td>no/mono</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>+</td>
<td>1</td>
<td>qhov/kuam/ (has)tas</td>
</tr>
<tr>
<td>Lango</td>
<td>+</td>
<td>1, 3</td>
<td>nî,-(kk)ø</td>
</tr>
<tr>
<td>Koasati</td>
<td>+</td>
<td>2</td>
<td>NMLZ (various forms)</td>
</tr>
<tr>
<td>Language</td>
<td>Rigid complement clauses</td>
<td>Type</td>
<td>Structural coding</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------</td>
<td>------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Itelmen</td>
<td>+</td>
<td>1, 2</td>
<td>Ø, INF (various forms)</td>
</tr>
<tr>
<td>Thai</td>
<td>+</td>
<td>1</td>
<td>wâa</td>
</tr>
<tr>
<td>Basque</td>
<td>+</td>
<td>2</td>
<td>-t(z)e</td>
</tr>
<tr>
<td>Abun</td>
<td>+</td>
<td>1</td>
<td>do/Ø</td>
</tr>
<tr>
<td>Bambara</td>
<td>+</td>
<td>1</td>
<td>ka</td>
</tr>
<tr>
<td>Georgian</td>
<td>+</td>
<td>3</td>
<td>-a</td>
</tr>
<tr>
<td>Bukiyip</td>
<td>+</td>
<td>1</td>
<td>Ø</td>
</tr>
<tr>
<td>Abkhaz</td>
<td>+</td>
<td>1, 3</td>
<td>NFIN/-ra</td>
</tr>
<tr>
<td>Polish</td>
<td>+</td>
<td>1, 2, 3</td>
<td>ze, INF, -nie</td>
</tr>
<tr>
<td>Burushaski</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>+</td>
<td>2</td>
<td>-e/-i</td>
</tr>
<tr>
<td>Alamblak</td>
<td>+</td>
<td>3</td>
<td>-nef/-(kfë)t</td>
</tr>
<tr>
<td>Pipil</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wambon</td>
<td>+</td>
<td>1</td>
<td>-e</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>+</td>
<td>3</td>
<td>-i/-an</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>+</td>
<td>1, 2</td>
<td>bifi/dati, fu/fi</td>
</tr>
<tr>
<td>Babungo</td>
<td>+</td>
<td>1</td>
<td>lâa</td>
</tr>
<tr>
<td>Nama</td>
<td>+</td>
<td>1</td>
<td>Ixâis-à</td>
</tr>
<tr>
<td>Hdi</td>
<td>+</td>
<td>1</td>
<td>kâ</td>
</tr>
<tr>
<td>Mandarin Ch</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil</td>
<td>+</td>
<td>2</td>
<td>-atu,-(kk)â</td>
</tr>
<tr>
<td>Kisi</td>
<td>+</td>
<td>1, 2</td>
<td>(m)âà, Ø (INF)</td>
</tr>
<tr>
<td>Nung</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garo</td>
<td>+</td>
<td>1, 3, 2</td>
<td>in-e, -a, -a-nî, -na, -kan-a</td>
</tr>
<tr>
<td>Krongo</td>
<td>+</td>
<td>1,2/3</td>
<td>âni tînn, (t)-</td>
</tr>
<tr>
<td>Hixkaryana</td>
<td>+</td>
<td>3</td>
<td>ni/-thi/-hito + ri</td>
</tr>
<tr>
<td>Slave</td>
<td>+</td>
<td>1</td>
<td>gü, ghâ</td>
</tr>
<tr>
<td>Nivkh</td>
<td>+</td>
<td>1</td>
<td>-vut/-vur, Ø</td>
</tr>
<tr>
<td>Greenlandic</td>
<td>+</td>
<td>3</td>
<td>-niq</td>
</tr>
<tr>
<td>Nunggubuyu</td>
<td>none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuscarora</td>
<td>none</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7.43: Availability and type of relative clauses in languages without adjectives, but with a flexible lexical strategy for the function of modifier in a referential phrase

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid relative clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Tagalog</td>
<td>–</td>
<td></td>
<td></td>
<td>contentives</td>
</tr>
<tr>
<td>2 Kharia</td>
<td>+</td>
<td>2, 3, 1</td>
<td>na-wala, -al, je, a/i</td>
<td>contentives</td>
</tr>
<tr>
<td>3 Samoan</td>
<td>+</td>
<td>1</td>
<td>-e</td>
<td>contentives</td>
</tr>
<tr>
<td>4 Guarani</td>
<td>+</td>
<td>1</td>
<td>va</td>
<td>contentives</td>
</tr>
<tr>
<td>5 Warao</td>
<td>+</td>
<td>1</td>
<td>kotai</td>
<td>non-verbs</td>
</tr>
<tr>
<td>6 Turkish</td>
<td>+</td>
<td>2</td>
<td>-An</td>
<td>non-verbs</td>
</tr>
<tr>
<td>7 Quechua</td>
<td>–</td>
<td></td>
<td></td>
<td>nominals</td>
</tr>
<tr>
<td>8 Ma’di</td>
<td>+</td>
<td>2</td>
<td>-re/-bá</td>
<td>nominals</td>
</tr>
<tr>
<td>9 Gooniyandi</td>
<td>–</td>
<td></td>
<td></td>
<td>nominals</td>
</tr>
<tr>
<td>10 Hungarian</td>
<td>+</td>
<td>2, 3, 1</td>
<td>-ó, ótt, andó, RelPron</td>
<td>nominals</td>
</tr>
<tr>
<td>11 Japanese</td>
<td>+</td>
<td>1</td>
<td>Ø (RSP)</td>
<td>nominals</td>
</tr>
<tr>
<td>12 Hmong Njua</td>
<td>+</td>
<td>1</td>
<td>kws</td>
<td>modifiers</td>
</tr>
<tr>
<td>13 Lango</td>
<td>+</td>
<td>1</td>
<td>à-mé</td>
<td>modifiers</td>
</tr>
<tr>
<td>14 Ket</td>
<td>+</td>
<td>1</td>
<td>RelPron, -s/-bes</td>
<td>modifiers, S/D adjectives</td>
</tr>
<tr>
<td>15 Koasati</td>
<td>+</td>
<td>2</td>
<td>various forms</td>
<td>D modifiers, S adjectives</td>
</tr>
</tbody>
</table>

Table 7.44: Availability and type of relative clauses in languages with an open class of simple adjectives

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid relative clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Kayardild</td>
<td>+</td>
<td>1, 2, 3</td>
<td>Ø, -Thirí-n, -n-garrba</td>
<td>adjectives</td>
</tr>
<tr>
<td>2 Thai</td>
<td>–</td>
<td></td>
<td></td>
<td>adjectives (S. modifiers)</td>
</tr>
<tr>
<td>3 Basque</td>
<td>–</td>
<td></td>
<td></td>
<td>adjectives (S/modifiers)</td>
</tr>
<tr>
<td>4 Abun</td>
<td>+</td>
<td>1</td>
<td>gato</td>
<td>adjectives (S. modifiers)</td>
</tr>
</tbody>
</table>
## Table 7.45: Availability and type of relative clauses in languages without an open class of simple adjectives, but with derived adjectives (and a small class of simple adjectives)

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid relative clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Georgian</td>
<td>+</td>
<td>1, 2, 3</td>
<td>Several participles, REL PRON</td>
<td>adjectives</td>
</tr>
<tr>
<td>6 Bukiyip</td>
<td>+</td>
<td>1</td>
<td>(ú)li</td>
<td>adjectives</td>
</tr>
<tr>
<td>7 Abkhaz</td>
<td>+</td>
<td>1</td>
<td>REL.PRON + NFIN</td>
<td>adjectives</td>
</tr>
<tr>
<td>8 Polish</td>
<td>+</td>
<td>1, 2</td>
<td>Several participles, REL PRON</td>
<td>adjectives</td>
</tr>
<tr>
<td>9 Burushaski</td>
<td>+</td>
<td>2</td>
<td>-im/-um/-am</td>
<td>adjectives</td>
</tr>
<tr>
<td>10 Lavukaleve</td>
<td>+</td>
<td>1</td>
<td>Person suffix</td>
<td>adjectives</td>
</tr>
<tr>
<td>11 Alamblak</td>
<td>+</td>
<td>2</td>
<td>ind</td>
<td>adjectives</td>
</tr>
<tr>
<td>12 Pipil</td>
<td>+</td>
<td>1</td>
<td>ne, ke</td>
<td>adjectives</td>
</tr>
<tr>
<td>13 Wambon</td>
<td>+</td>
<td>1</td>
<td>-a + o</td>
<td>adjectives</td>
</tr>
<tr>
<td>14 Dhaasanac</td>
<td>−</td>
<td></td>
<td></td>
<td>adjectives</td>
</tr>
<tr>
<td>15 Berbice Dutch</td>
<td>+</td>
<td>1</td>
<td>Wh</td>
<td>adjectives</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid relative clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Itelmen</td>
<td>+</td>
<td>1</td>
<td>min</td>
<td>D adjectives, S modifiers</td>
</tr>
<tr>
<td>2 Nama</td>
<td>+</td>
<td>1</td>
<td>Ø no INDIC (RSP)</td>
<td>S/D adjectives</td>
</tr>
<tr>
<td>3 Babungo</td>
<td>−</td>
<td></td>
<td></td>
<td>S/D adjectives</td>
</tr>
<tr>
<td>4 Kisi</td>
<td>+</td>
<td>1</td>
<td>CL</td>
<td>S/D adjectives</td>
</tr>
</tbody>
</table>
Table 7.46: Availability and type of relative clauses in languages without an open class of simple/derived class adjectives, but with a small class of adjectives

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid relative clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paiwan</td>
<td>+</td>
<td>2</td>
<td><em>in</em>- <em>an</em> + <em>a</em></td>
<td>S adjectives</td>
</tr>
<tr>
<td>Bambara</td>
<td>+</td>
<td>2, 1</td>
<td><em>le/-ne</em>, <em>min(u)/mun(u)</em></td>
<td>S adjectives</td>
</tr>
<tr>
<td>Hdi</td>
<td>+</td>
<td>1, 2</td>
<td><em>tà, tà</em> + NMLZ, <em>-a</em></td>
<td>S adjectives</td>
</tr>
<tr>
<td>Mandarin Ch</td>
<td>−</td>
<td></td>
<td></td>
<td>S adjectives</td>
</tr>
<tr>
<td>Tamil</td>
<td>+</td>
<td>2</td>
<td><em>-a</em></td>
<td>S adjectives</td>
</tr>
<tr>
<td>Nung</td>
<td>+</td>
<td>1</td>
<td><em>(ti-və) (+DEM) (+FOC)</em></td>
<td>S adjectives</td>
</tr>
</tbody>
</table>

Table 7.47: Availability and type of relative clauses in languages without a lexical strategy for the function of modifier in a referential phrase

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid relative clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garo</td>
<td>+</td>
<td>3</td>
<td><em>-gip-a</em></td>
<td>X</td>
</tr>
<tr>
<td>Krongo</td>
<td>−</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Hixkaryana</td>
<td>none</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Slave</td>
<td>+</td>
<td>1</td>
<td><em>i, ŝi, l̂i</em></td>
<td>X</td>
</tr>
<tr>
<td>Nivkh</td>
<td>+</td>
<td>1</td>
<td>NONFIN</td>
<td>X</td>
</tr>
<tr>
<td>West Greenlandic</td>
<td>+</td>
<td>3</td>
<td><em>-ta/-sa</em></td>
<td>X</td>
</tr>
<tr>
<td>Nunggubuyu</td>
<td>none</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Tuscarora</td>
<td>none</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Table 7.48: Availability and type of adverbial manner clauses in languages without an open class of simple/derived manner adverbs, but with a flexible lexical strategy for the function of modifier in a predicate phrase (and in some cases also a small class of rigid manner adverbs)

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid adverbial manner clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagalog</td>
<td>−</td>
<td></td>
<td></td>
<td>Flex (contentives)</td>
</tr>
<tr>
<td>Kharia</td>
<td>+</td>
<td>2, 3</td>
<td>-ker, -ga, -ta</td>
<td>Flex (contentives)</td>
</tr>
<tr>
<td>Samoan</td>
<td>none</td>
<td></td>
<td></td>
<td>Flex (contentives) + S MADv</td>
</tr>
<tr>
<td>Guaraní</td>
<td>+</td>
<td>1</td>
<td>vo</td>
<td>Flex (contentives)</td>
</tr>
<tr>
<td>Warao</td>
<td>none</td>
<td></td>
<td></td>
<td>Flex (non-verbs)</td>
</tr>
<tr>
<td>Turkish</td>
<td>+</td>
<td>2</td>
<td>-(y)ArAK, -(y)A</td>
<td>Flex (non-verbs, D modifiers)</td>
</tr>
<tr>
<td>Kayardild</td>
<td>−</td>
<td></td>
<td></td>
<td>Flex (non-verbs / predicatives) + S MADv</td>
</tr>
<tr>
<td>Paiwan</td>
<td>+</td>
<td>1</td>
<td>a parhu</td>
<td>Flex (predicatives) + S MADv?</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>none</td>
<td></td>
<td></td>
<td>Flex (modifiers)</td>
</tr>
<tr>
<td>Ket</td>
<td>+</td>
<td>1</td>
<td>eta qor’a, ñsqà,</td>
<td>Flex (modifiers) + S MADv</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-PROSEC</td>
<td></td>
</tr>
</tbody>
</table>
### Table 7.49: Availability and type of adverbial manner clauses in languages with an open class of simple manner adverbs

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid adverbial manner clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Gooniyandi</td>
<td>+</td>
<td>2</td>
<td>-wadda, -mawoo, -bari, -ya/-gowaa</td>
<td>MAdv</td>
</tr>
<tr>
<td>2 Hungarian</td>
<td>+</td>
<td>1, 2</td>
<td>úgy (a-)hogy, -vá/-vé, -vén</td>
<td>MAdv</td>
</tr>
<tr>
<td>3 Japanese</td>
<td>+</td>
<td>2</td>
<td>-te/-de/-ite, -i/-Ø</td>
<td>MAdv</td>
</tr>
<tr>
<td>4 Lango</td>
<td>none</td>
<td></td>
<td></td>
<td>MAdv</td>
</tr>
<tr>
<td>5 Abun</td>
<td>+</td>
<td>1</td>
<td>sa gato</td>
<td>MAdv</td>
</tr>
<tr>
<td>6 Bambara</td>
<td>+</td>
<td>2</td>
<td>-tò</td>
<td>MAdv</td>
</tr>
<tr>
<td>7 Georgian</td>
<td>+</td>
<td>1</td>
<td>ra, rogorc</td>
<td>MAdv</td>
</tr>
<tr>
<td>8 Bukiyip</td>
<td>+</td>
<td>1</td>
<td>bwidou -(u)mu</td>
<td>MAdv</td>
</tr>
</tbody>
</table>

### Table 7.50: Availability and type of adverbial manner clauses in languages without an open class of simple manner adverbs, but with derived manner adverbs (and in some cases a small class of simple manner adverbs)

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid adverbial manner clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Ma’di</td>
<td>+</td>
<td>2</td>
<td>-zi+ si</td>
<td>S/D MAdv</td>
</tr>
<tr>
<td>2 Koasati</td>
<td>+</td>
<td>2</td>
<td>-n, -k, -t</td>
<td>S/D MAdv</td>
</tr>
<tr>
<td>3 Itelmen</td>
<td>none</td>
<td></td>
<td>QUATZ</td>
<td>D MAdv</td>
</tr>
<tr>
<td>4 Thai</td>
<td>none</td>
<td></td>
<td>en bezala</td>
<td>D MAdv</td>
</tr>
<tr>
<td>5 Basque</td>
<td>+</td>
<td>1</td>
<td>en bezala</td>
<td>D MAdv</td>
</tr>
<tr>
<td>6 Abkhaz</td>
<td>+</td>
<td>1</td>
<td>-š- + -NFIN</td>
<td>S/D MAdv</td>
</tr>
<tr>
<td>7 Polish</td>
<td>+</td>
<td>1, 2</td>
<td>PRS.PL-c-, jak (gdy)by</td>
<td>S/D MAdv</td>
</tr>
<tr>
<td>8 Burushaski</td>
<td>+</td>
<td>2</td>
<td>n-STEM-(a)n</td>
<td>S/D MAdv</td>
</tr>
<tr>
<td>9 Nama</td>
<td>+</td>
<td>2</td>
<td>se/l’aatnsii (no INDIC)</td>
<td>D MAdv</td>
</tr>
<tr>
<td>10 Hdi</td>
<td>none</td>
<td></td>
<td></td>
<td>S/D MAdv</td>
</tr>
<tr>
<td>11 Garo</td>
<td>+</td>
<td>2</td>
<td>-e/-e-min/-e-r</td>
<td>S/D MAdv</td>
</tr>
</tbody>
</table>
### Table 7.51: Availability and type of relative clauses in languages without an open class of simple/derived class manner adverbs, but with a small class of manner adverbs

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid adverbiacl manner clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Quechua</td>
<td>+</td>
<td>2</td>
<td>-shpa, -y RDP</td>
<td>S Madv</td>
</tr>
<tr>
<td>2 Lavukaleve</td>
<td>-</td>
<td></td>
<td></td>
<td>S Madv</td>
</tr>
<tr>
<td>3 Alamblak</td>
<td>-</td>
<td></td>
<td></td>
<td>S Madv</td>
</tr>
<tr>
<td>4 Babungo</td>
<td>+</td>
<td>2</td>
<td>ki(i)/Ø</td>
<td>S Madv</td>
</tr>
<tr>
<td>5 Mandarin Ch.</td>
<td>+</td>
<td>2</td>
<td>zhe</td>
<td>S Madv</td>
</tr>
<tr>
<td>6 Tamil</td>
<td>-</td>
<td></td>
<td></td>
<td>S Madv</td>
</tr>
<tr>
<td>7 Krongo</td>
<td>-</td>
<td></td>
<td></td>
<td>S Madv</td>
</tr>
<tr>
<td>8 Hixkaryana</td>
<td>-</td>
<td></td>
<td></td>
<td>S Madv</td>
</tr>
</tbody>
</table>

### Table 7.52: Availability and type of adverbiacl manner clauses in languages without a lexical strategy for the function of modifier in a predicate phrase

<table>
<thead>
<tr>
<th>Language</th>
<th>Rigid adverbiacl manner clauses</th>
<th>Type</th>
<th>Structural coding</th>
<th>Lexical strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Pipil</td>
<td>+</td>
<td>1</td>
<td>ke:n-aken, ADV</td>
<td>X</td>
</tr>
<tr>
<td>2 Wambon</td>
<td>+</td>
<td>1, 2</td>
<td>-no/-o, ka</td>
<td>X</td>
</tr>
<tr>
<td>3 Dhaasanac</td>
<td>-</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>4 Berbice Dutch</td>
<td>-</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>5 Kisi</td>
<td>-</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>6 Nung</td>
<td>+</td>
<td>1</td>
<td>bat</td>
<td>X</td>
</tr>
<tr>
<td>7 Slave</td>
<td>+</td>
<td>1</td>
<td>ghărē</td>
<td>X</td>
</tr>
<tr>
<td>8 Nivkhh</td>
<td>+</td>
<td>2</td>
<td>r/-t/-f/-n</td>
<td>X</td>
</tr>
<tr>
<td>9 West Greenlandic</td>
<td>-</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>10 Nunggubuyu</td>
<td>none</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>11 Tuscarora</td>
<td>none</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
8 Discussion: Flexibility and Functional Transparency

8.1 Introduction

The results presented in the previous chapter suggest that there is a dependency relation between the amount of flexibility displayed by the parts of speech classes of a particular language, and the amount of flexibility displayed by the dependent clause constructions of that language. More specifically, the data show that, as expected, the possibility of having (pervasively) flexible deranked DC constructions in a language is dependent of the presence of (pervasively) flexible PoS classes in that language.

However, there were also a number of unexpected findings. First, not all languages with flexible PoS classes also have flexible deranked DCs. Moreover, those that do have flexible deranked DCs combine them with rigid (deranked or balanced) DCs. Second, there are no dependency relations between lexical and (deranked) clausal constructions with less pervasive flexibility, i.e. between nominals and nominal clauses, and between modifiers and modifier clauses. Third, the presence of specific rigid DCs is not dependent upon the presence of parallel types of rigid PoS classes. Finally, there are in general no correlations between the functional possibilities of PoS and balanced DCs.

The aim of the present chapter is to interpret the results of Chapter 7 from a specific functionalist perspective. In particular, recent functional-typological research has advanced the idea that flexibility and rigidity are relative rather than absolute notions, which may be applicable in various degrees to specific construction types at different levels of grammar. These
studies suggest that the categorial specificity of linguistic constructions increases – or their flexibility decreases – when they become structurally more complex (Haig 2006, Lehmann 2008). This generalization can be regarded as a specific instance of a more general functional principle regarding complexity in language systems, namely that flexibility or multifunctionality in one area of the grammar must be counterbalanced or ‘traded off’ by rigidity or categorial specificity in another area, in order to guarantee the functional identifiability of linguistic units within an actual utterance.

The structure of this chapter is as follows. I start out, in section 8.2, with a more extensive discussion of the relevant literature. Subsequently, in section 8.3, I apply the hypothesis that categoriality increases with grammatical complexity to the case at hand: (simple) PoS classes and (complex) dependent clauses. It will be shown that this approach accounts for the functional behaviour of virtually all deranked DCs attested in the sample. In section 8.4 I turn to balanced DCs, and show that, even though they do not behave in accordance with the specific hypothesis of increased complexity/categoriality, they do fit into a larger explanatory picture based on the principle of functional transparency. I conclude, in section 8.5, that the functionalist perspective developed in the present chapter provides a complete account of the data.

8.2 Theoretical background:

Trade-off effects between flexible and rigid constructions

Recently, the debate in the functional-typological literature about the existence or non-existence of ‘truly’ flexible languages (see Chapter 2, section 2.5) has taken a new turn. In particular, it has been suggested that flexibility (or pre-categoriality) and rigidity (or categorial specificity) are not absolute notions that do or do not apply to complete language systems, or even sub-systems. Rather, they are viewed as gradable notions that may apply in certain degrees to individual constructions, pertaining to specific levels of grammar.

A first proposal to this effect is Haig’s (2006). He distinguishes four grammatical levels at which any language may display a certain amount of categorial specificity: First, there is the basic level of so-called radical elements. At this level, languages may differ in their distinctions of major lexical categories. Second, at the level of input of lexical derivation, languages may display variation in the extent to which their derivational processes select certain base categories but not others. At the third level we find output of
lexical derivation, which concerns the degree to which derivational processes result in categorially specified forms. At the final, highest level of complexity, languages may vary in terms of the categorial distinctions drawn in inflectional and other syntactic categories (Haig 2006: 44). In short, languages differ in the amount and type of categorial specificity (or lack of it) displayed by their constructions at different levels of grammatical complexity.

Haig (2006) puts forward the hypothesis that this variation is not random, but rather constrained by what he terms the Principle of Successively Increasing Categorisation (henceforth PSIC), and formulates as follows:

“The extent to which different levels of linguistic structure impose categorial distinctions increases monotonically as we move to the right of the following hierarchy:
Radical elements > Derivational input > Derivational output > Inflection/Syntax”
(Haig 2006: 45)

Haig discusses a number of languages that conform to the PSIC, and also points out some possible counterexamples to it115. In general, however, he suggests that cross-linguistic differences in ‘flexibility’ can be captured as points on a cline defined by the hierarchical complexity level at which maximal categorial specificity is reached in a language. At one extreme of this cline, we find early-categorizing languages, in which categorial distinctions are built into the inventory of radical elements116, and are observed throughout the morphology and projected into the syntax. At the other extreme we find late-categorizing languages, in which the categorization of linguistic elements happens only through their syntactic or phrase-structural configurations. In between these two extremes we may find languages in which categorial

115 Interestingly, the counterexamples that Haig (2006) mentions are all flexible languages: Tagalog, Mundari, and Riau Indonesian (Gil 2000). As Haig (2006: 46) puts it, in these languages “the syntax unpacks categorial distinctions drawn in morphology, rather than refining them”. However, as I have argued in Chapter 2, the morphological categories of flexible languages are not defined in terms of phrase structure. Therefore, categorization at the morphological level is not only of a different degree than categorization at the syntactic level, but also of an entirely different nature (see also Himmelmann (2007) on Tagalog).

116 Here the question becomes relevant whether or not one assumes that, at some level of abstraction, all languages have a stock of un-categorized roots (cf. Arad 2003). For the present purposes, it is not necessary to take a specific stance in this respect; the general point is that the degree of categorial distinctiveness is supposed to be lower at the level of roots than at the level of stems, words, and phrases. For more discussion see also Haig (2006), who provides an overview of relevant literature since Aronoff (1976).
distinctions are introduced at intermediate (morphological) levels, and/or in which lower level categorial distinctions become more fine-grained when moving upwards along the complexity hierarchy.

Lehmann (2008) arrives at a very similar generalization on the basis of data from a small sample of six languages (German, English, Spanish, Latin, Yucatec Maya, and Mandarin Chinese). In particular, Lehmann finds that, for each language in this sample, the degree of categoriality at the level of stems is higher than at the level of roots (cf. Haig's 'radical elements'). In addition, Lehmann argues that every linguistic unit has a unique category at the sentence level, so that “we can safely generalize that categoriality increases with the grammatical levels.” (Lehmann 2008: 561) The term that Lehmann uses for this generalization is staggering level-dependent categoriality, which I will abbreviate as SLC.

Interestingly, Lehmann (2008: note 17, p. 565) points out a possible contradiction between his SLC principle and Dik's (1985, 1997) Principle of Formal Adjustment (PFA), the latter of which has been discussed in Chapter 3 (section 3.3.3.3). Lehmann argues that if, on the one hand, categoriality increases with grammatical levels, “this would invite the inference that higher levels dictate the necessary categories, and categorizations made at lower levels follow that model”. Dik's PFA, on the other hand, claims that derived, secondary constructions are coined on the basis of less complex, primary constructions. This, as Lehmann puts it, “seems to entail that categories of lower levels serve as models for categorization at higher levels”. Lehmann does not propose a solution to this problem, but I believe that, from the perspective of the present chapter, there are two points worth making.

First, consider the possibility that higher level categories determine lower level ones. Lehmann argues that the ultimate level of categorization of linguistic structures is determined by the propositional functions that must be performed within an actual utterance, i.e. reference, predication, and modification. According to him, this level of categorization constitutes the syntactic level. However, it seems that Lehmann confuses propositional functions with syntactically complex (rather than morphologically complex or lexically simplex) linguistic constructions that may express these functions. In other words, reference, predication, and modification are not constructions themselves; they are functions that may be expressed by constructions of different degrees of structural complexity.

Note that it is not necessary to measure the level of categorial specificity of linguistic constructions relative to some universal set of functions, which
would determine the maximal degree of categorical distinctiveness in any language. It is quite possible to assume that the ultimate set of relevant functional distinctions is a language-specific one. This still allows one to compare the set of function(s) that can be expressed by one construction with the functional possibilities of another construction that has a higher or lower degree of structural complexity.

A second issue concerns the status of derived, secondary constructions as they figure in Dik’s PFA. The PFA predicts that derived, secondary constructions will be formally modelled on their non-derived, primary counterparts. However, it is not always obvious to what extent a more complex construction of a particular type can be regarded as derived from or secondary in relation to a simpler construction. The soundest basis on which to decide this would probably be a functional one: two constructions can be regarded a primary and secondary instantiation of the same category if they perform the same function(s). Obviously, however, this undermines the possibility to compare the functional possibilities of the two construction types, and as such leaves the Principle of Increasing Categoriality toothless. Therefore, a formal criterion is required in order to establish a derivational (primary-secondary) relation between two constructions, before the nature of their functional connection can be the subject of any testable prediction.

Haig (2006) and Lehmann (2008) have tested the hypothesis of Increasing Categoriality on data that involve derivations from simple roots to (derived) stems, and from stems to (inflected) words. In these cases, the formal connection between the simpler and the more complex form is quite clear. In the case of DCs, in contrast, such a derivational connection is less straightforwardly established. In particular, while deranked DCs are formally modelled on lexical or phrasal expressions, balanced DCs are not; the latter retain the formal properties of independent clauses. In view of this difference, I will assume that deranked but not balanced DCs can be regarded as secondary constructions in the sense of Dik’s PFA.\(^{117}\)

117 This difference between deranked and balanced DCs was already built into the hypotheses formulated in Chapter 4 and tested in Chapter 7, to the extent that a weaker functional match was expected between PoS and balanced DCs than between PoS and deranked DCs. In fact however, no significant correlations at all were found between PoS and balanced DCs.
expected that the functional possibilities of secondary, derived constructions (deranked DCs) are not necessarily identical to those of their primary bases (PoS classes), but may also be – and usually are – more restricted. In other words, Lehmann’s observation that the PFA suggests an influence of the degree of categoriality of lower-level (or primary) constructions on the categoriality of higher-level (secondary) constructions seems justified, but this influence is not expected to take the form of direct modelling or copying. Rather, the simple, primary construction determines the minimal level of categoriality (or maximal level of flexibility) of the secondary, more complex construction.

Thus, in general, the degree of categorial distinctiveness displayed by a particular language at lower levels of complexity determines the amount of categorization that needs to be implemented at higher levels, in order to achieve maximal categorial distinctiveness at the level of the utterance. Lehmann describes this as follows:

“In a language with low stem categoriality, the speaker’s task of assigning the words to categories is fully achieved only at the level of syntax. This is done by inserting words into certain syntactic templates, which force a syntactic category on them. In a language with high stem categoriality the syntax contributes nothing to the categorization of words, which means that such templates play a minor role in constructions.” (Lehmann 2008: 557)\(^{118}\)

In fact, this idea is advanced already in an earlier study, namely Hengeveld et al. (2004). In particular, they predict that:

“Languages with flexible lexemes will have certain morpho-syntactic strategies at their disposal that provide the hearer with clues as to the correct interpretation of the flexible lexeme. Morpho-syntactically, there are basically two possibilities to guide the hearer as regards the intended function of a flexible lexeme in the actual linguistic expression:

\(^{118}\) Lehmann describes the implementation of categorical specificity as a task for the speaker. In fact, however, in order for this functional motivation to work, one should also take into account the perspective of the hearer: He or she should be able to identify the right function of each form. This point is acknowledged in the work of Hengeveld et al. (2004) (see below).
(i) morphological markers in the immediate linguistic context; and (ii) syntactic patterning.”
(Hengeveld, Rijkhoff & Siewierska 2004: 547-548)

This prediction is strongly supported by data from a balanced sample of 50 languages. In particular, it is shown that languages with flexible PoS resolve the functional ambiguity of their lexemes either by means of word order restrictions or by means of special markers indicating the functions in which lexemes appears. Note that this is an implicational universal in the weak sense, since languages with rigid lexeme classes may also make use of fixed word order or function-indicating markers. Thus, Hengeveld et al. (2004) interpret their results as evidence for a trade-off effect between the lexical, morphological, and syntactic coding means that a language has at its disposal. In particular, these coding means must together ensure that the function of any linguistic units is identifiable within an utterance. While this means that functional ambiguity must be resolved, it does not mean that redundant marking is necessarily excluded (see also Van Lier 2006, and cf. Chapter 9).

Functional trade-off effects have also been studied for more specific functional sub-domains. For instance, Sinnemäki (2008) finds such an effect between three types of strategies for core argument encoding: fixed word order, head marking, and dependent marking. In addition, and in line with the findings of Hengeveld et al. (2004), Sinnemäki suggests that languages adhere more strongly to distinctiveness than to economy, which is another way of saying that they allow for redundancy more than for ambiguity. Finally, Sinnemäki emphasizes that complexity trade-off should not be regarded as an all-encompassing principle, but rather as one of several factors that motivate certain aspects of the encoding of functions in a particular domain.

Similarly, Frajzyngier & Shay (2003) present their *Principle of Functional Transparency* as an explanatory mechanism in the division of labour between coding devices pertaining to different grammatical levels, in the expression of a certain functional domain. In particular, their principle states that:

“Every constituent must have a transparent function within the utterance. […] This is achieved either through the inherent properties of lexical items in the clause or through the system of grammatical means, which may include affixes, adpositions, linear order and free morphemes with grammatical functions.” (Frajzyngier & Shay 2003: 4-5)
Frajzyngier & Shay (2003: 7) explicitly do not assume a universal level of maximal categorial specificity. Rather, in their view, functional transparency is determined relative to the set of functional distinctions that are relevant in a particular language. As mentioned above, such a stance does not preclude the possibility to find cross-linguistic generalizations with respect to the way in which the workload of categorization is divided over the available grammatical resources.

To sum up, recent functional-typological literature suggests that there is a relationship of interdependence or trade-off between different construction types, pertaining to different levels of grammatical complexity, as regards their contribution to the ultimate goal of establishing functional transparency at the utterance level. In the next two sections (8.3 and 8.4) I will discuss the results from Chapter 7 against this theoretical background.

8.3 Applying the complexity hypothesis to PoS and deranked DCs

8.3.1 Hypothesis

As explained in the previous section, deranked DCs can be regarded as secondary constructions, derived from primary, underived lexical constructions, in terms of the Principle of Formal Adjustment. Secondary, derived constructions are assumed to be more complex than primary, underived ones. Therefore, the Principle of Increasing Categorization or Staggering Level-dependent Categoriality, which states that categoriality increases with complexity, predicts that:

(1) In any language, the functional flexibility of a PoS class is either greater than or equal to the flexibility of a deranked DC construction that can be used in at least one of the same function(s).

This hypothesis may be formally represented as a constraint, in the following way:

(2) $[\text{flex PoS} \geq \text{flex deranked DC}]$

There are three possibilities concerning the functional possibilities of a particular deranked DC that are in accordance with the constraint in (2). They are listed in (3):
A deranked DC has the same amount of flexibility as the relevant PoS class;

b. A deranked DC has less flexibility than the relevant PoS class;

c. A deranked DC has no flexibility (i.e. it is a rigid deranked DC).

The constraint in (2) is violated when a deranked DC has more flexibility than the relevant PoS class, i.e. when the DC can express more propositional functions than the PoS class.

Note that the situations described in (3b-c) would constitute counterexamples to some of the hypotheses tested in Chapter 7, but do not count as such in view of the prediction in (1)/(2). Consider first the situation in (3b): It would not constitute a counterexample to the hypotheses about global matches between flexibility in PoS systems and deranked DCs, since for a global match it is not necessary that flexible PoS and DCs can express exactly the same type and amount of functions. However, it would indeed constitute a counterexample to the hypotheses in Chapter 7 about one-to-one functional matches between specific types of flexible PoS classes and flexible deranked DC constructions (see the predictions in (18b-d), and (21b-d) of Chapter 7). Furthermore, situations like those described in (3c) would constitute counterexamples to all hypotheses in Chapter 7, whether they make reference to global matches between flexibility in PoS and DC systems, or to specific matches between particular flexible PoS classes and flexible deranked DC constructions.

Note further that in cases where the relevant flexible PoS class can express two functions (i.e. when it is a class of nominals or modifiers) and the DC only one, the difference between situation (3b) and (3c) becomes irrelevant: When the DC can express one function less than the PoS class, then there is only one function left, so that the DC construction is automatically rigid rather than flexible. I will return to this issue in more detail below.

Furthermore, recall that virtually all languages with a flexible PoS system have one or more rigid PoS classes. It is predicted that, whenever a language has a deranked DC construction that is used in a function for which a rigid PoS class is available, this deranked DC will be rigid too. In other words, if there is already maximal categoriality at the level of the simple, primary construction (PoS), then this situation should be retained at the level of the derived, secondary construction (deranked DC).
The same prediction, of course, applies to languages with a rigid PoS system, i.e. a system consisting of rigid classes only: Their deranked DCs, if they have any, are expected to all be rigid too. Note that, whenever a language lacks a lexical category for the expression of one or more particular function(s), there is no basis on which to predict the functional pattern of DC(s) appearing in any of those functions.

In the next sub-section I present the results for the hypothesis formulated in (1)/(2). I first discuss languages with flexible PoS systems (section 8.3.2.1), followed by languages with rigid PoS systems (section 8.3.2.2).

8.3.2 Results

8.3.2.1 Distribution of deranked DCs in languages with flexible PoS systems

Introduction

Table 8.1 below lists all languages in the sample with flexible PoS classes, and compares the functional possibilities of these PoS classes with the functional possibilities of deranked DCs (if the language has any) occurring in one or more of the same function(s). The rightmost column indicates either that a particular DC has the same functional distribution as the relevant flexible PoS (=), or that it can express a smaller amount of functions than the flexible PoS (>). The data in Table 8.1 thus show that the hypothesis in (1)/(2) above is confirmed in all cases (with the possible exception of one case that will be discussed further below), since there are no deranked DCs with a higher degree of flexibility than the relevant PoS class (<). The figures between the brackets in the rightmost column of Table 8.1 indicates the difference in flexibility between the DC and the PoS, in terms of the amount of expressible functions: (1) means that the DC can express one function less than the PoS, (2) means two functions less, and (3) means three functions less.

In short, there are some cases in which a flexible PoS class and a flexible DC construction can express exactly the same set of functions. This corresponds to the situation described under (3a) above. In other cases, the DC can express a subset of the functions that are expressed by the PoS class. These latter cases can be subdivided into two groups. In accordance with (3b) above, there is one group of cases in which the DC can express less functions
than the PoS class, but still remains flexible. And, in accordance with (3c) above, there is another group of cases in which the subset of functions that the DC can express is a singleton set, so that the DC is in fact a rigid rather than a flexible construction. In what follows I discuss and illustrate these three types of cases in turn.

Table 8.1: Languages with a flexible PoS class, and the distribution of deranked DCs that appear in at least one of the same function(s)

<table>
<thead>
<tr>
<th>Language</th>
<th>Relevant flexible PoS</th>
<th>Deranked DCs in the same function(s)</th>
<th>Form</th>
<th>Function</th>
<th>Type</th>
<th>Flex PoS ≥ Flex DC?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tagalog</td>
<td>contentives</td>
<td>pag-</td>
<td></td>
<td>contentive/</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td>Kharia</td>
<td>contentives</td>
<td>RDP/Ø contentive</td>
<td>3</td>
<td>=</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-na multi-functional</td>
<td>3</td>
<td>&gt; (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-na-wala relative</td>
<td>2</td>
<td>&gt; (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-al relative</td>
<td>3</td>
<td>&gt; (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-ker relative</td>
<td>3</td>
<td>&gt; (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-ga adverbial manner</td>
<td>2</td>
<td>&gt; (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-ta + RDP adverbial manner</td>
<td>2</td>
<td>&gt; (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samoan</td>
<td>contentives</td>
<td>-ga/Ø complement</td>
<td>3</td>
<td>&gt; (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ona/ina complement</td>
<td>2</td>
<td>&gt; (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warao</td>
<td>non-verbs</td>
<td>-kitane complement</td>
<td>2</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td>non-verbs</td>
<td>-DIK/ -AcAk, nominal</td>
<td>3</td>
<td>&gt; (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-mA complement</td>
<td>3</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-mAk complement</td>
<td>2</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-An relative</td>
<td>2</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-(y)ArAk adverbial manner</td>
<td>2</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-(y)A adverbial manner</td>
<td>2</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kayardild</td>
<td>non-verbs</td>
<td>-n- multi-functional</td>
<td>3</td>
<td>=</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Thirri-n relative</td>
<td>3</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-ngarrba relative</td>
<td>2</td>
<td>&gt; (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Relevant flexible PoS</td>
<td>Deranked DCs in the same function(s)</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>
</tr>
<tr>
<td></td>
<td>Form</td>
<td>Function</td>
<td>Type</td>
<td>Flex PoS ≥ Flex DC?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Quechua</td>
<td>-shka</td>
<td>nominal</td>
<td>2</td>
<td>=</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ngapaj</td>
<td>complement</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-chun</td>
<td>complement</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-y</td>
<td>complement</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Ma’di</td>
<td>-lɛ̄,-dʒɔ</td>
<td>nominal</td>
<td>2/3</td>
<td>=</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ka</td>
<td>complement</td>
<td>3</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-rɛ̄-bá</td>
<td>relative</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Gooniyandi</td>
<td>-woo</td>
<td>complement</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Hungarian</td>
<td>-ni</td>
<td>complement</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-á/-és</td>
<td>complement</td>
<td>3</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ó</td>
<td>relative</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ótt</td>
<td>relative</td>
<td>3</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-andó/ -endó</td>
<td>relative</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Japanese</td>
<td>no / mono</td>
<td>complement</td>
<td>3</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Ket</td>
<td>Ø (bare INF)</td>
<td>nominal</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Koasati</td>
<td>-n</td>
<td>adverbial manner</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-k</td>
<td>adverbial manner</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-t</td>
<td>adverbial manner</td>
<td>2</td>
<td>&gt; (t)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Deranked DCs with the same distribution as flexible PoS**

Table 8.1 shows that there are five cases in which a flexible PoS class and a flexible deranked DC construction can express exactly the same type and amount of functions (as signalled by the = sign). An example of such a case is Kayardild, which has lexical non-verbs and deranked multi-functional clauses. The latter are demonstrated in (4a-c, cf. (33a-c) in Chapter 6), where the same DC construction, marked by -n-, appears in the functions of head of a referential phrase (a), modifier in a referential phrase (b), and modifier in a predicate phrase (c):
Kayardild (Evans 1995: 472, 474)

(4) a. Ngada kurri-ja [ki-l-wan-ji]
    1SG.NOM see-ACT 2-PL-POSS-MLOC
dalwani-n-ki thawal-urrk]
dig.up-NMLZ-MLOC yam-MLOC:A.OBL
‘I saw you digging up yams.’

    1-INC-PL-NOM dance-PTC-M.PROP man-M.PROP
kurri-ju
    see-POT
‘We will watch the dancing man.’

c. [Bilaangka-nurru kari-i-n-da] ngada warra-j
    blanket-ASSOC cover-M-CONV-NOM 1SG.NOM go-ACT
‘I went along, covering myself in a blanket.’

Deranked DCs with less flexibility than flexible PoS

In all other cases listed in Table 8.1, the DC can express a subset of the functions of the PoS class (as indicated by the > signs). Two of these cases correspond to the situation described in (3b) above: They involve DCs that can express less functions than the relevant PoS classes, but are nevertheless still flexible (i.e. can be used in at least two functions). These cases are multifunctional -na clauses in Kharia and nominal -DIK/-AcAK clauses in Turkish. The former DC construction can express three functions, which is one function less than lexical contentives in Kharia. This distribution of the -na construction is illustrated in (5a-c) below, where it appears in the functions of head of a referential phrase (a), modifier in a referential phrase (b), and modifier in a predicate phrase (c). The second DC construction, Turkish nominal clauses, can express two functions, which is one function less than the lexical non-verbs in this language. This DC has already been illustrated several times (see Chapter 3 examples (23a-b); Chapter 6 example (37a-b); and Chapter 7 examples (8a-b)).
Kharia (Peterson 2006: 259, 243, 249)

(5) a. [\textit{ip} u \\textit{iku\dagi} sundar \textit{konthe=d=kite}] 
\begin{align*}
1sg & \text{this very beautiful bird=}_{\text{PL}=\text{OBL}} \\
\text{satay=na} & \text{um=}_{\text{ip}} \text{lam}=_{\text{te}} \\
\text{torment=}_{\text{INF}} \text{NEG=}_{1sg} \text{want-}{}_{\text{ACT.PRS}}
\end{align*}

‘I don’t want to torment these beautiful birds.’

b. [\textit{doli=te} \textit{dqi=d=na}] bhere 
\begin{align*}
\text{palanquin=}_{\text{OBL}} \text{take=}_{\text{INF}} \text{time}
\end{align*}

‘the time [they were] to take away the palanquin’

c. …lekin [\textit{lam=na} \textit{lamna}] sou\dip=te \textit{ikud} jughay 
\begin{align*}
\text{but search=}_{\text{INF}} \text{RDP all=}_{\text{OBL}} \text{very much}
\end{align*}
\begin{align*}
\text{qa=piyas} & \text{la=k}\text{i.}
\end{align*}

‘But searching and searching, [they] all became very thirsty.’

**Full and partial retention of flexibility in DCs: Chapter 7 revisited**

Above I discussed the cases in Table 8.1 that represent the situations described in (3a) and (3b) above: Deranked DCs that have the same amount of flexibility as the relevant flexible PoS classes, and deranked DCs that are less flexible than the relevant PoS classes, but still flexible. I want to emphasize here that these are the cases behind the general finding of Chapter 7 that flexibility in the domain of deranked DCs is dependent upon flexibility in the domain of PoS classes. This finding can now be reinterpreted as follows: The maximal amount of flexibility (or the minimal amount of categorial specificity) of deranked DCs in a language is constrained by the amount of flexibility (or categorial specificity) displayed by the PoS classes in that language.

Moreover, and as already pointed out in section 8.3.1, those cases in Table 8.1 which involve a PoS class that expresses two functions and a DC that expresses only one function can be interpreted as instances of the situation in (3b) as well, since the DCs express one function less than the PoS classes. This holds for all DCs in Table 8.1 that are marked with > (1) from Quechua down to Koasati.

If these cases are accordingly counted as representing partial loss of flexibility as compared to the relevant PoS, then the following picture emerges: Among the total 13 languages in Table 8.1 with flexible PoS classes and at least one deranked DC (as opposed to
balanced clauses only), there are 11 in which this deranked DC can either express all the functions that can be expressed by the relevant PoS class, or one function less\textsuperscript{120}. The situation in these 11 languages is summarized in Table 8.2 below. In terms of PoS classes (see the two leftmost columns), a distinction is made between languages with very flexible PoS (contentives or non-verbs) and languages with less flexible PoS (nominals or modifiers). In terms of deranked DCs (see the two top rows), a distinction is made between cases of equal flexibility (\(=\)), and cases in which the DC expresses one function less than the PoS class (\(>1\)).

<table>
<thead>
<tr>
<th>PoS</th>
<th>Deranked DCs</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>contentives/non-verbs</td>
<td>(=)</td>
<td>5</td>
</tr>
<tr>
<td>Tagalog</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kharia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kayardild</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>nominals/modifiers</td>
<td>(&gt;1)</td>
<td>6</td>
</tr>
<tr>
<td>Quechua</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ma’di</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gooinyandi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koasati</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket*</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>11</td>
</tr>
</tbody>
</table>

*Table 8.2: The Principle of Increasing Categorization: Flexible PoS and deranked DCs*

Note that there is one somewhat quirky case in Table 8.2, marked by an asterix: Ket has lexical modifiers and deranked nominal clauses. This means that, while both the lexical and the clausal construction can be used in two functions, there is overlap in only one of these functions, namely modifier in a referential phrase. Therefore, the DC is counted as expressing one function less than the PoS class\textsuperscript{121}.

In short, in the large majority of languages with flexible PoS systems, deranked DCs either display full retention or partial loss of the functional possibilities available at the lexical level. This is in accordance with the hypothesis in (1)/(2).

\textsuperscript{119} These cases stand in contrast to rigid deranked DCs that can express two or three functions less than the relevant PoS. This is indicated in Table 8.1 as \(>2\) or \(>3\). These cases will be discussed shortly.

\textsuperscript{120} As can be seen in Table 8.1 (and see also example (5)), Kharia has both a deranked DC that displays the same flexibility as lexical contentives and a DC that can express one function less than lexical contentives. In order to avoid double values, in Table 8.2 Kharia is mentioned as possessing only the former, most flexible DC.
Before turning to the remaining cases in Table 8.1 – those that involve complete loss of flexibility in the DC (cf. (3c) above) – I will consider flexible deranked DCs from the perspective of functional transparency, as discussed in the previous section. Put simply, any flexible construction, including a flexible deranked DC, presents a potential problem of functional ambiguity. As in the case of flexible lexeme classes (cf. Hengeveld et al. 2004), it is expected that languages will have certain morpho-syntactic means to resolve this ambiguity. The data from my sample confirm this prediction. In particular, the relevant languages employ the same phrase-structural devices that are used for lexical or phrasal constituents to unambiguously mark the functions of flexible deranked DCs. Consider for instance the examples in (6a-b) from Tagalog. Example (6a) illustrates the use of the flexible gerund construction in a functional slot marked by the locative marker sa. Example (6b) shows the use of a simple lexical construction in the same functional slot, marked by the same element:

Tagalog (Himmelmann 2005: 373; Himmelmann 2007: 252)

(6)  a. pag-bawal-an mo ang bata-ng iyó
    sf-forbidden-lv 2sg.poss spec child-lk dist
    sa [pag-la-laró? sa lansangan]
    loc ger-rdp-play loc street
    ‘Forbid that child to play in the street.’

   b. um-sakay silá sa bangká
    av-passenger 3pl loc boat
    ‘They got on the boat.’

Other strategies used to resolve the functional ambiguity of flexible deranked DCs include case marking and rigid word order. For instance in Quechua, -shka clauses are case-marked when they function as the head of a referential

121 It is possible, however, that the bare infinitival form in what I have classified as a nominal clause construction is in fact a lexical derivation when it appears in the function of modifier in a referential phrase. This is hard to say, since in this function the patient argument of the dependent predicate is gapped, and the agent is demoted, so that there are no overtly expressed arguments (cf. example (16) in Chapter 6).
phrase, but not when they modify a referential head. Nominal clauses in Ma’di follow the main predicate when functioning as object complements, while they follow their nominal head when functioning as relative clauses. I will return to the issue of functional transparency in section 8.4, where I discuss how languages may disambiguate the multiple functions of flexible balanced DCs.

Complete loss of flexibility in the DC

I now turn to the cases in Table 8.1 involving DCs that (i) can express a subset of the functions of the relevant PoS class, and (ii) where this subset is a singleton set, so that the DC is a rigid rather than a flexible construction. This corresponds to the situation described in (3c) of section 8.3.1: complete loss of flexibility in the deranked DC. As explained above, cases that involve a PoS class that can express two functions and a DC that can express only one are regarded as instances of partial loss of flexibility. Thus, the cases that remain to be discussed here involve rigid deranked DCs that can express two or three functions less than the relevant PoS. These cases are indicated in Table 8.1 above as > (3) or > (2).

Two languages in Table 8.1 only have deranked DCs that display complete loss of flexibility, i.e. loss of two or three functions compared to the relevant flexible PoS classes in these languages. The first case is Samoan, which has lexical contentives and rigid deranked clausal nominalizations; the second is Warao, with lexical non-verbs and rigid deranked infinitival clauses. The other languages with deranked DCs that show complete loss of flexibility (Kharia, Turkish, and Kayardild) also have deranked DCs that show no loss or only partial loss of flexibility.

Obviously, the categorial distinctiveness of any rigid deranked DC construction is by definition higher than that of any flexible PoS class. As such, these DCs all conform to the \([\text{flex PoS} \geq \text{flex deranked DC}]\) constraint, and do not require further explanation. Nonetheless, there seem to be additional motivations for the attestation of many rigid deranked DCs in languages with flexible PoS. A first reason may be borrowing. This occurs in Kharia, which has borrowed the rigid participial forms \(-na-wala, -al,\) and \(-ker\) from Indo-Aryan. A second reason – more relevant from the present perspective – may be the presence of an additional (possibly derived) rigid PoS class in an otherwise flexible system, serving as the primary category on which the rigid deranked DC is modelled. Warao, for instance, has
simple non-verbs plus a class of derived rigid nouns (see also Chapter 7, section 7.2.2.3).

The rigid deranked complement clause construction in Samoan also presents an interesting case. These nominalized clauses may occur with a special marker on the dependent predicate, namely the suffix \(-ga\). In Chapter 2 (section 2.5.2.2), I have shown that this suffix is also used for lexical derivation. I explained there that lexically derived \(-ga\) formations remain flexible (see example (39) in Chapter 2). In contrast, when \(-ga\) formation applies at a higher level of grammatical complexity, i.e. when \(-ga\) is a marker of syntactic derivation, then its output is apparently no longer flexible. This is in accordance with the \([\text{flex PoS} \geq \text{flex deranked DC}]\) constraint. An example of a syntactic \(-ga\) nominalization is given in (7). Note that this construction can be distinguished from a regular NP (with a \(-ga\)-derived head) by the fact that the second argument of the dependent predicate \(\text{apu ma moli ‘apples and citrus’}\) remains expressed as in an independent clause. Note also that the term ‘NP’ is not meant here to refer to a phrase with a nominal head, to the extent that Samoan has no lexical nouns (neither simple nor derived).

\[\text{Samoan (Mosel & Hovdhaugen 1992: 575)}\]
\[(7) \quad \text{A le faalavelave le tupu i le ai-ga}\]
\[\text{PST NEG trouble ART king LD ART eat-NMLZ}\]
\[\text{apu ma moli a le pipili ma le tauaso}\]
\[\text{apple and citrus POSS ART lame and ART blind}\]
\[\text{The king was not troubled that the lame and the blind ate the apples and oranges.}\]

Interestingly, Samoan rigid nominalized clauses can also occur \textit{without} any structural coding on the dependent predicate. This is illustrated in (8):

\[\text{Samoan (Mosel 1992: 267)}\]
\[(8) \quad \text{E lelei [I=a=u tunu ia]}\]
\[\text{GENR good ART=POSS=2SG roast fish}\]
\[\text{Your fish roasting is good.}\]

Arguably, these unmarked rigid nominalizations are indeed modelled on a rigid primary construction, but on a phrasal rather than a lexical one: an NP. Samoan NPs are characterized by the possibility to take determiners/case-
markers, and to contain possessive modifiers. Since in Samoan any lexeme can function as the head of an NP, the dependent predicate does not need to be re-categorized when it comes to function as the nucleus of a clausal nominalization. This explains why no special marker appears on the dependent predicate.

Similar types of unmarked nominalized clauses are attested in other Central-Eastern Malayo-Polynesian languages with a high degree of lexical flexibility, such as Tuvaluan (see example (9)) and Kambera (see example (10)). In the latter language nominalized clauses are cross-referenced with a dative marker on the matrix predicate, just like lexical and phrasal objects.

**Tuvaluan** (Besnier 2000: 512, 513)

(9) Koe na e poto i [te fakkaa o \\
You there NPAST clever in ART CAUS-burn \\
o te kasa], nee? \\
POSS ART pressure-lamp TAG \\
‘You know how to light a pressure-lamp, don’t you?’

**Kambera** (Klamer 1998: 96)

(10) Nda ku-pi-a-nya [na ngandi-mu \\
eg 1sg.nom-know-mod-3sg.dat ART take-2sg.gen \\
rú kuta] \\
leaf pepper plant \\
‘I didn’t know that you would bring kuta.’ \\
(lit.: ‘I didn’t know (of) your bringing kuta.’)

In short, there may be independent motivations, apart from the [flex PoS ≥ flex deranked DC] constraint, for the attestation of rigid deranked DCs in functions that can be expressed by means of a flexible PoS class, including borrowing and modelling on an additional rigid PoS class or a rigid phrasal construction.

**Rigid PoS and deranked DCs in languages with flexible PoS systems**

Finally, consider rigid PoS classes in languages with flexible PoS systems, and their clausal counterparts. It was hypothesized that any deranked DC appearing in a function for which a rigid PoS class is available, should be a rigid construction. In other words, this DC should not be able to express
any more function(s), since this would obviously violate the \([\text{flex PoS} \geq \text{flex deranked DC}]\) constraint. The data in Table 8.3 make clear that this prediction is born out: For every rigid PoS class in a language with a flexible PoS system it is true that, if there is a deranked DC available to express the same function, then this DC is also a rigid construction.

**Summary**

In sum, in this section I have tried to reinterpret three types of results from Chapter 7 in terms of the Principle of Increasing Categorization. One was an expected result, in terms of the hypotheses in Chapter 7, namely that flexibility in the domain of deranked DCs is dependent on flexibility in the domain of PoS classes. I have shown that this dependency relation can be interpreted as a constraint, imposed by the amount of flexibility attested in the PoS system of a particular language, on the maximal amount of flexibility (or minimal amount of categorial specificity) that can be displayed by the deranked DC(s) of that language.

Furthermore, I accounted for two results that were unexpected in terms of the predictions in Chapter 7. First, I explained the lack of correlations between lexical nominals and modifiers on the one hand, and deranked DCs with the same functional patterns on the other hand (see Tables 7.16-7.18 and Table 7.19 in Chapter 7). This finding can be understood as involving loss of flexibility in the DCs as compared to the PoS classes, and thus as supporting the Principle of Increasing Categorization. The same holds for the finding that virtually all languages with very flexible PoS classes have at least one rigid deranked DC (whether or not in combination with flexible deranked DC(s)).

**Table 8.3: Rigid PoS classes in languages with flexible PoS systems, and the distribution of deranked DCs that appear in the same function as these rigid PoS classes**

<table>
<thead>
<tr>
<th>Language</th>
<th>Relevant Rigid PoS</th>
<th>Deranked DCs in the same function(s)</th>
<th>Form</th>
<th>Function type</th>
<th>Type</th>
<th>Flex PoS (\geq) Flex DC?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Warao</td>
<td>D Noun</td>
<td>-kitane</td>
<td>2</td>
<td>complement</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td>2 Kayardild</td>
<td>Adj</td>
<td>-Thirri-n</td>
<td>3</td>
<td>relative</td>
<td>=</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj</td>
<td>-n-garrba</td>
<td>2</td>
<td>relative</td>
<td>=</td>
<td></td>
</tr>
</tbody>
</table>
8.3.2.2 Distribution of deranked DCs in languages with rigid PoS systems

This section is concerned with deranked DCs in languages with rigid PoS only. It was predicted that in such languages any deranked DC that appears in a function for which a lexical class is available should be a rigid construction too. The relevant data are presented in Table 8.4. These results make clear that the prediction is confirmed, except for the two counterexamples that were already discussed in Chapter 7: In Burushaski and Hdi we find flexible deranked nominal clauses, while these languages both have rigid lexical
nouns, and Burushaski also has rigid lexical adjectives. These cases violate the \[\text{flex PoS} \geq \text{flex deranked DC}\] constraint, since the degree of flexibility of the deranked DCs is higher (or their categorical distinctiveness lower) than that of the corresponding PoS classes.

Table 8.4: Rigid PoS classes in languages with rigid PoS only, and the distribution of deranked DCs that appear in the same function

<table>
<thead>
<tr>
<th>Language</th>
<th>Relevant rigid PoS</th>
<th>Deranked DCs in the same function(s)</th>
<th>Form</th>
<th>Function</th>
<th>Type</th>
<th>Flex PoS ≥ Flex DC?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Itelmen</td>
<td>Noun INF (various forms)</td>
<td>complement 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Basque</td>
<td>Noun -t(2)e, complement 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj -tu/-du/-i/-Ø relative 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Bambara</td>
<td>Adj -le/-ne relative 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Madv -tò adverbial manner 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Georgian</td>
<td>Noun -a complement 3</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj m- (-a) (-el/-al) relative 3</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj -ul/-il/m- -ar/-al relative 3</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj sa- (-el/-al/r) relative 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Abkhaz</td>
<td>Noun -ra complement 3</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Polish</td>
<td>Noun -nie complement 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Noun INF (various forms) complement 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj -c relative 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj -any etc. relative 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj (PST.PASS) PTC (various forms) relative 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S/D Madv PRS.PL-c adverbial manner 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Burushaski</td>
<td>Noun -(a)as nominal 2</td>
<td>&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adj -im/-um/-am relative 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S/D Madv n-STEM-(a)n adverbial manner 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Lavukaleve</td>
<td>Noun -e/-i complement 2</td>
<td>=</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Relevant rigid PoS</td>
<td>Deranked DCs in the same function(s)</td>
<td>Form</td>
<td>Function</td>
<td>Type</td>
<td>Flex PoS ≥ Flex DC?</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
<td>--------------------------------------</td>
<td>------------</td>
<td>----------------</td>
<td>------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Alamblak</td>
<td>Noun</td>
<td>-(nef) complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td></td>
<td>Noun</td>
<td>-(kfé)t complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td></td>
<td>Adj</td>
<td>ind relative 2</td>
<td></td>
<td>relative</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>Noun</td>
<td>-n/-an complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td>Berbice Dutch</td>
<td>Noun</td>
<td>fu/fi/Ø complement 2</td>
<td></td>
<td>complement</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td>Nama</td>
<td>D MAdv</td>
<td>-se!/’aa/tsii adverbial manner (no INDIC)</td>
<td></td>
<td>adverbial manner</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td>Hdi</td>
<td>Noun</td>
<td>tá nominal 3</td>
<td></td>
<td>nominal</td>
<td>3</td>
<td>&gt;</td>
</tr>
<tr>
<td>Tamil</td>
<td>Noun</td>
<td>-atu complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td></td>
<td>Noun</td>
<td>-(kk)a complement 2</td>
<td></td>
<td>complement</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td>Kisi</td>
<td>Noun</td>
<td>Ø complement 2</td>
<td></td>
<td>complement</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td>Garo</td>
<td>Noun</td>
<td>-a complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td></td>
<td>Noun</td>
<td>-a-ni complement 2</td>
<td></td>
<td>complement</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td></td>
<td>Noun</td>
<td>-na complement 2</td>
<td></td>
<td>complement</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td></td>
<td>Noun</td>
<td>-kan/-kan-a complement 2</td>
<td></td>
<td>complement</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td></td>
<td>S/D MAdv</td>
<td>-e/-e-min/-e-r adverbial manner</td>
<td></td>
<td>adverbial manner</td>
<td>2</td>
<td>=</td>
</tr>
<tr>
<td>Krongo</td>
<td>Noun</td>
<td>-(t) complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td>Hixkaryana</td>
<td>Noun</td>
<td>ni/-thí/-hitó + ri complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
<tr>
<td>W. Greenlandic</td>
<td>Noun</td>
<td>-niq complement 3</td>
<td></td>
<td>complement</td>
<td>3</td>
<td>=</td>
</tr>
</tbody>
</table>

Recall that in Chapter 7 no correlations were found between specific rigid PoS and rigid deranked DCs, i.e. between nouns and deranked complement clauses (see Tables 7.22–7.24 in Chapter 7), between adjectives and deranked relative clauses (see Tables 7.27–7.29), and between manner adverbs and deranked adverbial manner clauses (see Table 7.32). This is because, instead of deranked DC’s, languages may employ balanced DC constructions, the formal and functional properties of which are unrelated to PoS (cf. section 8.4 below). Nonetheless, as the data in Table 8.4 show, if a language does employ a deranked DC, then this construction displays the same functional specialization as the relevant rigid PoS class in that language.
8.3.3 Summary

The findings of this section can be summarized as follows:

(i) In languages with flexible PoS systems, there are no cases of deranked DCs that are more flexible than the relevant flexible PoS classes;

(ii) In languages with flexible PoS systems, deranked DCs occurring in a function that can be expressed by a rigid PoS class are rigid too;

(iii) In languages with rigid PoS systems, there are only two cases of flexible deranked DCs occurring in functions that are expressed by rigid PoS classes.

Thus, except for the two counterexamples mentioned under (iii), the functional patterns of all deranked DCs in the sample language are in accordance with the [Flex PoS ≥ Flex deranked DC] constraint. In the next section, I will consider the functional possibilities of balanced DCs.

8.4 Balanced constructions: functional transparency and other factors

8.4.1 Introduction

The results presented in Chapter 7 make clear that the functional patterns of balanced DCs, unlike those of deranked DCs, are not dependent on the degree of flexibility displayed by the PoS classes of the same language. This holds for global matches as well as for specific matches; and for languages with flexible PoS systems as well as for languages with rigid PoS systems. In Chapter 7 and in section 8.2 I have argued that this absence of a functional connection ties in with the absence of any formal similarity between lexical expressions and balanced DCs. In the present section, I attempt to further interpret the results for balanced DCs, taking into account the Principle of Functional Transparency as well as other functional factors.

8.4.2 Functional Transparency

From the point of view of functional transparency, rigid balanced DCs, both in languages with and languages without lexical flexibility, do not present any problem: they are functionally transparent by definition. Flexible balanced clauses, in contrast, are functionally ambiguous by definition. In the remainder of this sub-section I try to explain the distribution of flexible balanced DC
from the perspective of functional transparency; first in languages with flexible PoS systems, and then in languages with rigid PoS systems.

8.4.2.1 Flexible balanced DCs in languages with flexible PoS systems

The data presented in Chapter 7 show that a number of languages with flexible PoS systems display flexible balanced DCs (even though no dependency relation was established between these two construction types; cf. Tables 7.3, 7.11, 7.15, and 7.33 in Chapter 7). The relevant cases are listed in Table 8.5:

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
<th>Function</th>
<th>Type</th>
<th>Flexible PoS class in same functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Tagalog</td>
<td>Ø</td>
<td>multi-functional</td>
<td>1</td>
<td>contentives</td>
</tr>
<tr>
<td>2 Samoan</td>
<td>Ø</td>
<td>nominal</td>
<td>1</td>
<td>contentives</td>
</tr>
<tr>
<td>3 Turkish</td>
<td>ki</td>
<td>nominal</td>
<td>1</td>
<td>non-verbs</td>
</tr>
<tr>
<td>4 Kayardild</td>
<td>-ntha-</td>
<td>nominal</td>
<td>1</td>
<td>non-verbs</td>
</tr>
<tr>
<td>5 Paiwan</td>
<td>a (Ø)</td>
<td>nominal</td>
<td>1</td>
<td>predicatives</td>
</tr>
<tr>
<td>6 Gooniyandi</td>
<td>Ø</td>
<td>multi-functional</td>
<td>1</td>
<td>nominals</td>
</tr>
<tr>
<td>7 Ket</td>
<td>Ø</td>
<td>nominal</td>
<td>1</td>
<td>modifiers</td>
</tr>
</tbody>
</table>

Interestingly, in many of these cases, the flexible balanced DC is not marked by means of a dedicated subordinating conjunction. Rather, it makes use of the same function-indicating marking, at the level of phrase-structure, as constituents with a different internal structure, such as simple lexemes, phrases, or deranked DCs (see for instance the Tagalog examples in (6) of section 8.3.2.1 above). This fits in with the idea that languages with (more or less pervasive) lexical flexibility use a kind of ‘grid’ of morpho-syntactically marked slots to compensate for the functional ambiguity of their lexical material (cf. Hengeveld et al. 2004, Lehmann 2008). Apparently, these functional slots may not only accommodate lexical and phrasal constituents, but also complete clausal structures, which do not structurally deviate from independent clauses.

We may again take Tagalog as an example: In this language a balanced DC is marked with the same function word that would be used for any other linguistic unit expressing the same function. In (11), for instance, the ‘specifier’ ang is used to mark a balanced object complement clause, while in (12) ang marks a simple lexical object.
Tagalog (Palmer 2003: 268)

(11) **Moo-ng una bindi ko ma-alam-an**

\begin{verbatim}
REM-LG first NEG 1SG:GEN IRR:NC-knowledge-LOC
ang [akin-g ga-gaw-in]
SPEC 1SG-LG RDP-do-IRR
\end{verbatim}
‘At first I didn’t know what to do.’

(12) **I-in-abót ng manggagamot sa sundalo ang itlóg**

\begin{verbatim}
CV-REAL-reach GEN doctor LOC soldier SPEC egg
\end{verbatim}
‘The physician handed the egg to the soldier.’

Similarly, example (13) shows a balanced clause in relative function that is marked by the linking element *na*. The same marker is used for a simple lexical attribute in (14):

Tagalog (Himmelmann 2005: 368)

(13) **Sa mga lalaki na [maN-ibig nung**

\begin{verbatim}
LOC PL man LK AV-love DIST.GEN:lk
kanyá-ng anák]
3SG.DAT-LK child
\end{verbatim}
‘(So he held a contest) between the men who courted his child.’

(14) **ulól na unggó**

foolish LK monkey
‘foolish monkey’

Samoan also has the possibility to insert balanced DCs into functional slots without a dedicated subordinating conjunction. Example (15) shows a complement clause that is zero-marked for absolutive (the subject *Tigilau* is marked for ergative). This can be compared with (16), which shows that simple lexical arguments in the same functions are marked in the same way.

Samoan (Mosel & Hovdhaugen 1992: 589, 635)

(15) **Na iloa e Tigilau [úa sau Sina]**

\begin{verbatim}
PST know ERG Tigilau PFV comes Sina
\end{verbatim}
‘Tigilau knew that Sina had come.’

338 | Parts of Speech and Dependent Clauses
(16) Sā fasi le maile e le teine
   PST hit ART dog ERG ART girl
   ‘The girl hit the dog.’

Example (17) shows a balanced relative clause, which is marked by means of rigid word order: it obligatorily appears to the right of the head. As is shown in (18), the same order is used for lexical modifiers.

Samoan (Mosel & Hovdhaugen 1992: 635, 296)
(17) Ua tu le alii lea [na ua e Popi]
   PFV stand.up ART man that PST bite ERG Popi
   ‘That man who was bitten by Popi stood up.’

(18) le ta’avale mumu
   ART car red
   ‘the red car’

As a final example, consider balanced nominal clauses in Paiwan. They are marked with the particle a, as is shown in (19a-b), where the construction functions as a subject complement clause and as a relative clause, respectively. In (20) it can be seen that the same element is used with lexical constituents as a marker of the argument in focus (the agent in (20a)), and as an attributive marker in NPs (see (20b)):

Paiwan (Egli 1990: 230, 178)
(19) a. na’nguaq a [ma-ngetjez sun]
    good FOC pass-come you
    ‘It is good that you have come.’

    b. qala a [na tem-ker tua vaaua]
    stranger LK PFV drink-AG OBL wine
    ‘he stranger who has drunk wine’

Paiwan (Egli 1990: 180, 181)
(20) a. na qem-aung a alak
    PFV cry-AG FOC child
    ‘The child has cried.’
b. *vavayan a burai*
   woman LIK beautiful
   ‘beautiful woman’

There are two flexible balanced clauses in Table 8.5 that do not conform to the general picture of using the same phrase-structural coding for lexical and clausal constructions, but these both represent rather special cases. First, the balanced nominal clause construction in Kayardild displays so-called complete concord (see Dench 2006): All elements of the DC are separately marked with the oblique complementizer case -ntba. Example (21) shows this construction in the function of object complement:

*Kayardild* (Evans 1995: 516)

(21) *Ngada kamburri-ja niwan-ji [walbu-ntba dathin-inja*
   1SG.NOM say-ACT 3SG-MLOC raft-C.OBL that-C.OBL
   *barji-nyarra-ntb]*
   capsize-APPR-C.OBL
   ‘I told him that the raft would capsize.’

Second, the balanced nominal clause construction in Turkish, marked by the general subordinating conjunction *ki*, is special to the extent that it is borrowed from Persian.

In sum, the majority of flexible balanced DCs in languages with flexible PoS systems conform to the Principle of Functional Transparency by making use of the same function-indicating morpho-syntactic markers that are also employed for the disambiguation of non-clausal constituents.

### 8.4.2.2 Flexible balanced DCs in languages with rigid PoS systems

The data presented in Chapter 7 make clear that flexible balanced clauses also occur quite often in languages without any lexical flexibility. The relevant cases are listed in Table 8.6 below (cf. Table 7.33 in Chapter 7)\(^{122}\).

---

\(^{122}\) In fact, Thai and Basque both have small classes of flexible modifiers; this marginal amount of lexical flexibility is ignored here.
These DCs are again inherently functionally ambiguous. Therefore, on the basis of the Principle of Functional Transparency, it is expected that they combine with additional morpho-syntactic means that distinguish between their multiple functions. This expectation is confirmed. Some of the flexible balanced DCs in Table 8.6 employ disambiguating strategies that pertain exclusively to the domain of subordination. Other DCs make use of more general strategies that are also used with other types of constructions (cf. Hengeveld & Van Lier 2008).

As a first example of the former, subordination-specific type of strategy, consider Thai. This language has a balanced nominal clause construction marked with \(\text{th}ii\), as illustrated in (22a-b) below. When functioning as a relative clause, as in (22b), this construction can be combined with a resumptive pronoun. Even though this grammatical device does not count as structural coding in the sense of a relativizer, it does mark the function of the relativized item in the DC, and as such helps to show that we are dealing with the modifying use of the \(\text{th}ii\)-construction, rather than with its referential use.

Thai (Iwasaki & Ingkaphirom 2005: 245, 255)

(22) a. \(\text{t}e-\text{w}a\ dii \ n\acute{a} \ [\text{th}ii \ \text{m}a\text{y} \ \text{m}i\text{i} \ \text{kh}ay \ \text{pen} \ \text{alay}]\)

   but good **prt** **comp** **neg** have who **cop** what

   ‘But it was good that no one was hurt.’

---

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
<th>Function</th>
<th>Type</th>
<th>Rigid PoS class(es) in same functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Thai</td>
<td>(\text{th}ii)</td>
<td>nominal</td>
<td>1</td>
<td>nouns, adj</td>
</tr>
<tr>
<td>2 Basque</td>
<td>(-en)</td>
<td>nominal</td>
<td>1</td>
<td>nouns, adj</td>
</tr>
<tr>
<td>3 Georgian</td>
<td>(\text{rom})</td>
<td>nominal</td>
<td>1</td>
<td>nouns, adj</td>
</tr>
<tr>
<td>4 Burushaski</td>
<td>(\text{ke/ki})</td>
<td>multi-functional</td>
<td>1</td>
<td>nouns, adj, S/D Madv</td>
</tr>
<tr>
<td>6 Pipil</td>
<td>(\text{ko(h)})</td>
<td>nominal</td>
<td>1</td>
<td>nouns, adj</td>
</tr>
<tr>
<td>7 Dhaasanac</td>
<td>DET (+ DEM)</td>
<td>nominal</td>
<td>1</td>
<td>nouns, adj</td>
</tr>
<tr>
<td>8 Mandarin</td>
<td>(\emptyset/\text{de})</td>
<td>multi-functional/ modifier</td>
<td>1</td>
<td>nouns, S. Adj, S. Madv</td>
</tr>
<tr>
<td>9 Nung</td>
<td>(\emptyset)</td>
<td>multi-functional</td>
<td>1</td>
<td>nouns, S. Adj, X</td>
</tr>
<tr>
<td>8 Greenlandic</td>
<td>PTC mood</td>
<td>nominal</td>
<td>1</td>
<td>nouns, X</td>
</tr>
</tbody>
</table>

Table 8.6: Flexible balanced DCs in languages with rigid PoS systems
b. *Khon* [thii khâw pay yiu kan taam rogrian]
   
   people REL RSP\(^3\) go stay RECIP school
   
   ‘people who want to stay at school’

A second example is the Georgian DC construction marked with *rom*. This is also a nominal clause construction: it can be used as a complement clause and as a relative clause, as is illustrated in (23a-b) below. Example (23a) shows that the *rom*-construction, when functioning as a complement clause, can be combined with a pronominal element in the main clause. This is even obligatory when the complement clause is dependent on a postposition or functions obliquely (i.e. follows a verb that takes instrumental case). Moreover, the examples in (23) show a contrast in terms of the position of the conjunction: Very often (although not obligatorily) complement clauses start with *rom*, while in relative clauses *rom* avoids the first position.

**Georgian** (Hewitt 1995: 613, 607)

(23) a. *(is) uk’ve (that(NOM)) already

   \[\text{še-v-a-mčn-i-e } \text{rom es} \]
   
   \[\text{prev-1sg-nv-notice-Ths-aor.ind comp this} \]
   
   \[\text{xalx-i } \text{sando ar ar-i-s} \]
   
   people-nom trustworthy not be-prs-it
   
   ‘I have already noticed that this people is not trustworthy.’

b. *gušin rom m-a-čuk-g.*

   yesterday REL me-lv-present-aor.ind
   
   is beč’ed-I sad ar-i-s?
   
   that(NOM) ring-nom where be-prs-it
   
   ‘Where is that ring which you presented to me yesterday?’

There are also languages that use more generally applicable strategies to indicate the function of flexible balanced DCs. An obvious example of such a strategy would be fixed word order\(^{123}\). Consider for instance Pipil: the complementizer *ka(h)* in this language is sometimes also used to

\(^{123}\) See Hengeveld et al. (2004), who show that disambiguating morpho-syntactic measures, such as rigid word order, are not confined to languages with flexible PoS systems. While such measures are always present in the case of lexical flexibility, they may be present in languages with rigid PoS systems. Cf. section 8.2 above.
introduce relative clauses (Campbell 1985: 128). The two functions of the *ka(h)*-construction are illustrated in (24a-b). These examples show that complement clauses are post-verbal, while relative clauses must follow their nominal head. Since Pipil has rigid classes of verbs and nouns, these ordering restrictions resolve the functional ambiguity of this DC construction.

**Pipil** (Campbell 1985: 126, 129)

(24) a. *Tesu ki-mati [ka ne i-siwa:-w no it-know COMP the his-wife-poss se: bru:hab]*
   a witch
   ‘He didn’t know that his wife is a witch.’

   b. *Ni-k-miktih ne mistun [ka ki-kwah ne tu:tut]*
      I-it-killed the cat REL it-ate the bird
      ‘I killed the cat that ate the bird.’

The *-en* construction in Basque can also be used as a complement clause and as a relative clause. This is illustrated in (25a-b). The first example shows that this DC combines with a determiner when used as a complement clause. In contrast, when the construction is used as a relative clause, the determiner appears on the modified head:

**Basque** (Hualde & Ortiz de Urbina 2003: 646, 764)

(25) a. *Entzun dut [Amaiaren neba bil d-en]-a hear aux Amaia.gen brother die AUX-COMP-ART*
   ‘I heard that Amaia’s brother died.’

   b. *[Pellok ekarri du-en] diru-a galdu dut Peter.erg bring AUX-REL money-DET lose AUX*
   ‘I lost the money Peter brought.’

Finally, consider the somewhat exceptional case of the West Greenlandic participial mood construction. This is again a construction that can be used as a complement clause and as a relative clause, but, as already discussed in Chapter 6 (see example (57a-b)), it behaves quite differently in each of these two functions: Intransitive participial forms are fully balanced when functioning as complements, whereas in relative clause function they are not
marked for person other than third. Also, it is not possible to create relative clauses with corresponding transitive participial mood forms.

In sum, flexible balanced DCs in languages with rigid PoS systems conform to the Principle of Functional Transparency; their multiple functions can be disambiguated by means of several strategies. Some of these strategies, such as the use of resumptive pronouns, are specific to subordinate clauses, while others, such as word order constraints, are used more generally in the language system as function-indicating devices.

8.4.3 Other factors: diachrony and cognition

Tables 8.5 and 8.6 above show that many flexible balanced DCs display a particular pattern of flexibility: They are nominal clauses constructions, i.e. they can be used as complement clauses and as relative clauses. There appear to be two functional factors that motivate this pattern.

The first factor is a diachronic one. According to Cristofaro (1998), relative clause markers may develop complementizer use along the following pathway: The original construction involves a predicate meaning ‘see’, ‘think’, ‘know’ etcetera (i.e. a potentially complement-taking predicate) with a (pro-)nominal object followed by a non-restrictive relative clause. Cristofaro illustrates this type of construction with an example from Biblical Hebrew (taken from Givón 1991):

\begin{quote}
\textit{Biblical Hebrew (Givón 1991: 289)}
\end{quote}

\begin{verbatim}
(26) \textbf{2al} tir\textbf{\-\-ni} [she-\textbf{\-\-ani} shaxoret]
\end{verbatim}

\texttt{neg see:IPFV3SG:masc me REL-I be:dark:SG:F}

‘Don’t see me that I am dark-skinned.’

As Cristofaro (1998: 65) explains, in such constructions “the relative clause […] has high communicative value: it conveys a specification about the main clause object, and this specification represents the communicative focus of the sentence”. As a result of this pragmatic saliency, the relative clause is reanalyzed as the object of the matrix clause, so that the relative pronoun becomes a complementizer, and the (pro-)nominal head is dropped. Example (27) shows that the original relative marker \textit{she} in Biblical Hebrew is used as a complementizer:

124 The diachronic scenario to be described is probably most relevant for balanced DCs that are marked by means of subordinating conjunctions, as opposed to zero-marked constructions.
Second, Croft (2001: 346-349) suggests that there is also a cognitive connection between complementation and relativization, which seems to be in line with Cristofaro’s point about pragmatic saliency. In particular, Croft makes use of Langacker’s (1987) distinction between dependent and autonomous concepts: A dependent concept elaborates on (i.e. makes more specific) a certain aspect of an autonomous concept. According to Croft, complement clauses and relative clauses are alike in that they are both dependent concepts that elaborate on an aspect of the main clause event. However, complement clauses and relative clauses don’t do this in exactly the same way. Whereas complement clauses elaborate on the main clause event by filling an argument position, relative clauses rather represent an elaboration of a participant of the main clause event. Nonetheless, Croft argues that the sharing of structural properties between complement clauses and relative clauses in the same language can be explained in terms of their shared cognitive function of main clause elaboration.

In short, the fact that many flexible balanced DCs involve nominal clause constructions may be motivated in terms of diachronic and cognitive factors. These motivations make reference to a similarity between the functions of complementation and relativization constructions: both specify a discursively important aspect of the matrix clause.

8.4.4 Summary
In this section I have tried to show that the distributional behaviour of balanced DCs, even though it it is not related to the flexible or rigid functional properties of PoS classes in particular languages, does fit into a larger explanatory picture based on the Principle of Functional Transparency.

8.5 Summary
In this chapter I reconsidered the results obtained in Chapter 7 in light of two closely related functional principles proposed in recent typologically-based literature: the Principle of Increasing Categoriality (Haig 2006, Lehmann 2008), and the Principle of Functional Transparency (Frajzyngier & Shay 2003, Hengeveld et al. 2004, Sinnemäki 2008). In addition, a categorical
distinction was implemented between deranked and balanced dependent clauses, based on Dik’s (1997) Principle of Formal Adjustment. In particular, I hypothesized that deranked DCs, which are formally modelled on PoS, are also functionally constrained by the latter. Balanced DCs, in contrast, have no formal properties of PoS, and are not expected to display any functional similarities with lexical constructions either.

The hypothesized functional connection between PoS and deranked DCs was operationalized in accordance with the Principle of Increasing Categoriality, which states that the degree of categorial specificity of linguistic constructions increases with their structural complexity. In terms of the present study, this means that (secondary) deranked DCs are predicted to either display the same amount of flexibility as (primary) PoS classes, or less flexibility, but not more. This prediction was confirmed for virtually all deranked DC constructions attested in the sample.

Moreover, it was shown that flexible deranked DCs conform to the more general Principle of Functional Transparency, which claims that languages avoid ambiguity in the formal expression of a particular functional domain. Interestingly, most flexible deranked DCs in languages with flexible PoS systems do this by means of the same morpho-syntactic devices that are used in these languages to disambiguate the functions of lexical and phrasal constructions.

Regarding balanced DCs, it was shown that, even though their functional possibilities cannot be correlated with those of PoS classes in the same language, their distributional properties can be explained in terms of the Principle of Functional Transparency. Whereas rigid balanced DCs are by definition functionally transparent, flexible balanced DCs, like other flexible constructions, require additional morpho-syntactic means to compensate for their inherent functional ambiguity. I have illustrated several types of strategies that serve this function for balanced DC constructions. Some such strategies are confined to subordination constructions, while others are more general applicable. It was shown that balanced flexible clauses in languages with flexible PoS systems, like deranked ones, employ the same function-indicating devices that are used to compensate for functional ambiguity at the lexical/phrasal level.
In sum, the functional principles of Increasing Categoriality, Functional Transparency, and Formal Adjustment together provide a complete account of the results presented in Chapter 7. Moreover, the discussion in the present chapter confirms the general picture emerging from Chapter 7, namely that functional flexibility at the lexical level has pervasive repercussions at higher grammatical levels.
The aim of this study was to investigate the functional relationship between a language’s parts of speech classes and its dependent clauses, in terms of the set of propositional functions that both construction types can express. First, in Chapters 2 and 3, I discuss various functionalist approaches to the theory and typology of parts of speech and dependent clauses, respectively. It is shown that both parts of speech and dependent clauses can be defined as formal mappings onto a space consisting of four propositional functions. A basic distinction is adopted between rigid and flexible constructions: The former type can express only a single propositional function, while the latter can be used in two or more functions, without any difference in structural coding.

Whereas both lexical and clausal constructions can thus be defined in terms of their functional possibilities, only the latter are configurational, and can as such also be classified according to their internal morpho-syntactic properties. On the basis of earlier functional-typological studies, a distinction is made between balanced and deranked dependent clauses. Balanced clauses, which represent one extreme of a scale, are characterised by the fact that they express the same set of categories as independent clauses. Towards the other extreme we find various types of deranked clauses, the internal structure of which is increasingly deviant from the structure of independent clauses. This deviation may show in two ways: First, it may take the form of de-categorization, i.e. (partial) loss of ‘verbal’ features, such as TAM distinctions
and person marking, and/or non-expression of arguments. Second, deranked dependent clauses may show formal reflections of re-categorization, i.e. they may acquire (some) ‘nominal’ features, such as the expression of case markers and/or determiners and possessive coding of argument(s).

Against this theoretical backdrop, Chapter 4 formulates a set of predictions concerning the expected similarity between the functional possibilities as displayed by the parts of speech classes of a particular language and by its dependent clause constructions. Some predictions are formulated in general terms, i.e. they make reference to languages with some flexibility versus no flexibility in their parts of speech system, and the expected reflection of this difference on the functional properties of dependent clause constructions in these languages. Other predictions are more specific; they aim at the identification of one-to-one matches between the functional possibilities of particular types of part of speech classes and dependent clause constructions. In addition, all predictions are formulated first without differentiation for dependent clauses in terms of their internal morpho-syntactic properties, and then in sets of sub-predictions that make specific reference to balanced clauses versus (various types of) deranked clauses.

Chapters 5 and 6 present the classifications of parts of speech classes and dependent clauses in a balanced sample of 50 languages in terms of the typological frameworks developed in Chapters 2 and 3. Subsequently, in Chapter 7, the two data sets of Chapter 5 and 6 are linked in order to identify dependency relations between the functional patterns displayed by the parts of speech classes of particular languages, and their (different structural types of) dependent clause constructions. The analyses reveal that the presence of flexible deranked dependent clauses in a language is dependent upon the presence of flexible parts of speech classes in that language. This does not mean, however, that all languages with flexible parts of speech systems also display flexible deranked dependent clauses.

Furthermore, it is shown that pervasively flexible parts of speech classes and deranked dependent clauses are both, i.e. independently of each other, cross-linguistically rare phenomena. Therefore, the generalization that maximally flexible deranked dependent clauses do not occur in languages without the same degree of flexibility in the lexical domain does not have much explanatory power. More interestingly, it is shown that whenever maximal lexical flexibility does occur in a language, this strongly increases the chances of also finding the other rare phenomenon: maximally flexible deranked dependent clauses.
In contrast to the result obtained for pervasively flexible constructions, no dependency relations can be established between the presence of less flexible or rigid parts of speech classes and deranked dependent clauses with the same functional behaviour. Moreover, the functional patterns of balanced dependent clauses are not related to those of parts of speech classes either.

In Chapter 8 an attempt is made to provide a functionalist explanatory framework to account for these results. In particular, I suggest that the findings concerning parts of speech and deranked dependent clauses can be interpreted as conforming to the Principle of Increasing Categoriality (or Staggering Level-dependent Categoriality), which states that the categorial specificity of linguistic units increases – or their flexibility decreases – with increased structural complexity. Deranked dependent clauses, as opposed to balanced ones, are regarded as secondary constructions, derived from primary lexical or phrasal constructions, in terms of the Principle of Formal Adjustment. Deranked dependent clauses are more complex than their lexical counterparts, and therefore the Principle of Increasing Categoriality predicts that the former should be at least as categorially specific as the latter. In other words, the degree of categorial specificity attested at the lexical level is expected to determine the minimal degree of categorial specificity (or the maximal degree of flexibility) that should be maintained at the level of deranked dependent clauses. This principle explains the finding that flexible deranked dependent clauses are almost exclusively found in languages with (pervasively) flexible parts of speech systems.

Moreover, the Principle of Increasing Categoriality accounts for a number of other results of Chapter 7: First, it captures the finding that the functional matches between flexible parts of speech and flexible deranked dependent clauses need not be one-to-one, but may also involve a partial decrease in the flexibility of the latter construction type, as compared with the former. Second, the principle accounts for the finding that languages with pervasively flexible parts of speech systems almost always have rigid deranked dependent clauses (possible alongside their flexible deranked clauses). Third, the principle explains at least partly why a lower degree of flexibility at the lexical level does not correlate with flexibility at the level of dependent clauses: If a flexible part of speech class covers only two functions, then even a loss in flexibility of one function on the part of the dependent clause automatically results in a rigid construction.

Finally, the Principle of Increasing Categoriality and the Principle of Formal Adjustment shed light on the fact that no correlations are found
between particular types of rigid parts of speech classes and rigid deranked dependent clauses with the same distribution. The reason for this is that, instead of rigid deranked dependent clauses, languages may employ balanced constructions. The latter, since they are presumably not modelled on lexical categories, may be rigid as well as flexible. However, whenever a language has a deranked dependent clause construction to express a function for which a rigid parts of speech class is available, then this dependent clause construction is rigid as well. This holds both for languages with flexible parts of speech systems that include one or more rigid parts of speech classes, as well as for languages with rigid parts of speech classes only. These findings confirm the prediction that a deranked dependent clause cannot exhibit a greater degree of flexibility than a parts of speech class appearing in the same function.

Regarding balanced dependent clauses, I argue that even though their functional possibilities cannot be linked to those of parts of speech, their distributional behaviour is in accordance with the more general Principle of Functional Transparency. This principle predicts that every linguistic unit should ultimately, i.e. at the utterance-level, be endowed with a categorial value, in order to make its function identifiable for the hearer. As expected, the multiple possible functions of flexible balanced clauses, both in languages with rigid and with flexible parts of speech systems, can be disambiguated through additional morpho-syntactic strategies. Some of these strategies, such as the use of resumptive pronouns, are confined to the functional domain of subordination, while others, such as fixed constituent order, are used much more generally to establish functional transparency in grammatical (sub-)systems.

With regard to this last point, a particularly interesting finding is that the flexible dependent clauses – both balanced and deranked ones – in languages with very flexible parts of speech systems often make use of the same morpho-syntactic strategies that are used to indicate the functions of lexical and phrasal constituents. This supports a general pattern emerging from the present study concerning languages with pervasive lexical flexibility: They appear to have a particular type of phrase structure, consisting of a ‘grid’ of morpho-syntactically marked slots, which may be filled by formal units of any type of internal structure: not just single lexemes, but also complex phrasal and clausal constituents. These complex units can (but need not) retain maximal flexibility until the final level of the structure-building process, when they are inserted into a syntactic frame corresponding to a
particular propositional function. This pattern is in accordance with findings from other studies, which suggest that, in the encoding of a particular functional domain, flexibility in one grammatical area must be compensated with rigidity in another area.

Both the Principle of Increasing Categoriality and the Principle of Functional Transparency are ultimately motivated in terms of the functional principle of Economy. In particular, languages either retain the functional specialization present in simple units in the process of building more complex ones, or gradually increase it, until the ultimate level of functional specialization required at the utterance level is reached. If, in contrast, the functional specialization of less complex units would be decreased or altogether lost along the way towards more complexity, this would require additional and as such uneconomical processing on the part of the speaker. Moreover, if the loss of categorial specificity would not be repaired at some later stage in the derivational process, this would imply an additional processing load for the hearer, who would ultimately have to solve the problem of functional ambiguity that he or she would be presented with. Therefore, the Economy principle predicts that complexity-increasing linguistic processes will produce output structures that are at least as categorially specific as their input structures, in order to maximize processing ease for both participants of a communicative situation.

On a final note, I would like to point out that redundancy seems to be a common phenomenon in languages, while ambiguity is not. In other words, languages often encode a particular functional distinction within a certain domain by means of more than one formal device, while they apparently tend to prevent situations in which there is too little grammatical coding to unambiguously identify the function of a linguistic unit within an utterance. Viewed from the perspective human communication, this suggests that languages strongly adhere to processing ease for the hearer, although not necessarily for the speaker.


---

125 References with an asterix are grammars or other data sources used for the sample languages. Therefore they appear also in the list of ‘References for the sample languages’ below.


References


Haig, Geoffrey 2006. ‘Word-class distinctions and morphological type: agglutinating and fusional languages reconsidered.’ Unpublished manuscript, University of Kiel.


Berlin: Mouton de Gruyter. pp. 325-357.


* Peterson, John 2005. There’s a grain of truth in every “myth”, or, Why the discussion of lexical classes in Mundari isn’t quite over yet. Linguistic Typology 9:3, 391-441.
Rijkhoff, Jan 2007. Word classes. Language and Linguistics Compass, 1:6, 709-726


References for the Sample Languages

Abkhaz

Abun

Alamblak

Babungo
Bambara


Basque


Berbice Dutch Creole


Bukiyip


Burushaski


Dhaasanac

Garo


Georgian


Gooniyandi


Guaraní


Hdi


Hixkaryana

Hmong Njua


Hungarian


Imbabura Quechua


Itelmen


Japanese


Kambera

**Kayardild**


**Ket**


**Kharia**


**Kisi**


**Koasati**

Krongo


Lango


Lavukaleve


Ma’di


Mandarin Chinese


Nama


Nunggubuyu

Nivkh


Nung


Paiwan


Pipil


Polish


Samoan


Santali


Slave


Tagalog


Tamil


Thai


Turkish

Tuscarora

Wambon

Warao

West Greenlandic
### Appendix 1: Language Sample

<table>
<thead>
<tr>
<th>Language</th>
<th>Family (Ethnologue)</th>
<th>Subfamily (Ethnologue)</th>
<th>Genus (WALS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hdi</td>
<td>Afro-Asiatic</td>
<td>Chadic</td>
<td>Biu-Mandara</td>
</tr>
<tr>
<td>Dhaasanac</td>
<td>Afro-Asiatic</td>
<td>Cushitic</td>
<td>Eastern-Cushitic</td>
</tr>
<tr>
<td>Turkish</td>
<td>Altaic</td>
<td>Altaic</td>
<td>Turkic</td>
</tr>
<tr>
<td>Pipil</td>
<td>Uto-Aztecan</td>
<td>Uto-Aztecan</td>
<td>Aztecan</td>
</tr>
<tr>
<td>Hixkaryana</td>
<td>Carib</td>
<td>Carib</td>
<td>Cariban</td>
</tr>
<tr>
<td>Tuscarora</td>
<td>Iroquoian</td>
<td>Iroquoian</td>
<td>Northern Iroquoian</td>
</tr>
<tr>
<td>Koasati</td>
<td>Moskogean</td>
<td>Moskogean</td>
<td>Muskogean</td>
</tr>
<tr>
<td>Guaraní</td>
<td>Tupi</td>
<td>Tupi</td>
<td>Tupi-Guaraní</td>
</tr>
<tr>
<td>Warao</td>
<td>Isolate</td>
<td>Isolate</td>
<td>Warao</td>
</tr>
<tr>
<td>Imbabura Quechua</td>
<td>Quechuan</td>
<td>Quechuan</td>
<td>Quechuan</td>
</tr>
<tr>
<td>Gooniyandi</td>
<td>Australian</td>
<td>Bunaban</td>
<td>Bunaban</td>
</tr>
<tr>
<td>Kayardild</td>
<td>Australian</td>
<td>Tangkic</td>
<td>Tangkic</td>
</tr>
<tr>
<td>Nunggubuyu</td>
<td>Australian</td>
<td>Gunwinyungan</td>
<td>Nunggubuyu</td>
</tr>
<tr>
<td>Tagalog</td>
<td>Austronesian</td>
<td>Western Malayo-Polynesian</td>
<td>Meso-Philippine</td>
</tr>
<tr>
<td>Kambera</td>
<td>Austronesian</td>
<td>Central Malayo-Polynesian</td>
<td>Central Malayo-Polynesian</td>
</tr>
<tr>
<td>Samoan</td>
<td>Austronesian</td>
<td>Eastern Malayo-Polynesian</td>
<td>Oceanic</td>
</tr>
<tr>
<td>Paiwan</td>
<td>Austronesian</td>
<td>Paiwanic</td>
<td>Paiwanic</td>
</tr>
<tr>
<td>Garo</td>
<td>Sino-Tibetan</td>
<td>Tibeto-Burman</td>
<td>Baric</td>
</tr>
<tr>
<td>Thai</td>
<td>Tai-Kadai</td>
<td>Kam-Tai</td>
<td>Kam-Tai</td>
</tr>
<tr>
<td>Santali</td>
<td>Austro-Asiatic</td>
<td>Munda</td>
<td>Munda</td>
</tr>
<tr>
<td>Kharía</td>
<td>Austro-Asiatic</td>
<td>Munda</td>
<td>Munda</td>
</tr>
<tr>
<td>Hmong-Njuá</td>
<td>Hmong-Mien</td>
<td>Hmong-Mien</td>
<td>Hmong-Mien</td>
</tr>
<tr>
<td>Abkház</td>
<td>North-Causasian</td>
<td>North-Causasian</td>
<td>North-west Caucasian</td>
</tr>
<tr>
<td>Georgian</td>
<td>Kartvelian</td>
<td>Kartvelian</td>
<td>Kartvelian</td>
</tr>
<tr>
<td>Itelmen</td>
<td>Chukotko-Kamchatkan</td>
<td>Chukotko-Kamchatkan</td>
<td>Southern Chukotko-Kamchatkan</td>
</tr>
<tr>
<td>Tamil</td>
<td>Dravidian</td>
<td>Southern Dravidian</td>
<td>Southern Dravidian</td>
</tr>
<tr>
<td>West Greenlandic</td>
<td>Eskimo-Aleut</td>
<td>Eskimo-Aleut</td>
<td>Eskimo-Aleut</td>
</tr>
<tr>
<td>Polish</td>
<td>Indo-European</td>
<td>Slavic</td>
<td>Slavic</td>
</tr>
<tr>
<td>Wambon</td>
<td>Trans-New Guinea</td>
<td>Awju-Dumut</td>
<td>Awju-Dumut</td>
</tr>
<tr>
<td>Alamblak</td>
<td>Sepik-Ramu</td>
<td>Sepik Hill</td>
<td>Sepik Hill</td>
</tr>
<tr>
<td>Lavukaleve</td>
<td>East Papuan</td>
<td>Central Solomons</td>
<td>Solomon East Papuan</td>
</tr>
<tr>
<td>Abun</td>
<td>West Papuan</td>
<td>North Central Bird’s Head</td>
<td>North Central Bird’s Head</td>
</tr>
<tr>
<td>Bukiyp</td>
<td>Torricelli</td>
<td>Arapesh</td>
<td>Kombio-Arapesh</td>
</tr>
<tr>
<td>Nama</td>
<td>Khoisan</td>
<td>Central Koisan</td>
<td>Central Koisan</td>
</tr>
<tr>
<td>Language</td>
<td>Family (Ethnologue)</td>
<td>Subfamily (Ethnologue)</td>
<td>Genus (WALS)</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>36 Slave</td>
<td>Na-Dene</td>
<td>Athapaskan</td>
<td>Athapaskan</td>
</tr>
<tr>
<td>37 Ket</td>
<td>Yeniseian</td>
<td>Yeniseian</td>
<td>Yeniseian</td>
</tr>
<tr>
<td>38 Burushaski</td>
<td>Isolate</td>
<td>Isolate</td>
<td>Burushaski</td>
</tr>
<tr>
<td>39 Basque</td>
<td>Basque</td>
<td>Basque</td>
<td>Basque</td>
</tr>
<tr>
<td>40 Nivkh (Gilyak)</td>
<td>Isolate</td>
<td>Isolate</td>
<td>Nivkh</td>
</tr>
<tr>
<td>41 Babungo</td>
<td>Niger-Congo</td>
<td>Benue Congo</td>
<td>Bantoid</td>
</tr>
<tr>
<td>42 Kisi</td>
<td>Niger-Congo</td>
<td>Atlantic</td>
<td>Southern-Atlantic</td>
</tr>
<tr>
<td>43 Bambara</td>
<td>Niger-Congo</td>
<td>Mande</td>
<td>Western-Mande</td>
</tr>
<tr>
<td>44 K戎ongo</td>
<td>Nilo-Saharan</td>
<td>Kadugli</td>
<td>Kadugli</td>
</tr>
<tr>
<td>45 Lango</td>
<td>Nilo-Saharan</td>
<td>Eastern-Sudanic</td>
<td>Nilotic</td>
</tr>
<tr>
<td>46 M戎di</td>
<td>Nilo-Saharan</td>
<td>Central Sudanic</td>
<td>Moru-M戎di</td>
</tr>
<tr>
<td>47 Berbice Dutch Creole</td>
<td>Creole</td>
<td>Creole</td>
<td>Creoles and Pidgins</td>
</tr>
<tr>
<td>48 Nung</td>
<td>Tai-Kadai</td>
<td>Kam-Tai</td>
<td>Kam-Tai</td>
</tr>
<tr>
<td>49 Mandarin Chinese</td>
<td>Sino-Tibetan</td>
<td>Chinese</td>
<td>Chinese</td>
</tr>
<tr>
<td>50 Hungarian</td>
<td>Uralic</td>
<td>Finno-Ugric</td>
<td>Ugric</td>
</tr>
</tbody>
</table>
Appendix II: Excluded parts of speech systems

The following parts of speech systems are predicted not to occur in any language, because they violate one or more of the constraints developed in Chapter 2 as part of the implicational map model of parts of speech. These constraints are repeated below for convenience. Note that the constraints are numbered in the same way as in Chapter 2, in order to facilitate comparison with the predicted systems discussed there. The PoS systems are simply numbered (1)* through (34)*. The asterisks are added to indicate that these are ‘forbidden’ systems.

(11) Predication \( \subseteq \) Reference
   a. If a language has a rigid class of lexemes that can be used as the head of a referential phrase, it must also have a rigid class of lexemes that can be used as the head of a predicate phrase.
   b. If a language has a flexible class of lexemes that can be used as the head of a referential phrase (but not as the head of a predicate phrase, since then the restriction becomes irrelevant), it must also have a flexible or rigid class of lexemes that can be used as the head of a predicate phrase.

(12) Head \( \subseteq \) Modifier
   a. If a language has a rigid class of lexemes that can be used as the modifier within a phrase, it must also have a rigid class of lexemes that can be used as the head of that phrase.
   b. If a language has a flexible class of lexemes that can be used as the modifier within a phrase (but not as the head of that phrase, since then the restriction becomes irrelevant), it must also have a flexible or rigid class of lexemes that can be used as the head of that phrase.

(13) \( (\text{Predication/Reference}) \subseteq (\text{Head/Modifier}) \)
   If a language has distinct (rigid or flexible) classes of lexemes for heads and modifiers within any phrase, then it must also have distinct (rigid or flexible) classes of lexemes for heads of predicate versus referential phrases.

The system in (1)* has a flexible class of lexemes (non-verbs) that can be used as the head of a referential phrase, but no class of lexemes that can be used as the head of a predicate phrase (verbs). Thus, it violates constraint (11b) above. Furthermore, it has a flexible class of lexemes that can be used as a modifier in a predicate phrase, but no class of lexemes that can be used as the head of a predicate phrase. Thus, it also violates constraint (12b).

(1)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>–</td>
</tr>
<tr>
<td>Modifier</td>
<td>Non-verb</td>
</tr>
</tbody>
</table>

The system in (2)* is excluded on the basis of constraint (12b): It has a flexible class of lexemes that can be used as the modifier (but not the head) in a referential phrase, and no corresponding class of heads.

(2)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>–</td>
</tr>
<tr>
<td>Modifier</td>
<td>Flex</td>
</tr>
</tbody>
</table>

The system in (3)* is excluded because it has a rigid head class in the referential, but not in the predication domain. As such, it violates constraint (11a).

(3)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Flex</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
</tr>
</tbody>
</table>

The system in (4)* involves a rigid modifier class in the referential domain without the corresponding rigid head class and thus violates (12a). In addition, it has distinct classes for heads and modifier in the referential domain, without distinct classes of heads of predicative versus referential phrases. Thus, it also violates (13).
The system in (5)* is excluded because it violates constraint (12a) in the predication domain, where is has a rigid class for modifiers, but no rigid class for heads. This system also violates constraint (13), since it has distinct classes for heads and modifiers in the referential domain, but no distinct classes for heads of predicate versus referential phrases.

The system in (6)* is excluded because it has a class of flexible lexemes that can be used as the head of a referential phrase, but no class of lexemes that can be used as the head of a predicate phrase. Thus, it violates constraint (11b).

The system in (7)* violates two constraints: (11b) because it has a flexible class of lexemes that can be used as the head of a referential phrase, but no class of lexemes that can be used as the head of a predicate phrase. And (12a) because it has a rigid class of modifiers in the predication domain, but not the corresponding heads.

The system in (8)* violates constraint (12b): It has a flexible class of modifiers, but no heads.

The system in (9)* is excluded because it violates constraint (13): There is a lexical distinction between heads and modifiers in the predication domain, but there are no distinct classes for heads of predicative versus referential phrases. In addition, it violates constraint (12b) in the referential phrase, where modifiers can be expressed lexically, but not heads.

The system in (10)* is also excluded because it violates constraint (13) and constraint (12b) (but this time in the predication domain). In addition, constraint (11a) is violated: this system has a rigid class of heads for referential but not for predicate phrases.
(10)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Noun</td>
</tr>
<tr>
<td>Modifier</td>
<td></td>
</tr>
</tbody>
</table>

The system in (11)* violates constraint (13): There is lexical distinction between heads and modifiers, but there are no distinct classes for heads of predicate versus referential phrases.

(11)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Modifier</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The system in (12)* is excluded because it has two rigid classes of modifiers, without the corresponding rigid classes of heads. As such it violates constraint (12a) in both the predication and the reference domain.

(12)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Manner adverb</td>
</tr>
<tr>
<td></td>
<td>Adjective</td>
</tr>
</tbody>
</table>

The system in (13)* is excluded because it violates constraint (12a) in the predication domain: It has a class of rigid lexemes that can be used as the modifier in a predicate phrase, but no rigid class for heads of predicate phrases.

(13)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Manner adverb</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The system in (14)* also violates constraint (12a), but in the referential domain, where it has a rigid class of modifiers, but no rigid class for heads.

(14)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Adjective</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The system in (15)* violates constraint (11a): It has a rigid class of heads in the referential but not in the predication domain.

(15)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicative</td>
<td>Noun</td>
</tr>
<tr>
<td>Predicative</td>
<td>Adjective</td>
</tr>
</tbody>
</table>

For the same reason, the system in (16)* is excluded:

(16)*

<table>
<thead>
<tr>
<th>Predication</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicative</td>
<td>Noun</td>
</tr>
<tr>
<td>Predicative</td>
<td></td>
</tr>
</tbody>
</table>

The system in (17)* violates constraint (12a): It has a class of rigid modifiers in the reference domain, but not the corresponding rigid class of heads.
(17)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Predicative</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

The system in (18)* involves a flexible class of lexemes that can be used as the head of a referential phrase and as a modifier in a predicate phrase (Flex C). It is excluded because it violates constraint (12a) in the reference domain, where it has a rigid modifier class without a rigid head class.

(18)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

The system in (19)* is excluded because it violates four constraints. It violates constraint (11b) because it has a flexible lexeme class that can be used as the head of a referential phrase, but no lexeme class for heads of predicate phrases. In addition, this system violates constraint (12a) in the reference domain, where it has a specialized class of lexemes for referential modifiers, but no specialized class for referential heads. It also violates constraint (12b), because it has a flexible class of lexemes that can be used as the modifier (but not the head) of a predicate phrase, without a lexical class for heads of predicate phrases. Finally, this system violates constraint (13): it has distinct classes for heads and modifiers in referential phrases, but no distinct class for heads of predicate phrases as opposed to referential phrases.

(19)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

The system in (20) is excluded because it violates (11b): it has a flexible lexeme class that can be used as the head of a referential phrase, and as a modifier in a predicate phrase, but no lexical means to express the head of a predicate phrase. In addition, this system violates constraint (12b) in the predication domain, where it has a flexible lexeme class that can be used for modification, but no class for heads.

(20)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

The system in (21)* involves a flexible lexeme class that can be used as the head of a predicate phrase and as a modifier in a referential phrase. It is excluded because it violates constraint (11a). It has a rigid class of heads in the referential, but not in the predication domain. In addition, it violates constraint (12a) in the predication domain, where it has a rigid modifier class but no rigid head class.

(21)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>

The system is (22)* is excluded because it violates constraint (11a): It has a rigid class of heads in the referential, but not in the predication domain.

(22)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>
The system in (23)* is excluded because it violates constraint (12a) in the predication domain, where it has a rigid class of modifiers, but a flexible class for heads. In addition, this system violates constraint (12a) in the reference domain: it has a flexible lexeme class that can be used as a modifier in a referential phrase, but no lexical means to express the function of head of that phrase. Finally this system violates constraint (13): it has distinct classes for heads and modifiers in the predication domain, but no distinct classes for heads of referential phrases, as opposed to predicate phrases.

(23)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Flex D</td>
</tr>
<tr>
<td>Reference</td>
<td>-</td>
</tr>
</tbody>
</table>

The system in (24)* violates constraint (12b), since it has a flexible lexeme class that can be used as a modifier in a referential phrase, but no lexeme class that can be used as the head of that phrase.

(24)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Flex D</td>
</tr>
<tr>
<td>Reference</td>
<td>-</td>
</tr>
</tbody>
</table>

The system in (25)* is excluded since it has only one rigid class of heads, but in the reference rather than in the predication domain. As such it violates constraint (11a).

(25)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>-</td>
</tr>
<tr>
<td>Reference</td>
<td>Noun</td>
</tr>
</tbody>
</table>

The system in (26)* has a rigid class of modifiers in the referential domain, without the corresponding rigid class of heads. As such it violates constraint (12a).

(26)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>-</td>
</tr>
<tr>
<td>Reference</td>
<td>-</td>
</tr>
</tbody>
</table>

The system in (27)* also violates constraint (12a), since it has a rigid class of modifiers without the corresponding rigid class of heads, but this time in the predication rather than the reference domain.

(27)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>-</td>
</tr>
<tr>
<td>Reference</td>
<td>-</td>
</tr>
</tbody>
</table>

The system in (28)* has distinct classes of heads and modifiers in the predication domain, while it does not have distinct classes for heads of predicative versus referential phrases. Thus, constraint (13) is violated.

(28)*

<table>
<thead>
<tr>
<th>Head</th>
<th>Modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predication</td>
<td>Verb</td>
</tr>
<tr>
<td>Reference</td>
<td>-</td>
</tr>
</tbody>
</table>

The system in (29)* violates constraint (12a), since is has a rigid class of modifiers in the reference domain, without the corresponding rigid class of heads. In addition, this system violates constraint (13) because it has distinct classes of heads and modifiers in the predicate domain, but no distinct classes for heads of referential phrases, as opposed to predicate phrases.
The system in (30)* is excluded because it violates (12a) in the reference domain, where it has rigid modifiers without any class for heads.

The system in (31)* violates two constraints: (11a), since it has a single rigid class of heads, but not in the predication domain. And (13), since it has a head-modifier distinction in the referential domain, but no distinct classes of lexemes for heads of predicative versus referential phrases.

The system in (32)* also violates two constraints: It has a single rigid class of heads, but not in the predication domain, so that (11a) is violated. Furthermore, (12a) is violated in the predication domain, for which the system has a rigid modifier class without the corresponding rigid head class.

The system in (33)* violates constraint (12a), since it has a rigid class for heads in the referential but not in the predication domain. It violates (12a), since it has a rigid modifier class in the predication domain but no rigid head class in that domain. And it violates (13), because it has a head-modifier distinction in the referential domain, but no distinct classes for heads of predicative versus referential phrases.

The system in (34)* has two rigid classes of modifiers, without the corresponding rigid classes of heads, so that (12a) is violated in both the predication and the reference domain.
Appendix III: Dependent Clause Constructions Key Examples

Tagalog

pag-Pred ('gerund')

Functional distribution: Flex: Pred Head?, Ref Head, Ref Mod, Pred Mod = PoS minus Pred Head (contentives) Himelmann (2005: 372): "As with all Tagalog content words, gerunds can be used in any syntactic function, provided their meaning fits."

Structural type: 3 (D-ALT)

Verbal categories: No voice/mood marking, aspect can be expressed.

Nominal categories: Combines with the same phrase-marking function words as all other content items. (Pred Head function is marked by sentence-initial position.)

Argument encoding: POSS - SENT1

Examples:

Pred Head: [pag-la-luto? ng pagkain] ang trababs niyá
GER-COF-cook GEN food SPEC work 3SG.POSS
'His/her job is cooking food.' (Himmelmann 2005: 372)

Ref Head: pag-bawal-an mo ang bata?sa [pag-la-laró? sa laangan]
SF-FORBIDDEN-LV 2SG.POSS SPEC child-LK DIST LOC NMNZ-RDP-play LOC street
'Forbid that child to play in the street.' (Himmelmann 2005: 373)

Ref Mod: No example available

Pred Mod: [pag-dating naming don in-iwan] namin don ang bangká
CONV-ARRIVAL 1PL.EXCL.POSS DIST.LOC REAL.UV-ABANDON 1PL.EXCL.POSS DIST.LOC SPEC boat
'When we arrived there we abandoned the boat, ….' (Himmelmann 2005: 373)

na/-ng/(kung) + clause (=Ø)

Functional distribution: Flex: Ref Head, Ref Mod, Pred Mod = PoS minus Pred Head (contentives)

Structural type: 1 (Balanced)

Verbal categories: All retained

Nominal categories: Combines with the same phrase-marking function words as all other content items. (Pred Head is marked by sentence-initial position.)

Argument encoding: SENT- SENT

Examples:

Ref Head: Sinabi [kung maganda si Maria]
tell -LK beautiful nom Maria.
'I said that Maria was beautiful.' (Schachter & Otanes 1972: 173)

Gusto ni Pepito na [sagip-in ang dabon]
liking PN.POSS Pepito LK salvage-PV SPEC leaf
'Pepito wanted to catch the leaf.' (Himmelmann 2005: 364)

Nakita ni Pedro-(ng)/Manuel na [puno na ang bus]
see NOM Pedro-LK/Manuel LK full LK RM bus
'Pedro/Manuel saw that the bus was already full.' (Schachter & Otanes 1972:177)

Ref Mod:

Sa mga lalaki na [maN-ibig nung kasyu-na anak]
LOC PL man LK AV-LOVE DIST.GEN.LK 3SGDAT-LK child
'So he held a contest' between the men who courted his child.' (Himmelmann 2005: 368)

1 The following symbols are used in this appendix: '=' means 'same functional possibilities as a PoS class in the same language. This PoS class is added between brackets. '≠' means 'different functional possibilities than any PoS class in the same language'. The relevant PoS classes, i.e. those that express the function(s) in which the DC is used, are added between brackets. When there is no lexical class available for the relevant function(s) this is also indicated. For the meaning of other abbreviations concerning the functions, the expression, and the classification of the DCs, see Chapters 3 and 6.

2 As Koptjevskaja-Tamm (1993: 119-120) explains, the SENT classification of coding of the second argument is not entirely straightforward, because there is no difference between the marking of the second argument of an actor-voice predicate and the possessor in Tagalog; both are marked by ng. However, there is a second type of possessive construction in which the possessor is expressed as a sa-phrase. Since the first argument in a gerund construction can be both a sa and a ng-phrase (just like possessors), while the second argument can only be a ng-phrase, Koptjevskaja-Tamm argues in favour of SENT expression of the second argument.

3 The status of kung is unclear: it may be a combination of –ng with some other element.
Pred Mod:

[Bíglá] rýl]-ng nagbangon

sudden 3SG-LK REALAV:rising

'She got up quickly' (Himmelmann 2005: 360)

**KHARIA**

**Pred-Ø/RDP** *(Freestanding form/masdar)*

Functional distribution: Flex: Pred Head (HAB), Ref Head, Ref Mod, (+ case) Pred Mod = PoS (contentives)

Structural type: 3 (D-ALT)

Verbal categories: No voice/tense marking and no Person agreement; retains valency-related marking such as causative and passive/reflexive marking. In Pred Head function obligatorily combined with the middle voice, indicating habituality.

Nominal categories: May take case and number

Argument encoding: POSS - SENT, occasionally also POSS - POSS

**EXAMPLES:**

**Pred Head:**

\[ iɲ dařa biɖa\-biɖa \] = ki

1SG water pour.out-REDP=EMOT=1SG

'I used to pour water out.' (i.e. that was my job). (Peterson 2006: 74)

Ref Head:

\[ OɁ=ya=Ɂbay-bay\]

house=GEN build-REDP NEG=1SG like=M.FRS

'I don’t like (the act of) building houses.' (Peterson 2006: 73)

Ref Mod:

\[ iɲ =a=Ɂdura=te \] ruɖa=te

1S=gen door-obl open-REDP

'The key I open the door with.' (Peterson 2006: 73)

Pred Mod, without case-marking (with reduplication):

\[ nata\-te \] [ɖa=te \] lemen\-Ɂ=te

Rate=obl sit.down RDP sleep =EMOT=M.PST

'while he was seated, Rata became tired.' (Peterson 2006: 249)

Pred Mod, with oblique case-marker:

\[ khõ=ɖa=te \] [ɖa=te \] ko=ki=ya=Ɂjha\-i\-ya\]

a.while=CNTR sister= obl=FOC doze.off=EMOT=M.PST

'Sitting for a while, he dozed off (…)' (Peterson 2006: 295)

**Pred=na** *(Infinitive)*

Functional distribution: Flex: Ref Head, Ref Mod, (+case/postposition) Pred Mod = PoS minus Pred Head (contentives).

Structural type: 3 (D-ALT)

Verbal categories: No Voice/Tense/Person agreement

Nominal categories: Case

Argument encoding: POSS - SENT or Ø - SENT

**EXAMPLES:**

Ref Head:

\[ in \] [u iku=Ɂɖ sundar konthe=ki=te bajhay=kon\]

1SG this very beautiful bird=pl-obl trap=SEQ.CONV

satay=na um=in lam=te

torture=INF NEG=1SG want-APT.M.FRS

'I don’t want to trap and torment these beautiful birds.' (Peterson 2006: 259)

Ref Mod:

\[ Ho=ka \] [ɖa=te \] buli=na modhe buŋ kha=Ɂ=ki=ya? jha=Ɂ\]

that=s.HUM [village.section wander=PTC] means INSTR Kharia=pl=GEN all

bahbah=te erikhudi koŋ may=sik=ko=a

mannerisms=obl from.bottom.to.top know total=FFV=APT.M.PST

'She had learned by wandering through the village (= through the in the village wandering means) all of the mannerisms of the Kharia inside out. (Peterson 2006: 306-307)

Pred Mod, without case marking (with reduplication):

\[ …lekin lam=na lamna \] sou=te ikau jughay Ɂ=Ɂpiyas la=r=ki.\]

but search=INF RDP all=obl very much water-thirst =EMOT=M.PST

'But searching and searching, [they] all became very thirsty.' (Peterson 2006: 248)
Appendix iii: Dependent Clause Constructions Key Examples

Pred Mod, with oblique case-marking (and reduplication):

\[ [\text{Aw}=\text{na} \text{ avwa}]=\text{te} \ 	ext{khari}=\text{ki} \ 	ext{be}=\text{te} \ 	ext{stay}=\text{INF} \ 	ext{RDP}=\text{OBL} \ 	ext{Kharia}=\text{PL} \ 	ext{that}=\text{OBL} (=\text{there}) \ 	ext{again} \ 	ext{house} \ 	ext{dura} \ 	ext{bay}=\text{kon} \ \text{ikug}=\text{ga} \ \text{memon} \ =\text{aw}\text{ki}=\text{may} \ 	ext{door} \ 	ext{build}=\text{SEQ} \ 	ext{conv} \ 	ext{very}=\text{foc} \ 	ext{year} \ 	ext{up} \cdot \text{to} \ \text{stay}=\text{M.PST}=\text{3PL} \]

‘Staying, the Kharia stayed there for several years, again building homes (houses and doors).’ (Peterson 2006: 249)

**Pred=na-wala** *(participle)*

Remark: Borrowed from Hindi. Structural coding consists of the infinitive =na followed by =wala.

It denotes iterativity and habituality.

Functional distribution: Rig: Ref Mod ≠ PoS (contentives)

Structural type: 2 (D-SENT)

Verbal categories: No Voice/Tense/Person agreement

Nominal categories: None (no case agreement)

Argument encoding: The relativized argument is gapped, other argument(s) are SENT.

**Example:**

Ref Mod:

\[ [\text{Jharkan}=\text{te} \ \text{aw}=\text{na-wala}] \ \text{lebu}=\text{ki} \ \text{iku}=\text{dr}=\text{gu} \ \text{milansar} \ =\text{aw}=\text{ta}=\text{ki} \ 	ext{Jharkhand}=\text{obl} \ 	ext{live}=\text{inf} =\text{ptc} \ 	ext{person}=\text{PL} \ 	ext{very much} \ 	ext{friendly} \ \text{cop} \ =\text{m.prs} =\text{PL} \]

‘The people from Jharkhand are very friendly.’ (Peterson 2006: 307)

**Pred-al** *(participle)*:

Remark: The suffix attaches only to lexical predicates of Sadani origin which end in -a or -ay.

This suffix has thus been borrowed with the root.

Functional distribution: Rig: Ref Mod ≠ PoS (contentives)

Structural type: 3 (D-ALT)

Verbal categories: No Voice/Tense/Person agreement

Nominal categories: None (no case agreement)

Argument encoding: POSS - SENT

**Example:**

\[ \text{Muda} \ 	ext{mo}=\text{ɲ} \ 	ext{Brahman} \ [\text{ho}=\text{ko} =\text{a} \ 	ext{daru}=\text{te} \ 	ext{ʈ}=\text{aŋ}-\text{al}] \ \text{janew}=\text{te} \ \text{yo}=\text{yo} \]

‘But one brahman that=sg. hum=gen tree=obl hang=ptc holy.thread=obl see=act pst’

‘But a Brahman saw the holy thread which he had hung on a tree.’ (Peterson 2006: 307)

**Pred-ker(r) / -kon / -kan** *(converb)*

Remark: “The first two of these markers are direct borrowings from Sadani. […] These markers denote, among other things, that the two or more (sub-)predicates are portrayed by the speaker as being directly related to one another in some way, combining to form a larger, more complex event. […] =kon […] appears to be a calque from the Sadani form =ker. Like the cognate form =kar in Hindi, =ker in Sadani appears to derive from the root kar ‘do’. The sequential converbal marker =kon in Kharia apparently derives from a similar lexeme, i.e., ikon ‘make, do’. The core function of these forms is to denote the completion of one action before another begins. [But they] are also often used to denote the manner in which an action is carried out. In these cases, the action denoted by the converb is generally a more exact specification of that of the morphologically finite predicate.” (Peterson 2006: 243-244)

Functional distribution: Rig: Pred Mod ≠ PoS (contentives)

Structural type: 2 (D-SENT)

Verbal categories: No Voice/Tense/Person agreement

Nominal categories: None.

Argument encoding: Ø - SENT

**Examples:**

Pred Mod:

\[ \ldots [\text{lay} \ 	ext{koj}=\text{kon}] \ 	ext{go}=\text{ju} \ 	ext{bay}=i=\text{may} \]

‘…they have built the path by digging an scraping [the dirt away].’ (Peterson 2006: 244)

\[ \text{rakin} \ 	ext{o}=\text{ph} \ 	ext{jugbay} \ 	ext{khia}=\text{ti} =\text{ta} =\text{ro} \ [\text{gone} \ 	ext{keb}=\text{kon}] \ 	ext{gam}=\text{te} \ldots r\]

‘The witch grows even angrier and … grinding her teeth, says:…’ (Peterson 2006: 244)

**Pred-ge + RDP** *(imperfective converb)*

Remark: The =ge form is primarily a focus marker.

Functional distribution: Rig: Pred Mod ≠ PoS (contentives)

Structural type: 2 (D-SENT)

Verbal categories: No Voice/Tense/Person agreement

Nominal categories: None.

Argument encoding: Ø - SENT
Example:
Pred Mod
ro  ho  kuda  kholag?  daru  samhle=to  [\\[num=ga  \[numga]  goj  jem=ta
and that millet bread=GEN tree base=OBJ cry=CONV RDP die AP=MP.PRS
‘And crying and crying, she just died at the base of that millet bread tree.’ (Peterson in prep: 248)

Pred=ta ('imperfective converb')
=ta is homophonous with the general imperfective middle marker.
Functional distribution: Rg: Pred Mod ≠ PoS (contentives)
Structural type: 2 (D-SENT)
Verbal categories: No Voice/Tense/Person agreement
Nominal categories: None
Argument encoding: Ø - SENT

Example:
Pred Mod:
ele  \[am=pe-te  go=ta  go=ta]  ban=ti/ti  \[iɲam  =\[iɲamga]
[1pl.excl 2=2pl=obl carry.on.shoulders=conv RDP that=side this=side take=act.ibr=3pl.excl
‘We will carry you around on our shoulders.’(=we will take you, carrying you on our shoulders)’ (Peterson 2006: 248)

no/Ø + clause
Remark: Used for object complement clauses, especially with utterance predicates.
Functional distribution: Rg: Ref Head ≠ PoS (contentives)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
ap=gb,mo  rat= ae=ki  remakh-oj?  ro  gam-o?  [no  babu  musa  iɲ  kimir
father=3ross Rata=obl call=act.pst and say=act.pst comp child today 1sg forest
\[onna  \[onm=ip  pal=]  go=inf neg=1sg be.able=act.ibr
‘His father called Rata and said “child, to day I will be unable to go to the forest.’ (Peterson 2006: 298)

gam=kon + clause
Remark: ‘The form gam=kon is the sequential converb of gam ‘to say’. It is occasionally found instead of no
as a kind of quotative form.
Functional distribution: Rg: Ref Head ≠ PoS (contentives)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
[je  jansar  taz=ki  ke  jansar=ya?  gbos  al=er=ki]\[gam=kon]  gam=o?
cr animal kill=act.ibr=pl that animal=gen meat take=act.ibr quot say=act.pst
‘Whatever animal they kill, that animal’s meat they should bring, he said.’ (Peterson 2006: 299)

Correlative construction
Remark: There are two types of correlative constructions:
(i) With je-class markers: all correlative forms begin with j- and have been borrowed from
Indo-Aryan.
(ii) With a/i/other question particle-class markers: all correlative forms are homophonous with
interrogatives. This construction is not borrowed from Indo-Aryan, although it could be an older
calque of the Indo-Aryan correlative construction, using purely language-internal means.
In both construction types, the head is usually repeated in the main clause, preceded by a demonstrative.
Alternatively, the head is not repeated and only the demonstrative is there.
Functional distribution: Rg: Ref Mod ≠ PoS (contentives)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT (optional gapping)
Appendix iii: Dependent Clause Constructions Key Examples

Examples:
Ref Mod:
(ʃi-class)
\[ \text{adj [ie bhere] epski se bhere adj=ya pajo=te soreŋ kui=ki} \]
\[ \text{anaph CR time return=m.pst dem time anaph=gen bundle=obl stone find=m.pst} \]

‘Which time he returned, (at) that time he found a stone in the bundle.’ (Peterson 2006: 302)

(ʃi-i-class)
\[ \text{a=bo=Ɂ=te pujapaʈh karay=na au=ki,} \]
\[ \text{ho=bo=Ɂ=te ɖam=ke, ...} \]
\[ \text{q=place=obl sacrifice do=inf cop=m.pst dem place=obl arrive=seq.conv} \]

‘Having arrived at the place where the sacrifice was to be done ...’ (Peterson 2006: 302)

Unmarked gapped relative clauses with a (partially) finite predicate

Remark: This construction is balanced, except that person marking may be lost on the dependent predicate (see second example).

Functional distribution: 1 (Balanced)

Structural type: 1 (Balanced)

Verbal categories: All retained

Nominal categories: None

Argument encoding: SENT - Ø / Ø - SENT (gapping)

Examples:
Ref Mod:
\[ \text{[iɲ yo=yo Ɂj lebu=ki iɲ=aɁ hoʈel=te aw=ta=ki} \]
\[ \text{1sg see=act.pst person=pl hotel=obl live=m.prs=pl} \]

‘The people I saw live in my hotel.’ (Peterson 2006. 303)

\[ \text{[iɲ =te yo=yo Ɂj lebu=ki ula Ɂlikha=yo Ɂ}=ki} \]
\[ \text{1s=obl see=act.pst person=pl letter write=act.pst=pl} \]

‘The people who saw me wrote a letter.’ (Peterson 2006. 303)

Kambera

Pa-deranked clause

Remark: In combination with the prepositional verb wàngu ‘use’ this construction can be used as an adverbial clause with an interpretation of simultaneity or immediate sequence (see Pred Mod example below).

Functional distribution: 1 (Balanced)

Structural type: 2 (D-SENT)

Verbal categories: No aspect, no mood marking. (Kambera has no tense marking.)

Nominal categories: DET + number agreement

Argument encoding: Ø - SENT / POSS - Ø In Ref Head function the subject remains unexpressed under co-referentiality; the object is SENT and cross-referenced on the dependent predicate, in the DAT form. In Ref Mod function the subject is POSS, and the object is gapped, but remains cross-referenced on the dependent predicate in the DAT form.

Examples:
Ref Head:
\[ Ta-pakiring [pa-tinus-nya na lau] haromu \]
\[ 1pl.nom-start comp-weave-3sg.dat art sarong tomorrow \]

‘We will start to weave the sarong tomorrow.’ (Klamer 1998: 338)

Ref Mod:
\[ Ta-pakiri-nja da lau [pa-tinus-nda] \]
\[ 1pl.nom-start-3sg.dat art sarong rel-weave-1pl.dat \]

‘We start (with) (them) the sarongs woven by us.’ (Klamer 1998: 338)

\[ na kalembi na [pa-kei wā-nggu-nya] \]
\[ art shirt art rel-buy use-1sg.gen-3sg.dat \]

‘the shirt that I bought’ (Klamer 1998: 326)

Pred Mod:
\[ Patiанг ana mandài-ndài [wàngu pa-buta ana rumba] \]
\[ wait dim rop-belong use comp-pick dim grass \]

‘(We) wait a while weeding some grass in the meantime.’ (Klamer 1998: 240)

ma-deranked clause

Functional distribution: 2 (D-SENT)

Structural type: 2 (D-SENT)

Verbal categories: No aspect, no mood marking. (Kambera has no tense marking.)
Nominal categories: DET + number agreement (in Ref Mod function)
Argument encoding: Ø - SENT (gapping)

Examples:
Ref Mod:
Na-meti-ka na tau na [ma-piti-ya na kabela-nggu]
3SG.NOM-die-3PFV ART person ART REL-take-3SG.ACC ART machete-1SG.GEN
'The person that took my machete died already.' (Klamer 1998: 315)

Ita-nggu-nya na tau na [na-meti kuru uma-na].
see-1SG.GEN-3SG.DAT ART person ART REL-die wife-3SG.GEN
'I saw the man whose wife died.' (Klamer 1998: 320)

Unmarked nominalized clause
Remark: Nominal clauses can be dependent or independent. In combination with a conjunction, the construction can apparently also be used in adverbial function as a simultaneity clause (see Pred Mod examples below).
Functional distribution: Flex?: Pred Head/main clause, Ref Head (+ CONJ also Pred Mod, simultaneity).
Structural type: 3 (ALT-SENT)
Verbal categories: (Some) aspect and mood marking is retained (also lexically). (Kambera has no tense-marking).
Nominal categories: DET, when functioning as such construction is cross-referenced on the main predicate as an object, with a DAT form.
Argument encoding: POSS - SENT
The subject is expressed through a genitive pronominal enclitic.

Examples:
Pred Head:
[Na apu-mu, katuda-na] la pino bolak-ka una
ART grandmother-2SG.GEN sleep-3SG.GEN LOc top mattress-3PFV EMPH.3SG
'Your granny, she will sleep on a mattress.' (Lit. 'Your grandmother’s sleeping is on a mattress.' (Klamer 1998: 97)

Ref Head:
Nda ku-pí-anggau [na ngàndi-mu rà kuta]
NEG 1SG.NOM-know-MOD-2SG.DAT ART take-2SG.GEN leaf pepper plant
'I didn’t know that you would bring kuta.' (lit.: I didn’t know (of) your bringing kuta.)' (Klamer 1998: 97)

Nda ku-mbuti-nya [na taka-mu]
NEG 1SG.NOM-expect-3SG.DAT ART arrive-2SG.GEN
'I did not expect your coming.' (Klamer 1998: 315)

Pred Mod:
[Ba meu-meu-na,] ba na-imbu-ya
CONJ rdp-roar-2SG.GEN CONJ 3SG.NOM-search-3SG.ACC
'And it roared (a tiger), while it went after him.' (Klamer 1998: 100)

[ba jía-pa-ma na banggapi-3SG-DAT dà]
CONJ continuously-EMPH ART clasp-3SG.GEN-3SG.DAT inside
'while he kept clasping it inside…' (Klamer 1998: 97)

wà + clause
Remark: Quotative construction.
Functional distribution: Rig: Ref Head ≠ PoS (contentives, adverbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
Ka [na-ngàndi-ya na mbuku] wà-nga-nya làti
CONJ 3SG.NOM-take-3SG.ACC ART book say-1SG.GEN-3SG.DAT in fact
'In fact, I told him that he should take the book.' (Klamer 1998: 347)

Samoan
Pred-ya
Functional distribution: Rig: Ref Head, ≠ PoS (contentives, adverbs)
Structural type: 3 (D-ALT)
Appendix iii: Dependent Clause Constructions Key Examples

Verbal categories: No TAM marking
Nominal categories: DET/CASE
Argument encoding: POSS - SENT (occasionally also POSS - POSS)

Examples:
Ref Head:

A le faaaleleve le tupu i [le ai-ga
pst neg trouble art king ld art cat-nmlz
apa ma moli a le pipili ma le tauasa]
apple and citrus poss art lame and art blind
‘The king was not troubled that the lame and the blind ate the apples and oranges.’

At na oo lave in moumou mali se atu le pisa
But pst reach emph conj disappear gentle dir art noise
o [e sapini-ga o Pale ma Maria e o la Tina]
poss art whip-nmlz poss Pale and Maria erg poss 3du mother
‘But finally the noise of the whipping of Pale and Maria by their mothers gently faded away.’
(Mosel 1992: 279)

Unmarked nominalized clause
Functional distribution: Rig: Ref Head, ≠ PoS (contentives, adverbs)
Structural type: 3 (D-ALT)
Verbal categories: No TAM marking
Nominal categories: DET/CASE
Argument encoding: POSS - SENT

Example:
Ref Head:

E lelei [l-a-u tunu ia]
genr good art-poss-2sg roast fish
‘Your fish roasting is good.’ (Mosel 1992: 267)

ona / ina * deranked clause
Remark: Ona is used for core-arguments, ina for adjuncts.
Functional distribution: Rig: Ref Head ≠ PoS (contentives, adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No TAM marking
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:

ua taga [ona inu ava malosi tatou]
pfv allowed comp drink ‘kava’ strong 1.incl.pl
‘It is allowed that we drink alcohol.’ (Mosel & Hovdhaugen 1992: 599)

*e * clause
Functional distribution: Rig: Ref Mod ≠ PoS (contentives, adverbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: (DET)
Argument encoding: SENT - Ø/Ø - SENT (gapping) Gap can be filled with anaphoric element.

Example:
Ref Mod:

‘O lue o le tama’ita’i l-[i na tatou o i ai.]
pres that pres art woman art-rel pst 1.incl.pl go(PL) ld anaph
‘She is the woman we went to find.’ (Mosel & Hovdhaugen 1992: 635)

Unmarked clause
Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (contentives, adverbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: (DET)
Argument encoding: SENT - SENT
EXAMPLES:

Ref Head:

Na ioa e Tegilau [via sau Sina]
past know erg Tegilau pfv comes Sina
'T knew that S had come.' (Mosel & Hovdhaugen 1992: 589)

Ref Mod:

Ua tu le alii lea [na ua e Popi]
pfv stand.up art man that pst bite erg Popi
'The man who was bitten by Popi stood up.' (Mosel & Hovdhaugen 1992: 635)

GUARANI

course + /b Generated/

Functional distribution: Rig: Ref Head ≠ PoS (contentives, verbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: (DET)
Argument encoding: SENT - SENT

Examples:

Ref Head:

Rey-anú [íte is-raú hí]
you-hear I I-be.sick comp
'You heard that I was sick.' (Gregores & Suárez 1967:158)

ai-buá la [v s-où mošša i há]
I-know def(det) neg he-go mod neg comp
'I know that he does not intend to go.' (Gregores & Suárez 1967:158)

Rei-mošša [íte-tavi]
You-think I-be.silly
'You think that I am silly.' (Gregores & Suárez 1967: 157)

Pred-va + clause*

Functional distribution: Rig: Ref Mod ≠ PoS (contentives, verbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - Ø/Ø - SENT (gapping)
The relativized item is gapped, but there is a person prefix on the dependent predicate.

Example:

Ref Mod:

A-bechali a karai [o-jagua-va-ekue ka angu'a].
I-see def man 3-buy-rel-pst def mortar
'I saw the man who bought the mortar.' (Velázquez-Castillo 2002: 162)

Clause + ve:

Functional distribution: Rig: Pred mod ≠ PoS (contentives, verbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT- SENT

Pred Mod:

H-ua
[si-ö ve]
she-cry she-go.out adv
'She goes out crying.' (Gregores & Suárez 1967: 180)

SANTALI

Unmarked clause; pred without -a (IND)
Remark: In Ref Head subject function, the construction shows no subject marking, while middle voice markers and TAM can be expressed. In Ref Head object function subject marking and TAM are lost, while object markers are retained. Only with verbs of perception subject marking and TAM can be expressed. In Ref Mod function, subject pronominals are omitted, but all TAM suffixes can be expressed.

* The status of the relativizer is not entirely clear: it does not seem to change the internal syntax of the dependent clause, but does attach to the predicate, preceding the tense marker.
Appendix iii: Dependent Clause Constructions Key Examples

The relativized element is gapped. In Pred Mod function, with case-marking, there is neither person marking (subject/object) nor TAM marking, but middle voice can be expressed. In combination with a postposition (with simulative semantics), all verbal categories are expressed, except for the indicative marker.

Functional distribution: Flex: Ref Head, Ref Mod (+ case (LOC/INSTR)/postposition also Pred Mod) ≠ PoS (contentives, verbs)

Structural type: 2 (D-SENT)
Verbal categories: Variable (see above)
Nominal categories: CASE
Argument encoding: SENT - Ø/Ø - SENT (co-referentiality/gapping)

Examples:
Ref Head:
Subject:
[her bahun-kate dera-k'-'do] buy bes-a
sun set-conv camp-M-TOP NEG good-IND
'it is not good to camp after sunset.' (Neukom 2001: 181)

Object:
[onko a] mana-ko-m
those PL bring-3PL.OBJ forbid-3PL.OBJ-2SG.SBJ
'Forbid them to bring those.' (Neukom 2001: 182)

Perception predicate:
[mɔ̃ṛɛ̃-gɔṭɛ̃n iməi i daki lo=ko hijuk'kan-e]
[nin]-CL woman water fetch-3PL.SBJ come-M-IPFY-3SG.SBJ see-V-PST:ACT-3PL.OBJ-IND
'He saw five women come to fetch water.' (Neukom 2001: 183)

Ref Mod:
[Uni-y-[e]
[bujhə u-ɲɔ̃ k'-ket'-ko-a]
that-(anim)Y-3SG.SBJ understand.-little-PST:ACT person-3SG.SBJ beckon-V-APPL:PST:ACT-3SG.OBJ-IND
'He beckoned the man who had understood a little.' (Neukom 2001: 197)

Pred Mod:
[cala-k'-'calak'-te mit'-tay toyo-ko pel-tiok'-'ked-e-a]
go-M-RDP-INST one-CL jackal-3PL.SBJ see-reach-PST:ACT-3SG.OBJ-IND
'While they were walking along, they caught sight of a jackal.' (Neukom 2001: 187)

With postposition (simulative):
[one-y-e met'-'at'-ko leka-ge]
that-y-3SG.SBJ say-APPL:PST:ACT-3PL.OBJ like-FOC all person-3PL.SBJ go-M-IPFY-IND
'They all went as he had told him.' (Neukom 2001: 195)

Pred-kate:

Functional distribution: Rig: Pred Mod ≠ PoS (contentives, verbs)

Structural type: 2 (D-SENT)
Verbal categories: No Tense/Mood, no Person, voice can be expressed
Nominal categories: CASE
Argument encoding: Ø - SENT

Example:
Pred Mod:
[[nonka budi-kate] bəbərsə-kate-ko] cala-k'-'kan-a
like.this think-CONV be.hopeless-CONV-3PL go-M-IPFY-IND
'Thinking so and being hopeless, they walked on.' (Neukom 2001: 186)

1 This construction is described as a *converb* but its status is not completely unambiguous: "-'kate occurs elsewhere in isolated position as ‘then’ or together with deictic elements such a nit’soum’ (cf. nit-kate ‘nowadays’) or ona that’ (cf. ona-kate ‘thereupon’), or in combination with numerals, e.g. ponea-kate ‘(give them) four each.’ (Neukom 2001: 185)

2 Note that the subject pronominal is nevertheless retained in the second converbal form of the example.
Correlative construction

Remark: Correlative constructions are probably an influence from Indo-Aryan languages. These constructions make use of various types of pronouns: interrogative, indefinite, demonstrative. The dependent predicate often lacks the indicative marker -a, but the pronominal subject clitic is always present.

Functional distribution: Rig: Ref Mod = PoS (contentives, verbs)
Structural type: 1 (Balanced)
Verbal categories: Allmost all retained (see above)
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Mod:
ona dare [oka-m mak-at']
that(inanim) tree cut-pfv:act
‘the three which you have cut’ (Neukom 2001: 200)

Clause + mente

Remark: The complementizer/quotative is a lexicalized instrumental case-marked form of ‘to say’.

Functional distribution: Rig: Ref Head = PoS (contentives, verbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
Ba-kin pel-thik-e-kan-a, nai-dɔ
NEG-3DUAL.SG see-correct-3SG.OBJ-IPPFV-IND this(AN)-TOP
[skin-ren apa-t kan-a-ɛ mente]
they(DUAL) father-3POS poss cop-ind-3SG.SBJ comp/quot
‘They did not recognize that he was their father.’ (Neukom 2001: 183)

WARAO

Pred-kitane ('infinitive'):

Remark: According to Romero-Figeroa (1997) this construction is used for same-subject complements and for purpose-clauses, but no example is available of the former use.

Functional distribution: Rig: Pred Head ≠ PoS (non-verbs)
Structural type: 2 (D-SENT)
Verbal categories: None
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

Clause + kotai:

Functional distribution: Rig: Ref Mod ≠ PoS (non-verbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - Ø/Ø - SENT (gapping)

Example:
Ref Mod:
Ima-ya domu [nari-te kotai] mi-kitane nas-kotu
night-ALL bird fly-N.PST REL see-INF come-2PL.IMP
‘You all, come to see the bird that flies at night.’ (Romero-Figeroa 1997: 42)

TURKISH

Pred-DIK/-(y)AcAK

Remark: In combination with the postposition gibi this construction can be used as a simulative adverbial clause.

Functional distribution: Flex: Ref Head, Ref Mod (non-subject/possessor clauses), (+ postposition also Pred Mod, simulative). ≠ PoS (non-verbs, derived modifiers)
Structural type: 3 (D-ALT)
Verbal categories: No Aspect and Mood marking, (relative) tense is expressed by the choice of marker: -DIK for past and present, -(y)AcAK for future.
Nominal categories: CASE, nominal agreement
Argument encoding: POSS - SENT/Ø - POSS
(In Ref Mod function, the relativized element is gapped; the subject is POSS.)
Appendix iii: Dependent Clause Constructions Key Examples

Examples

Ref Head: [Ahmed-in öl-duc-ün]-ü duy-du-m 'I heard that Ahmed died.' (Kornfilt 1997: 50)

[Orhan-gen bir şey yap-ma-yec-i] kelliydı. Orhan-gen anything do-NEG-NMLZ-3SG.Poss it.was.obvious 'It was obvious that Orhan wouldn’t do/wasn’t going to do anything.' (Göksel & Kerslake 2005: 423)

Ref Mod: [adam-in git-tiq-i] okul man-gen go-PTC-3SG school 'the school that the man goes/went to' (Kornfilt 1997: 50)

[Pastay [anne-m-in analat-tiq-ti] ýapmaya çalı ştı-m] mother-1SG.Poss-GEN describe-NMLZ-3SG.Poss like 'I tried to make the cake [as my mother had described].' (Göksel & Kerslake 2005: 477)

Pred-mAK

Remark: "The crucial difference between -mAK clauses and those with -mA is that -mA clauses in the majority of cases contain their own subject, whereas -mAK clauses do not." (Göksel & Kerslake 2005: 413)

Functional distribution: Rig: Pred Head ≠ PoS (non-verbs)

Structural type: 2 (D-SENT)

Verbal categories: No TAM and Person agreement

Nominal categories: CASE (usually when functioning as a direct object complement, except with the verb iste- 'to want', see third example)

Argument encoding: Ø - SENT (co-referentiality)

Examples:

Ref Head: [Lütfen pencere-yi aç-mağ]-ı unut-ma please window-ACC open-INF-ACC forget-NEG 'Please, don’t forget to open the window!' (Kornfilt 1997: 51)


[Patreş-an yarın gör-eec-i] film Fatma-gen tomorrow see-PTC-3SG.Poss film 'the film that Fatma is going to/will be seeing tomorrow' (Göksel & Kerslake 2005: 442)

Pred-mA

Remark: In general terms noun clauses formed with -mA are less abstract in meaning than those formed with -mAK.

Functional distribution: Rig: Pred Head ≠ PoS (non-verbs)

Structural type: 3 (D-ALT)

Verbal categories: No TAM and Person agreement

Nominal categories: CASE, nominal agreement

Argument encoding: POSS - SENT (co-referentiality)

Examples:

Ref Head: [Kerkes-in birier kikaye anlat-ma-s] iste-n-yor-mu everyone-gen one.each story tell-NMLZ-3SG.Poss want-PASS-IPPV-EV.COP 'It seems they want [everyone to tell a story].' (Göksel & Kerslake 2005: 420)

Ahmed-t [ben-i bekle-me-sin]-i işle-di-m a.-DAT I-ACC wait-NMLZ-3SG.Poss-ACC tell-PST-1SG 'I told Ahmet to wait for me.' (Kornfilt 1997: 53)

Pred-An

Functional distribution: Rig: Pred Mod (subject/possessor clauses) ≠ PoS (non-verbs, derived modifiers)
Structural type: 2 (D-SENT)
Verbal categories: No TAM, no Person Agreement (verbal)
Nominal categories: None.
Argument encoding: Ø - SENT (gapping)

Examples:
Ref Mod:
[burada sat-ə-an] kitap-lar
here sell-PASS-PTC book-PL
‘the books that are sold here’ (Göksel & Kerslake 2005: 440)
[əgreten ol-an] haydar
teacher be-PTC Haydar
‘Haydar, who is a teacher’ (Göksel & Kerslake 2005: 440)
[araba-si (əl-in-an)] komşu-muz
car-3SG.Poss steal-PASS-PTC neighbour-1PL.Poss
‘our neighbour, whose car was stolen’ (Göksel & Kerslake 2005: 440)

Pred-(y)ArAk
Remark: Normally, the subject is unexpressed under co-referentiality.
Functional distribution: Rg: Pred Mod ≠ PoS (non-verbs, derived modifiers)
Structural type: 2 (D-SENT)
Verbal categories: No TAM / Person agreement
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)
Example:
Pred Mod:
Ben [etraf-im-a bak-arak] yür-ər-əm
I around-1SG.dat look-conv(manner) walk-aor-1SG
‘I walk looking around (myself).’ (Kornfilt 1997: 73)

Pred-(y)/d... Pred-(y)/d
Remark: “This construction occurs either with identical verb stems or with different ones. Its use is less widespread than that of -(y)ArAk, and its meaning is more emphatic, stressing the continuous or repeated nature of the action it expresses. The forms involving two different verb stems are for the most part lexicalized items.” (Göksel & Kerslake 2005: 476)
Functional distribution: Rg: Pred Mod ≠ PoS (non-verbs, derived modifiers)
Structural type: 2 (D-SENT)
Verbal categories: No TAM/Person agreement
Nominal categories: None
Argument encoding: Ø - Ø (/Ø - SENT?)

Pred Mod:
Geç kade [əğala-yəa agala-yəa] hikayeini anlattı
The young woman told her story [continuously weeping]’ (Göksel & Kerslake 2005: 476)
Adam [it-e kək-ə] oне geşmeye calışıyordu
man [pushing shoving] was trying to get to the front
‘[Pushing and shoving,] the man was trying to get to the front.’ (Göksel & Kerslake 2005: 476)

ki + clause
Remarks: This construction is borrowed from Persian.
In relative clause function, the construction is mostly non-restrictive (the head is almost always the subject of the main clause, and 3rd person singular or plural). Normally, the relativized item is gapped, but under certain circumstances, it may or must be reiterated in the dependent clause, by means of a resumptive pronoun.
Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (non-verbs, derived modifiers)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping in Ref Mod function)

Examples:
Ref Head:
İstiyar-sum [ki yarın ben-ile sinema-ya gel-ein]
want-PRES.PROGR.1SG comp tomorrow I-gen-with cinema-dat come-sg.opt
‘I want you to come to the movies with me tomorrow.’ (Kornfilt 1997: 46)
Appendix iii: Dependent Clause Constructions Key Examples

Sanşorum [ki iy-in-i brarak-mak isti-yor]
I think COMP job-sg.poss-acc leave-nmlz want-IPFV

‘I think that s/he wants to quit his/her job.’ (Göksel & Kerslake 2005: 409)

Ref Mod:
Bir adam [ki s embark-in-2 sev-me-z] yahni ya-yasa-mak-dir
a man REL child-pl3sg-acc love-neg-aor alone live-neg-ep/cop

‘A man who does not love his children must live alone.’ (Kornfilt 1997: 60)

bi ahs [ki baklava yap-mey-i bil-me-sin]
a cook REL baklava make-nmlz-acc know-neg-3sg-opt

‘a cook who doesn’t know how to make baklava’ (Göksel & Kerslake 2005: 459)

Clause + diye

Remark: diye is the converbal (-yA) form of the verb de ‘to say’. The construction is used for complements of predicates of speech other than de, and for complements of predicates of cognition, perception, and emotion.

Functional distribution: Rig: Ref Head ≠ PoS (non-verbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
Meral [Turgut onu Selim’le gör-ür-se diye] kork-yor-du
Meral Turgut with Selim see-aor-con.cop COMP/quot be.afraid-IPFV-PST.cop

‘Meral was afraid that Turgut would see her with Selim.’ (Göksel & Kerslake 2005: 409)

Unmarked clause

Remark: Used for complements of de ‘to say’, and for complements of predicates of believe and desire.

Functional distribution: Rig: Ref Head ≠ PoS (non-verbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
Herkes [sen sinema-ya git-ti] san-yor
everybody you(nom) cinema-dat go-PST-SG believe-PRES.PROGR

‘Everybody believes that you went to the movies.’ (Kornfilt 1997: 47)

Kayardild
Pred-n(-marri) (‘plain/privative nominalization’)

Remark: Active plain nominalizations may, apart from their dependent uses, function as main clauses describing ongoing, uncompleted actions (Evans 1995: 472).

In Ref Head function, the construction can occur only as the complement of a perception predicate.
In Ref Mod function, it is used when the subject of the dependent clause is relativized. This subject is gapped.

Functional distribution: Flex: Ref Head (perception complements), Ref Mod (subject clauses), Pred Mod = PoS (non-verbs).
Structural type: 3 (D-ALT)
Verbal categories: No TAM
Nominal categories: Nominal agreement
Argument encoding: All overt arguments take the associating oblique case (A.OBL) or proprietive/locative modal case (MPROP/MLOC).

Examples:
Ref Head:
Ngada kurri-ja [ki-l-san-ji dalwani-n-hi thawal-urrk]
1sg.nom see-act 2pl.poss-mloc dig.up-nmlz-mloc yam-mloc:a.obl

‘I saw you digging up yams.’ (Evans 1995: 472)

Ngada kurri-ja [nica-w-an-ji budii-n-marri]
1sg.nom see-act 3sg.poss-mloc run-nmlz-priv

‘I saw that he was not running.’ (Evans 1995: 476)
Ref Mod:
*Nga-ku-l-da [wurr-n-ku] dangka-cu kurri-ju*
1-INC-PL-NOM dance-NMLZ-MPROP man-MPROP see-POT
‘We will watch the dancing man.’ (Evans 1995: 474)

Pred Mod:
*[Bilaangka-nurrk kari-i-n-da] ngada warra-j*
blanket-ASSOC cover-M-NMLZ-NOM 1SG.NOM go-ACT
‘I went along, covering myself in a blanket.’ (Evans 1995: 474)

*Deja-ja karun-ki*[kenau-n-marri]
eat-ACT food-M.LOC tell-NMLZ-PRIV
‘(He) eats food without telling (anyone).’ (Evans 1995: 475)

**Pred-Thirri-n (resultative nominalization)**

Remark: This construction can also be used as a main clause, and as an adverbial clause expressing temporal sequence.

Functional distribution: Rig: Ref Mod = PoS (adjectives)
Structural type: 3 (D-ALT)
Verbal categories: No TAM
Nominal categories: Nominal agreement
Argument encoding: Ergative alignment: Objects of transitive verbs and subjects of intransitives (S and P) take nominative case; (demoted) subjects of transitive verbs (A) take oblique case.

**Examples:**
Ref Mod:
*Bath-in-ki bal-umban-ji*[niwan-jarrng-niaba-ya
west-from-MLOC west-orig-MLOC 3SG.Poss-DU-ABL-MLOC
jibarna-jarrng-kinaba-ya bidiru-thirri-n-ji*
uncle.in.law-du-ABL-MLOC miss-RES-NMLZ-MLOC
‘One coming from the west, that had been missed by his two uncles-in-law.’ (Evans 1995: 480)

*Ningka kada buru-tharra*[wangi-jirri-n-jina] mala-na
2SG.NOM again get-PST steal-RES-NMLZ-ABL beer-M.ABL
‘Did you get some stolen beer again?’ (Evans 1995: 479)

**Pred-n-ngarrba (consequential nominalization)**

Remark: The construction is rarely used as a main clause describing actions preceding the temporal reference point. (Evans 1995: 481) All arguments are marked with a consequential suffix. This kind of marking resembles so-called ‘complete concord’ (all elements of a constituent are marked for case; see Dench 2006). This means that the consequential form looks like a case-marker with the function of a complementizer. Therefore, the construction is classified as a D-SENT, rather than a D-ALT construction.

Functional distribution: Rig: Ref Mod = PoS (adjectives)
Structural type: 2 (D-SENT)
Verbal categories: No TAM
Nominal categories: Resultative case (see above)
Argument encoding: Resultative case (see above)

**Example:**
Ref Mod:
*Nyinka kamburi-ja dathin-a dangka-a*[yarbu-ngarrba balangkali-ngarrba ba-yii-n-ngarrbi]
2SG.NOM speak-IMP that-NOM man-NOM snake-CONS brown.snake-CONS bite-M-NMLZ-CONS
‘You speak to that man who was bitten by a brown snake!’ (Evans 1995: 481)

Clause + -ntha (oblique complementizer case)

Remarks: This construction closely resembles normal finite clauses, permitting almost the full range of verb inflections. (…) Commonly a Complementizing Oblique or Locative case appears after all other inflections, usually on all constituents.” (Evans 1995: 488)

In Ref Head function, this construction is used for complements of predicates of perception,

1 There are no examples available of the privative nominalization in Ref Mod function, but according to Evans (p.c.) this is possible: “I’m sure you can say it, but it’s a gap – probably accidental – in my data.”

2 The tense-system deviates slightly from independent clauses:
   Independent clause: - ACTual, which covers present, past and immediate future, the latter two of which can be marked if extra precision is desired.
   - POTential
   Dependent clause: - PAST
   - IMMED (= present and immediate past
   - POTential
   - ACT can not be expressed

Modal case marking is the same in independent and dependent clauses.
knowledge and speech. The morpheme -(u)rrka is a special portmanteau for LOC + C.OBL. It is used for locative complement clauses (Evans, p.c.). In Ref Mod function, this construction is used only when the relativized item is not the subject of both clauses.

Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (non-verbs, adjectives)
Structural type: 1 (Balanced)
Verbal categories: Largely retained (see above)
Nominal categories: Complementizer case (see above)
Argument encoding: Complementizer case (see above)

Examples:
Ref Head:
Ngada mungurru [ngijjewa] kada-ntba thaa-thaa-ntb
1sg.nom know.nom again-c.obj return-pot-c.obj
'I know that I will come back again.' (Evans 1995: 490/491)

Ngada kamburri-ja niwan-ji [wathu-ntba dathin-inja barji-nyarra-ntb]
1sg.nom say-act 3sg-mloc raft-c.obj that-c.obj capsize-appr-c.obj
'I told him the raft would capsize.' (Evans: 516)

Ngada marin-marri-i-jarr [dathin-kurrka thungalurrka kamburri-jurrk]
1sg.nom self-hear-M-pst that-LOC-c.obj thing-LOC-c.obj speak-IMMED:c.obj
'I heard myself speaking on that thing (the radio).' (Evans 1995: 491)

Ref Mod:
nyinka kurri-jarra dathin-kina dangka-na
2sg.nom see-pst that.mabl man-mabl
[thawurr-inaa-ntba raa-jarra-ntha niwan-jinaa-ntb]
throat-mabl-c.obj spear-pst-c.obj 3sg-mabl-c.obj
'Did you see the man whom (he) speared in the throat?' (Evans 1995: 490)

Unmarked clause
Remark: This construction is used in Ref Mod function, in cases where no complementizer case appears, i.e. when the relativized element is the subject of the relative clause. Usually, the relativized element is gapped, but it may also be retained.

Functional distribution: Rig: Ref Mod = PoS (adjectives)
Structural type: 1 (Balanced)
Verbal categories: Largely retained (see above)
Nominal categories: None
Argument encoding: SENT - SENT/Ø - SENT (gapping)

Example:
Ref Mod:
jina-a dathin-a dangk-a, [dan-kina yii-jarrma-tharra wangal-kina]
where-nom that-nom man-nom, here-mabl put-caus-pst boomerang-mabl
'Where is the man, who left the boomerang here?' (Evans 1995: 489)

Paiwan
tu(a) / tjai + clause
Remark: The oblique marker tu(a) / tjai is also used for non-clausal arguments. Semantically, it is used for patients, beneficiaries, instruments, goals, objects of comparison etc. With DCs it marks "less integrated" complements.

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

Examples:
Ref Head:
mavhekatu ti-naju [tu laqlag-en ni-a-maja]
fear foc-he ob! tickle-pat defoc-ag-pl-he
'He is afraid that she will tickle him.' (Egli 1990: 177)

na na'e'uru a rhavats [tu riyu ti Yohan],
PFV believe lex very ob! prophet foc John
'He believed strongly that John was a prophet.' (Egli 1990: 192)
patelip [tua ini pu'valav-an ni-maju tua qatjuvi]  
regret obl not marry-loc defoc-ag-he obl snake  
'She regretted that she hadn't married the snake.' (Egli 1990: 202)

**a + clause**

Remark: The focus marker/linking element a is also used for non-clausal constituents. This construction is used for subject clauses, and for 'more integrated'object complements, such as with modal predicates.

Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (nouns, small/derived adjectives)

Structural type: 1 (Balanced)

Verbal categories: Retained

Nominal categories: None

Argument encoding: SENT/Ø - SENT/Ø (gapping)

**Examples:**

Ref Head:

nānguaq [a ma-nejtez sun]  
good foc pass-come you  
'It is good that you have come.' (Egli 1990: 230)

ini'ka masqati [a tja-parb-patson-en]  
not can lk we-one.another-see-pat  
'We cannot see one another.' (Egli 1990: 230)

Ref Mod:

qala [a na tem-ker tua vaua]  
stranger lk psv drink-ag obl wine  
'the stranger, who has drunk wine' (Egli 1990: 178)

[α zu' a i-vetiik a] kai  
foc those lk psv-write lk word  
'the word that I have written' (Egli 1990: 271)

[pin-atsay an nua culau a] curay-an  
die-psv loc defoc-ag husband lk woman  
'the woman whose husband has died' (Egli 1990: 183)

**a parhu + clause**

Functional distribution: Rig: Pred Mod ≠ PoS (no manner adverbs)

Structural type: 1 (Balanced)

Verbal categories: Retained

Nominal categories: None

Argument encoding: SENT/Ø - SENT/Ø (gapping)

**Example:**

Pred Mod

sa ringul-i sun [a parhu qemlee]  
and be.around-pron.pat you lk like put.a.cover.on  
'And she will be around you as if she would want to cover you.' (Egli 1990: 209)

((-)in-Pred -an + a**

Remark: This is a perfective participle construction, formed with the perfective affix (-)in (infix for verbs with a consonant in the Anlaut, prefix for verbs with a vowel in the Anlaut), which is also used with independent verb forms, and the participial suffix -an. Since there is no inflectional verbal morphology, it is hard to say whether this is a deranked construction. However, participles are restricted to the extent that they cannot occur with other tenses (marked with particles), nor with focus markers and they cannot be transitive. The participial form must be combined with the linker a.

Functional distribution: Rig: Ref Mod = PoS (small/derived adjectives)

Structural type: 1 (Balanced)

Verbal categories: Restricted tense

Nominal categories: None

Argument encoding: Ø (gapping, intransitive)

9 Paiwan also has a present participle form, with -an and without the perfective marker, but this form is hardly ever used. Like the perfective participle, it can be used as a modifier in a referential phrase. In addition, it can be used for manner expressions, i.e. as a modifier in a predicate phrase:

rhenakej-an a masengsen  
to.do.all.day.long-tpc lk work  
'to work steadily.' (Egli 1990: 124)

401 | Parts of Speech and Dependent Clauses
Appendix iii: Dependent Clause Constructions Key Examples

**Example:**
Ref Mod: 
$kā-in-a'ũy-a$-$an$ $a$ impits
$ʔ-yyv$-point(pass)-ptc $lk$ pencil
‘the pointed pencil’ (Egli 1990: 122)

**Imbabura Quechua**

**Pred-$j/-shka/-na**

Functional distribution: Flex: Ref Head (different-subject), Ref Mod = PoS (nominals)
Structural type: 2 (D-SENT)
Verbal categories: The different forms indicate different relative tense values: $-j$ for present, $-shka$ for past, $-na$ for future. Progressive aspect is retained. No subject agreement.
Nominal categories: CASE in Ref Head function; in Ref Mod function only when the relative clause is extraposed.
Argument encoding: SENT/$Ø$ - SENT/$Ø$ (gapping in Ref Mod function) The object can remain without accusative case (noun-stripping).

**Examples:**

**Ref Head:**

Marya nin-$n$ $[Juζi jatun wasi-ta chari-$j$]-$ta$
María say-3 José big house-acc have-nmlz:prs-acc
‘Maria says that José has a big house.’ (Cole 1982: 14)

$tuka-ka$ $[Juan kay-pi ka-shka]-tu$ $yu-ni$
I-top Juan this-in be-nmlz:ptc-acc think-I
‘I think that Juan was here.’ (Cole 1982: 33)

$Juζi-ka$ $[tuka kaya llama-tar andi-$na$]-$ta$ $kri-n$
José-top I tomorrow sheep-acc buy-nmlz:fut-acc believe-3
‘José believes that I will buy a sheep tomorrow.’ (Cole 1982: 37)

With noun-stripping:

$Juζi-ka$ $[tuka kaya llama-$Ø$ randi-$na$]-$ta$ $kri-n$
José-top I tomorrow sheep buy-nmlz:fut-acc believe-3
‘José believes that I will buy a sheep tomorrow.’ (Cole 1982: 37)

**Ref Mod:**

$[Marya riku-$f$]$ runa
María see-ptc:prs man
‘the man whom Maria sees’ (Cole 1982: 47)

$[Marya riku-shka]$ runa
María see-ptc:prs man
‘the man whom Maria saw’ (Cole 1982: 47)

$[Juζi kulkı-ta kara-$nu$]$ warmi
José silver-acc give-ptc:fut woman
‘the woman to whom José gave money’ (Cole 1982: 54)

**Extraposed:**

$Kuζiša-ta jiya-$ni$ $[Juan-wan tushu-shka ka-shka]-ta$
girl-acc love-1 Juan-with dance-ptc:pst be-ptc:pst-acc
‘I love the girl who had danced with Juan.’ (Cole 1982: 51)

**Pred-$ngapaj$ (subjective)**

Remark: Subjunctive forms are used for the complements of manipulative and desiderative predicates; $-ngapaj$ is used for same subject.

Functional distribution: Rig: Ref Head, same subject ≠ PoS (nominals)
Structural type: 2 (D-SENT)
Verbal categories: No tense and subject agreement. Aspect can be retained.
Nominal categories: None
Argument encoding: $Ø$ - SENT (coreferentiality) The object can remain without accusative case (noun-stripping).

**Example:**

Ref Head, same subject:

$muna-y-man$ $[tuka mama-ta riku-ngapaj]$
want-1-cond my mother-acc see-sbjv
‘I want to see my mother.’ (Cole 1982: 37)
**Pred-chun** *(subjunctive)*

Remark: Subjunctive forms are used for the complements of manipulative and desiderative predicates; -chun is used for different subject.

Functional distribution: Rig: Ref Head, different subject ≠ PoS (nominals)

Structural type: 2 (D-SENT)

Verbal categories: No tense and subject agreement. Aspect can be retained.

Nominal categories: None

Argument encoding: SENT - SENT

**Example:**

Ref Head:

\[
\text{muna-ni } [\text{Juzi pay-paj mama-ta } \text{rika-chun}]
\]

want-1 José he-poss mother-acc see-sbjv

'I want José to see his mother.' (Cole 1982: 37)

**Pred-y**

Remark: In Pred Mod function, the form is reduplicated.

Functional distribution: Flex: Ref Head, Pred Mod ≠ PoS (nominals, small manner adverbs)

Structural type: 2 (D-SENT)

Verbal categories: No tense/aspect/subject agreement

Nominal categories: CASE (in Ref Head function)

Argument encoding: Ø - SENT (coreferentiality) The object can remain without accusative case (noun-stripping).

**Examples:**

Ref Head:

\[
\text{Juzi-ka } [\text{llama-ta/Ø randi-y]-ta usha-n}
\]

José-top sheep-acc/-Ø buy-inf-acc can-3

'José is able to buy sheep.' (Cole 1982: 40)

\[
\text{muna-ka } [\text{shuj ali wagra-ta-mi randi-y]-ta muna-ni}
\]

I-top one good cow-acc-val buy-inf-acc want-1sg

'I want to buy a good cow.' (Cole 1982: 40)

**Pred-shpa**

Functional distribution: Rig: Pred Mod = PoS: small manner adverbs

Structural type: 2 (D-SENT)

Verbal categories: No tense/aspect/subject agreement

Nominal categories: None

Argument encoding: Ø - SENT (coreferentiality)

**Example:**

Pred Mod:

\[
[\text{Kanda-y kanda-y} ] \text{shamu-rka-ni}
\]

sing-inf sing-inf come-pst-1

'I came singing.' (Cole 1982: 62)

**Ma’di**

Pred-\text{i}́

Remarks: Apart from the subordinating suffixes \text{i}, \text{k}, \text{r}, \text{b}, and \text{dʒɔ́}, Ma’di dependent predicates can only take a low-tone prefix, which in independent clauses expresses non-past tense. It is not clear whether the prefix on dependent predicates is the same, since it is compatible with any tense interpretation (Blackings & Fabb 2003: 192).

In Ref Head function this construction is used for complements of desiderative predicates. The subject is unexpressed under co-referentiality; the object is SENT.

In Ref Mod function this construction is used for object relative clauses. The object is either gapped, or expressed with the postposition \text{nā}, meaning ‘aforementioned’ (AFR), and interpreted as the possession of the modified noun. The subject is either left unexpressed, or expressed with the possessive postposition.

Functional distribution: Flex: Ref Head (desiderative), Ref Mod (object) = PoS (nominals)

Structural type: 2/3 (D-SENT/D-ALT)

Verbal categories: No tense

Nominal categories: None

Argument encoding: Ø - SENT/POSS - Ø/POSS - OBL

404 | Parts of Speech and Dependent Clauses
Appendix iii: Dependent Clause Constructions Key Examples

Examples:
Ref Head:
Má  lɛ̀ ā[èɓí´ɲ â] lá
1sg (n)want-obj fish n-cat-nmlz aff
'I certainly want to eat fish.' (Blackings & Fabb 2003: 202)

Ref Mod:
àràb [ɔ́pɨ́ɗà Ɂà dƷɨ̄-lɛ́] rɨ lɪff
The car which Opi took certainly has a flat tyre.' (Blackings & Fabb 2003: 22)

ílí [ŋɔ̀-lɛ́] rɨ lɔ́tʃɨ̄
The knife which was broken is sharp.' (Blackings & Fabb 2003: 201)

ág [tín āɓ ā râ nà] Ɂà āgù rā
The man whose cow that child stole…’ (Blackings & Fabb 2003: 201)

Pred-dʒɔ̀
Remarks: In Ref Head function this construction is used for the complements of phrasal predicates.
The subject remains unexpressed under co-referentiality. In Ref Mod function this construction is used for relative clauses, in which the relativized element is a source. This relativized argument is gapped; the subject is either unexpressed or possessive, and the object is SENT.

Functional distribution: Flex: Ref Head (phasals), Ref Mod (source clauses) = PoS (nominals)
Structural type: 2 / 3 (D-SENT / D-AL T)
Verbal categories: No tense
Nominal categories: None
Argument encoding: Ø - SENT / POSS - SENT

Examples:
Ref Head:
ɔ́ pɨ́ ɛ̄ɗɔ́ sī dƷɔ́ râ
Opi has certainly started to build (with) it.' (Blackings & Fabb 2003: 22 / 207)

Ref Mod:
bélè ág rɨ̀ pɨ̄ Ɂmgbā dƷɔ́ rɨ̀ dɨ̀ Ɂɨ̄
The stick with which the man and his associates was beaten is this one.' (Blackings & Fabb 2003: 22)

Pred-hₕₐ
Functional distribution: Rig: Ref Head ≠ PoS (nominals)
Structural type: 3 (D-ALT)
Verbal categories: No tense
Nominal categories: None
Argument encoding: POSS - SENT / Ø - SENT

Examples:
Ref Mod:
má  bũ mₕₐ-kā kₐrₕu
1sg try [(n)-go-nmlz] neg(pst)
'I have not tried/did not try to go/going.' (Blackings & Fabb 2003: 22)

Má ndrē ârângwâ rɨ̀ ɗ-dē-kā rā
1sg see [bird def poss (n)-ve-fall-nmlz] aff
'I saw the bird's falling/fall.' (Blackings & Fabb 2003: 21)

Î-łb tîhā sê-kā dîn kî
ind-want [cigarette (n)-smoke-nmlz] here neg
'Smoking is not permitted here.' (Blackings & Fabb 2003: 213)

Pred-r̥⁄ã́dₕₐ
Remark: r̥ is the plural equivalent of r Only in non-active cases can an overt subject appear, which is then marked with a postposition (see second example).

Functional distribution: Rig: Ref Mod (subject/possessive clauses) ≠ PoS (nominals)
Structural type: 2 (D-SENT)
Verbal categories: No tense

Appendix iii: Dependent Clause Constructions Key Examples | 405
Nominal categories: None
Argument encoding: Ø - SENT

Examples:
Ref Mod:
\( \text{ág } \text{ɔ́́ } \text{àm-à } \text{dƷɔ́̀ } \text{ǹ} \text{ɨ̄ } \text{́s} \text{ī-ɓ} \text{á} \text{r} \text{ɨ̀} \text{ } \text{man } \left[ \text{1PL-poss } \text{house } \text{SPEC } \text{PRON } \text{n-build-PTC(PL)} \right] \text{DEF} \)

‘one of the men who built/are building our house’ (Blackings & Fabb 2003: 193)

\( \text{ɗɨ } \text{ág } \text{ɔ́ ti } \text{ǹ } \text{égwè-r} \text{ɗ } \text{r} \text{ɨ̀} \text{ } \text{this } \text{man } \left[ \text{cow } \text{AFR } (\text{S})-\text{get.lost-PTC COM} \right] \text{DEF FOCS} \)

‘This is the man whose cow got lost.’ (Blackings & Fabb 2003: 195)

Unmarked clause

Remark: Used for complements of utterance, propositional attitude, knowledge, manipulatives and desideratives.

Functional distribution: RIG: Ref Head ≠ PoS (nominals)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
\( \text{ɲɨ́ } \text{ʄō } \text{k-ē } \text{mú } \text{ɔ̀ɓʊ́} \text{2sg } \text{say } \left[ \text{dir-ve-go } \text{tomorrow} \right] \)

‘You said that she should come tomorrow.’/‘You told her to come tomorrow.’ (Blackings & Fabb 2003: 21)

Clause with \( z \text{̄ } + s \text{ɨ̀} \)

Remark: This construction is marked by the grammatical verb \( z \text{̄} \) and the source propositiona \( s \text{ɨ̀} \). It expresses simultaneous action.

The subject is co-referential but overtly expressed (SENT). (Blackings & Fabb 2003: 421).

Functional distribution: RIG: Pred Mod = PoS (small and derived manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: CASE (postposition)
Argument encoding: SENT - SENT

Example:
Pred Mod:
\( \text{ká } \text{sɨ́ } \text{tà } \text{ɲ } \text{ā } \text{míndr } \text{ā } \text{n } \text{ā } \text{kó } \text{́-z } \text{ɨ̄ } \text{r } \text{ā } \text{s } \text{ɨ̀} \text{ } \text{3 } \text{pepper } (\text{n})\text{eat } \left[ \text{tears } \text{AFR } 3-N-SIM \text{ leak} \right] \text{SR} \)

‘She was eating pepper as her eyes were (continuing) running.’ (Blackings & Fabb 2003: 440)

Hungarian

Pred-\( ni \)

Remarks: When this construction has subject function, agreement (in person and number) with the subject is optionally retained. In object function, subject agreement is always lost. Only one argument can be overtly expressed. If the subject is overt, it takes DAT; if the object is overt, it is SENT.

Functional distribution: RIG: Ref Head ≠ PoS (nominals)
Structural type: 3 (D-ALT)
Verbal categories: No tense/mood. Causative and frequentative affixes can be expressed. Agreement is sometimes retained (see above).

Nominal categories: None
Argument encoding: Ø - SENT/DAT - Ø

Examples:
Ref Head:
\( \text{Fontos } \text{volt } \left[ \text{Péter-nek } \text{olvas-ní-f-ul} \right] \)

‘It was important for Peter to read.’ (Kenesei et al. 1998: 35)

\( \text{Anna } \text{most } \text{akar } \left[ \text{olvas-ní} \right] \)

‘Anna wants to read now.’ (Kenesei et al. 1998: 33)

\( \text{Anna } \text{meg-próbál-t-a } \left[ \text{meg-tanul-ní } \text{a } \text{verset} \right] \)

‘Anna tried to learn the poem.’ (Kenesei et al. 1998: 33)
Appendix iii: Dependent Clause Constructions Key Examples

**Pred-áss / -és**
Remark: The subject of a transitive nominalized verb is optionally expressed as an attributive adjectival phrase involving the adjectival általi form of the postposition által ‘by’, or involving the postposition által and the active participial form való of van ‘to be’. The subject of an intransitive predicate and the object of a transitive predicate trigger nominal agreement on the dependent predicate.

Functional distribution: Rig: Ref Head ≠ PoS (nominals)
Structural type: 3 (D-ALT)
Verbal categories: No tense/aspect/mood/agreement
Nominal categories: CASE, nominal agreement
Argument encoding: POSS/OBL – POSS

**Examples:**
Ref Head:
János [a kincs el-rejti-és-és-javasol-t-a] ‘John the treasure pref-hide-nmlz-poss.3sg-acc suggest-pst-def.3sg
‘John suggested to hide the treasure.’ (Kenesei et al. 1998: 207)

Örülök [Pál váratlan meg-érkezésének] ‘I am happy about Paul’s unexpected arrival.’ (Kenesei et al. 1998: 207)

**Pred-ó** (**active present participle**):

Functional distribution: Rig: Ref Mod (subject clauses) ≠ PoS (nominals)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect/mood/agreement
Nominal categories: None
Argument encoding: Ø - SENT (gapping)

**Example:**
Ref Mod:
Az [Anna által tegnap olvas-ott könyv] ‘the book pref-read-pst.ptc.3sg that was read by Anna yesterday.’ (Kenesei et al. 1998: 45)

**Pred-ótt** (**passive, past participle**)
Remark: Used in prenominal relative clauses where the relativized item is the undergoer/patient of the DC. This element is gapped; the agent is marked OBL with the agentive postposition által ‘by’.

Functional distribution: Rig: Ref Mod (patient clauses) ≠ PoS (nominals)
Structural type: 3 (D-ALT)
Verbal categories: No tense/aspect/mood/agreement
Nominal categories: None
Argument encoding: OBL - Ø

**Example:**
Ref Mod:
Az [Anna által tegnap olvas-ott könyv] ‘the book pref-read-pst.ptc.3sg that was read by Anna yesterday.’ (Kenesei et al. 1998: 45)

**Pred-andó / -endő** (**future participle**)

Functional distribution: Rig: Ref Mod (subject and object clauses) ≠ PoS (nominals)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect/mood/agreement
Nominal categories: None
Argument encoding: Ø-SENT/SENT-Ø/Ø-Ø

**Examples:**
Ref Mod:
az [el-jövendő kor] ‘the age to come’ (Kenesei et al. 1998: 319)

a [ki-javítandó dolgozat-ak] ‘the papers to be corrected’ (Kenesei et al. 1998: 320)

**Pred-év / -vé** (**simple converb**)

Functional distribution: Rig: Pred Mod ≠ PoS ((derived) manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect/mood/agreement
Nominal categories: None
Argument encoding: Ø - SENT

Example:
Pred Mod:
A gyerek-wk [kiabd-va] szalad-t-ak végig az utca-n.
the child-pl shout-conv run-pst-indef.3pl along the street-superess
The children ran down the street shouting.' (Kenesei et al. 1998: 320)

Pred-vén (‘perfective converb’):
Remark: This form is very infrequent in spoken Hungarian; it is used only in formal and ceremonious style in writing. The perfective converb has historically been used to refer to an action preceding that of the finite verb. Nowadays it is sometimes used synonymously with the simple converb.
Functional distribution: Rig: Pred Mod = PoS ((derived) manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect/mood/agreement
Nominal categories: None
Argument encoding: Ø - SENT

Example:
Pred Mod:
Ez-t mond-t-a nek-em [az asztalfő-n ül-vén].
this-acc say-pst-def.3sg dat-1sg the table.head-superess sit-pfv.conv
‘Sitting at the head of the table s/he said this to me.’ (Kenesei et al. 1998: 321)

Clause + hogy
Remarks: The construction can be combined with the expletive pronominal az, which takes case according to the function of the DC. In nominative and accusative function case can be omitted; in other functions it cannot. When functioning as the complement of a manipulative or evaluative predicate, the dependent predicate takes the subjunctive marker -j- and a prefix meg-. In the case of a manipulative predicate the complementizer can be omitted; in the case of an evaluative predicate it cannot.
Functional distribution: Rig: Ref Head ≠ PoS (nominals)
Structural type: 1 (Balanced)
Verbal categories: All retained (SBJV)
Nominal categories: (CASE)
Argument encoding: SENT - SENT

Examples:
Ref Head:
Anna elmondta nekünk (azt), [hogy Péter beteg volt].
Anna told.def to.us it.comp Peter sick was
‘Anna told us that Peter had been sick.’ (Kenesei et al. 1998: 31)

Anna azt mondta, [(hogy) tanul-ja-d, meg a versez]
Anna it.acc said.def comp learn-imp/sbjv-def.2sg pref the poem.acc
‘Anna told you to learn the poem.’ (Kenesei et al. 1998: 32)

Nem szükséges, [*hogy] Péter meg-tanul-j-a a versez
Not necessary comp Peter prefix-learn-sbjv-def the poem-acc
‘It is not necessary for Peter to learn the poem.’ (Kenesei et al. 1998: 32)

Clause úgy + hogy
Remark: úgy is the adverbial form of the pronominal az in the main clause (most often ‘thus’). Optionally a relative pro-adverb is attached to the subordinator hogy.
Functional distribution: Rig: Pred Mod = PoS ((derived) manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Examples:
Pred Mod:
Péter úgy akadt el, [hogy olvastat].
Peter adv/pron slept pref sub read.3sg
‘Peter fell asleep in such a manner that he was reading.’ 2 ‘Peter fell asleep reading.’ (Kenesei et al. 1998: 50)
Appendix III: Dependent Clause Constructions Key Examples

Peter slept in his childhood.

Remarks: The relative pronoun takes case-marking according to the function of the head noun in the dependent clause. Optionally, the construction can be combined with a clause-initial demonstrative along with the lexical head noun.

Functional distribution: Rig: Ref Mod ≠ PoS (nominals)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: Nominal agreement
Argument encoding: SENT - SENT/Ø - SENT/SENT - Ø

Examples:
Ref Mod:

(Az) a könyv, [amely-et Anna olvas-ott] érdekes volt
dem the book which-acc Anna read-pst interesting was

‘The book that Anna was reading was interesting.’ (Kenesei et al. 1998: 38)

Japanese

Clause + no, mono, koto, wake, yoo

Remarks: The ‘nominalizers’ are in fact case-marked nouns meaning ‘thing’. The dependent predicate is tensed, but the subject can optionally appear in the genitive. Therefore this construction has a double classification 1/3.

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)/3 (D-AL T)
Verbal categories: Tense retained
Nominal categories: CASE
Argument encoding: POSS-SENT/SENT-SENT

Example:
Ref Head:

[Ano hito ga/no hon o kai-ta koto ga] yoku sira-re-te iru
that person nom/gen book acc write-past thing(nmlz) nom well know-pass-ger be

‘It is well known that that person wrote a book.’ (the fact that that person wrote a book is well-known./ the fact of that person’s having written a book is well known.’) (Lombardi Vallauri 1997: 497)

Unmarked REL clause

Remark: There is no relative marker, but resumptive pronouns (in the form of demonstrative, personal, or reflexive pronouns) can be optionally used.

Functional distribution: Rig: Ref Mod ≠ PoS (nominals)
Structural type: 1 (Balanced)
Verbal categories: Tense retained
Nominal categories: None
Argument encoding: SENT-SENT/Ø - SENT/SENT-Ø (gapping)

Examples:
Ref Mod:

Dakara, [tamago motte-ru] hito mo ita shi,
so eggs hold-pst person too were and
[juusu mo motte-ru] hito mo ita shi…
juice too hold-pst person too was and

‘So, there were people holding eggs, and people holding juice, and…’ (Hinds 1986: 59)

With resumptive pronoun:

[Sono mae ni kuruma ga tomar-te-iru] mitie
dem front dir car nom stopped shop
‘the shop in front of which a car is parked’ (Hinds 1986: 61)

Predicate Te-/de/-ite

Remark: Overt subjects can take the topic marker in stead of the nominative.

Functional distribution: Rig: Ref Mod = PoS (manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No Tense
Nominal categories: None
Argument encoding: SENT - SENT/Ø - SENT (co-referentiality)

Examples:
Pred Mod (see for more instances the next example):
Son san wa [hoka no nihonjin sutaffu o kun-de] shigoto o shi-te i-ru
Song Mr. top other attr Japanese staff obj unite-conv job obj do-conv be-prs
‘Mr Song is working together (in a united manner) with other Japanese staff.’
(Alpatov & Podlesskaya 1995: 469)

Different subject:
Yasuko wa jaugoroku no koru ibiki no kuse ga
Yasuko top fifteen:sixteen attr time snore attr habit subj
at-te, oya wa kyoei ni kushin shi-ta-soo da
be-conv parents top correction tobj efforts do-psr-likely cop-prs
‘They say that Yasuko snored when [she was] fifteen or sixteen [and her] parents did their best to get rid
[of this habit].’ (Alpatov & Podlesskaya 1995: 469)

Pred-i-/Ø (‘infinitive’)
Remark: Overt subjects can take the topic marker in stead of the nominative.
Functional distribution: Rig: Ref Mod = Pos (manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No Tense
Nominal categories: None
Argument encoding: SENT - SENT/Ø - SENT (co-referentiality)

Pred Mod:
Same subject:
Ogata Shingo wa [sukoshi mayu o yose-Ø] [sukoshi kuchi o
Ogatta Shingo top slightly eyebrow obj moved.together-inf slightly mouth obj
ake-te] [nanila kange-te] i-ru fuu datta
open-conv something think-conv be-prs look cop.pst
‘Ogata Shingo looked as if he was thinking (about) something, bringing his eyebrows slightly together
and slightly opening his mouth.’ (Alpatov & Podlesskaya 1995: 468)

Different subject:
Shingo wa kao o shikame-Ø, Shuuichi wa yoi ga same-ta daroo
Shingo top face d$obj frown-inf Shuuichi top drunkenness subj abate-pst tent
‘Shingo frowned, [and] it seemed that Shuuichi got sober.’ (Alpatov & Podlesskaya 1995: 468)

Hmong Njua
qhoe + clause
Remarks: This construction can be a complement clause or a nominalization, depending on the scope of
its structural coding, as shown by the position of the marker, which may either precede the dependent
predicate, or the whole DC. Since there is no difference in terms of verbal/nominal categories and/or
argument encoding, this construction is classified as a balanced one only.
This construction is used for subject clauses.
Functional distribution: Rig: Ref Head = Pos (nouns)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
Qhoe cov zej zog tiw tacom wa koro zok sok
comp cl neighbour move out do 1sg happy
‘It makes me happy that the neighbour moves out.’ (Harriehausen 1990: 200)

kuam/(has)tas + clause
Remarks: The choice between the two forms depends on the type of matrix predicate. Kuam is used with
complements of desideratives and manipulatives; (has)tas with complements of perception, knowledge,
propositional attitude, and utterance. Both forms are used for object complements only.
Functional distribution: Rig: Ref Head = Pos (nouns)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT - SENT
Appendix iii: Dependent Clause Constructions Key Examples

Examples:

Ref Head:
Peter xaav [kuam Paul yuav lab fais]
Peter want comp Paul buy cl car
‘Peter wants Paul to buy a car.’ (Harriehausen 1990: 220)

Kuv paub [(has)tas nwg yuav npta]
1sg know comp 3sg buy flowers
I know that he has bought flowers.’ (Harriehausen 1990: 22)

Unmarked clause
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
Kuv xaav [kuw moog tsev sai]
1sg hope 1sg go house soon
‘I hope to go home soon.’ (Harriehausen 1990: 221)

kws (REL) + clause:
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT - Ø/Ø - SENT (gapping)

Example:
Ref Mod:
Tag tsev neej [kws kuw saib sab]
clo man rel 1sg see big
‘The man that I saw was tall.’ (Harriehausen 1990:141)

Lango
Pred-(kk)3 (‘infinitive/nominalization’)
Remarks: Overt subjects are marked as attributive modifiers.
The construction is used for complements of phasal, modal, desiderative, and achievement predicates.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 3 (D-ALT)
Verbal categories: No aspect; no subject agreement. Object agreement, voice and (benefactive/ventive) valency retained.
Nominal categories: None
Argument encoding: Ø - SENT/POSS - SENT

Example:
Ref Head:
ámɪ̀ttɔ [gwɛ̀ɛ̀ yòò dìlò]
1sg want progr kick.inf ball
‘I want to kick the ball.’ (Noonan 1992: 213)

àpwọ̀ɲ ńtọ̀wàd à cèck [kè́er à kwàm]
teacher 3sg.tell.REFV.1PL ATTR.PRT little.bit good.NMLZ ATTR.PRT read-INF
‘The teacher told us briefly about the benefits of reading.’ (The teacher told us briefly the being good of reading.) (Noonan 1992: 213)

[Ni-khɔ̀g wɔ̀k à dàkɔ̀] ràc
kill-NMLZ dog ATTR.PRT woman 3sg-bad-HAB
‘The killing of a dog by a woman is bad.’ (Noonan 1992: 214)

ni + clause
Remark: The dependent predicate is indicative when there is independent time reference and subjunctive when there is dependent time reference.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained (see remark)
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Ref Head:
\[ \textit{Agbó} \quad [nì \quad lócò \quad dágó \quad tìc] \]
1sg-know-HAB comp man 3sg-hate-HAB work
'I know that the man hates work.' (Noonan 1992: 191)

\[ \textit{ámítò} \quad [nì \quad dákô \quad ìbùñì \quad àtìn \quad bìù] \]
1sg-want-PROGR comp woman 3sg-buy-REN-SBJV child book
'I want the woman to buy the child a book.' (Noonan 1992: 191)

\textit{ámè} \* clause
Remark: The marker \textit{ámè} is a combination of the attributive particle \textit{à} and the relative particle \textit{mè}. The relativized item is gapped. A resumptive pronoun must be used when the relativized item is a benefactive, associative, or object of preposition.

Functional distribution: Règ: Ref Mod ≠ PoS (modifiers)
Structural type: 1 (Balanced)
Verbal categories: Retained (see remark)
Nominal categories: None
Argument encoding: SENT - SENT/SENT - Ø/Ø - SENT (gapping)

Examples:
Ref Mod:
\[ lócò \quad [\textit{ámè} \quad måró \quad qùì] \]
man REL-ATTR.PRT 3sg-like-HAB dog
'The man that likes the dog.' (Noonan 1992: 215)

\[ àtìnn \quad [\textit{ámè} \quad dákô \quad ìbùñì \quad bìù] \]
child REL-ATTR.PRT woman 3sg-buy-REN-PFV-3sg book
'The child for whom the woman bought a book.' (Noonan 1992: 215)

\textbf{KET}

Bare infinitive
Remarks: The bare form is used in Ref Head function for the complements of phasal and modal (ability) predicates. The object, if expressed, is incorporated. For the complement of desiderative or modal (necessity) predicates, the infinitive is marked for translative case. The bare infinitive is also used in Ref Mod function, but this may involve lexical derivation, since the infinitive apparently cannot take any arguments.

Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (nouns, modifiers, small/derived adjectives)
Structural type: 2 (D-ALT)
Verbal categories: None (i.e. no tense/mood, no subject-object agreement)
Nominal categories: (CASE)
Argument encoding: Ø - INC/POSS - INC

Examples:
Ref Head:
\[ [\textit{àk-nà} \quad qùì-bèt \quad qùì] \]
pl-anim.pl.gen tent-making,INF it.ended
'We finished making the tent.' (Lit: 'Our tent-making ended.') (Vajda 2004: 78).

\[ \textit{át} \quad \textit{ìpë} \quad \textit{ìtìpàram} \quad 1sg \quad sing,INF \quad \textit{I know} \quad 'I know how to sing/I can sing.' (Vajda 2004: 78) \]

\[ [\textit{Ab-àŋ} \quad \textit{dàsànũ̩-nàat}] \quad måró \quad 1sg-DAT \quad hunt,INF-TRL need \quad 'I need to hunt' (Vajda 2004: 77) \]

Ref Mod:
\[ bây \quad ìdègò \quad \textit{find} \quad \textit{boot} \quad 'A boot that is found' (Vajda 2004: 79) \]

Unmarked clause
Remark: This construction is apparently infrequent. Werner (1997: 355): "It happens in Ket that a finite verb stands before a noun in attributive position." (emphasis added, EvL)

Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (nouns, modifiers, small/derived adjectives)
Appendix iii: Dependent Clause Constructions Key Examples

Structural type:   1 (Balanced)
Verbal categories:   All retained
Nominal categories:  (CASE)
Argument encoding:  SENT - SENT

Examples:
Ref Head:

dótám-bám ɪyōcide [b̩əniin ḵûg̱iːnaŋtoŋəŋ]
dotam-old.woman she.heard brother.PL orphans.they.became
‘Old Dotam Woman (a forest witch) heard that the brothers had become orphans.’ (Vajda 2004: 93)

Ref Mod:

tuɁɛ; [ital’em] keŋ
dem he.has.knowledge person
‘This knowledgeable person’ (lit: this person who has knowledge) (Werner 1997: 355)

Clause + ɛta ɣw’ə
Functional distribution: Rig: Pred Mod = PoS (small/derived adjectives) ≠ PoS (modifiers)
Structural type:   1 (Balanced)
Verbal categories:   All retained
Nominal categories:  None
Argument encoding:  SENT- SENT

Example:
Pred Mod
TajɁəm, [ɛta ɣw’ə ke’eta]
It has become cold, as if it cuts
‘It has become piercingly cold.’ (Werner 1997: 348)

Clause + ɑsqà
Functional distribution: Rig: Pred Mod = PoS (small/derived adjectives) ≠ PoS (modifiers)
Structural type:   1 (Balanced)
Verbal categories:   All retained
Nominal categories:  None
Argument encoding:  SENT- SENT

Example:
Pred Mod
bū  tøn ðiɁləq bɪdə [deɁp dölən ɑsqà].
3masc so he.lived all people they.lived like
‘He lived like everyone (else) lived.’ (Vajda 2004: 87)

Clause + PROS(ecutive) CASE
Functional distribution: Rig: Pred Mod = PoS (small/derived adjectives) ≠ PoS (modifiers)
Structural type:   1 (Balanced)
Verbal categories:   All retained
Nominal categories:  None
Argument encoding:  SENT - SENT

Example:
Pred Mod:
ū-k hɛɁp [dùrèn-bes] tiːsə
2sg.-gen son he.cries-PROS he.sits
‘Your son sits (there) crying.’ (Vajda 2004: 87/Werner 1997: 354)

REL PRON + clause
Functional distribution: Rig: Ref Mod = PoS (small/derived adjectives)
Structural type:   1 (Balanced)
Verbal categories:   All retained
Nominal categories:  Class/number agreement
Argument encoding:  SENT - SENT

Examples:
Ref Mod:
āt  qım dɪəɾən [ɡə-Ɂə ʃʊə dʊɾəɾəŋ]
1sg woman I.see.her who-FEM there she.lives
‘I am looking at the woman who lives there.’ (Vajda 2004: 30)
All the people who lived here have died. (Vadja 2004: 30)

Pred-s/-bes clause

Remark: The -s suffix, and its allomorph -bes (used when the object of the dependent clause stands between the dependent predicate and the head noun) are called 'nominalizers' (and glossed as such), but they appear on finite predicates, and arguments remain SENT.

Examples:
Ref Mod:
[mámúl dóbá-s] dil
milk he.drink.it-NMLZ child
'a child who drinks milk' (Vajda 2004: 79)

[dóbá-bes mámúl] dil
he.drink.it-NMLZ milk child
'a child who drinks milk' (Vajda 2004: 79)

Itelem

Infinitives (various forms)

Examples:
Pred-s (INF I) (Used for complements of modal, phasal and manipulative predicates.)

Ref Head:
Utu-z-en [agga bolse fowale-s.]
can.not-PRS-3.SG something more-say-INF.I
'He cannot say anything anymore.' (Georg & Volodin 1999: 168)

Komma t-iti-çen [p'z-jh no-ka-s.]
1.SG 1.SG-force-3.SG.PAT child-DIM eat-CCM-INF.I
'I forced the child to eat.' (Georg & Volodin 1999: 169)

Pred-kilh-kalh/-kila'n/-kala'n (PL) (INF II) (Used for complements of direct perception predicates.)

Ref Head:
Truk sin 'agueq k-ekua-ýn [t'it'im es-kalh.]
suddenly S. INF.II-see-INF.II smoke come.out-INF.II
'Suddenly S. saw that smoke was coming out.' (Georg & Volodin 1999: 173)

Pred-ki/ka (INF V); (Used for complements of phasal and modal predicates and some manipulative predicates. Applies only to those verbs that take -k in Infinitive I.)

Ref Head:
A sin'agueq-n nita aggar-kit k-azu-knen [aggar-kí.]
INTERJ S.-POS soul was-CAUS INF.II-start-INF.III hurt-INF.IV
For some reason S. became sad. (lit.: S's soul began to hurt.) (Georg & Volodin 1999: 180)

Pred-l (INF VI) (Same function as Infinitive V, but it is used with those verbs that do NOT take k- in Infinitive I.)

Ref Head:
T-utu-r-en [Ememqut met'le-4.]
1.SG-can.not-PRS-3.SG.PAT e. kill-INF.IV
'I cannot kill E.' (Georg & Volodin 1999: 182)

Unmarked clause

Remark: Used for subject clauses and sometimes for object complements of perception predicates.
Appendix iii: Dependent Clause Constructions Key Examples

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT (co-referentiality)

Examples:

Ref Head:
Qetew: [ki-yiçmen qefqef].
go:find-3.sg.3.sg
strength
'It is good that he has found the strength.' (Georg & Volodin 1999: 204)

Utre: n-exsfl-k t-laxl-kiçen [mum-kexsfl-kiçen muza-ʔn].
in.the.morning 1.pl-wake.up-1.pl 1.sg-see-1.sg 1.pl-be.dry-1.pl 1.pl-pl
'In the morning we woke up, and I saw that we were dry.' (Georg & Volodin 1999: 205)

min + gapped clause
Functional distribution: Rig: Ref Mod = PoS (derived adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping)

Example:

Ref Mod:
T-çki-kiçen nu Nwilwej-ŋ [min k-çil-knen kulaka-ʔn].
1.sg-find-1sg dem N. rel inf.red-choose-inf.red Kulak-loc
'I found this Nwilwejngen, whom the Kulaken had chosen.' (Georg & Volodin 1999: 203)

qatz + clause
Functional distribution: Rig: Pred Mod = PoS (derived manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT- SENT

Example:

Pred Mod:
Ewnun fçora celoj ql [qatz k-nig-ʔn] t-laxl-s-ŋ.
thus yesterday whole day as.if inf.red-loaded-inf.red 1.sg-go-prs-1.sg
'Thus I went around all day, heavily loaded.' (Georg & Volodin 1999: 213)

Koasati

'Nominalizations' (various forms)
Remarks: For all verb classes, except one ('class 2A'), the nominalization is formed form the 1st person plural affirmative, without phrase-terminal markers. Other nominalizations are formed by replacing the element -li with the element -k. (Kimball 1991: 273-274). Nominalization cannot take any other verbal morphology than categories related to voice/valency (reciprocal, reflexive, locative and instrumental prefixes).

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No TAM/phrase-terminal marker; only voice/valency can be retained.
Nominal categories: None
Argument encoding: Ø - SENT

Examples:

Ref Head:
[Ici sammici-t inkab-6s] vehayili-
deer do:how-CONN to:shoot&hit-NMLZ know-1sg.sbj
'I know how to shoot deer.' (Kimball 1991: 280)

[Aybacikka awb-ka] ca-baw-ka-V
law ruin-NMLZ 1sg.sbj.stat-want-neg-phtm
'I don't want to break the law.' (Kimball 1991: 275)

Participles (various forms)
Remark: Participles can express subject/object cross-reference, but apparently with special forms. (Kimball 1991: 288): “It seems likely that the participle suffixes are added to an already nominalized verb. This supposition is strengthened by the fact that all the participial suffixes (with the exception of the future participle, which is transparently derived from the present participle) also appear as article suffixes on nouns.”
Parts of Speech and Dependent Clauses

Functional distribution: Rig: Ref Mod = PoS (small adjectives)
Structural type: 2 (D-SENT)
Verbal categories: No TAM, no phrase-terminal marker; only voice/valency can be retained.
Nominal categories: CASE agreement (ACC is zero-marked)
Argument encoding: Ø - SENT/SENT - Ø (gapping)

Pred-

Remark: Subjects of these participles are usually stripped of their case marking. The focus forms are used for subject/object marking.

Ref Mod:

that  person  eat-1HAR-PTC-SS.FOC  be.fat
A person who eats all the time is very fat." (Kimball 1991: 290)

Pred-

Remark: Focus forms are used for subject/object marking.

Ref mod:

Á:ti  [í:pa-li:t]
person  see-1SG.SBJ-IPFV-PTC-TOP  see-1SG.SBJ-PST
I saw the person that I used to see." (Kimball 1991: 292)

Pred-

Remark: This formation is a combination of the present participial suffix -:sáya with the verbal suffixes
-:lab- (IRR), and -:t-, a hearsay suffix with the meaning 'deduction' (Kimball 1991: 292):

Ref Mod:

orange  give.to.me-2SG.SBJ-IRR-PTC-OBJ.FOC  eat-1SG.SBJ-IRR-PST
I intend to eat the orange that you will give me." (Kimball 1991: 292)

Pred-

Remark: – is the different-subject switch-reference (SW) marker. It can be combined with the diminutive/ intensive suffix –o:si/-ö:si.

Functional distribution: Rig: Pred Mod = PoS (small/derived manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No TAM, no phrase-terminal marker
Nominal categories: None
Argument encoding: SENT - SENT/Ø - SENT

Examples:

Pred Mod:

that  person  eat-1HAR-PTC-SS.FOC  be.fat
A person who eats all the time is very fat." (Kimball 1991: 290)

[í:pa-li:t]
person  see-1SG.SBJ-IPFV-PTC-TOP  see-1SG.SBJ-PST
I saw the person that I used to see." (Kimball 1991: 292)

Pred-

Remark: The suffix is a connector. This construction is used when the verb modifier can be considered as

416 | Parts of Speech and Dependent Clauses
an action that takes place at the same time as the action of the matrix verb (= simultaneity clause).
Functional distribution: Rig: Pred Mod = PoS (small/derived manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No TAM, no phrase-terminal marker
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

Example:
Pred Mod: [Fololohkáci-ɭ] cokkíolollí-n hí:ca-li:p im-ca-matááli-Vho-k
be.coiled.up-CONN sit(3SG)-SW see-1SG-IRRIV 3STAT.OBJ-1SG,STAT-be.afraid(3SG)-HAB-SS
‘If I see one sitting coiled up, I am afraid of it.’ (Kimball 1991: 489)

Pred-ɭ
Remark: The suffix is a same-subject marker
This construction is used when the adverbial action can be applied as much to the subject of the sentence as to the verb (i.e. secondary predication).
Functional distribution: Rig: Pred Mod = PoS (small/derived manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No TAM, no phrase-terminal marker
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

Example:
Pred Mod: [Wayókhák-ɭ] ho-pálki-pálámmi-n
fly(PL)-SS DISTR-be.fast-ADV-SW
‘They all fly very fast.’ (Kimball 1991: 490)

Thai
Unmarked clause
Remarks: This construction is used for subject clauses and for object complements of desiderative and achievement predicates. Under special conditions it also occurs in Ref Mod function, namely expressing a subject relative clause that gives a general description of the head noun. In such cases, the relativized item is gapped.
Functional distribution: Flex: Ref Head, Ref Mod PoS (nouns, adjectives, small modifiers)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT - SENT/Ø - SENT (gapping)

Examples:
Ref Head:
[ɔ̀ɔk-kamla ŋ thúk wan] dii tɔ̀ɔ râa-kaay
exercise every day good towards body
‘Doing exercise every day is good for your body.’ (Iwasaki & Ingkaphirom 2005: 253)

Khów yiak [ɛɔ bûu khà]
3 want pierce ear PREP
‘She wants to have her ears pierced.’ (Iwasaki & Ingkaphirom 2005: 231)

Ref Mod:
ɛk pen dûk [rian kêt]
Ek cop child study well
‘Ek is a child who studies well.’ (Iwasaki & Ingkaphirom 2005: 250)

thîi + clause:
Remark: This construction is used in Ref Head function, as the complement of predicates expressing evaluation and emotion, and sometimes (optionally) of desiderative predicates. In Ref Mod function it is used in combination with a gapping strategy or a resumptive pronoun.
Functional distribution: Flex: Ref Head, Ref Mod PoS (nouns, adjectives, small modifiers)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT - SENT/Ø - SENT/SENT - Ø (gapping)
Examples:
Ref Head:
tx-wâa dîi ná [thîi mây mii khay pen alay]
but good PRT COMP NEG have who COP what
‘But it was good that no one was hurt.’ (Iwasaki & Ingkaphirom 2005: 255)

Ref Mod:
Khon [thîi duulee] ni pen pen acaan lâ
person REL take.care PRT COP COP teacher Q
‘Is the person who takes care [of the students] a teacher?’ (Iwasaki & Ingkaphirom 2005: 243)

Khon [thîi khâe pay yau kan taam raprain]
person REL REF3 go stay rcp school
‘people who want to stay at school’ (Iwasaki & Ingkaphirom 2005: 245)

(thîi+) wâa/câ + clause
Remark: In some cases this construction combines with the flexible subordinator thîi, while in other cases this element is optional and only wâa (‘say’) and/or the ‘challengeable marker’ (CM) câ remain.

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: No Tense, mood and agreement
Nominal categories: DET
Argument encoding: SENT - SENT

Example:
Ref Head:
khît [wâa câ hâa qaam tham thîi-nil]
think say/COMP CM look.for work do here
‘I think that I will look for work here.’ (Iwasaki & Preeya Ingkaphirom 2005: 262)

Basque
Pred-t(z)e
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No Tense, mood and agreement
Nominal categories: DET
Argument encoding: SENT - SENT

Examples:
Ref Head:
Damu dut [zuri gezzura esu-te]-a
regret have YOU.DAT lie say-NMLZ-DET
‘I regret telling you a lie.’ (Hualde & Ortiz de Urbina 2003: 656)

[haurrek etxean liburuak sari irakurtze]-a
children.ERG home.LOC books often read-NMLZ-DET
‘children’s often reading books at home’ (Hualde & Ortiz de Urbina 2003: 666)

Pred-tu/du/-i/-Ø (‘perfective participle’)
Remark: this is the perfective counterpart of the nominalization with -tze (see above). In combination with instrumental or partitive case10, or with a postposition such as gabe ‘without’, this construction can also be used in Pred Mod function.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No Tense, mood and agreement
Nominal categories: DET
Argument encoding: SENT - SENT

Examples:
Ref Head:
Damu dut [zu irain-du]-a
regret aux you offend-PPV.NMLZ-DET
‘I regret having offended you.’ (Hualde & Ortiz de Urbina 2003: 668)

10 In Eastern dialects, in stead of –rik, the morpheme –ta is used, which is probably related to the conjunction eta and as such seems to form a dedicated adverbial construction in combination with the participial form (see Hualde & Ortiz de Urbina 2003: 745-746). In the classification of Chapter 6, this is not taken into account as a separate coding strategy.
Appendix III: Dependent Clause Constructions Key Examples

Dependent Clause Constructions

Example 1: Children often read books at home.
- **Children** (subject) have read **books** at home. (Hualde & Ortiz de Urbina 2003: 666)

Example 2: I need/want to buy a house in Bilboa.
- **I** (subject) need/want to buy **a house** in Bilboa. (Hualde & Ortiz de Urbina 2003: 694)

Predicative Modifiers

Example 3: Money is earned by working, not by being lazy.
- **Money** is **earned** by working, not by being lazy. (Saltarelli 1988: 55)

Subjunctive Clause -(en)

Remarks: This construction is used for subjunctive complements of desiderative, manipulative, and emotive predicates. Although -(en) is the most common complementizer in such cases, the complementizer -(a) seems to be more frequent in western dialects. In the western dialects -(a) (complementizer + determiner -a) is also used to mark factive complements. The conjunction -(en) is the counterpart of -(en) used for the factive complement of a negated matrix verb (Hualde & Ortiz de Urbina 2003: 646).

Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (nouns, adjectives, small modifiers)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: (DET/CASE)
Argument encoding: SENT/Ø - SENT/Ø (gapping)

Examples:

Ref Head:

Example 4: Many Basque speakers want that the Basque language be official in the entire Basque Country.
- **Many Basque speakers** want that **the Basque language** be official in the entire Basque Country. (Hualde & Ortiz de Urbina 2003: 640).

Example 5: I heard that Amaia's brother died.
- **I** heard that **Amaia's brother** died. (Hualde & Ortiz de Urbina 2003: 646)

Example 6: I wanted to speak about matters that I don't know.
- **I** wanted to speak about **matters** that **I** don't know. (Hualde & Ortiz de Urbina 2003: 763)

Ref Mod:

Example 7: I lost the money that Peter brought.
- **I** lost the money that **Peter** brought. (Hualde & Ortiz de Urbina 2003: 764)
Clause + –en bezala

Functional distribution: Ríg: Pred Mod = PoS (derived manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping)

Example:
Pred Mod:

\[Zeuk ean didaz\-en bezala egin dut lana\]

\[you.EMPH say AUX.ADV do AUX job\]

'I did my job the way you told me.' (Hualde & Ortiz de Urbina 2003: 722)

(subjunctive) clause –(e)la

Remarks: With desiderative and manipulative main predicates, the auxiliary in the dependent clause takes subjunctive form. –(e)la can be used in Pred Mod function for adverbial manner clauses, either by itself or combination with the partitive case –rik. In its basic form –ela has a modal/temporal (simultaneous) meaning1. The combination with –rik is a dialectal variant, separating the Bizkaian and Gipuzkoan area from the eastern dialects.

Functional distribution: Flex: Ref Head, Pred Mod (+ partitive case) ≠ PoS (nouns, derived manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: (CASE)
Argument encoding: SENT - SENT

Examples:
Ref Head:

\[batzuek uste dute \[hauk oro kazeten eta kazeta-egileen egitekoak dir\]-ela\]

\[some.ERG think AUX these all journals.GEN and journal-makers.GEN duties are-COMP\]

'Some think that all these are duties of journals and journalists.' (Hualde & Ortiz de Urbina 2003: 635)

\[Unibertsitateak \[agiri guztiak euskaraz eta gaztelaniaz egin daitz\]-ela\]

\[University.ERG document all.DEF.PL Basque.INSTR and Spanish.INSTR do aux(subj)-COMP\]

'The University has demanded that all documents be written in Basque and Spanish.' (Hualde & Ortiz de Urbina 2003: 461)

Pred Mod:

\[Zer egin ez neki\-ela geratu nintzen\]

\[what do not knew-AUX stay AUX\]

'I stood there not knowing what to do.' (Hualde & Ortiz de Urbina 2003: 712)

Jaikitzen da, \[jauzi egiten du\]-ela-rik

\[rise.IPFV AUX, jump do.IPFV AUX-ADV-PART\]

'(S)he gets up, jumping.' (Hualde & Ortiz de Urbina 2003: 713)

bait-clause

Remark: The conjunction bait can be used in Ref Head function, but this is uncommon. In Ref Mod function it is used frequently, namely for extraposed relative clauses. In this function, a resumptive pronoun can optionally be added.

Functional distribution: Flex: Ref Head, Pred Mod ≠ PoS (nouns, adjectives, small modifiers)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT/Ø - Ø/SENT (gapping)

Examples:
Ref Head:

\[Hau da haren abanrailik bandiena \[ez baitu ainitz xahutzen\]\]

\[this is this advantage.PART biggest.DEF not conj.aux much spend.ipfv\]

'That's the main advantage, that he doesn't spend much.' (Hualde & Ortiz de Urbina 2004: 648)

Ref mod:

\[Landibarren badira lau kartier\, \[\text{beriak} \text{bait-\text{ira}}\]\]

\[Landibaba.loc ba.are four neighbourhood those(RSP) conj-are\]

11 Saltarelli (1988) seems to analyze the suffix –(e)la in MP function as a real adverbializing suffix. Hualde & Ortiz de Urbina (2003: 712) also allude to a difference between the complementizer and the adverbializer function: "Mitxelena points out that this modal –ela is not exactly homophonous with the completive –ela since they have a different accentual pattern (in some Gipuzkoan and Bizkaian dialects)."
Behaunem Dona Martine, Donoztia eta Azkonbegi
Behaune Dona Martine, Donoztia and Azkonbegi
‘There are in Landibarre four neighbourhoods, which are Behaune, Dona Martine, Donoztia and Azkonbegi.’ (Hualde & Ortiz de Urbina 2004: 816)

Abun
do/Ø + clause
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT/Ø - SENT

Examples:
Ref Head:
An jam [do an karowa ne nde]
3sg know comp 3sg close.to there neg
‘He knew not to [go] close to there.’ (Berry & Berry 1999: 189)

An syogat pa [jegru san].
3sg order child take.off clothes
‘He ordered the girl to take off (her) clothes.’ (Berry & Berry 1999: 187)

Gato + clause
Remark: The relative clause can be followed by a determiner.
Functional distribution: Rig: Ref Mod ≠ PoS (modifiers)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT/Ø - Ø/SENT (gapping)

Examples:
Ref Mod:
Men mu gu ye [gato man siri su men bi nggon].
1pl go kill person rel do wrong with 1pl poss woman
‘We will go and kill the person who committed adultery with our (clans) woman.’ (Berry & Berry 1999: 146)

Men mu dw syur mo syur wak [gato nje ben] ne
1pl go go.in water at water hole rel people make det
‘We went and washed at the well that people had made.’ (Berry & Berry 1999: 146)

Sa gato + clause
Functional distribution: Rig: Pred Mod ≠ (modifiers)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: none
Argument encoding: SENT- SENT

Example:
Pred Mod:
An da ben mó sarewo an ye ben kete
3sg actual do exist however 3sg neg do too.much
budo yo teker [sa gato nyim ne nde re.]
maybe neg too.much adv earlier det neg pfv
‘Although she does [these things] she does not do [them] very much, I mean, not like [she did them] before.’ (Berry & Berry 1999: 158)

Bambara
Clause + ka
Remark: The exact coding details are unclear; the literal translation suggests that this is a balanced construction.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained?
Nominal categories: None
Argument encoding: SENT- SENT
Examples:

Ref Head:
*M'b'a fè i ka tags*
I want that you go ⇒ I want you to go.

*M'b'a fè au ka ka foro cikè*
I want that you cultivate your field ⇒ I want you to cultivate your field. (Brauner 1974: 80)

**Pred-le/-ne (‘perfective participle’):**

Remark: Apparently, the dependent predicate is non-finite and the relativized item is gapped.

Functional distribution: Rig: Ref Mod = PoS (small adjectives)
Structural type: 2 (D-SENT)
Verbal categories: None
Nominal categories: None
Argument encoding: Ø - SENT/Ø?

Example:
Ref Mod:
*Mëgb [pasa-le]*
person loose.weight-pfv.ptc
‘a thin person’ (‘a person who has lost weight’) (Brauner 1974: 73)

**min(u)/mun(u) + clause**

Functional distribution: Rig: Ref Mod = PoS (small adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping)

Example:
Ref Mod:
*Du nan [min nana Kališkoro] oyu tubahu yu*
stranger rel come K. cop European cop
‘The stranger who has come to K is a European.’ (Brauner 1974: 82)

**Pred-tò (‘present participle’ ⇒ ‘converb’)**

Remark: Apparently, the dependent predicate is non-finite, and the co-referential subject is coded twice.
It is not clear how other argument(s) are coded.

Functional distribution: Rig: Pred Mod = PoS (manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: None
Nominal categories: None
Argument encoding: SENT- SENT

Example:
Pred Mod:
*[A kasi-tò seginas] a ka dagu*
3sg cry-prs.ptc. go.back.2 3sg pst? place/village
‘He went back to his village, crying.’ (Brauner 1974: 72)

**Georgian**

**Pred-a:**

Remark: Argument coding is in ergative alignment: subjects of intransitive and objects of transitive verbs (S and P) are POSS. Transitive subjects (A) are typically OBL: accompanied by mier ‘by’ (or -gan ‘from, by’) (Hewitt 1995: 542).

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 3 (D-ALT)
Verbal categories: No tense/mood/person-number agreement; aspect is retained.
Nominal categories: CASE
Argument encoding: POSS/OBL - POSS

Examples:
Ref Head:
*minda [st'at'iis ts'er-a]*
1.3.want.prs article.gen write-nmlz(-nom)
‘I want to write an article.’ (Vamling 1989: 35)
Appendix iii: Dependent Clause Constructions Key Examples

### 1.3. prefer.

- **Example:**
  
  *I prefer to do this.* (Vamling 1989: 99)

### m-Pred(-el / al) (active participle):

Remarks: Indicates the actor of the relative clause predicate. The negative (privative) counterpart of this participle is formed with the prefix *u-.*

- **Example:**
  
  *Nuclear energy, which is subordinate to no control...* (Hewitt 1995: 609)

### Pred-ul / il / Pred-ar / al (past participle):

Remark: When derived from a transitive verb, this construction has a passive interpretation; the agent is marked oblique and the patient is gapped. When derived from an intransitive verb, the constructions has perfective interpretation and the relativized argument is gapped.

- **Example:**
  
  *Georgia was a country (that had been) occupied by the Arabs.* (Hewitt 1995: 609)

### sa-Pred(-el / al / r) (future participle)

Remark: The meaning of this construction is ‘that which is to be V-ed’. The agent is demoted and as remains unexpressed; the patient is gapped. It is not clear whether this construction can take an oblique agent argument.

- **Example:**
  
  *Legends linked to the golden ram* (Hewitt 1996: 609)
**Parts of Speech and Dependent Clauses**

**Example:**

Ref Mod:  
\[ \text{Ager} \quad \text{m-i-k'au-i-a} \quad \text{[xel-mo-sa-v'er]-i} \quad \text{kaggald-eb-i} \]  
here  \( \text{1-ov-} \)  
\( \text{hand-prev-fut.ptc-sign}\)-nom  
\( \text{paper-pl-nom} \)  

'I am holding here the papers which are to be signed.' (Hewitt 1995: 609)

**rom + (subjunctive) clause**

Remarks: In Ref Head function, the strategy can be combined with a suitable correlative in the main clause. This is obligatory when the clause is dependent on a postposition or functions obligatorily (i.e. follows a verb that governs instrumental case). (Hewitt 1987: 218)

Complements of modal and desiderative predicates combine with the subjunctive mood (optative when the matrix verb is present or future tense and pluperfect when the matrix verb is past). The complementizer is often omitted with subjunctive complements, whereas it is usually obligatory with indicative complements (Vamling 1989: 32/33). With subjunctives, the complementizer *rom* may also be replaced by *titko(s)* or any of its synonyms *vitom(c)* and *vitom(c)da*.

In Ref Mod function either the head or the co-referential noun can be deleted, or both may be retained.

(Hewitt 1995: 606-607) When the co-referential NP is deleted and it does not function as a subject or direct object in the DC, a resumptive pronoun is often used. In the Ref Mod function, the relativizer itself avoids clause-initial position and usually appears between the first constituent and the verb. (Hewitt 1987: 187)

**Functional distribution:**

<table>
<thead>
<tr>
<th>Flex:</th>
<th>Ref Head, Ref Mod ≠ PoS (nouns, adjectives)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural type:</td>
<td>1 (Balanced)</td>
</tr>
<tr>
<td>Verbal categories:</td>
<td>Retained</td>
</tr>
<tr>
<td>Nominal categories:</td>
<td>Case agreement on the REL.PRON</td>
</tr>
<tr>
<td>Argument encoding:</td>
<td>SENT - SENT (gapping)</td>
</tr>
</tbody>
</table>

**Examples:***

Ref Head:  
\( (\text{i}) \quad \text{uk'we} \quad \text{ie-o-a-m'en-i-e} \)  
\( \text{(that(nom))} \)  
\( \text{already prev-i-nv-notice-ths-aor.ind} \)  
\( \text{rom es xalx-i} \quad \text{sando} \quad \text{ar ar-i-s} \)  
\( \text{comp} \)  
\( \text{this people-nom trustworthy(nom) not be-prs-it} \)  
'I have already noticed that this people is not trustworthy.' (Hewitt 1995: 613)

Ref Mod:  
\( \text{Minda} \quad \text{[gamovatsxo namtsxvari]} \)  
\( \text{1.3.want.prs 1.3.bake-opt cake.nom} \)  
'I want to bake a cake.' (Vamling 1989: 33)

Ref Mod:  
\( \text{[gušin rom beč'ed-i m-a-čuk-a,]} \quad \text{is (beč'ed-i) sad ar-i-i'?} \)  
\( \text{yesterday rel ring-nom me-ly-present-aor.ind} \quad \text{that(nom) (ring-nom) where be-prs-it} \)  
'Where is that ring which you presented to me yesterday?' (Hewitt 1995: 607)

Ref Mod:  
\( [\text{K'ino-ši rom (ma-s) e-lap'arak'-eb-o-i}] \quad \text{is megreli-i sad} \quad \text{cinema-in rel rsp-dat iffy-speak.to-ths-iffy-ind} \quad \text{that Mingrelian-nom where ga-i-can-i?} \)  
\( \text{prev-sv-get.to.know-aor-ind} \)  
'Where did you get to know that Mingrelian to whom you where speaking in the cinema?' (Hewitt 1995: 607)

**REL PRON + clause**

Remark: The relative pronouns are *cin-c* ‘who’, *ra-c* ‘which’, and *ro-mel-i-c* ‘who, which’.  

Functional distribution: Rig: Ref Mod = PoS (adjectives)  

**Structural type:**  
1 (Balanced)  

**Verbal categories:**  
Retained  

**Nominal categories:**  
Case agreement on the REL.PRON  

**Argument encoding:**  
SENT - SENT  

Ref Mod:  
\( [\text{g'ut'la, cin-c nacionalizm-s}] \quad \text{[nacionalizm-s who(nom) rel nacionalism-dat} \)  
\( \text{a-h-q'v-eb-a samisbo-s u-galat'eb-i} \)  
\( \text{prev-it-follow-ths(fut)-he homeland-dat ov-betray-ths(fut)-he} \)  
'Everyone who follows nationalism will betray this country.' (Hewitt 1995: 601)
Appendix iii: Dependent Clause Constructions Key Examples

**ra in clause**

Functional distribution: Rig: Pred Mod = PoS (manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

**Examples:**
Pred Mod:
[r-a-xemar-eb-i-an ra saxalxep armi-is iet`eun-a-

ov-help-ths-prs-they ADV people's army-gen attack-dat

part`izan-eb-i a-flet-eb-en…

partisan-pl-nom NV-strengthen-ths-they(prs)

‘Helping the people’s army to attack, the partisans strengthen …’ (Hewitt 1995: 600)

Bukiyip

Unmarked clause

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT/Ø - SENT

**Examples:**
Ref Head:
Nénan n-a-kli [yek i-wich umu énaniny moul]

He he-real-say I I-IRR-enter ben his work

‘He said that I would have his job.’ (Conrad & Wogiga 1991: 179)

**Clause + (a)li**

Remark: The relativized item is gapped, but it is cross-referenced (with a prefix denoting class and number) on the dependent predicate.

Functional distribution: Rig: Ref Mod = PoS (adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT
**Example:**

Ref Mod: *H-a-agé-xagú agú-dak*

3PL-MASC.SUBJ-REAL-SEP CL3.SG.DEM-this

*nehe-puli* *trag [g-a-lahe-mu bhensin uli]

large-CL3.SG truck CL3.SG.SUBJ-REAL-TRAVEL-BEN gasoline REL

‘They repaired this big truck which transports gasoline.’ (Conrad & Wogiga 1991: 103)

**bdidauk + clause + -(u)mu**

Functional distribution: RIG: Pred Mod = PoS ((small) manner adverbs)

Structural type: 1 (Balanced)

Argument encoding: SENT - SENT

**Abkhaz**

**Pred-ra**

Functional distribution: RIG: Ref Head = PoS (nouns)

Structural type: 3 (D-ALT)

Argument encoding: POSS/OBL - POSS

**Examples:**

Ref Head:

[Ara *iaq-a-zau-ra*]

here 1SG.PESS-be-this-NMLZ it-2PL-KNOW-DYN-FIN

‘They know that I was here.’ (Hewitt 1979: 31)

[dará *và-la wàc˚ Čerkèссk-qà há-it-ra*]

them them-by(instr) today Čerkessk-to our-send-NMLZ it-prev-1SG-be.surprised.at-FIN

‘I am surprised at their sending us to Čerkessk today.’ (Hewitt 1979: 31)

**Pred-N.FIN**

Remark: This construction makes use of a special non-finite paradigm. It expresses nearly all verbal categories that are also expressed in independent clauses, but in a different form.

Functional distribution: RIG: Ref Head = PoS (nouns)

Structural type: 1 (Balanced)

Argument encoding: SENT - SENT

**Example:**

Ref Head:

[d-xaa-wà] *(O-)*xà-s-c’a-wa-yt’.  
he-come-DYN(N.FIN.PRS) it-head.in-I-put-DYN-FIN(PRS)

‘I believe that he will come.’ (Hewitt 1987: 238)

**REL.PRON + Pred-N.FIN**

Remarks: This construction makes use of a special non-finite paradigm. It expresses nearly all verbal categories that are also expressed in independent clauses, but in a different form.

The person affix that is used in independent clauses is replaced with one of two relative affixes. *-y(ə)* is used in stead of all person affix of the first declination, regardless of the person, class and number of the head noun, whereas *-ɀ(ə)* performs this function for all person affixes of declinations 2 and 3. (Hewitt 1987: 200)

Functional distribution: RIG: Ref Mod = PoS (adjectives)

Structural type: 1 (Balanced)

Argument encoding: SENT - SENT
Examples:
Ref Mod: 
[da-xàc' d-aa-yt' ] det man he-come-fin
'det man who killed the woman came.' (Hewitt 1987: 201)

[da-xàa yà-dàw(ak˚a-)-z ] à-h˚sa (Ø)-z-dw-wa-yt'
the-man whom(=) he-see-pl.n-fin the-women them-I-know-dyn-fin
'I know the women whom the man saw.' (Hewitt 1987: 201)

ş-Pred-N.FIN
Remark: This construction makes use of a special non-finite paradigm. It expresses nearly all verbal categories that are also expressed in independent clauses, but in a different form.
Functional distribution: Rig: Pred Mod = PoS (small/derived manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Almost all retained, but dependent paradigm
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Pred Mod: 
[Yo-ś-w-à-s- ħ̊a-z ] yə-q'a-c'à
it-adv you-to-I-say-nfin(pst.indf ) it-prev-do.imp
'Do it as I told you.' (Hewitt 1987: 120)

Polish
Pred-nie
Remark: Occurs with A-argument unexpressed under co-referentiality and P-argument SENT, but also with an A-argument oblique A and a P-argument POSS.
Functional distribution: Rig: Pred Head = PoS (nouns)
Structural type: 3 (D-ALT)
Verbal categories: No TAM/agreement
Nominal categories: CASE
Argument encoding: OBL - POSS/Ø - SENT

Examples:
Ref Head:
'I request the dismissal from work of the two women workers.' (Comrie 1976: 191)

[kupowa-nie mies-a prez Hanke] buy-nmlz meat:gen by Hanka
'the buying of meat by Hanka' (Koptjevskaja-Tamm: 293)

Pred-INF (various forms)
Remark: There is a number of different infinitival endings:
-uc' -uc'
-ić -iç
-ić -iç
-yć -owanć
-uć -owanć
-t ' -towanć -towanć
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No tense, aspect, agreement
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

Example:
Ref Head:
Postaniwi ɫ em [kupić dom].
I.decided buy.inf house
'I decided to buy a house.' (Bielec 1998: 19)
**Pred-c-- (present active participle)**

*Functional distribution:* Rig: Ref Mod = PoS (adjectives)
- Structural type: 2 (D-SENT)
- Verbal categories: No tense, aspect, valency, person marking
- Nominal categories: Adjectival agreement (number, gender, and case)
- Argument encoding: Ø - SENT (gapping)

*Example:*

Ref Mod:

*Widz chłopę [słuchaj-go radio]*

'I see a boy (who is) listening to the radio.' (Bielec 1998: 170)

**Pred-anya/-anal-(SG)/-anl-(PL) (present passive participle)**

*Remarks:* Used with imperfective verbs only. Can possibly be regarded as derived adjectives, since there are no overt arguments: The relativized item is gapped, and the demoted agent argument remains unexpressed.

*Functional distribution:* Rig: Ref Mod = PoS (adjectives)
- Structural type: 2 (D-SENT)
- Verbal categories: No tense, aspect, valency, person marking
- Nominal categories: None
- Argument encoding: Ø - Ø (gapping)

*Example:*

Ref Mod:

*Odziaż [sprzedaw-ana]*

The clothes (being) sold there are cheap. (Bielec 1998: 171)

**Pred-PASS.PST.PTC (various forms)**

*Remarks:* There are several formation strategies, depending on the form of the infinitive. This construction is used with perfective verbs only. These participles can possibly be regarded as derived adjectives, since they take no overt arguments: the relativized item is gapped, and the demoted agent argument remains unexpressed.

*Functional distribution:* Rig: Pred Mod = PoS (adjectives)
- Structural type: 2 (D-SENT)
- Verbal categories: No tense, aspect, valency, person marking
- Nominal categories: None
- Argument encoding: Ø - Ø (gapping)

*Example:*

Ref Mod:

*Mikołaj ma [złamać-nąleg]*

'Nicolas has a broken leg.' (Bielec 1998: 171)

**Pred-PRS.3rdPL + -c (adverbial participle)**

*Remark:* The (present) converb (called adverbial participle in Bielec 1998) is formed by adding the suffix -c to the 3rd person plural of the present tense.

*Functional distribution:* Rig: Pred Mod = PoS (small/derived manner adverbs)
- Structural type: 2 (D-SENT)
- Verbal categories: No tense, aspect, person marking
- Nominal categories: None
- Argument encoding: Ø - SENT (co-referentiality)

*Example:*

Pred Mod

*[słuchaj-c]*

'Listening to music, I got dressed.' (Bielec 1998: 71)

**ze + clause**

*Functional distribution:* Rig: Ref Head = PoS (nouns)
- Structural type: 1 (Balanced)
- Verbal categories: Retained
- Nominal categories: None
- Argument encoding: SENT - SENT
Appendix iii: Dependent Clause Constructions Key Examples

Example:
Ref Head: Myślę, [że ona jest mężatką].
'I think that she is married.' (Bielec 1998: 239)

który/które/co (REL.PRON) + clause
Functional distribution: Ríg: Ref Mod = PoS (adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: Adjectival declination of relative pronoun
Argument encoding: SENT - SENT

Example:
Ref Mod: Chata, [gdzie mieszkali,] była mała. [= w którym]
The cottage in which they lived was small.' (Bielec 1998: 155)

Jak (gdyby) + clause
Functional distribution: Ríg: PredMod = PoS (small/derived manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
Pred Mod: Agata spojrzała na mnie, [jak gdyby chcieła o coś zapytać].
Agatha looked at me as if she wanted to say something,' (Bielec 1998: 238)

Burushaski
Pred-(á)as (INF)
Remarks: Used as the complement of modal, phasal, desiderative, and manipulative predicates. With phasal and desiderative predicates, the dependent verb form it is declined like a regular noun. When combined with an inherent (dative, inessive, adessive, superessive or locative) case marker, this construction can be used as an adverbial simultaneity clause.

Functional distribution: Flex: Ref Head, Ref Mod, + CASE also Pred Mod. ≠ PoS (nouns, adjectives, small/derived manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: Tense, mood and subject agreement are lost. Aspect and object-agreement are retained.
Nominal categories: (CASE), see above
Argument encoding: SENT - SENT / Ø - SENT

Examples:
Ref Head:
but mukil bila [gía-lir-as]
very difficult be.IV 2-show-INF
'It is very difficult to show (it) to you.' (Anderson 2002: 545)

Ref Mod:
[Chá-áte oóo-rv-sá-as] bok
post-superess neg-sit-INF dog
'a dog which doesn’t sit at its post' (Anderson 2002: 545/Berger 1998: 171)

Pred Mod:
[svin-á-á]-ar
say-INF-dar
'when he said' (Anderson 2002: 547/Berger 1998: 190)

[du-ús-ar]-ulb
SUBJ.VERS-come.out-INF-INESS
'when he came out' (Anderson 2002: 547/Berger 1998: 190)

Pred-im/-um/-am (aorist participle)
Remarks: In combination with a dative, inessive, adessive or superessive (privileged) case marker this construction can be used as an adverbial simultaneity clause. The durative form of the aorist participle can also function adverbially, in combination with locative or superessive case, or with the comitative marker kha and the genitive case. These constructions seem to have ‘true’ manner semantics.
Functional distribution: Rig: Ref Mod (+ CASE also Pred Mod) = PoS (adjectives)
Structural type: 2 (D-SENT)
Verbal categories: Tense, mood and subject agreement are lost. Aspect and object-agreement are retained.
Nominal categories: (CASE), see above
Argument encoding: SENT - SENT Ø- SENT

Examples:
Ref Mod:
[iŋ barum-man-um] mapéer-an
he-beard white be(come)-aor.ptc elder-sg.art
'an old man with a white beard.' (Anderson 2002: 546 / Berger 1998: 166)

Pred Mod:
[juliz hay-a-m]-ulu K'udé-e-re duá et-a-m
ill be-1-aor.ptc-INESS God-obl-dat prayer aux-tr-1.aor.ptc
'When I was ill I prayed to God.' (Anderson 2002: 549)

[kos-čerum]-ate
NEG-cry-aor.ptc-superess

Clauses
Remark: This is the only primary converbal form, i.e. one that is not derived from another non-finite form by case-marking. It normally expresses anteriority but it may also express manner of action or an accompanying circumstance, in which case the conjunctive participle is often reduplicated.
Functional distribution: Rig: Pred Mod = PoS (small/derived manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT / Ø - SENT

Examples:
Pred Mod:
Aya máma [ji-ar taklíiff n-á-či-n] uyóŋ-ko u-yú-ar
father(erg) mother(erg) 1-dat trouble conv-1sg.io-give-conv big-pl.3cl.poss-son(pl)-dat
šuá n-e gar-íŋ um-an-um.
good cp-do marriage-pl do-stat-plsubj
'Father and mother arranged good marriages for their big sons, giving me a lot of trouble.' (Tikkanen 1995: 502)

Clauses + ke/sí
Remark: This construction is borrowed from Persian/Urdu. It can be used as the object complement of utterance and cognition predicates (no example available), as a relative clause (no example available), and as a simultaneity clause.
Functional distribution: Flex: Ref Head, Ref mod (Simultaneity clause) ≠ PoS (nouns, adjectives, small/derived manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT
Example simultaneity clause:
[Zungatinge gisam-ular aspar-így d-sí-mi
bẹ-3gen. village-ill. flower-PL PREP-3CL.SG-{come}-3CL.SG
ke] ban dastúr-an b-li-üm
SUB one:cl. custom-sing be-3CL.SUBJ-ST
‘The apricots coming into blossom, when the flowers came to the village of Zungating, there was a custom …’ (Tikkanen 1998: 498)

Clause + -sén / -ét (QUOT)
Remark: The quotative marker takes the form of the anterior converb of the quotative verb sén- ‘to say’ or d’ét- ‘to do/to speak’. The construction can be used in Ref Head function for complements of utterance and cognition predicates. No examples.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: All retained
Nominal categories: None
Argument encoding: SENT - SENT

Lavukaleve
Pred-e-/i
Remark: The so-called “Agreement Suffix” (which is lost in this construction) marks gender and number of a core argument (which argument that is depends on various factors, such as focus). The possessive paradigm is identical to the verbal subject prefix paradigm except for one form: the first person singular subject prefix is a-, while the possessive form is nga-. However, the fact that nominalizations take a- is taken as evidence that the other forms are also subject prefixes rather than possessive markers.
The construction can take a determiner and can also be dependent on a postposition.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No TAM, no Agreement Suffix
Nominal categories: (CASE / DET)
Argument encoding: SENT - SENT

Example:
Ref Head
Ngai [ngu-bolea la o-ma-e
1sg. 1sg.poss -arrow sg.fem.art 3sg.fem.obj -take-nmlz
e-li-li-re ta a-lei
3sg.neut.obj want-n.fin just 1sg.subj-exist
‘No! I just want to take my arrow.’ (Terrill 2003: 352)

Clause + AGR + DET
Remark: The Agreement Suffix is obligatorily used on the (final) dependent predicate to cross-reference the head of the construction, which can be the subject, the object, or a postpositional object. Relative clauses are internally headed.
Functional distribution: Rig: Ref Mod = PoS (adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: DET (+ adjectival gender and number agreement)
Argument encoding: SENT - SENT

Examples:
Ref Mod:
[Ali nga-fósal a-u-m na] a-le-m fin.
man 1sg.poss-fish 3sg.masc.obj-catch-sg.masc ART 1sg.subj-see-sg.masc 3sg.masc.loc
I saw the man who ate my fish.’ (Terrill 2003: 442)

[Ali a-na so-oks-ne a-ngooa-m na] o-foc
man 3sg.masc.obj-in rdp-laugh-ipfv 1sg.subj-stay-sg.masc sg.masc.art 3sg.poss-pig
na o-lufu-m
sg.masc.art 3sg.subj-leave-sg.masc
‘The man who I laughed at had lost his pig.’ (Terrill 2003: 448)

Alamblak
Pred-nef
Remark: The subject of intransitives may be incorporated or POSS. With transitive predicates, either the subject or the object can be POSS.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 3 (ALT-SENT)
Verbal categories: No agreement (actor and undergoer)
Nominal categories: Gender and number marking on predicate
Argument encoding: INC/POSS/SENT - SENT/POSS

Examples:
Ref Head:
Yifémr pëthawonmëanr nan-bo yi-nef-t
father talk.try.I.him [I-GEN go-NMLZ-3SG.FEM]
‘I tried talking to Father (about) my going.’ (Bruce 1984: 124)

Yifémr pëthawonmëanr [vir-bo yak-ni-nef-t na]
father talk.try.I.him he-GEN get-go-NMLZ-3SG.FEM me
‘I tried talking to Father (about) his taking me.’ (Bruce 1984: 125)

Pred-/kfët (INF)
Remark: Unlike the nominalization in -nef, the infinitive does not take gender/number markers.
Functional distribution: Rlg: Ref Head = PoS (nouns)
Structural type: 3 (ALT-SENT)
Verbal categories: No tense/agreement (actor and undergoer)
Nominal categories: None
Argument encoding: INC/POSS - INC/POSS

Examples:
Ref Head:
Na  iñji wañf
ɨ
nahmëanr [baripat rhu-haku-t]
1SG thus hear.arrive.REM.PST.1SG.3SG.FEM lake.only remain-always-INF
‘Thus I have heard it up to now (that) it is always only a lake.’ (Bruce 1984: 284)

Pred-/+ Pred + clause
Remark: Relative clauses are optionally marked with ɨnd, but mostly unmarked. The function of the head noun is not marked, except when a possessor is relativized.
Functional distribution: Rlg: Ref Head = PoS (nouns, adjectives)
Structural type: 2 (D-SENT)
Verbal categories: Aspect and mood are retained; agreement (subject and object) is lost
Nominal categories: (CASE), see above
Argument encoding: SENT/Ø - Ø/SENT (gapping)

Examples:
Ref Mod:
[met-t maroba-bayme] yima-r
woman-3SG.FEM money-gave person-3SG.MASC
‘a man (who) gave money to a woman / a man (to) whom a woman gave money’ (Bruce 1984: 106)

Pipil
ka(h) + clause
Functional distribution: Flex: Ref Head, Ref Mod (occasionally) ≠ PoS (nouns, adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT- SENT
Appendix iii: Dependent Clause Constructions

Key Examples

Examples:

Ref Head:

- **ya klaroh k-itau-k [kah witi na siua-t].**
  - He saw clearly that the woman (was) coming.' (Campbell 1985: 126)

- **he ne i-siua-w se brushah].**
  - (He) didn't know before that his wife (was) a witch.' (Campbell 1985: 126)

Ref Mod:

- **Ni-k-miktih ne mitsun [ka ki-kwah ne tu:tu-t].**
  - I killed the cat which ate the bird.' (Campbell 1985: 129)

- **kunih abhi-k-t a-i-ban ne taka-t [kah ki-kutan-k].**
  - Then they arrived (at) the house of the man who threw it down.' (Campbell 1985: 129)

**ne (REL. PRON) + clause**

Functional distribution: Rig: Ref Mod = PoS (adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:

Ref Mod:

- **Naha ni-k-tau-k ne siua-t [ne ki-pak ne kwah-kwach-ti].**
  - I saw the woman who washed the clothes.' (Campbell 1985: 129)

**ke (REL) + clause**

Remark: This construction is borrowed from Spanish (*que*).
Functional distribution: Rig: Pred Mod = PoS (adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:

Ref Mod:

- **Kunih ne taka-t [ke ki-kutan] k-its-ki ne chumpipi k-wi:ka ka i-ban.**
  - Then the man who threw it down grabbed the turkey (and) took it to his house.' (Campbell 198: 129)

**ke:ken-aken (‘just as’) + clause**

Remark: This construction has simulative semantics.
Functional distribution: Rig: Pred Mod ≠ PoS (no lexical strategy)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:

Pred Mod:

- **Ah, ni-paxa:lua [ken-aken taha ti-kiv-tuk ti-paxa:lua].**
  - ‘Oh, I'm taking a walk just as you have come out to take a walk.' (Campbell 1985: 289)

**ADV + clause**

Remark: This construction has simultaneity semantics.
Functional distribution: Rig: Pred Mod ≠ (no lexical strategy)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT
**Examples:**

**Pred Mod:**

*Kuman*  
**[yaba yavvi katka derechob ne kayeh, yah kis-ki i-xpan ivomawozal]**  
when he go before straight the street then leave-pret his-before Siguanaba

'When he was going (down) the street, he came out in front of (the) Shaunaba.' (Campbell 1985: 131)

**Wambron**

- **-e (CONN) + clause**
  
  *Remark:* The connector -e links all preverbal NPs, irrespective of their functional or structural specification, so this includes object clauses.
  
  **Functional distribution:** Rig: Ref Head = PoS (nouns)
  
  **Structural type:** 1 (Balanced)
  
  **Verbal categories:** Retained
  
  **Nominal categories:** None
  
  **Argument encoding:** SENT- SENT/Ø - SENT

**Example:**

*Ref Head:*  
 **Keno**  
**[beta-khe-rr-[e] lassa-tnabo...]**

sand-pret [CONN]

*and see-3sg.prs-tr-CONN trap-3sg.pst*

'And he saw that he had trapped it (...)’ (De Vries 1986: 29)

- **-a + clause-ø**
  
  *Remark:* The suffix -o links all types of pre-nominal modifiers to the head noun, including relative clauses, which are additionally marked by the relativizer -a, which delineates the relative clause by cliticizing to its first constituent. Relative clauses can also be head-internal. If that case, the construction is not pre-nominal, so the connector -o does not appear.
  
  **Functional distribution:** Rig: Ref Mod = PoS (adjectives)
  
  **Structural type:** 1 (Balanced)
  
  **Verbal categories:** Retained
  
  **Nominal categories:** None
  
  **Argument encoding:** Ø/SENT - Ø/SENT

**Examples:**

**Ref Mod:**

**Nuk**  
**[oy-a temke-n-ø]**  
kev-e  
**betak-nok-ndep**

I pig-rel shoot-3sg.prs-tr-CONN man-CONN see-neg-1sg.prs

'I do not see the man who shoots the pig.' (De Vries & De Vries-Wiersma 1992: 56)

**[Aliv-a ndu-n-ø takhima-lopo-n-eve]**  
setelep-ke-khe

yesterday-rel sago-tr-CONN buy-1sg.pst-tr-top delicious-superl-3sg.prs

'The sago I bought yesterday is delicious.' (De Vries & De Vries-Wiersma 1992: 57)

**Pred stem-mo**

*Remark:* This is a *medial verb form*, which takes switch-reference marking for same subject (SS).

The interpretation is temporally neutral (but conceptually close).

**Functional distribution:** Rig: Pred Mod ≠ PoS (no lexical strategy)

**Structural type:** 2 (D-SENT)

**Verbal categories:** No tense, mood, and subject agreement

**Nominal categories:** None

**Argument encoding:** Ø - SENT (co-referentiality)

**Example:**

**Pred Mod:**

**Jakbov-e**  
**[matet-mo] ka-lombo?**

they-CONN good-ss go-3pl.pst

'Did they travel without problems?' (lit.: ‘Did they go well?’) (De Vries & De Vries-Wiersma 1992: 19)

**Pred-ø**

*Remark:* This is also a *medial verb form* with same-subject switch reference marking. The interpretation is not temporally neutral: it expresses simultaneity.

**Functional distribution:** Rig: Pred Mod ≠ PoS (no lexical strategy)

**Structural type:** 2 (D-SENT)

**Verbal categories:** No tense, mood, and subject agreement. Transitivity markers can be retained.

**Nominal categories:** None

**Argument encoding:** Ø - SENT (co-referentiality)
Appendix iii: Dependent Clause Constructions Key Examples

Example:
Pred Mod: Jakhov-e [ko-n-o] nggom bi-knde
they-conn go-tr-ss song sing-3PL-FRS
‘While travelling they are singing.’ (De Vries & De Vries-Wiersma 1992:20)

clause-ka
Remark: The suffix -ka (and its allomorphs) is a circumstantial suffix that also occurs with inanimate NPs that specify the circumstances of the event denoted by the verb, including time, location, instrument, and manner.
Functional distribution: Rig: Pred Mod ≠ PoS (no lexical strategy)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

Example:
[Nuub-e ande-lepo-ngga] ev-o kav-e nde-tmbo
I-conn cut-1SG.PST-ADV that-conn man-conn come-3SG.PST
‘When I ate, that man came.’ (de Vries 1986: 41)

Dhaasanac
Clause (+DEM) + DET
Remarks: Used for the complements of desiderative and perception predicates, and for all types of relative clauses. With subject relative clauses the subject is repeated in the matrix clause with a pronoun.
The construction further takes the determiner ka or its clitic form =a; in the Ref Mod function often preceded by a deictic.
There is only one difference with independent clauses: for constructions in non-past tense, imperfective aspect is excluded; instead, the dependent form must be used.
Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (nouns, (small) adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: DET
Argument encoding: Ø/SENT - Ø/SENT

Example:
Ref Head:
Yú [kúun kosolonni=] ku ɲargi
I  you laugh.pfv=DET you.obj see.pfv
‘I saw that you were laughing.’ (Tosco 2001: 287)

yú [kúun hí fišši=] fedé
I  [you 3(verb pron) open,DEF=DET want,IPFV
‘I want you to open it.’ (Tosco 2001: 288)

Ref Mod:
Máa [ɠuo ɗɠo ɠɠ=ti=a]
Man  trees dig,RPD,DEF=that=DET 3,SUBJ me know
‘That man who is digging trees (he) knows me.’ (Tosco 2001: 283)

Pred Mod:
Yú [ɠúo ɠii ka] máa pigɛ=ati= a du seół
I  cattle have,PFV DET person young=that=DET upon go,PFV
‘I, having the cattle, went to the boy.’ (2001: 286)

Pred-ɲ/-an
Remarks: The suffix -ɲ is used for the basic and causative/factive paradigms, -an for the middle and inchoative paradigms. The dependent predicate has feminine nominal gender.
The object is unmarked (noun-stripping).
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 3 (D-ALT)
Verbal categories: Tense and aspect are lost
Nominal categories: (DET), gender marking
Argument encoding: Ø/POSS - unmarked

Examples:
Ref Head:
[vúb mümür-an] ɬɛ beyyi
sorghum cut,RPDM,NMLZ 3,SUBJ finish,PFV
‘The harvesting of the sorghum is finished.’ (Tosco 2001: 121)
Berbice Dutch Creole

**(fu/fi) + deranked clause**

Remark: The complementizer can be deleted.

Functional distribution: Rig: Ref Head = PoS (nouns)

Structural type: 2 (D-SENT)

Verbal categories: No TAM

Nominal categories: None

Argument encoding: Ø - SENT

**Examples:**

Ref Head:

Eke \( \text{timi} \) \([\text{(Ø)} \tan \text{langgi ka}]\)

1sg be.able stand long NEG

‘I am not able to stand up for very long.’ (Kouwenberg 1994: 249)

Governor \( \text{ma} \) \( \text{sub-a} \) \([\text{fi} \text{nimi bosa ju krik-te-ke}]\)

Governor irr want-IPFV comp know how.FOC 2sg get-PPFV.1sg

‘The governor will want to know how you got me.’ (Kouwenberg 1994: 249)

\((\text{bi}(/\text{fi})) / (\text{QUOT}) + (\text{dati}) / (\text{COMP}) + \text{clause}\\)

Remark: The complementizer dati can be deleted. Dati-clauses can also be preceded by the quotative marker \(\text{bi}(/\text{fi})\) (from the verb for ‘to say’), resulting in a ‘double’ complementizer sequence.

Functional distribution: Rig: Ref Head = PoS (nouns)

Structural type: 1 (Balanced)

Verbal categories: Retained

Nominal categories: None

Argument encoding: SENT - SENT

**Examples:**

Ref Head:

Bat \( \text{now eke boro [dat-o mwa-a]} \)

but now 1sg hear COMP-3sg go-IPFV

‘But now I hear that he is leaving.’ (Kouwenberg 1994: 244)

Bateni \( \text{nim-te} \) \([\text{dat-ek ma ku-mona-ngga}]\)

but.3pl acquire.knowledge.FFV COMP-1sg irr come-Mona-LOC

‘But they knew that I would come to Mona’s.’ (Kouwenberg 1994: 241)

Eke \( \text{pan-te ju [date ju mu bifi ka]} \)

1sg tell-PPFV 2sg COMP 2sg must speak NEG

‘I told you that you must not talk.’ (Kouwenberg 1994: 238)

Ek \( \text{glof ka} \) \([\text{(Ø)} \o \text{nin-te musa}]\)

1sg believe NEG 3sg know-PPFV much

‘I don’t believe he knows much.’ (Kouwenberg 1994: 242)

Eni \( \text{pama-te-ke [bifi date-ke ma mu danga ka]} \)

3pl tell-PPFV.1sg QUOT COMP-1sg irr go there NEG

‘They told me that I should not go there.’ (Kouwenberg 1994: 238)

**Wh-word) + clause:**

Remark: The Wh-element can be deleted (only in restrictive clauses).

Functional distribution: Rig: Ref Mod = (adjectives)

Structural type: 1 (Balanced)

Verbal categories: Retained

Nominal categories: None

Argument encoding: SENT/Ø - SENT/Ø (gapping)

**Example:**

Ref Mod:

Di \( \text{jugu manj-ap [swat fol me hari]} \)

the big-big man-PL what full with hair

‘the giants that are covered with hair’ (Kouwenberg 1994: 268)
Appendix iii: Dependent Clause Constructions Key Examples

**Di** bwa:, di en [wari wa siki,] ori doto-te esti
The leg the one what pst ill, 3sg die-pfv first
The leg, the one that was ill, it became numb first. (Kouwenberg 1994: 271)

**Di** sem jermatoko [(Ø) che pan-te ju abut]
the same woman.child 1sg tell-pfv 2sg about
'the same girl I told you about' (Kouwenberg 1994: 268)

### Barungo

**lāa + clause:**
Remark: The subject of the matrix clause is repeated in the dependent clause as a logophoric pronoun.

**Functional distribution:**
- Flex: Ref Mod, Pred Mod (similative)
- PoS (small/derived adjectives, small manner adverbs)

**Structural type:**
- 2 (D-SENT)

**Verbal categories:**
- Restricted aspect

**Argument encoding:**
- SENT/Ø - SENT/Ø (gapping)

**Examples:**

**Ref Mod:**

mə̀ yə wə ntis [fáŋ/yiiu (gwə) shaw ngə yə]  
I see-pfv person that REL he steal pfv bow your  
'I have seen the man who has stolen your fowl.' (Schaub 1985: 32)

lwəŋ ngwə [fáŋ rii gwə nə juwə nə]  
hammer that [REL father my pst come pfv with gwənə,]  
I see, gəli liu tua gii  
h(ANAPH) hammer that.anaph be in house  
'That hammer which my father came with (it), that hammer is in the house. (Schaub 1985: 33)

**Pred Mod:**

gwə sə səy  
she plant pfv corn ADV mother her tell pfv to her  
'She has planted the corn like her mother told her.' (Schaub 1985: 39)

**kii(I)/Ø + clause**

Remark: There are two types of these "simultaneous aspect constructions": One relating to a preceding verb (anaphoric), and the other relating to a following verb (kataphoric). With stative verbs the anaphoric construction is marked by the simultaneity particle kii. With other verbs, there is no overt coding. In both cases, the verb following the particle can only be in the imperfective aspect. The other type (kataphoric) is used when event B occurs while A is still in process. It is marked by the marker kii in the first verb phrase, which expresses the event in process. The verb with which the marker kii occurs, is in the imperfective aspect, while the verb of the following clause is in the perfective aspect.

**Functional distribution:**
- Rig: Pred Mod = PoS (small manner adverbs)

**Structural type:**
- 2 (D-SENT)

**Verbal categories:**
- Restricted aspect
Nominal categories: None
Argument encoding: Ø-SENT

Examples:
Pred Mod:

"He was asleep, resting his body." (Schaub 1985: 220)

"When he was walking on the road, he suddenly fell down." (Schaub 1985: 221)

Nama

Clause without ke (= DECL) + !xáis-a

Remark: This construction is marked by the nominal class marker !xáis (SG. FEM), which can be shortened to -s, and the object marker -à.

Apart from the declarative marker ke, no verbal categories are lost.

Functional distribution: Rig: Ref Head = PoS (small manner adverbs)

Structural type: 1 (Balanced)

Verbal categories: Retained except declarative marker

Nominal categories: Nominal class

Argument encoding: SENT - SENT

Examples:

Ref Head:

"And we told him that we were going." (Hagman 1974: 257)

Unmarked clause without ke (= DECL) (+ RSP)

Remark: When the head of a relative clause functions as an oblique argument, the construction is usually marked with an appropriate postposition or associative particle, and a resumptive pronoun may be added which agrees with the head in person/number/gender.

Functional distribution: Rig: Ref Mod = PoS (small/derived adjectives)

Structural type: 1 (Balanced)

Verbal categories: Retained except declarative marker

Nominal categories: (Nominal agreement of RSP)

Argument encoding: SENT/Ø - SENT/Ø (gapping)

Examples:

Ref Mod:

"the town I am going to teach (in)." = ‘the town where I am going to teach’ (Hagman 1974: 231)

Deranked clause without ke (= DECL) + -se/!aai/tsií/Ø

Remark: The declarative marker ke cannot be expressed and the marking of aspect is restricted: the dependent clause must have non-punctual aspect. Usually, the subject is unexpressed under coreferrentiality.

The adverbial manner suffix -se and the general adverbial subordinating conjunction !aai are completely interchangeable, but -se is by far the most frequently used. The construction with tsií requires punctual aspect in the dependent clause.

The unmarked variant has the same meaning, but can only be used with a special set of verbs in the
dependent clause. Tense cannot be expressed in this variant (but aspect is retained).
Functional distribution: Rig: Pred Mod = PoS (derived manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No declarative marker, restricted tense/aspect
Nominal categories: None
Argument encoding: Ø - SENT

Examples:
Pred Mod:

![ái 'a ≠ áó]
ra-se-p ke kè pèé
rejoicing-ADV-he DECL left
'Rejoicing, he left.' (Hagman 1974: 244)

![xóókxoeku kè !nàeí'úhè]
![kxétehǹ/xáa !áéhè]
tsíí
the.prisoners were transported bound with chains ADV (Hagman 1974: 246)

Hdi
Clause + ka (+ low tone subject pronouns)
Remark: All subject pronouns following the complementizer, except the 3rd person singular a have low tone, whereas subject pronouns in independent clauses, except for the 3rd person plural, have high tone.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained (special subject forms)
Nominal categories: None
Argument encoding: SENT - SENT

Examples:
Ref Head:

![mnd-á ráyá Mbákà ka]
bá mbítsá mná-tà
'mbítsa said that Mbaka is a hunter.' (Frajzyngier & Shay 2002: 422)

Ndá sn-í
[žá-ná-p-zá-xə̀n tá pítsákw-á-tán ká-ú]
stat know-1 sg forget-out-forget-3pl.pl obj
'I know that they forgot their hoe.' (Frajzyngier & Shay 2002: 450)

tá (OBJ/COMMENT) + deranked clause:
Remarks: The marker tá is an object / comment marker: It is also used to mark lexical objects and comments in focus constructions (cf. glosses in examples).
In Ref Head function the subject is POSS. In Ref Mod function it is gapped, but marked with a possessive pronoun on the dependent predicate. In Ref Mod function, this construction is only for perfective subject relative clauses.
Functional distribution: Flex: Ref Head, Ref Mod ≠ PoS (nouns, small adjectives)
Structural type: 3 (D-ALT)
Verbal categories: No person marking
Nominal categories: Possessive person marking
Argument encoding: POSS - SENT

Examples:
Ref Head:

Sí tá !ɗv-áy-xə̀n [tá hliy-á-mú]
pst ipfy want-po-3pl obj leave-gen-1pl.incl
'They wanted us to leave.' (lit. They wanted our leaving.) (Frajzyngier & Shay 2002: 480)

Snà-n-sn-í
[tá gbw-ży-tán tá ngårílýay-á blá.]
hear-3-hear obj slaughter-po-3pl.poss obj neck-gen cow
'I heard them slaughter cattle.' (Frajzyngier & Shay 2002: 470) (lit. I heard their slaughtering the cow’s neck?)

Ref Mod:
ghùrùm [tá lâ-gbw-i ndá mà xàdák]
hole COM go-dso-ref assoc in ground
'a hole that went deep into the ground'
màmú sán ngèbhám [tá kl-df-tá màràkw xìs]
exist certain chief COM take-up-ref wife two
'There was once a chief who married two wifes.' (Frajzyngier & Shay 2002: 470 406)

tá (IPFV) + verbal root form (+ ka (COMP)-subject pronoun)
Remark: The predicate takes dependent aspect marking, but is otherwise balanced.
This construction is used for (imperfective) object relative clauses only.
Functional distribution: Rig: Ref Mod = PoS (small adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained (dependent imperfective aspect)
Nominal categories: None
Argument encoding: SENT - Ø (gapping)

Example:
Ref Mod:

Wúyá s kwì [tà klá-ghá-tà-ŋnì, ká-ə̀ŋ]
here thing take-2SG-REF-1PL.EXCL comp-3PL

‘Here is the thing that we give you.’ (Frajzyngier & Shay 2002: 409)

tà (IPFV) + clause with nominalized verb
Remark: Used for subject relative clauses only.
Functional distribution: Rig: Ref Mod = PoS (small adjectives)
Structural type: 2 (D-SENT)
Verbal categories: Only imperfective aspect retained, no other affixes indicating role/number of argument(s).
Nominal categories: None
Argument encoding: Ø - SENT (gapping)

Example:
Ref Mod:

Màmú ts ə̀m ə́k-xà k̄érí [tà irí ndà tsì]
exist enemy-pl:gen Kderi ipfv envy.
nmlz assoc 3sg

‘There were enemies of Kderi who envied him.’ (Frajzyngier & Shay 2002: 404)

Pred-a (dependent perfective aspect):
Remark: Used for (perfective) object relative clauses only.
Functional distribution: Rig: Ref Mod = PoS (small adjectives)
Structural type: 1 (Balanced)
Verbal categories: Retained (dependent perfective aspect)
Nominal categories: None
Argument encoding: SENT - Ø (gapping)

Example:
Ref Mod:

Gítà kàm ná dzà’á plá-ghá-m-plá-ŋnì tà dà-ghá [dzá-ə̀ŋ]
today then dem fut return-2SG-REF-1PL.EXCL OBJ father-2SG kill:dep.PFV-3PL

Today, we will avenge your father whom they have killed. (Frajzyngier & Shay 2002: 410)

Mandarin Chinese
Unmarked clause
Remarks: In Ref Mod function, the relativized element is gapped; in Pred Mod function it can be left unexpressed under co-referentiality.
In Ref Mod function, this construction is called a "descriptive clause". It is semantically similar to a relative clause marked with de (see below), although according to Li & Thompson it constitutes a separate assertion: "Semantically, a descriptive clause simply adds another assertion to the first one. A relative clause, on the other hand, is a part of the noun phrase naming the item in question, and as such expresses a pre-established class of items with the property it names." (Li & Thompson 1981: 618)
Functional distribution: Flex: Ref Head, Ref Mod, Pred Mod ≠ PoS (nouns, small adjectives, small manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping/co-referentiality)

Examples:
Ref Head:

Wō pànwàng [nǐ kuài yìdiàn bìye]
I hope you soon a.little graduate
‘I hope you’ll graduate a bit sooner.’ (Li & Thompson 1981: 599)

Ref Mod:

wó mài-le yì-jìàn yìfu [tài dà]
I buy-PFV one-cl outfit too big
‘I bought an outfit that turned out to be too big.’ (Li & Thompson 1981: 614)
Appendix iii: Dependent Clause Constructions: Key Examples

**Tā yǒu yì-ge mèimei [běn xǐhuan kàn diànyǐng]**
3sg exist one-cl younger.sister very like see movie
'S/he has a younger sister who like to see movies.' (Li & Thompson 1981: 611)

**Pred Mod:**
Tāmen [yòng shǒu] bǐ-fān
they use hand eat-food
'they eat with their hands.' (Li & Thompson 1981: 597)

**Clause + de**
Functional distribution: Flex: Ref Mod, Pred Mod ≠ PoS (small adjectives, small manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping/co-referentiality)

**Examples:**
**Ref Mod:**
[Jīntiān yīng de] qián fù fāng-zū
today win conn money pay house-rent
'The money we won today goes to pay the rent.' (Li & Thompson 1981: 581)

**Pred Mod:**
[nǐ pāo [de běn kuài]]
you run conn very quick
'You run very quickly.' (Li & Thompson 1981: 625)

**Pred-zhe (DUR)**
Remarks: The durative aspect marker -zhe can be used in the first of two clauses to signal that one event provides a durative background for another event.
Since this construction occurs only with (unexpressed) co-referential subjects, it is analyzed as a Ø-SENT construction, even though there is no independent evidence for deranking (since there is no verbal inflectional morphology).
Functional distribution: Rig: Pred Mod = PoS (small manner adverbs)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

**Examples:**
**Pred Mod:**
Xiǎo gǒu [yáo-zhe wěi] pāo le
small dog shake-DUR tail run crs
'The small dog ran away wagging its tail.' (Li & Thompson 1981: 223)

**Tamil**
Pred-atu
Remark: With locative case or postposition poola 'like' the construction can be used in Pred Mod function.
Functional distribution: Rig: Ref Head (+ case/postposition also Pred Mod: simultaneity and simulative)
= PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: Agreement is lost; tense is retained; some but not all aspect and mood distinctions are retained.
Nominal categories: CASE
Argument encoding: SENT - SENT

**Examples:**
Ref Head:
[Tiruṇa nakayellaam eputukkiṭirukkaratē] naan patteen
thief jewellery-all take-progr-pres-nmlez/acc 1 sec-pst-1sg
'I saw the thief taking the jewels.' (Asher 1982: 20)
The children were busy building a sand-house on the beach.’ (Asher 1982: 21)

My wife makes coffee just as my mother makes it.’ (Asher 1982: 48)

‘Raman told me to come to his house.’ (Asher 1982: 22)

‘They say and it’s no lie, that your people are your people.’ (Childs 1995: 278)

‘They say it’s no lie, that your people are your people.’ (Childs 1995: 278)

‘He really thinks that if he cries he will be free.’ (Childs 1995: 280)

‘You should tell him (that) he should come quickly!’ (Childs 1995: 281)
Appendix iii: Dependent Clause Constructions Key Examples

Unmarked deranked clause:
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No TAM
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

Examples:
Ref Head:
`í ló [kísìé pë́kùò].`
I stay Kisi study
'I continued to study Kisi.' (Childs 1995: 281)

ò dèmál [sàá wàlló tòsàl kpì́ŋ]
he fail Saa work do ideoph
'He failed completely to do the work for Saa.' (Childs 1995: 281)

Noun class pronoun + clause (+ RSP) + noun class suffix
Remarks: The head noun loses its noun class suffix, which is supplanted by a noun class pronoun (unless the noun belongs to the o-class) and the noun class suffix appears at the end of the relative clause. The relativized item is gapped, but when the syntactic role of the head noun is oblique, or when the head noun is distant, a resumptive pronoun can appear in the relative clause.
Functional distribution: Rig: Ref Mod = PoS (small/derived adjectives)
Structural type: 2 (D-SENT)
Verbal categories: No TAM
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping)

Examples:
Ref Mod:
mɛ̀ŋ mà [é kòlf-áŋ mà nìŋ lè]
water pron he drink-suff pron good neg
'The water that he drank wasn't good.' (Childs 1995: 286)

When the head noun belongs to the o-class, there is no pronoun, only a suffix:
Wàndá [tòsìí cìí́lì]-ó b cò lè háńɔ̀
person inspect towns-suff pron aux again come
'The person who inspects towns will return.' (Childs 1995: 286)

Nung
Unmarked clause
Functional distribution: Flex: Ref Head, Ref Mod, Pred Mod ≠ PoS (nouns, small adjectives, no adverbs)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping/co-referentiality)

Examples:
Ref Head:
Mù̀hng [càng slóng vám đày lè]
you speak two sentences be.able all
'You can speak both sentences.' (Saul & Wilson 1980: 55)

lè̀ mu’hn ca tóo só́nj [bë́hì vú]
than he about again desire do king
'Then he wanted to be king again.' (Saul & Wilson 1980: 52)
Lào hăn mbhn [kém mi lài] then see PL gold be much
‘Then they saw that there was much gold.’ (Saul & Wilson 1980: 111)

Mu’hn vả [pi kẹọ mu’hn ma] he say older brother his come
‘He said that his older brother came.’ (Saul & Wilson 1980: 115)

Ref Mod: mu’hn tẹọ pẹn pẹn tẹ đêk-slào [dày-slào lài] she again become like CL child-girl good-girl much
‘She then became a very pretty girl.’ (Saul & Wilson 1980: 33)

Pred Mod: Tú mā pày [chăhm lài] CL horse go slow very
‘The horse goes very slowly.’ (Saul & Wilson 1980: 95)

Mu’hn càng cỡ [heng lài] cạ técn He speak story strong very all day
‘He spoke loudly all day.’ (Saul & Wilson 1980: 95)

(REL) + gapped clause + DEM/FOC
Remarks: Relative clauses are formed with a demonstrative or a focus marker or both added to the end of the clause and/or a relativizer at the beginning of the clause. When the function of subject/source/locative-directional is relativized, the relativizer is optional. When the object is relativized, the DEM/FOC marking is optional. When the indirect object/beneficiary are relativized, the marker bọ ‘to, for’ is added at the end of the relative clause. With all types of relative clauses, an anaphoric mu’hn is often inserted before the main predicate.

Functional distribution: Rig: Ref Mod ≠ PoS (small adjectives)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT/Ø - SENT/Ø (gapping)

Examples:
Ref Mod:
Lào wang [khi bẹ tẹ] chinh du tẹ bẹ hụ’ kẹ Hong [khi mā tẹ]
Then boy ride goat DEM then take CL goat give man Hong ride horse DEM
‘Then that boy riding the goat gave the goat to that man Hong riding the horse.’ (Saul & Wilson 1980: 16)

kẹ [ti-sọ tọc cọn thọn tẹ nị] man REL lost CL stone DEM FOC
‘The man who had lost the stone.’ (Saul & Wilson 1980: 78)

Áhn bẹn [câu-nị-ọe-kbọi nị] CL house I.run.out DEM
‘The house (that) I was running from.’ (Saul & Wilson 1980: 80)

kẹ [ti-sọ ụọ ụọ ụọ ụọ hụ’ nị] mu’hn man REL I.gave CL book to FOC he
‘The man I have the book to, he ….’ (Saul & Wilson 1980: 79)

Ông kẹ [câu ụọ ạhn tu hụ’ nị] CL man I.bought CL hat for FOC
‘The man for whom I bought the hat.’ (Saul & Wilson 1980: 79)

Kat + clause
Remark: This construction is a simultaneity clause.
The subject is expressed in both the matrix and the dependent clause.
Functional distribution: Rig: Pred Mod ≠ PoS (no lexical strategy)
Structural type: 1 (Balanced)
Verbal categories: Not applicable
Nominal categories: None
Argument encoding: SENT - SENT/Ø
Example:
Pred Mod:
Lō [mu'hn kät ḥabn cáb sl'u] ma'hn láo lái
then he when see cl. tiger he fear much
‘Then when he saw the tiger he was very afraid.’ (Saul & Wilson 1980: 110).

Garo
Clause + in-e
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT- SENT

Examples:
Ref Head:
[Pir-ga-cha song ni-to-a in-e] ang-a chanchi-a
Pirgacha village beautiful-neut COMP I-NOM think-neut
‘I think Pirgacha village is beautiful.’ (Burling 2004: 318)

[Bi-a sok-ba-ku-ja in-e] ang-a chanchia
he-nom arrive-here-not-yet COMP I-NOM believe
‘I believe that he has not yet arrived here.’ (Burling 2004: 319)

Pred-a
Remark: With the suffix git-a ‘like’, this construction can be used as a simulative clause.
Functional distribution: Rig: Ref Head (+ extra suffix also simulative) = PoS (nouns)
Structural type: 3 (D-ALT)
Verbal categories: No tense/aspect
Nominal categories: CASE
Argument encoding: POSS - SENT

Examples:
Ref Head:
ang-a [bi-ni giit ring'-a]-ko kin-a-a
I-nom he-gen song sing-nmlz-acc hear-neut
‘I hear him singing songs.’ (Burling 2004: 294)

Simulative:
Jal  dong-kan, [bi-ni ha'-rik-a-git].
whatever be.at-imp he-gen wish-nmlz-like
‘Let (him) be according to his wishes.’ (Burling 2004: 297)

Pred-a-ni
Remark: Even though the suffix –ni is homophonous with the genitive case marker, the construction takes (a second) case-marker according to function.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect
Nominal categories: CASE
Argument encoding: Ø - Ø?

Example:
Ref Head:
[agen-a-ni]-ko seng-ja, an-tang dra-erina kam-ko dák-a
talk-nmlz-gen-acc wait-NEG own force-ADV work-ACC do-neut
‘Not waiting for talk (instructions), (he) does the work forcefully himself.’ (Burling 2004: 296)

Pred-na (INF) (+ in-e)
Remark: Used as the same-subject complement of modals, desideratives and predicates of achievement and emotion. With achievement and emotion predicates, the complementizer in-e can optionally be added.
Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect
Nominal categories: None
Argument encoding: Ø-SENT (co-referentiality)
Examples:
Ref Head:
Ang-a  [mi cha-na] man-a
I-nom rice eat-inf can-neutral
'I can eat rice.' (Burling 2004: 320)

Bi-a  [kat-ang-na] ba-sik-ing-a
he-nom run-away-inf want-prog-neutral
'He wants to run away.' (Burling 2004: 320)

Ang-a  [bik-ko nam-et-na (in-e)] cheu-ta dak-no-a
I-nom you-acc good-caus-inf comp try-fut
'I will try to improve you.' (Burling 2004: 321)

Ang-a  [nang-ko nik-na (in-e)] ken-a
I-nom you-acc look-at-inf comp fear-neutral
'I am afraid to look at you.' (Burling 2004: 321)

Pred-kan/ka-na (IMP-INF) (+ in-e)
Remark: -ka is the 3rd person imperative suffix, used here as a subordinating suffix with manipulative predicates. The complementizer in-e is optional with this construction.

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

Example:
Ref Head:
Nang-ko  [boi bre-kan-na (in-e)] ang-a bit-a
you-acc book buy-comp comp I-nom order
'I order you to buy a book.' (Burling 2004: 322)

Pred-gip-a
Functional distribution: Rig: Ref Mod = PoS (no lexical strategy)
Structural type: 3 (D-ALT)
Verbal categories: No tense/aspect
Nominal categories: Case agreement
Argument encoding: Ø/POSS - Ø/SENT (gapping)

Examples:
Ref Mod:
[nok-o pi-sa-ko nang-gip-a] me-tra
house-loc child-acc see-ptc woman
'the woman who saw the child at the house.' (Burling 2004: 301)

[me-ebik-ni skang-o den-gip-a] a-bol
women-gen previously-loc cut-ptc firewood
'firewood that the women chopped previously' (Burling 2004: 299)

Pred-e/-e-ning/-e-min-a
Functional distribution: Rig: Pred Mod = PoS (small/derived manner adverbs)
Structural type: 2 (D-SENT)
Verbal categories: No tense/aspect
Nominal categories: None
Argument encoding: Ø - SENT (co-referentiality)

Example:
Pred Mod:
Na-tok-ko  [nam-e] [ni-e] cha’ja-ad-de, (…)
fish-acc good-conv watch-conv eat-subj-if
'If you don’t watch out well when eating the fish, (…)' (Burling 2004: 314)
lit.: ‘If you don’t eat the fish in a good and watchful manner, …’

[Klemen-ko rim-jol-e] ba-ti-cha re-ang-a
Clement-acc take-along-conv market-loc go-neutral
'(I) go to the market, taking Clement along.' (Burling 2004: 315)
Appendix iii: Dependent Clause Constructions Key Examples

**Krongo**

*ânti (QUOTE) - tî ŋ clause*

Remark: This construction is used only with utterance verbs. The quotative marker is part of the matrix clause; the particle *tî* is suffixed to the first word of the dependent clause.

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 1 (Balanced)
Verbal categories: Retained
Nominal categories: None
Argument encoding: SENT - SENT

**Example:**
Ref Head:  
*g-ânâ Lâwâla ânti tî áanây ânti*
conn:masc-cop Luwaala inf: say-tr dat: 3. pl QUOT  
FL-pst-make-prt you.pl. plan conn:masc-IPFV:good

‘And Luwaala says to them, that they have made a good plan.’ (Reh 1985: 354)

*(t-)*Pred (*nominalization/infinitive*)

Remark: For class II verbs, the nominalization and the infinitive have the same form; for class I verbs, the infinitive is formed without the prefix. The construction (NMLZ and INF) takes case marking when not in object function (the accusative is zero-marked). Infinitives in object function with subject-to-object-raising are locative case-marked. When an overt subject is a full noun, the subject takes a possessive prefix. If not, it is expressed by a dependent person suffix.

Functional distribution: Rig: Ref Head = PoS (nouns)
Structural type: 3 (D-ALT)
Verbal categories: No agreement/ mood, tense, aspect and valency/voice are retained
Nominal categories: CASE
Argument encoding: POSS - SENT/Ø - SENT

**Examples:**
Ref Head:  
n-âtaasà àʔ àŋ [t-óshó-ókò-tú tâamâ àʔ àŋ]  
1/ 2- impf: want I nmlz-IPFV: cook-BEN-TR.2SG(DEP) thing dat:1

‘I wanted you to cook something for me.’ (lit. I wanted your cooking something for me.) (Reh 1985: 333)

n-âtaasà àʔ àŋ [t-iminî kà- Sârrà àʔ àŋ]  
1/ 2- IPFV: want I nmlz-IPFV: help poss-Sarah me

‘I want Sarah to help me.’ (lit. I wanted Sarah’s helping me.) (Reh 1985: 333)

Ácáamà-ânti kâaw àʔ àŋ  
masc: impf: speak-BEN person dat:1 gen-nmlz-fut-3SG(DEP) loc-INF:go dat: Tabanya

‘The man says to me that he will go to Tabanya.’ (lit. The man speaks me of his (future) going to Tabanya.) (Reh 1985: 334)

n-âtaasà àʔ àŋ [âkî âuada] n-âtaasà àʔ àŋ  
1/ 2- impf: want I inf: eat.meat 1/ 2- impf: want I

‘I want to eat meat.’ (Reh 1985: 335)

ù àŋ [k-âlalâdà àŋ kí-ninî mó-ânti]  
you(acc) loc-INF: teach me loc-language gen-homeland

‘I want you to learn me Krongo.’ (Reh 1985: 335)

**CONN-Pred (+ RSP)**

Remarks: The construction is marked by a connector-prefix, which is fused with morphemes encoding agreement.

In Ref Mod function, if the relativized item is not the subject of the DC, an anaphoric pronoun is used, which takes case according to function.

In Pred Mod function, the DC obligatorily has imperfective aspect.

Functional distribution: Flex: Ref Mod, Pred Mod ≠ PoS (no adjectives, small mAdverbs)
Structural type: 2/3 (D-SENT/D-ALT)
Verbal categories: No agreement/mood, tense, aspect and valency/voice are retained.
Nominal categories: Gender and number agreement (+ case agreement of RSP)
Argument encoding: POSS - SENT/Ø - SENT

---
Examples:
Ref Mod:
káaw [m-úú]  
.person conn FEM IPFV walk  
‘the woman who walks’; ‘the walking woman’ (Reh 1985)

1/2-IPFV:love I loc sg clothes conn neut IPFV sew tr poss mother my  
‘I love the dress that my mother is sewing.’ (Reh 1985: 256)

Káaw [m-asålú-ti]  
.person conn FEM IPFV look at 1SG DEF she fem  
‘the woman that I looked at.’ (Reh 1985)

Káaw [m-àasàlàa-tɪ́̇]  
.person conn fem IPFV: look at 1sg dep she fem  
‘the woman that I looked at.’ (Reh 1985)

nóoni  ì́ kí-tùlùnkwáaná  
.know I conn: neut IPFV: laugh dtr I  
‘I laugh joylessly.’ (Reh 1985: 302)

Hixkaryana
Pred nɨ/-thɨ/-hɨ to + -rɨ  
Remarks: The construction is formed with the nominalizer -nɨ, followed by the inflectional suffix -rɨ ‘possessed item’, glossed as POSSD. For past tense, -nɨ is replaced by -thɨ; for negation, the form -hɨ is used. The arguments are marked ergatively: the first argument of a transitive verb (A) is marked by a postpositional phrase with wya (‘to, by’, normally the indirect object marker). The only argument of an intransitive (S) and the second argument of a transitive verb (P) are marked as possessives. Since Hixkaryana is head-marking, the possessive appears on the predicate. In combination with extra elements, the nominalization construction is used in adverbial functions. The relevant elements are the “de-nominalizing relator” me for manner, and the postposition wyaro ‘like’ for simile constructions.

Functional distribution: Rig: Ref Head (+ extra element me, wyaro ‘like’ also Pred Mod) = PoS (nouns)

Structural type: 3 (D-AL T)

Verbal categories: Aspect is lost; 3 out of 7 tense distinctions (non-past, simple past, remote past) are retained.

Nominal categories: Possessive prefixes

Argument encoding: POSS - OBL

Examples:  
Ref Head:
ra-wanota-nɨ-ri  
1 POSS sing NMLZ POSSD  
‘my singing’ (Derbyshire 1979: 165)

Amna-y-omok-ɨ-thɨ-ri  
1-3 POSS come PST NMLZ POSSD  
‘our coming (in the past)’ (Derbyshire 1979: 165)

i-ntki-bto-ri-komo  
3 POSS sleep NEG NMLZ POSSD COLL  
‘their not going to sleep’ (Derbyshire 1979: 166)

Karyhe kamara Q-to-bti-ri  
quickly jaguar 3 POSS go PST NMLZ POSSD  
‘the going quickly of the jaguar.’ (Derbyshire 1979: 23)
Appendix iii: Dependent Clause Constructions Key Examples

**Koseryebaha, [r-eburka-nt-r] bona**

I.am.afraid, 1.Poss-fall-NMLZ-Possd to
‘I am afraid of falling.’ (Derbyshire 1979: 24)

**Pred Mod:**

*[Terrureyero ro-czavasta-nt me] kwezeyaha*

loudly my-singing-NMLZ DNMLZ I.take.a.bath
‘I take a bath singing loudly.’ (Derbyshire 1979: 28)

**Uro ehtxemako,** [o whya Waraka yahxeman-ti-ri uyaro] me treat.with.medicine by.you Waraka treat-pst.NMLZ-Posssd like
‘Treat me with medicine just like you treat Waraka.’ (Derbyshire 1979: 31)

---

**Slave Clause**

With the verbs ‘want’ and ‘say’ the complementizer can be omitted.
The construction is used in Ref Mod function when the relativized item expresses a time or location.

**Functional distribution:** Flex: Ref Head, Ref Mod ≠ PoS (nouns, no adjectives)

**Structural type:** 1 (Balanced)

**Verbal categories:** Retained

**Nominal categories:** None

**Argument encoding:** SENT - SENT

**Examples:**

**Ref Head:**

[Ɂelá táh Ɂa] kodeyihshá yíle
boat 3.is comp 1sg.knew neg
‘I didn’t know that the boat came in.’ (Rice 1989: 1245)

**Ref Mod:**

hej Ɂs Ɂa [Jim Ɂeghálayeda Ɂa] sa
here probably Jim 3.work relat foc area.exist
‘This must be the place where Jim works.’ (Rice 1989: 1317)

---

**Clause + gha/gú**

**Functional distribution:** Rig: Ref Head = PoS (nouns)

**Structural type:** 1 (Balanced)

**Verbal categories:** Retained

**Nominal categories:** None

**Argument encoding:** SENT - SENT

**Examples:**

**Ref Head:**

[Ɂihalá lits’e’Ɂa gha] kegőfa
canvas one.folds comp is.difficult
‘It is difficult for me to fold canvas.’ (Rice 1989: 1246)

**Clause + il/s’il/i**

**Functional distribution:** Rig: Ref Mod ≠ PoS (no adjectives)

**Structural type:** 1 (Balanced)

**Verbal categories:** Retained

**Nominal categories:** None

**Argument encoding:** SENT/Ø - SENT/Ø (gapping)

**Rõli** [bõhá Ɂ]
boat 3.is.big relat
‘the big boat’ (Rice 1989: 1309)
Parts of Speech and Dependent Clauses

 Clause + $\text{gháré}$

Remark: The meaning of the adverbial subordinator is instrumental ‘by means of’.

Functional distribution: Rig: Pred Mod ≠ PoS (no manner adverbs)

Structural type: 1 (Balanced)

Verbal categories: Retained

Argument encoding: SENT - SENT

Example:

Pred Mod:  
$dene \ [\text{gháts'eyeda} \ \text{gháré}] \ yá \ gá \ Ɂ \ \text{ets'ederet} \ ę$

‘one watches top one teaches refl’

‘One learns through watching.’ (1063)

Nivkh

Pred-vut / vur (‘retelling converb’):

Used only for the complements of utterance predicates.

The allomorphy is conditioned by person agreement: -t for 1st person (SG/PL), -r for 2nd or 3rd person.

Functional distribution: Rig: Ref Head = PoS (nouns)

Structural type: 1 (Balanced)

Verbal categories: Retained

Nominal categories: None

Argument encoding: SENT - SENT

Example:

Ref Head:  
$\text{If} \ [\text{imn-ay} \ \text{als} \ p'\text{-ny-vur}] \ \text{it-}t'$

‘He they-dat/acc berry pick-fut-comp say-fin’

‘He said [that] they would pick berries.’ (Gruzdeva 1998: 57)

Unmarked clause

Functional distribution: Rig: Ref Head = PoS (nouns)

Structural type: 1 (Balanced)

Verbal categories: Retained

Nominal categories: None

Argument encoding: SENT/Ø - SENT (co-referentiality)

Examples:

Ref Head:  
$\text{Rask} \ [\text{čo} \ \text{nI-}d] \ \text{eska-}d$

woman fish eat-fin dislike-fin’

‘The woman does not like to eat fish.’ (Gruzdeva 1998: 49)

Pred without FIN SUFFIX (‘participles’)

Functional distribution: Rig: Ref Mod ≠ PoS (no adjectives)

Structural type: 1 (Balanced)

Verbal categories: Retained except for FIN suffix

Nominal categories: None

Argument encoding: SENT/Ø - SENT/Ø (gapping)

Example:

Ref Mod:  
$\text{[NI - zosk]} \ t'a \ Ɂ \ \text{tyr} \ t'sy \ p'\text{-i-}d'$.  
I break:ptc knife table on be-fin

‘The knife which I have broken is on the table.’ (Gruzdeva 1998: 50)
Appendix III: Dependent Clause Constructions Key Examples

**Pred -t/-t/-r**

**Remark:** The allomorphy is triggered by person agreement: -t for 1st person (SG/PL) subject/agent, and -r for 2nd or 3rd person. The other variants are dialectological.

**Functional distribution:** Rig: Pred Mod ≠ PoS (no manner adverb)
**Structural type:** 2 (D-SENT)
**Verbal categories:** No tense/agreement, aspect and mood can be retained
**Nominal categories:** None
**Argument encoding:** Ø - SENT (co-referentiality)

**Examples:**

Pred Mod:

Haimam t’axkyt n’a -kav -tUr
old man straight eye instr blink -prt -neg conv
t’yr-d
‘The old man looked straight, not blinking [his] eyes.’ (Gruzdeva 1998: 54)

hoor [en-gu-r] mzy-d
therefore fast-caus -conv:2/3 SG descend -IND
‘Therefore he descended quickly.’ (Matissen & Drossard 1998: 44)

**WEST GREENLANDIC**

**Pred -niq**

**Remark:** With intransitive predicates, the S takes relational case (RELC) and triggers possessive agreement. With transitive predicates, there are several possibilities. When only the P is present, the dependent predicate can be unmarked or marked with the passive suffix. The P appears in relative case and triggers possessive agreement. When both arguments are present, the dependent predicate is marked with a semi-transitivizing suffix (SEMTR). The A appears in the relative case, and the predicate bears the corresponding possessive suffix, while the P is in the instrumental. Thus, the possessive argument in the nominalization always corresponds to the absolutive argument in the corresponding finite clause\(^{12}\).

**Functional distribution:** Rig: Ref Head = PoS: nouns
**Structural type:** 3 (D-ALT)
**Verbal categories:** No mood, person/number agreement. Tense can be retained, but is usually not expressed.
**Nominal categories:** Case, nominal agreement
**Argument encoding:** RELC/POSS - Ø/INSTR

**Examples:**

Ref Head:

[umiarsu-up qassi-nut tikin-ni-ssa-a] nalungil-ara
ship-rel how-many all arrive-nmlz-fut-3 SG POSS know-1 SG 3 SG IND
I know when the ship will arrive. (Fortescue 1984: 115)

Anguitq tuqunnira
man:reLC kill:nmlz:3SG POSS
‘the killing of the man’ (i.e. he is killed) (Fortescue 1984: 46)

nalu-aa [qinnuta-ata qanuq naammassi-niqar-ni-ssa-a]
not know-3 SG:3 INC how implement-pass-nmlz-fut-3 SG POSS
‘She didn’t know how his request would be implemented.’ (Fortescue 1984: 45)

[tkinnergum-mi-nik tuqut-i-ir-a] tusan-para
friend-his refl instr kill:seMT -NMLZ:3 SG POSS hear-1 SG 3 PL IND
‘I heard of his killing his friend.’ (Fortescue 1984: 45)

**Pred -ta/sa (passive participle)**

**Remark:** Overt subjects are in relative or ablative case, non-gapped objects remain SENT.

**Functional distribution:** Rig: Ref Mod ≠ PoS (no lexical strategy)
**Structural type:** 3 (D-ALT)
**Verbal categories:** No tense, mood, verbal agreement.
**Nominal categories:** Nominal agreement (number and case)
**Argument encoding:** RELC/ABL/Ø - SENT/Ø (gapping)

**Examples:**

Ref Mod:

angum-mut [qippasuq naspi-ta-e] -nut tunniup-para
man-all SG yesterday meet-pass -PTC -1 SG POSS-all SG give-1 SG 3 SG IND
‘I gave it to the man I met yesterday.’ (Fortescue 1984: 49)

\(^{12}\) Strictly speaking, therefore, West Greenlandic has no transitive nominalizations. (see Koptjevskaja-Tamm 1993).
angut  [aallasa-a  tigu-sa-ni]  nigur-niar-paa
man  gun-his  take-pass.ptc-3poss.refl.  avoid-try-3sg.3sg.ind
‘He tried to avoid the man whose gun he had taken.’ (Fortescue 1984: 53)

Participial mood clause
Remarks: The participial mood endings are parallel to the indicative ones, but built up on mood-marker su
in the intransitive and gi in the transitive.
In Ref Head function, the construction is used for different-subject complements. When it functions as
such, the dependent clause is morphologically marked as an object.
In Ref Mod function, the construction must be intransitive and can only take 3rd person inflection. Relative
clauses that modify the subject of the matrix clause agree in (relative) case with their head.
When marked for instrumental case the participial mood can also be used in Pred Mod function.
Functional distribution: Flex: Ref Head, Ref Mod, + CASE also Pred Mod ≠ PoS (nouns, no lexical
strategy for modifier functions)
Structural type:  1 (Balanced)
Verbal categories: Retained
Nominal categories: Nominal agreement in Ref Mod function
Argument encoding: SENT/Ø - SENT/Ø (gapping)

Examples:
Ref Head:
Ilisima-vaa  [urni-ssa-giga]
know-3sg.3sg.ind  come.to-fut-1sg.3sg.ptc
‘He1 knew I would come to him2.’ (Fortescue 1984: 36)
pasii-nggil-lainnar-para  [ilaas-juma-sutit]
understand-not-completely-1sg:3sg.ind  come.along-wany-2sg.ptc
‘I didn’t understand at all that you wanted to come along.’ (Fortescue 1984: 36)

Ref Mod:
Niviariag  [kalalitsut  ilinnia-lir-suq]
girl    Greenlandic  learn-begin-intr.ptc
‘the/a girl who has begun learning Greenlandic.’ (Fortescue 1984: 49)

Pred Mod (with case):
[Kulturi-kuut  sunniuti-par-iar-tu]-mik
culture-pros  effect-have-well-intr.ptc-instr
‘having considerable cultural effect’ (Fortescue 1984: 56)

Contemporative mood clause
Remark: The contemporative or conjunctive mood for same-subject clauses. The contemporative is formed
with the mood-marker lu/lla followed by person markers (1st, 2nd, 4th) or 3rd person object markers.
Functional distribution: Flex: Ref Head, Pred Mod ≠ PoS (nouns, no lexical strategy for Pred Mod function)
Structural type:  1 (Balanced)
Verbal categories: Retained
Nominal categories: Nominal agreement in Ref Mod function
Argument encoding: SENT/Ø - SENT

Examples:
Ref Head:
Uqar-sinnaa-vunga  [tama-asa uuqattaar-sima-llugit]
say-can-1sg.ind  all-3pl  try-ppv-1sg-3pl.cont
‘I can say that I have tried them all.’ (Fortescue 1984: 40)

Pred Mod:
[Iririr-lunng]  isir-pung
do.hurriedly-1sg.cont  go.in-1sg.ind
‘He entered quickly/in a hurry.’ (Fortescue 1984: 55)

[uqa-nngua-rant]  ini-mi-nut  isir-pung
say-little-4sg.neg.cont  room-his.refl  go.in-3sg.ind
‘He entered his room without even speaking.’ (Fortescue 1984: 62)
Linguistic typology is concerned with cross-linguistic comparison. The basic problem encountered by typologists is that, in order to compare language-specific structures, a *tertium comparationis* is needed that is independent of structure. Therefore, typologists choose a particular *functional* domain, and then examine the *formal* structures that languages use to express certain distinctions within that domain. When they have identified the range of cross-linguistic formal variation, typologists search for *universals*: dependency relations between certain structural traits that apply to all human languages. Within the *functionalist* approach to typology, adhered to in the present study, it is assumed that such universals can ultimately be explained in terms of the function of languages, namely to encode communication (Croft 2003).

Three basic communicative or *propositional* functions may be distinguished: Reference, predication, and modification. *To refer* means to identify a referent, i.e. an entity that one wants to talk about. *To predicate* is to ascribe something to a referent, i.e. to report that this referent is involved in some state of affairs. *To modify* is to enrich either reference or predication, by means of expressing an additional feature of either a referent or a predication (Searle 1969, Croft 2001: 66).

The basic propositional functions of reference, predication, and modification may be expressed by linguistic forms or constructions of various types. Example (1) shows that a referential expression such as the
object argument of the predicate saw can take the form of a simple lexical construction (*a car* in (1a)), or of a more complex clause-like construction (*that John has bought a new car* in (1b)):

(1)  
   a. I saw a car.
   b. I saw that John has bought a new car.

This study is concerned with the two construction types illustrated in (1): lexical constructions or *parts of speech* and clause-like constructions or *dependent clauses*.

*Parts of Speech* (henceforth PoS) is the traditional term used to refer to the major classes of lexemes or content words that are distinguished in a particular language (Schachter & Shopen 2007: 1). Typically, the class of *nouns* is used to express the function of reference (cf. *car* in example (2a)); *verbs* are used for predication (cf. *wash* in (2b)); *adjectives* for modification of referential expressions (cf. *red* in (2c)); and *adverbs* for modification of predications (cf. *quickly* in (2d)).

(2)  
   a. I wash the *car*.
   b. I *wash* the car.
   c. I wash the *red* car.
   d. I *quickly* wash the car.

However, not all languages follow this particular pattern, in which every PoS class is dedicated to the expression of a single function. Alternatively, languages may display PoS classifications in which two or more functions can be expressed by members of the same class. In Turkish, for instance, there is one class of lexemes that may be used to express a referential unit (cf. example (3a) below), or to modify a referential or a predicative expression (cf. examples (3b) and (3c), respectively). The translations of examples (3b) and (3c) show that in English the use of the lexical noun *beautiful* in modifying functions requires derivation with the suffixes -ful and -ly.

*Turkish* (Göksel & Kerslake 2005: 49)

(3)  
   a. *güzel-im*
      
      beauty-1poss
      
      ‘my beauty’
b. güzel bir kopek
   beauty ART dog
   ‘a beautiful dog’

c. güzel konuştu
   beauty s/he.spoke
   ‘S/he spoke beautifully’

Adopting a terminology proposed by Hengeveld (1992), I call parts of speech that are functionally specialized rigid, and those that can express multiple functions flexible.

As regards dependent clauses (henceforth DCs), three types are traditionally distinguished: A complement clause functions as a referential unit, expressing an arguments of the main clause predicate (cf. (1b) above); A relative clause functions as a modifier of a main clause argument; and an adverbial clause functions as a modifier of the main clause predicate (Thompson et al. 2007: 238).

Like PoS classes however, DCs can be rigid, i.e. specialized for the expression of a single function, or flexible, i.e. able to express more than one function. For instance, the DC construction marked by do in the West-Papuan language Abun can only be used in referential function, i.e. as a complement clause (cf. example (4)). In contrast, the Basque DC construction marked by -en can be used in both referential and modifying function, i.e. as a complement clause and as a relative clause (cf. examples (5a) and (5b), respectively).

**Abun** (Berry & Berry 1999: 167)
(4) Ji nut [do an bi obat ne nggi wa
   1sg think COMP 3sg poss magic DET strong for
   be an gu bi ré.]
   later 3sg kill 1sg this
   ‘I thought that his magic was strong [enough] for him to kill me.’ (167)

**Basque** (Hualde & Ortiz de Urbina 2003: 646, 764)
(5) a. Banekien [Mikel berandu etoriko z-en]-a
   knew Mikel late arrive.FUT aux-COMP-DET
   ‘I knew that Mikel would arrive late.’
b. \[ Pellok \ ekarri \ du-en \] dirua \ galdu \ dut \\
Peter.erg bring aux-rel money.det lose aux \\
‘I lost the money that Peter brought.’

Apart from their possibilities with regard to the expression of propositional function(s), DCs – being complex constructions – can also be studied from the perspective of their internal morpho-syntactic properties. Usually, these properties are described in comparison to the properties of ordinary independent clauses in the particular language under analysis. Consider the difference between examples (6a) and (6b):

(6) a. John regrets [that Paul will move to Finland].  
    b. John regrets [Paul’s moving to Finland].

The dependent clause in (6a), marked with the subordinator that, expresses future tense and the subject Paul is encoded in the same way as it would be in an independent clause. This type of dependent clause, which structurally resembles an independent clause, is called balanced. In contrast, the dependent clause in (6b) is marked by the special -ing form of the predicate, it cannot express tense, and it has a subject that is encoded as a possessor (Paul’s). This second type of dependent clause, the structure of which deviates to a certain extent from the structure of an independent clause, is called deranked (Stassen 1985).

The aim of the present research is to investigate the relationship between the functional flexibility or rigidity as displayed by a particular language’s PoS classes on the one hand, and by its balanced/deranked DC constructions on the other hand. More specifically, this study tries to discover to what extent it is possible to predict, on the basis of a language’s PoS system, what the functional potential of its DC constructions will be. This goal is approached by means of an investigation of the PoS classes and DC constructions of a genealogically and geographically balanced sample of 50 languages (see Appendix 1).

The first part of this book (Chapters 2 – 4) provides the theoretical and methodological background of the study. Chapter 2 outlines a number of theoretical issues related to the definition and comparison of PoS classes within and across languages. It starts out with a characterization of possible approaches to PoS typology in terms of the type of criteria on which PoS class definitions are primarily based, distinguishing between semantic,
morphological, syntactic, and pragmatic criteria (Givón 2001: 49; Sasse 1993b:196).

Chapter 2 continues with a detailed discussion of two (relatively) recent approaches to PoS typology: Hengeveld's pragmatic-syntactic theory of parts of speech and Croft's universal-typological theory of parts of speech. The former defines PoS exclusively in terms of the propositional function(s) that they can express (Hengeveld 1992, Hengeveld et al. 2004, Hengeveld & Van Lier 2008, 2009). The Hengeveldian model defines a functional space in terms of two interacting dimensions. The first dimension involves the distinction between predication and reference; the second concerns the opposition between heads and modifiers. As regards the first parameter, the function of predication is assumed to be privileged in relation to the function of reference, since the communicative act of referring presupposes that something is predicated of a referent. With respect to the second parameter, the assumption is that obligatory heads are privileged in relation to optional modifiers. In addition, these two hierarchical parameters are ranked with respect to one another: Since all languages apparently have a predication-reference distinction at some level of the grammar, while the function of modification does not seems to be relevant in the grammar of all languages, it is expected that the predication-reference parameter is primary in relation to the head-modifier parameter. On the basis of this hierarchical model, a number of predictions are made concerning possible and impossible types of PoS classifications, in terms of the type and amount of propositional functions that they can express.

The Hengeveldian approach to PoS typology can thus be characterized as taking into account syntactic-pragmatic criteria only. This narrow approach has been criticized by Croft (2000, 2001, 2005), who defines PoS in terms of a broader set of criteria, including semantic meaning and the ability to express certain morphological categories. In particular, Croft defines PoS classes as prototypical combinations of semantic classes and propositional functions: A prototypical noun denotes an object and functions referentially; a prototypical adjective denotes a property and functions as a modifier; and a prototypical verb denotes an action and functions predicatively. These typological prototypes are claimed to correlate cross-linguistically with relative degrees of morphological markedness: Categories that represent prototypical combinations take less function-indicating markers and can express more morphological distinctions than non-prototypical combinations.
Croft’s approach yields a typology in which PoS classes are not directly comparable across languages. On the other hand, the cross-linguistic comparability of PoS classes defined according to Hengeveld’s approach is achieved at the cost of ignoring specific sets of linguistic facts. As such, the difference between these two approaches touches upon a fundamental issue in linguistic typology, namely the (non-)existence of universal categories (see Haspelmath 2007; Newmeyer 2007).

In addition, the difference between the Hengeveldian and the Croftian approaches has important repercussions for the debate about so-called *flexible languages*, i.e. languages that presumably lack all parts of speech distinctions, most significantly a basic distinction between nouns and verbs. The most thorny issue in this discussion probably concerns *semantic shift* in flexible lexemes, i.e. the phenomenon that such lexemes may have slightly different meanings in each of the multiple functions in which they can be used. According to some authors the interpretation of flexible lexemes should be fully compositional (Evans & Osada 2005; Croft 2005), while others contend that it may involve idiosyncrasies (Hengeveld & Rijkhoff 2005). At the end of Chapter 2 I propose that, in order to explain the relevant linguistic facts, one must take into account the mismatch between lexical and syntactic categorization in flexible languages: Such languages do have lexical categorization, but of a nature that is irrelevant to the phrase-structural distribution of lexical items (Himmelmann 2007; Don & Van Lier, forthcoming).

Chapter 3 is concerned with the theory and typology of dependent clauses. As already mentioned, DC typology can be approached from two perspectives. First, like PoS, DCs can be defined in terms of the propositional function(s) that they may express. In parallel with the PoS classes predicted by the Hengeveldian model in the previous chapter, in the first part of Chapter 3 I define a number of rigid and flexible DC constructions that are predicted to be possible in actual languages.

Second, DCs can be classified according to their internal morpho-syntactic properties. These properties can be characterized as involving specific combinations of ‘verbal’ features, which are associated with independent clauses, and ‘nominal’ features, which are typical of lexical expressions (cf. example (6) above). Earlier functional-typological studies have shown that both verbal and nominal features display certain hierarchical relations (Bybee 1985; Dik 1989, 1997; Rijkhoff 2002). This is reflected, among other things, by universal ordering patterns of these features relative
to their base unit; a verbal or nominal head. These patterns are presumably iconically motivated, to the extent that they reflect the nature and degree of relevance of the various features to the interpretation of the head: Features that occur closest to the head — such as aspect, tense, and mood (TAM) markers for verbs, and class and number markers for nouns — affect the head's semantics. In contrast, features that are expressed further away from the stem — in particular verbal agreement, and nominal case and definiteness — do not affect the meaning of the stem but are rather relevant to its syntactic and/or pragmatic function.

A number of functional-typological studies have investigated the way in which these verbal and nominal feature hierarchies are combined in the morpho-syntactic expression of DCs (Comrie 1976; Noonan 1985/2007; Lehmann 1988; Koptjevskaja-Tamm 1993; Mackenzie 1996; Croft 1991, 2001; Dik 1997; Cristofaro 2003; Malchukov 2004, 2006; Dixon & Aikhenvald 2006). In Chapter 3 I focus on two recent exponents of this literature: Cristofaro’s (2003) typology of subordination and Malchukov’s (2004) typology of nominalizations.

Cristofaro (2003) identifies a number of correlations between various structural phenomena in DCs that can be summarized as follows: (i) the loss of TAM marking implies the loss of verbal agreement; and (ii) the expression of nominal features implies the loss of verbal features. Cristofaro proposes three types of functional factors underlying these dependency relations, two of which are the well-known functional principles of Economy and Iconicity (Haiman 1983). These principles account for the non-expression of verbal features in DCs, including the non-expression of arguments. However, since neither Economy nor Iconicity can account for the expression of nominal features in DCs, Cristofaro proposes a third explanatory principle. In particular, she claims that DCs are cognitively not construed as independent processes, but rather as things or properties. Since verbal features are relevant for processes but not for things and properties, Cristofaro argues that cognitive thing/property-construal leads to loss of verbal features. In addition, thing/property-construal may lead to the expression of features associated with the grammatical entities that prototypically code things and properties, i.e. nouns and adjectives.

Broadly speaking, Malchukov’s (2004) typological study of clausal nominalizations identifies the same range of cross-linguistic generalizations as Cristofaro’s, but proposes a different functional explanation, which is adopted in the present study. Malchukov shows that verbal and nominal
features at the external end of their respective hierarchies are affected in DC constructions before internal features. According to Malchukov, the functional explanation for this generalization must be sought in the fact that external categories reflect the syntactic and/or pragmatic function of the linguistic unit on which they operate, while internal categories are relevant for the semantic interpretation of their base. DCs can have referential or modifying pragmatic functions, but they do not semantically denote things or properties (rather, they denote higher order entities such as states of affairs or propositions). This explains why external nominal (or adjectival) features, which are relevant to the function of reference (or modification), are expressed in DCs before internal ones, which are relevant to thing/property-denoting entities. In addition, this explains why external verbal features, which are relevant to pragmatic and syntactic aspects of (independent) predications, are the firsts to be lost in DCs.

At the end of Chapter 3, and taking the earlier typological studies as a basis, I define three types of DC constructions, in terms of their particular combination of verbal and nominal features. One type corresponds to balanced DCs, the other two represent different types of deranked DCs (see above). This three-way distinction is combined with the classification of DCs in terms of their functional possibilities, as defined earlier in the same chapter. Thus, I arrive at a typological framework for DC constructions that takes into account both their functional and their internal formal properties.

Chapter 4 rounds off the theoretical part of the book. In this chapter, I first present the composition of the sample of 50 languages that are investigated. Second, I formulate and operationalize the specific hypotheses that are tested on this sample. These hypotheses take the form of predictions about dependency relations between the functional patterns displayed by the PoS classes of particular languages, and the functional patterns of their (different structural types of) DC constructions. Some predictions are formulated in general terms, i.e. they make reference to languages with some flexibility versus no flexibility in their PoS system, and the expected reflection of this difference on the functional properties of DCs in these languages. Other predictions are more specific; they aim at the identification of one-to-one matches between the functional possibilities of particular types of PoS classes and DC constructions. In addition, both types of predictions are formulated first without differentiation for dependent clauses in terms of their internal formal properties, and then in sets of sub-predictions that
make specific reference to balanced clauses, as opposed to (different types of) deranked clauses.

Finally, Chapter 4 presents the method that is used to test these predictions. Each prediction involves two binary parameters; one concerning PoS, the other DCs. In order to test the predictions, the observed frequencies (the number of languages that display a particular value combination of the PoS and the DC parameter) are compared to the expected frequencies (the number of languages with this feature combination that would be expected if the co-occurrence of the PoS and DC features would be purely coincidental). Fisher’s Exact tests are used to calculate whether the deviations between observed and expected frequencies are statistically significant. Whenever a significant result is found, the particular nature of the dependency relation between the PoS and DC traits is further assessed by means of a statistical method developed by Maslova (2003).

The second part of the book (Chapters 5-7) contains the actual typological data. First, Chapter 5 presents the typology of PoS classes as displayed by the languages of the sample. The attested PoS systems are evaluated in light of the predictions made in Chapter 2. It is shown that most predicted PoS systems are found in actual languages, either in their ‘pure’ form, or in combination with another predicted system. In general, these data suggest a reasonable typological adequacy of the implicational map model of PoS as outlined in Chapter 2.

The remainder of Chapter 5 addresses a variety of issues regarding the identification of PoS classes in individual languages, and their cross-linguistic comparison. These issues, which are also touched upon in the theoretical discussion on PoS typology in Chapter 2, include the problem of fuzzy boundaries between PoS classes, the phenomenon of variable distributional patterns displayed by sub-groups of lexical items, and the occurrence of ‘restricted’ lexeme classes, i.e. small, closed classes and classes consisting of derived lexemes. I also discuss the expression of non-verbal predication in the sample languages. In the final section of Chapter 5 I apply the theoretical notions about lexical flexibility, as proposed in Chapter 2, to the relevant languages of the sample.

Chapter 6 presents the classification of the DC constructions attested in the sample languages, in terms of the typological framework developed in Chapter 3. First, the DCs are classified according to the propositional function(s) that they can express. The attested DC types are compared with the DC types predicted in Chapter 3. It is shown that all predicted types
occur in the languages of the sample, except those involving the functional slot for head of a predicate phrase. I suggest that this latter finding may be explained in terms of cognitive-semantic markedness of DCs with respect to the function of independent predication, and/or a structural constraint on the expression of finite verbal morphology on DCs. Second, the DCs of the sample languages are classified according to their internal morphosyntactic properties. On the basis of these data, every DC construction is assigned to one of the three structural DC types defined in Chapter 3. At the end of Chapter 6 these data are combined with the DC classification based on functional distribution into an integrated typology.

Subsequently, in Chapter 7, the data sets presented in Chapter 5 (PoS) and Chapter 6 (DCs) are linked in order to test the predictions about their functional interrelationship, as formulated in Chapter 4. The analyses reveal significant dependency relations between flexible (rather than rigid) PoS and DCs; between very flexible (rather than less flexible) PoS and DCs; and between PoS and deranked (rather than balanced) DCs. The relevant dependency relations are shown to be asymmetrical in nature, and can as such be interpreted as statistical bases for implicational universals.

More specifically, it is shown that the presence of a pervasively flexible deranked DC construction in a particular language requires the presence of a pervasively flexible PoS class in that language. This does not mean, however, that all languages with a large degree of flexibility in their PoS system also display flexible deranked DC constructions. In fact, the results make clear that pervasively flexible PoS classes and deranked DCs are both, i.e. independently of each other, cross-linguistically rare phenomena. Therefore, the generalization that maximal flexibility in the domain of deranked DCs depends upon maximal flexibility in the lexical domain does not have much explanatory power. More interestingly, it is shown that whenever maximal lexical flexibility does occur in a language, this strongly increases the chances of also finding the other rare phenomenon: maximally flexible deranked DCs. This shows that there is indeed a strong tendency for the PoS and DC parameters to have the same value.

In contrast to the result obtained for pervasively flexible PoS and deranked DCs, no significant dependency relations can be established between the presence of less flexible and rigid PoS classes on the one hand and deranked DCs with the same functional behaviour on the other hand. In addition, it is shown that the functional patterns of balanced DCs are not related to those of PoS classes.
In the third and final part of the book (Chapters 8 and 9) I further discuss the results of the study and present its conclusions. First, in Chapter 8, the findings of Chapter 7 are reconsidered, taking a specific functionalist perspective that makes reference to the overall complexity of language systems and the various ways in which this level of complexity can be attained. More specifically, the proposed explanatory framework builds upon recent functional-typological research, which has advanced the idea that flexibility and rigidity are relative rather than absolute notions that may be applicable in various degrees to specific construction types pertaining to different levels of the grammar. These studies suggest that the categorial specificity of linguistic constructions increases—or their flexibility decreases—when they become structurally more complex. This generalization has been termed the Principle of Increasing Categoriality or the Principle of Staggering Level-dependent Categoriality (Haig 2006, Lehmann 2008). This principle can be regarded as a specific instance of an even more general functional principle regarding complexity in language systems, namely that flexibility or multifunctionality in one area of the grammar must be counterbalanced or “traded off” by rigidity or categorial specificity in another area, in order to guarantee the identifiability of the function of any linguistic unit within an actual utterance. This has been termed the Principle of Functional Transparency (Frajzyngier & Shay 2003, cf. Hengeveld et al. 2004; Sinnemäki 2008).

In Chapter 8 I argue that deranked (but not balanced) DCs can be regarded as secondary constructions, derived from primary unerived lexical constructions. Under this assumption, the Principle of Increasing Categoriality predicts that the degree of functional flexibility of any deranked DC construction should be either equal to or smaller than the flexibility of a PoS class that can express at least one of the same function(s). This prediction is born out for virtually all relevant constructions in the sample languages. Thus, the Principle of Increasing Categoriality explains the finding of Chapter 7 that flexibility in the domain of deranked DCs is dependent on flexibility in the domain of PoS classes. This dependency relation can now be interpreted as a constraint, imposed by the amount of flexibility attested in the PoS system of a particular language, on the maximal amount of flexibility (or minimal amount of categorial specificity) that can be displayed by the deranked DC(s) of that language.

Furthermore, the Principle of Increasing Categoriality sheds light on the unexpected findings of Chapter 7: (i) the lack of correlations between the presence of weakly flexible PoS classes and deranked DCs with the
same functional possibilities in particular languages; and (ii) the lack of correlations between various types of rigid PoS classes and the corresponding rigid deranked DCs.

The first finding can be understood as involving a decrease in flexibility on the part of the DC construction, as compared to the PoS class. In particular, while the relevant PoS classes are weakly flexible to the extent that they can express two propositional functions, the DCs are rigid constructions; they can be used in just a single function. Notably, a similar explanation can be offered for the fact that even languages with pervasively flexible PoS classes display rigid deranked DC (alongside flexible deranked DCs). Both findings are in accordance with the prediction that deranked DCs can be either as flexible as or less flexible than PoS classes, but not more flexible.

Second, the lack of a correlation between rigid PoS classes and rigid deranked DCs can be interpreted as follows. As expected, whenever a language has a deranked DC construction that can appear in a function for which a rigid PoS class is available, then the DC construction in question is rigid as well. Again, this confirms that a deranked DC cannot exhibit a greater degree of flexibility than a PoS class. However, instead of rigid deranked DCs, languages may employ balanced DCs. The latter, which are presumably not derived from lexical categories, were shown in Chapter 7 to not display any functional relation with PoS classes.

Despite the fact that balanced DCs and PoS are thus neither formally nor functionally related, I argue at the end of Chapter 8 that the distributional properties of balanced DCs are in accordance with the more general Principle of Functional Transparency. Whereas rigid balanced DCs are functionally transparent by definition, flexible balanced DCs, like any other flexible construction type, are expected to require additional morpho-syntactic means in order to compensate for their inherent functional ambiguity. I show that balanced flexible DC constructions may indeed combine with several types of such disambiguating strategies. Some of these strategies, such as the use of resumptive pronouns, are confined to DC constructions, while others, such as constituent order constraints, are used more generally to establish functional transparency in linguistic (sub-)systems.

Regarding this last point, a particularly interesting finding is that flexible DCs (both balanced and deranked ones) in languages with very flexible PoS systems often make use of the same morpho-syntactic strategies that are used to disambiguate the functions of flexible lexical and phrasal constructions in these languages. This suggests that languages with pervasive
lexical flexibility make use of a kind of ‘grid’ of morpho-syntactically marked functional slots, which may be filled by formal units of any structural type: not just single lexemes, but also complex phrasal and clausal constituents. All these construction types can (but need not) retain maximal flexibility until they are inserted into a slot corresponding to a particular phrase-structural function.

Finally, Chapter 9 summarizes the main findings of the study. It concludes with the observation that the explanatory principles of Increasing Categoriality and Functional Transparency are ultimately motivated in terms of Economy: Languages divide the task of establishing maximal categorial specificity over the lexical, morphological, and syntactic devices available in their grammatical systems, and complexity-increasing linguistic processes typically produce output structures that are at least as categorially specific as their input structures, in order to maximize processing ease for the participants of a communicative situation.
Taaltypologie houdt zich bezig met taalvergelijking. Het primaire probleem waarmee typologen zich geconfronteerd zien, is het volgende: Om taalspecifieke structuren onderling te kunnen vergelijken, is een vormonafhankelijk *tertium comparationis* nodig. Om die reden kiezen typologen eerst een bepaald *functioneel* domein, waarna ze onderzoeken welke *formele* structuren er in verschillende talen bestaan om bepaalde onderscheidingen binnen dat domein uit te drukken. Nadat een typoloog het spectrum van structurele variatie tussen talen binnen een functioneel domein in kaart heeft gebracht, probeert hij of zij *universalia* te formuleren: afhankelijkheidsrelaties tussen bepaalde structuurkenmerken die gelden voor alle talen. Binnen de *functionalistische benadering* van taaltypologie, waaronder ook de onderhavige studie valt, wordt aangenomen dat dergelijke universalia verklaard kunnen worden in termen van de *functie* van menselijke taal, namelijk het coderen van communicatie (Croft 2003).

Deze drie propositionele functies – referentie, predicatie, en modificatie – kunnen in talen worden uitgedrukt door verschillende soorten *constructies*. Voorbeeld (i) laat zien dat een referentiële uitdrukking zoals het direct object van het werkwoord *zie* de vorm kan hebben van een simpele lexicale constructie (een *auto* (in 1a)), maar ook van een meer complexe, ingebedde *zin* (*dat Jan een nieuwe auto heeft gekocht* in (1b)). Dit onderzoek heeft betrekking op de twee constructietypes in voorbeeld (i): lexicale constructies of woordsoorten (*parts of speech*) en ingebedde of afhankelijke zinsconstructies (*dependent clauses*).

(i)  

a. Ik zie een *auto*.

b. Ik zie *dat Jan een nieuwe auto heeft gekocht.*

*Parts of speech* is de traditionele term voor *woordsoorten*, dat wil zeggen de klassen van inhoudswoorden die onderscheiden worden in de grammatica van een bepaalde taal (Schachter & Shopen 2007:1). De klasse van zelfstandig naamwoorden of nomina (*nouns*) wordt meestal gebruikt om te refereren (cf. *de auto* in (2a)); werkwoorden of verba (*verbs*) worden gebruikt voor predicatie (cf. *wast* in (2b)); bijvoeglijk naamwoorden of adjectieven (*adjectives*) voor het modificeren van referentiële uitdrukkingen (cf. *rode* in (2c)); en bijwoorden of adverbia (*adverbs*) voor het modificeren van predicatieve uitdrukkingen (cf. *grondig* in (2d)).

(2)  

a. Henk wast *de auto*.

b. Henk *wast* de auto.

c. Henk wast de *rode* auto.

d. Henk wast de auto *grondig*.

Echter, niet alle talen van de wereld vertonen dit specifieke woordsoortensysteem, waarbinnen elke klasse speciaal geschikt is voor het uitdrukken van één bepaalde functie. Sommige talen hebben een ander type woordsoortensysteem, waarbinnen twee of meer functies uitgedrukt kunnen worden door één en dezelfde woordklasse. Het Turks, bijvoorbeeld, heeft één klasse van woorden die zowel gebruikt kunnen worden voor referentie (zoals in voorbeeld (3a)), als voor modificatie van een referentiële of predicatieve uitdrukking (cf. respectievelijk voorbeeld (3b) en (3c)). De Engelse vertalingen van de voorbeelden in (3b) en (3c) laten zien dat in het Engels een nomen zoals *beauty* eerst afgeleid moet worden met de uitgangen *-ful* en *-ly*, voordat het gebruikt kan worden in modificerende functies.
(3)  a. güzel-im  
    beauty-1poss  
    ‘my beauty’

b. güzel bir kopek  
    beauty ART dog  
    ‘a beautiful dog’

c. güzel konuştu  
    beauty s/he.spoke  
    ‘S/he spoke beautifully’

Deze verschillende types van woordsoorten worden in dit boek aangeduid met twee termen die zijn voorgesteld door Hengeveld (1992): Woordklassen die gespecialiseerd zijn in het uitdrukken van één functie worden rigide genoemd, en woordklassen die zonder aanpassingen gebruikt kunnen worden in twee of meer functies worden flexibel genoemd.

Met betrekking tot ingebedde zinnen of dependent clauses (verderop afgekort als DCs) wordt meestal een drievoudig onderscheid gemaakt. Complementzinnen (complement clauses) worden gebruikt om te refereren: ze functioneren als het subject of object van een hoofdzin (zie voorbeeld (1b)). Betrekkelijke of relatieve zinnen (relative clauses) modificeren het subject of object van de hoofdzin. En bijwoordelijke of adverbiale zinnen (adverbial clauses) modificeren het predicaat van de hoofdzin (Thompson et al. 2007: 238).

Echter, net zoals er rigid en flexibele woordklassen zijn, bestaan er ook rigide en flexibele ingebedde zinnen. Het Abun bijvoorbeeld (een taal uit het westen van Papua) beschikt over een DC-constructie, gemarkeerd door het onderschikkende voegwoord do, die uitsluitend referentieel gebruikt kan worden, dat wil zeggen als complementzin (cf. voorbeeld (4)).

Abun (Berry & Berry 1999: 167)

(4) Ji nut [do an bi obat ne nggi wa  
    1sg think COMP 3sg poss magic DET strong for  
    be an gu bi ré.]  
    later 3sg kill 1sg this  
    ‘I thought that his magic was strong [enough] for him to kill me.’
De Baskische DC-constructie gemarkeerd door -en kan daarentegen zowel een referentiële als een modificerende functie vervullen. Dit blijkt uit voorbeeld (5), waar de constructie als complementzin wordt gebruikt in (5a) en als relatieve zin in (5b).

*Basque* (Hualde & Ortiz de Urbina 2003: 646, 764)

(5)  

a. *Banekien [Mikel berandu etoriko z-en]-a*
   
   knew Mikel late arrive.fut aux-comp-det
   
   ‘I knew that Mikel would arrive late.’

b. *[Pellok ekarri du-en] dirua galdu dut*
   
   Peter.erg bring aux-rel money.det lose aux
   
   ‘I lost the money that Peter brought.’

Behalve in termen van hun functionele mogelijkheden kunnen ingebedde zinnen – aangezien het complexe constructies zijn – ook worden bestudeerd vanuit het perspectief van hun interne morfo-syntactische eigenschappen. Meestal worden deze eigenschappen beschreven in relatie tot de eigenschappen van ‘gewone’, niet-ingebedde hoofdzinnen in de taal in kwestie. Dit kan worden geïllustreerd aan de hand van de verschillen tussen de Engelse DCs in voorbeeld (6). In de DC in (6a), gmarkeerd door het onderschikkende voegwoord dat, wordt de toekomende tijd uitgedrukt (will move) en ziet het subject (Paul) er net zo als in een hoofdzin. In de DC in (6b) daarentegen, staat het werkwoord in de speciale ing-vorm, kan er geen tijdsmarkerings worden uitgedrukt, en heeft het onderwerp de bezittelijke vorm (Paul’s). Het eerste type DC, dat structureel gezien lijkt op een gewone hoofdzin, wordt in de literatuur balanced genoemd; het tweede type, dat in bepaalde opzichten afwijkt van de structuur van een hoofdzin, wordt deranked genoemd (Stassen 1985).

(6)  

a. John regrets [that Paul will move to Finland].

b. John regrets [Paul’s moving to Finland].

In dit proefschrift onderzoek ik of er een relatie bestaat tussen de functionele flexibiliteit versus rigiditeit van de woordklassen in een bepaalde taal en de ingebedde zinnen in die taal. Meer in het bijzonder richt deze studie zich op de vraag in hoeverre het mogelijk is om op basis van het woordsoortensysteem van een taal te voorspellen wat de functionele mogelijkheden zullen zijn van
verschillende typen ingebedde zinsconstructies in die taal. Het antwoord op deze vraag is gebaseerd op een systematische analyse van de woordklassen en ingebedde zinnen in een groep van 50 talen met een maximale genealogische en geografische diversiteit (zie Appendix 1).

In het eerste gedeelte van dit boek (Hoofdstuk 2-4) wordt de theoretische en methodologische basis gelegd voor het eigenlijke onderzoek. Hoofdstuk 2 bespreekt een aantal vraagstukken met betrekking tot de definitie van woordklassen in talen, en de vergelijking van woordklassen tussen talen. Dit hoofdstuk begint met een kort overzicht van benaderingen in de typologie van woordklassen, aan de hand van verschillende soorten criteria die relevant zijn voor de definitie van woordklassen, te weten semantische, morfologische, syntactische, en pragmatische criteria (Givón 2001: 49; Sasse 1993b:196).

Uit het bovenstaande blijkt dat de Hengeveldiaanse benadering van woordsoortentypologie uitsluitend gebruik maakt van pragmatisch-syntactische criteria. Deze benaderingswijze is bekritiseerd door Croft (2000, 2001, 2005), volgens wie woordklassen gedefinieerd moeten worden op basis van een uitgebreidere set van criteria, waarbij ook de semantische betekenis van woorden en hun morfologische gedrag in aanmerking worden genomen. Binnen Croft’s theorie worden woordsoorten gezien als prototypische combinaties van een semantische betekenis en een propositionele functie. Een prototypisch nomen verwijst naar een object (een ding) en heeft een referentiële functie; een prototypisch verbum verwijst naar een actie en heeft een predicatieve functie; en een prototypisch adjectief verwijst naar een eigenschap en heeft een modificerende functie. In crosslinguïstisch perspectief correrlen deze typologische prototypen met bepaalde maten van morfo-syntactische markering: Als een woordklasse een prototypische combinatie van semantiek en pragmatiek representeert, dan gaat dit samen met relatief minder morfo-syntactische functie-indicatoren, en met meer mogelijkheden om morfologische onderscheidingen uit te drukken dan wanneer er sprake is van een niet-prototypische combinatie.

Croft’s benadering resulteert in een typologie waarbij woordklassen niet direct vergelijkbaar zijn tussen talen. Aan de andere kant, voor de crosslinguïstische vergelijkbaarheid van de woordklassen in de Hengeveldiaanse benadering wordt ook een prijs betaald, namelijk het buiten beschouwing laten van bepaalde taalkundige feiten. Dit onderscheid tussen de twee benaderingen raakt aan een fundamentele kwestie binnen de taaltypologie, namelijk de vraag of er zoiets bestaat als een set van universele categorieën in menselijke taal (zie Haspelmath 2007; Newmeyer 2007).

Bovendien hebben de verschillen tussen de Hengeveldiaanse en de Croftiaanse benaderingen belangrijke gevolgen voor de discussie over zogenaamde flexibele talen, dat wil zeggen talen waarvan gezegd wordt dat ze geen woordklassen hebben, met name geen onderscheid tussen nomina en verba. Het meest heikle punt in deze discussie betreft semantic shift: kleine verschillen in de interpretatie van flexibele woorden of lexemen, die samenhangen met de verschillende functies waarin zulke lexemen kunnen voorkomen. Volgens sommige onderzoekers vereist flexibiliteit dat de interpretatie van lexemen in verschillende functies volledig compositioneel is (Evans & Osada 2005; Croft 2005), terwijl anderen juist bewerken dat die interpretatie idiosyncratische kenmerken kan vertonen (Hengeveld & Rijkhoff 2005). In het laatste gedeelte van Hoofdstuk 2 betoog ik dat de
relevante taalkundige feiten verklaard kunnen worden door te onderkennen dat er in flexibele talen geen één-op-één-relatie bestaat tussen lexicale en syntactische categorisatie. Zulke talen hebben wel degelijk woordklassen, maar van een zodanige aard dat de klasse waartoe een woord behoort irrelevant is voor de functionele distributie van dat woord (Himmelmann 2007; Don & Van Lier, forthcoming).

Hoofdstuk 3 bespreekt een aantal theoretische en typologische aspecten van ingebedde zinnen (DCs). Zoals eerder aangegeven kan de typologie van DCs vanuit twee perspectieven worden bezien. Ten eerste kunnen DCs, net als woordklassen, worden geclassificeerd in termen van de propositionele functie(s) die ze kunnen uitdrukken. In het eerste deel van Hoofdstuk 3 gebruik ik het Hengeveldiaanse functionele model uit het vorige hoofdstuk om voorschepinglen te doen over welk soort rigide en flexibele DC constructies wel en niet in talen zouden moeten kunnen voorkomen.

De tweede manier waarop DCs kunnen worden geclassificeerd is op basis van hun interne morfo-syntactische eigenschappen. Deze eigenschappen kunnen worden gedefinieerd als specifieke combinaties van verbale features, die geassocieerd worden met hoofdzinnen, en nominale features, die typisch zijn voor lexicale uitdrukkingen (zie voorbeeld (6) hierboven). Functioneel-typologisch onderzoek laat zien dat zowel verbale als nominale features bepaalde hiërarchische patronen vertonen, wat onder meer tot uitdrukking komt in de volgorde waarin zulke features staan ten opzichte van het hoofd waar ze bij horen (een verbum of nomen). Deze vaste volgorde is iconisch gemotiveerd; het is een reflectie is van de manier waarop en de mate waarin de afzonderlijke features de interpretatie van het hoofd beïnvloeden. Features die dicht bij het hoofd worden uitgedrukt hebben betrekking op de semantiek van het hoofd. Dit zijn met name de verbale categorieën tijd, aspect, en modaliteit (tense, aspect, mood, afgekort als TAM) en de nominale categorieën getal en geslacht. Features die verder van het hoofd verwijderd zijn hebben daarentegen geen directe invloed op de betekenis, maar zijn relevant voor de syntactische en pragmatische functie van de constructie waarin ze verschijnen. Dit geldt voor verbale congruentie en voor nominale (on)bepaaldheid en casusmarkering.

Hoofdstuk 3 beperk ik mij tot het bespreken van twee recente studies op dit gebied: Cristofaro’s typologische studie van subordinatie (Cristofaro 2003) en Malchukov’s typologische studie van nominalizaties (Malchukov 2004).

Cristofaro (2003) stelt een aantal correlaties vast tussen structurele aspecten van DCs. Deze correlaties kunnen als volgt worden samengevat: (i) het niet kunnen uitdrukken van TAM features impliceert het niet kunnen uitdrukken van verbale congruentie; en (ii) het uitdrukken van nominale features impliceert het verlies van verbale features. Cristofaro draagt drie typen functionele verklaringen aan voor deze afhankelijkheidsrelaties. Twee van de drie verklarende factoren zijn de fundamentele functionele principes van Economie en Iconiciteit (Haiman 1983). Deze principes verklaren het verlies van verbale features in DCs, inclusief het achterwege blijven van subject en/of object argumenten. Echter, noch Economie noch Iconiciteit kan verklaren waarom er ook nominale features in een DC kunnen verschijnen. Daarom stelt Cristofaro een derde factor voor: Zij betoogt dat DCs in cognitief opzicht niet worden beschouwd als onafhankelijke processen, maar worden geconstrueerd als dingen of eigenschappen. Aangezien verbale features niet relevant zijn voor processen, kunnen ze in DCs achterwege blijven. Bovendien kan het cognitief construeren van een DC als een ding of eigenschap leiden tot het uitdrukken van features die normaal gesproken horen bij lexicale constructies die dingen en eigenschappen uitdrukken: nomina en adjectieven.

Malchukov’s (2004) studie levert in grote lijnen dezelfde generalisaties op als Cristofaro’s onderzoek, maar geeft er een andere functionele verklaring voor, waarbij ik mij in dit proefschrift aansluit. Malchukov’s onderzoek laat zien dat er in DC constructies vaker iets aan de hand is met verbale en nominale features die een externe positie innemen binnen hun respectievelijke hiërarchieën dan features met een meer interne positie. Om precies te zijn, het uitdrukken van een bepaalde nominale feature in een DC-constructie impliceert dat alle nominale met een meer externe positie ook kunnen worden uitgedrukt in die DC. Voor verbale features geldt dat het niet kunnen uitdrukken van een bepaalde feature impliceert dat alle meer externe features van de hiërarchie ook niet kunnen verschijnen. Volgens Malchukov moet de verklaring voor deze bevindingen gezocht worden in het feit dat externe features relevant zijn voor pragmatische en syntactische aspecten van de constructie waarin ze verschijnen, terwijl interne features betrekking hebben op de semantiek. Hoewel DCs de pragmatisch-syntactische functies van referentie en modificatie vervullen, hebben ze niet de semantische
betekenis van een ding of eigenschap (integendeel, ze drukken situaties of proposities uit). Dit verklaart waarom externe nominale (of adjectivische) features, die betrekking hebben op deze pragmatisch-syntactische functies, eerder uitgedrukt worden in DCs dan interne features. Bovendien verklaart dit waarom externe verbale features, die te maken hebben met de pragmatiek en syntax van *predicaties*, juist kunnen vervallen in DCs constructies.

Op basis van deze studies definieer ik in het laatste deel van Hoofdstuk 3 drie types van DC-constructies, in termen van hun specifieke combinaties van verbale en nominale features. Eén van deze types is vergelijkbaar met *balanced* DCs; de andere types zijn te beschouwen als twee verschillende soorten *deranked* DCs (zie boven). Deze classificatie wordt vervolgens gecombineerd met de DC-classificatie in termen van functionele mogelijkheden, zoals gedefinieerd in het eerste gedeelte van Hoofdstuk 3. Dit resulteert in een typologisch raamwerk voor DCs dat zowel hun functionele als hun vormelijke eigenschappen in aanmerking neemt.

Hoofdstuk 4 is het laatste hoofdstuk van het theoretische gedeelte van het boek. In dit hoofdstuk ga ik eerst in op de samenstelling van het sample van talen waarop mijn onderzoek gebaseerd is. Vervolgens formuleer en operationaliseer ik de specifieke hypotheses die in de rest van de studie worden getest. Deze hypotheses hebben de vorm van voorspellingen over afhankelijkheidsrelaties tussen de functionele mogelijkheden van woordklassen in specifieke talen en van verschillende structurele DC-types in die talen. Een aantal van deze voorspellingen is relatief breed van opzet; ze hebben betrekking op talen die enige flexibiliteit in hun woordsoortensysteem vertonen versus talen die geen enkele flexibiliteit vertonen, en de manier waarop dit verschil gereflecteerd wordt in de functionele eigenschappen van DCs. Andere voorspellingen zijn specifieker: ze zijn gericht op het identificeren van één-op-één relaties tussen woordsoorten met bepaalde functionele eigenschappen enerzijds en DCs met dezelfde eigenschappen anderzijds. Bovendien worden beide typen voorspellingen telkens op twee manieren geformuleerd: eerst zonder te differentiëren tussen verschillende structurele DC-types, en vervolgens in sets van van sub-predicties die speciaal betrekking hebben op *balanced* DCs, dan wel op (verschillende soorten) *deranked* DCs.

Tenslotte wordt in Hoofdstuk 4 de methode uitgelegd die ik gebruik om deze hypotheses te toetsen. Elke voorspelling bevat twee parameters: één met betrekking tot woordklassen, de ander met betrekking tot DCs. Elke voorspelling wordt getoetst door de *geobserveerde frequenties* (het aantal
talen dat een bepaalde combinatie van woordklasse- en DC-kenmerken vertoont) te vergelijken met de verwachte frequenties (het aantal talen met deze combinatie dat je op basis van toeval zou verwachten). Om te bepalen of het verschil tussen de geobserveerde en verwachte frequenties statistisch significant is, wordt een Fischer’s Exact test gebruikt. Wanneer er sprake is van een significante correlatie tussen een woordklasse-gerelateerd en een DC-gerelateerd structuurkenmerk, dan wordt de aard van deze afhankelijkheidsrelatie nader gespecificeerd door middel van een methode ontwikkeld door Maslova (2003).

Het tweede deel van dit proefschrift (Hoofdstuk 5-7) bevat de eigenlijke typologische gegevens. Om te beginnen wordt in Hoofdstuk 5 de classificatie van de woordklassen in de talen van mijn sample gepresenteerd. De aangetroffen woordsoortensystemen worden vergeleken met de woordsoortensystemen zoals voorspeld in Hoofdstuk 2. Hieruit blijkt dat de meeste voorspelde systemen ook inderdaad gevonden worden, in hun ‘pure’ vorm dan wel in een combinatie met een ander voorspeld systeem. In het algemeen suggereren deze data dat het in Hoofdstuk 2 gepresenteerde hiërarchische model in redelijke mate typologisch adequaat is.

De rest van Hoofdstuk 5 behandelt verschillende kwesties rondom de identificatie en vergelijking van woordklassen binnen en tussen talen. Deze problemen, die al in algemeen-theoretische zin aan orde kwamen in Hoofdstuk 2, hebben betrekking op het voorkomen van vage grenzen tussen woordklassen, van variabel distributioneel gedrag bij bepaalde subklassen, en van ‘beperkte’ woordklassen (kleine, gesloten klassen en klassen die bestaan uit afgeleide vormen). Ook wordt de uitdrukking van niet-verbale predicatie in de talen van het sample besproken. De laatste sectie van Hoofdstuk 5 betreft een toepassing van de theoretische noties over lexicale flexibiliteit, zoals voorgesteld in Hoofdstuk 2, op de relevant talen in het sample.

Hoofdstuk 6 beschrijft de classificatie van de DC-constructies in de sample talen, in termen van het in Hoofdstuk 3 ontwikkelde typologische raamwerk. Eerst worden de DC constructies geclassificeerd op basis van de propositionele functie(s) die ze kunnen uitdrukken. Het resultaat wordt vergeleken met de voorspellingen uit Hoofdstuk 3. De data laten zien dat alle voorspelde DC-types voorkomen, behalve degenen die (onder andere) de functie van hoofd van een predicatie zouden moeten kunnen uitdrukken. Deze bevinding duidt op een cognitief-semantisch gemarkeerde status van DCs ten opzichte van de functie van onafhankelijke predicatie, en/of
op een structurele beperking ten aanzien van het uitdrukken van finiete werkwoordsmorfologie op DCs.

In het tweede deel van Hoofdstuk 6 worden de interne morfo-syntactische eigenschappen van de DCs in de talen van het sample geïdentificeerd. Op basis van deze gegevens kan elke DC-constructie worden geclassificeerd als behorende tot één van de drie DC-types zoals gedefinieerd in Hoofdstuk 3. Aan het einde van Hoofdstuk 6 wordt deze classificatie geïntegreerd met de DC-classificatie op basis van functionele distributie.

Hoofdstuk 7 brengt de datasets uit Hoofdstuk 5 (woordklassen) en Hoofdstuk 6 (ingebedde zinnen) op systematische wijze met elkaar in verband, zodat de predicties uit Hoofdstuk 4 getest kunnen worden. De analyses wijzen uit dat er significante afhankelijkheidsrelaties bestaan tussen flexibele (in plaats van rigide) woordklassen en DCs; tussen zeer flexibele (in plaats van minder flexibele) woordklassen en DCs; en tussen woordklassen en *deranked* (in plaats van *balanced*) DCs. De correlaties in kwestie zijn asymmetrisch van aard, en kunnen als zodanig worden geïnterpreteerd als statistische fundamenten voor de formulering van implicationele universalia.

Meer in het bijzonder laten de resultaten in Hoofdstuk 7 zien dat de aanwezigheid in een bepaalde taal van een *deranked* DC-constructie met een hoge mate van flexibiliteit vereist dat die taal een zeer flexibel type woordklasse heeft. Dit betekent echter niet dat alle talen met veel flexibiliteit in hun woordklassensysteem ook flexibele deranked DC-constructies hebben. Bovendien wijzen de data uit dat zowel zeer flexibele woordklassen als zeer flexibele *deranked* DCs – onafhankelijk van elkaar – crosslinguïstisch zeldzame fenomenen zijn. Om die reden heeft de generalisatie dat maximale flexibiliteit bij *deranked* DCs afhankelijk is van maximale flexibiliteit in het woordklassensysteem niet veel verklarende potentie. Veel interessanter is de observatie dat, als er sprake is van zeer flexibele woordklassen in een taal, dit de kans aanzienlijk verhoogt dat die taal ook zoiets zeldzaams als flexibele *deranked* DC-constructies heeft. Dit betekent dat er wel degelijk een sterke tendens in talen bestaat om gelijke waarden te vertonen voor structurele parameters die betrekking hebben op het functionele gedrag van woordklassen en deranked DCs.

In tegenstelling tot de resultaten voor zeer flexibele woordklassen en *deranked* DCs, kunnen er geen significante correlaties worden vastgesteld tussen de aanwezigheid in talen van minder flexibele en rigide woordklassen enerzijds, en de aanwezigheid van *deranked* DC met hetzelfde functionele gedrag anderzijds. Bovendien wordt duidelijk uit de resultaten dat de
functionele patronen van *balanced* DCs in talen niet gerelateerd kunnen worden aan die van woordklassen in dezelfde talen.


In Hoofdstuk 8 bepleit ik dat *deranked* DCs, in tegenstelling tot *balanced* DCs, beschouwd kunnen worden als een secundair constructietype dat is afgeleid van een primaire lexicaal constructietype. Onder deze aanname voorspelt het *Principle of Increasing Categoriality* dat de mate van functionele flexibiliteit van een *deranked* DC-constructie in een bepaalde taal *kleiner of gelijk* moet zijn aan de flexibiliteit van een woordklasse in die taal die gebruikt wordt in tenminste één van dezelfde functies. Deze hypothese wordt bevestigd voor vrijwel alle relevante constructies in de sample talen. Dit betekent dat het *Principle of Increasing Categoriality* verklaart waarom flexibiliteit bij *deranked* DCs afhankelijk is van flexibiliteit in het woordsoortensysteem. Deze afhankelijkheidsrelatie, vastgesteld in Hoofdstuk 7, kan nu worden geïnterpreteerd als een beperking, opgelegd door de mate van flexibiliteit in het woordsoortensysteem van een taal, op de mate van flexibiliteit die is ‘toegestaan’ bij *deranked* DCs in die taal.
Bovendien werpt het Principle of Increasing Categoriality licht op de onverwachte resultaten uit Hoofdstuk 7, namelijk (i) het ontbreken van correlaties tussen minder flexibele woordklassen en deranked DCs met dezelfde functionele distributie, en (ii) het ontbreken van correlaties tussen verschillende rigide woordsoorten en rigide deranked DCs van hetzelfde type.

De eerste bevinding kan worden geïnterpreteerd als een geval van afnemende flexibiliteit bij DCs, in vergelijking met de woordklassen in kwestie. De relevante woordklassen zijn matig flexibel omdat ze (slechts) twee propositionele functies kunnen uitdrukken. Wanneer de DC-constructies minder flexibel zijn dan de woordklassen, dan resulteert dit automatisch in rigide constructies, die nog maar in één enkele functie gebruikt kunnen worden. Een gelijksoortige mechanisme verklaart waarom zelfs talen met zeer flexibele woordklassen toch rigide deranked DCs hebben (naast eventuele flexibele deranked DCs): dit is in overeenstemming met de voorspelling dat deranked DCs óf dezelfde flexibiliteit hebben als woordklassen, óf minder flexibiliteit (maar nooit méér).

Ten tweede kan het ontbreken van correlaties tussen rigide woordklassen en deranked DCs nu beter begrepen worden. Zoals verwacht blijkt in alle gevallen waar een taal een deranked DC-constructie gebruikt in een functie waarvoor een rigide woordklasse beschikbaar is, de DC-constructie óók rigide te zijn. Echter, in plaats van een deranked constructie, kan een taal ook een balanced DC gebruiken. Deze balanced DCs kunnen zowel rigide als flexibel zijn; uit Hoofdstuk 7 blijkt immers dat balanced DCs, die formeel niet afgeleid zijn van lexicale constructies, ook geen functionele verwantschap met woordklassen vertonen.

Hoewel balanced DCs en woordklassen dus formeel en functioneel ongerelateerd zijn, laat ik in het laatste deel van Hoofdstuk 8 zien dat de distributionele eigenschappen van balanced DCs wel in overeenstemming zijn met het meer algemene principe van Functional Transparency. Terwijl rigide balanced DCs per definitie functioneel transparant zijn, zijn flexibele balanced DCs – net als andere flexibele constructies – functioneel ambigu. De verwachting is daarom dat het gebruik van balanced flexibele DCs vereist dat er andere morfo-syntactische middelen in worden gezet om de inherente ambiguïteit van deze constructies te compenseren. Ik laat zien dat balanced flexibele DC constructies inderdaad samengaan met verschillende soorten disambiguërende strategieën. Soms zijn deze strategieën speciaal toegespitst op DC-constructies, zoals bij het gebruik van resumptieve pronomina, en in andere gevallen, zoals bij woordvolgorde-restricties, gaat het om strategieën
die een bredere functie vervullen bij het bewerkstellig van functionele transparantie binnen het taalsysteem.

Met betrekking tot dit laatste punt is het interessant om te zien dat flexibele DCs – zowel balanced als deranked – in talen met zeer flexibele woordsoortensystemen vaak gebruik maken van dezelfde morfo-syntactische strategieën als die ingezet worden bij het functioneel disambiguëren van flexibele lexicale en andersoortige constructies in die talen. Dit duidt op een patroon waarbij talen met een hoge mate van lexicaal flexibiliteit gebruik maken van een soort rasterwerk van morfo-syntactisch gemankeerde slots, die gevuld kunnen worden met constructies van allerlei aard: niet alleen simpele woorden, maar ook complexere frases en ingebedde zinnen. Al deze meer en minder complexe constructietypes hebben de mogelijkheid (maar niet de noodzaak) om hun functionele flexibiliteit te behouden totdat ze worden ingevoegd in een slot die correspondeert met een bepaalde zinsfunctie.

In Hoofdstuk 9, tenslotte, worden de belangrijkste resultaten van het onderzoek samengevat. Dit hoofdstuk sluit af met de concluderende observatie dat de verklarende principes van Increasing Categoriality en Functional Transparency gefundeerd zijn op het algemene Economie principe. De algemene taak om maximale categoriale specificiteit te bewerkstelligen wordt door talen verdeeld over de verschillende lexicaal, morfologische, en syntactische middelen die ze tot hun beschikking hebben. Bovendien produceren complexiteitverhogende grammaticale processen meestal output-constructies die een mindstens even hoge mate van categoriale specificiteit vertonen als hun input-constructies. Zo kan het verwerkingsgemak voor de taalgebruiker worden geoptimaliseerd.

480 | Parts of Speech and Dependent Clauses
Eva van Lier was born in The Hague (the Netherlands) in 1978. During her secondary education at the Gymnasium Haganum (The Hague) she studied classical Latin and Greek, as well as French, German, and English. After finishing high school in 1996, she spent a year in Granada and Sevilla (Spain) to learn Spanish and obtain a Diplóma Básico de Español como lengua extranjera (basic degree of Spanish as a foreign language). In 1997, she started studying Dutch Linguistics and Literature at the University of Groningen. She continued the following year at the University of Amsterdam, studying Spanish Linguistics and Literature and General Linguistics. She obtained her MA degree in the latter subject with honours in 2002, with a thesis on the relationship between linguistic typology and second language acquisition. Subsequently, she completed a Research Master in Linguistics and obtained her MPhil degree with honours in 2003. Her thesis was again concerned with typology and its connections with other linguistic sub-disciplines, including language contact and creolization, as well as first and second language acquisition. During this period Eva was also employed as a student assistant at the Steunpunt Nederlands as Vreemde Taal (Support Center for Dutch as a Foreign Language) of the University of Amsterdam and the Nederlandse Taalunie (Dutch Language Academy). After her university studies, she worked for one year as a researcher at the Meertens Institute for Dutch language and culture, on a project about multilingualism in the Netherlands. In 2004 she started on her PhD research project at
the department of Theoretical Linguistics of the University of Amsterdam. During this period she functioned occasionally as a teaching assistant (BA and MA courses), and as a freelancer for the Taalstudio, contributing material for linguistic education in Dutch high schools. She was a PhD representative at the Advisory Board of the ACLC (Amsterdam Center for Language and Communication) in 2006 and 2007. Eva has been involved in several national and international collaborative projects, which resulted in an international workshop in Amsterdam and a two-month stay at the Max Planck Institute for Evolutionary Anthropology in Leipzig in 2007, as well as several (co-authored) conference presentations, reviews, and journal papers. A co-edited volume on lexical flexibility is in preparation. The present thesis is the result of Eva’s doctoral studies.