



## UvA-DARE (Digital Academic Repository)

### The expression of modifiers and arguments in the noun phrase and beyond

*A typological study*

van Rijn, M.A.

**Publication date**

2017

**Document Version**

Other version

**License**

Other

[Link to publication](#)

**Citation for published version (APA):**

van Rijn, M. A. (2017). *The expression of modifiers and arguments in the noun phrase and beyond: A typological study*. [Thesis, fully internal, Universiteit van Amsterdam]. LOT.

**General rights**

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

**Disclaimer/Complaints regulations**

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

## 3

## The grammaticalization of possessive person marking: A typological approach<sup>1</sup>

### Abstract

This study focuses on the grammaticalization of agreement markers from possessive pronouns, which has two different dimensions: loss of referentiality (function) and loss of morpho-phonological independence (form). I examine the referential potential and formal expression type of possessive person markers in a worldwide sample of 39 languages with an alienability distinction. Referential potential is measured independently of expression type by applying a new typology of person markers. First, I demonstrate that inalienable possessive marking is at least as referential and formally independent as alienable possessive marking, and often less referential and less independent. Unlike explanations in terms of frequency and iconicity, I argue that this asymmetry is essentially semantics-based: the presence of a possessive relationship is inherent to the meaning of the inalienable noun, which is therefore in less need of expressive marking than alienable nouns. Second, I show that loss of referentiality correlates with loss in form, but in a relative rather than an absolute sense: in individual languages, higher referential markers never show a greater degree of bonding with the possessum than lower referential markers. These results suggest that function and form evolve in the same direction, but need not evolve at the same pace.

---

<sup>1</sup> This chapter appeared in *Transactions of the Philological Society*. The full reference for the article is: Van Rijn, Marlou. 2016. The grammaticalization of possessive person marking: A typological approach. *Transactions of the Philological Society* 114(2). 233–276. A number of modifications have been made to the chapter relative to the published article. First, Table 6 and 7 are published as part of an appendix but are given in the main text below. The total number of counts in Table 7 is not 61 but 77. Second, Nasioi was excluded from the published article but is included in this thesis, since it employs a combination of alienable and inalienable appositional referential (AR) markers. This raises the number of AR markers in Table 3 (Section 3.5.1.1) from 19 to 20 and the total number of combinations from 69 to 70. Third, the data contains not 11 but 13 contextual agreement markers that co-occur with possessor nouns, as shown in Table 5 (Section 3.5.4). This raises the total number of markers from 115 to 117. None of these modifications affect the results of the study.

### 3.1 Introduction

A well-studied grammaticalization process is the development of bound agreement markers from independent pronouns.<sup>2</sup> Most studies on this development focus on subject and object agreement markers on verbs (Givón 1976; Mithun 1986, 1991b, 2003; Siewierska 1999; Ariel 2000; Fuß 2005). However, agreement also obtains in other grammatical domains, a cross-linguistically frequent one being the possessive noun phrase. Possessive pronouns may develop into agreement markers much in the same way as subject/object pronouns do. Both types of developments are typically assumed to follow the grammaticalization cline in (1) below (see e.g. Lehmann 1982b: 39–42, 1988: 59–61; Hopper & Traugott 1993: 15; Fuß 2005: 4):

- (1) independent pronoun  
     > weak pronoun  
         > clitic pronoun  
             > affixal agreement marker  
                 > fused agreement marker  
                     > Ø

According to this cline, phonological erosion may turn an independent pronoun into a weak form, and subsequently into a clitic that attaches to the possessed noun (henceforth *the possessum*). Further loss of morpho-phonological substance causes the clitic to develop from a pronoun to a marker of agreement. While the former is traditionally considered to be a referential expression of the possessor, the latter is generally believed to lack any referential potential, as it only redundantly expresses features, such as person, number, gender and (genitive) case, in agreement with the possessor on the possessum.<sup>3</sup> The agreement marker eventually fuses with the possessum before disappearing altogether.

---

<sup>2</sup> But note that independent pronouns are by no means the only diachronic source of agreement marking (cf. Heine & Song 2011 for an overview).

<sup>3</sup> This paper focuses solely on markers expressing possessor features (such as person, number and gender) for which the shorthand term *person marker* is used. Agreement markers may alternatively express features (such as number and/or gender) of the possessum on the possessor, but such markers are not attested in languages with a distinction between alienable and inalienable possession (Lehmann 1983: 362) and are therefore not considered in this paper.

A presumed major factor in this development is the presence of an opposition between alienable and inalienable possession (Nichols 1988: 579, 1992: 117–122; Dahl 2004: 152; Haiman 1983: 793–795; Haspelmath 2008a: 15–16, 18–22). This opposition involves the use of dedicated possessive marking for nouns that are inherently related to their possessors, like kinship terms and body part terms (inalienable), versus nouns that are not (alienable). The opposition is generally assumed to develop in two ways: markers of inalienable relations progress further down the cline in (1), leaving markers of alienable relations behind (Nichols 1988: 589), or new alienable constructions arise on the left end of the cline, leaving the old possessive morphology to mark inalienable relations (Dahl & Koptjevskaja-Tamm 1998: 48). The result of the two diachronic scenarios is the same: cross-linguistically, inalienable possessive marking tends to be more grammaticalized, i.e. located further rightwards on the cline in (1), than alienable possessive marking.

It is argued here that the major problem with the cline in (1) is that it collapses developments along two different dimensions: a loss of referentiality, i.e. the functional transition of a referential pronoun to a marker of agreement, and a loss of morpho-phonological independence, i.e. the formal transition of a separate word to a fusional form via affixation and cliticization. Although it is widely presupposed that both dimensions tend to go hand in hand, a claim sometimes referred to as the ‘parallel path hypothesis’ (Bakker & Siewierska 2009), they are at least in principle independent of each other (see also Bisang 2004 and Narrog 2009: 697, 2012: 107–109). This has been demonstrated most clearly for the domain of verbal person marking. On the one hand, there are languages with so-called ‘pronominal affixes’ (Corbett 2003, 2006) or ‘bound/incorporated pronouns’ (Mithun 1991b; Kibrik 2011: 92–104; Bresnan 2001: 144): person markers that function as free pronouns, in that they can be used independently, like English *he*, but that formally resemble agreement markers, in being affixed to the verb, like English *-s*. Such person forms have received much attention in the literature, due to their common use among the world’s languages (Siewierska 1999: 238). On the other hand, there are markers that have the form of independent pronouns, but that function as bound agreement markers. An example is the set of clitics coding (main clause) subjects in Skou (Skou, Indonesia; Donohue 2004: 238, see also Corbett 2006: 75–76 and Hengeveld 2012: 475). Both types of person markers represent an intermediate stage in the development of free pronouns to bound agreement markers, which is not captured by the cline in (1). The rise of such forms can only be accounted for when referential potential (function) and expression type (form) are

represented as individual pathways along which person markers may mature over time. The present paper does precisely this.

This paper pursues two goals. First, it investigates the widely held claim that inalienable possessive person marking shows a higher degree of grammaticalization than alienable possessive person marking, for the functional dimension and the formal dimension in isolation. To this end, I study the referential potential and formal expression type of possessive person markers in a worldwide sample of 39 languages with a distinction between alienable and inalienable possession. The degree of referential potential of individual person markers is determined on the basis of a four-part typology, proposed by Hengeveld (2012). The major advantage of this typology is that it straightforwardly accounts for person markers 'in between' free pronouns and bound agreement markers, such as those mentioned above. The referential potential of such markers is determined by considering the distribution of (grammatical) feature information between the possessive person marker on the one hand and the co-occurring possessor (pro)noun on the other.

The second aim of this study is to investigate to what extent loss of referentiality correlates with reduction in morpho-phonological form. The relationship between functional and formal changes has been the subject of much debate in the literature. As mentioned above, the dominant position within grammaticalization theory is that both types of changes will tend to coincide. Although there is disagreement as to the presence and nature of a causal relationship, it is generally believed that an item will lose formal independence as a result of a gradual loss or shift in meaning, often mediated by an increase in usage frequency (e.g. Givón 1975: 96; Bybee et al. 1994: 20; Heine 1994: 269; Haspelmath 1999: 1050). Thus, functionally more grammaticalized markers are expected to also be formally more grammaticalized (e.g. Givón 1984: 416; Lehmann 1982a: 236; Dahl 2001: 118; Siewierska & Bakker 2005).

In this paper, I investigate the relationship between function (referential potential) and form (expression type) for the possessive person markers in the 39 languages in my sample. I test the hypothesis that, in any language with two or more sets of possessive person markers, markers of a lower degree of referentiality never have a more independent expression form than markers of a higher degree of referentiality. This hypothesis expects function and form to develop in parallel, but is more specific than the parallel path hypothesis in two respects. On the one hand, it does not preclude the possibility of highly referential markers that are formally bound (i.e. so-called 'pronominal affixes' or 'bound/incorporated pronouns') or less referential markers that are formally independent (as the subject clitics in Skou). Hence, the relationship between

function and form is expected to be relative, rather than absolute. From a diachronic perspective, this means that function and form are expected to develop in the same direction, but not necessarily at the same pace. On the other hand, my hypothesis not only predicts an asymmetry in the expression of more or less referential markers across languages, but also within languages, which presupposes a much stronger relationship between function and form than the parallel path hypothesis.

The paper is organized as follows. Section 3.2 provides the theoretical background for the study. In Section 3.2.1, I first discuss the semantic opposition between inherently relational and inherently non-relational nouns, which underlies distinctions in the expression of (in)alienable possession. I also illustrate the cross-linguistically most common types of alienability contrasts, and discuss the necessity of separating function and form in accounting for their diachronic development. The functional development of possessive person markers, following the four-part typology of referential markers and agreement markers by Hengeveld, is discussed in Section 3.2.2. Section 3.3 in turn presents the three hypotheses tested in this study. The 39-language sample on which these hypotheses are tested, and the types of possessive NPs and person markers taken into account are presented in Section 3.4. The results of testing each hypothesis are discussed in Section 3.5. In this section, I also propose an explanation for the findings building on the inherently relational and inherently non-relational nature of possessums. I conclude with a summary and directions for further research in Section 3.6.

## **3.2 Theoretical background**

### **3.2.1 The inherently (non-)relational nature of possessed nouns**

Following much previous work (e.g. Seiler 1983a; Lehmann 1985; Barker 1995; Taylor 1996; Partee 1997 and Partee & Borschev 2003), this paper starts from the observation that languages distinguishing alienable and inalienable possession grammaticalize a semantic opposition between two types of nouns: those encoding concepts that are inherently relational, and those encoding concepts that are not. Nouns of the former type by their nature presuppose a relationship to a possessor. At least two types of nouns fit this semantic profile: kinship terms and body part terms. Part of the meaning of nouns such as ‘father’ or ‘leg’ is their relationship to another entity. Moreover, while a phrase such as *John’s father* is normally interpreted as involving an inherent relationship between two

individuals, a phrase such as *John's leg* is normally interpreted as denoting the inherent part of one's body.<sup>4</sup> Both the presence and the semantic interpretation of the possessive relationship are thus inherent to the meaning of the relational noun.

Based on these semantic properties, languages may reserve special morphosyntactic treatment for a distinct class of nouns. This is traditionally referred to as the opposition between alienable and inalienable possession. As is widely known, the precise composition of the class of inalienable nouns varies from language to language (Chappell & McGregor 1996: 9; Heine 1997: 174; Stolz et al. 2008: 38–40). In addition to kinship and body part terms, individual languages often include spatial nouns, other parts of wholes, culturally basic items and property nouns such as 'beauty' or 'strength' in their class of inalienable nouns (cf. Nichols 1988: 572, Nichols & Bickel 2013b). Despite this cross-linguistic variation, however, a set of kinship terms and/or body part terms is always part of any inalienable class (see also Haiman 1985: 136 and Chappell & McGregor 1989: 26). Moreover, these terms are also the most clearly inherently relational in terms of their lexical semantics, and are therefore the relational items focused on in this study (for further discussion, see Section 3.4.2).<sup>5</sup>

Nouns encoding kinship and body parts are often likened to verbs and adpositions as items with an argument structure (or valency): each may be conceived of as opening up a semantic argument position, and determining the semantic (and syntactic) relationship with items filling this position (e.g. Seiler 1983a: 11–13). As demonstrated in the next subsection, this argument position can be filled by an NP, i.e. a lexical noun or an independent pronoun, a person marker or both. The extent to which either of such items is obligatorily expressed varies extensively cross-linguistically. In many of the world's languages, relational nouns are obligatorily affixed with a possessive person marker, while the expression of the possessor NP is optional (so-called 'obligatorily possessed nouns', Bickel & Nichols 2013). Other languages do not require an expression of the possessor altogether, such as Tidore (West Papuan, Indonesia; Van Staden

---

<sup>4</sup> In principle, a phrase such as 'John's leg' could alternatively be interpreted as referring to a leg detached from John's body, or to a leg of an animal that John is eating or that John has caught, for instance. In practice, however, such alienable interpretations occur only in certain pragmatic contexts, and are therefore far less natural (cf. the study by Lichtenberk et al. (2011) discussed below).

<sup>5</sup> Some languages show further variation in their class of inalienable nouns in that only a subset of kinship and/or body part terms is treated inalienably. For examples and a discussion of the methodological treatment of such languages in the present paper, I refer the reader to Section 3.4.2.

2000), exemplified in the next subsection. Thus, the fact that possessors of nouns such as kinship terms and body part terms are semantically obligatory, as presupposed by the meaning of the noun, is independent of whether such possessors are obligatorily expressed in individual languages.

Non-relational nouns encode concepts that do not entail the presence of a possessor, and therefore lack an inherently relational meaning. A possessive relationship is not inherent to such nouns, but must be established, via the presence of a possessor (Seiler 1983a: 62). This possessor merely provides a further characterization of the possessum, restricting the reference of the possessum to a specific (subset of) item(s). Typical examples of non-relational nouns are concrete (inanimate), countable objects, such as ‘book’, ‘chair’, ‘basket’ or ‘pot’. Since nouns of this type do not presuppose an inherent relationship to a possessor, the possessive relationship may receive a range of different interpretations. Examples include legal ownership (*John’s car*, the car that John owns), but also location (*the house’s garden*, the garden near the house) and control (*Mary’s employees*, the employees Mary is responsible for) (cf. Seiler 1983a: 40–41; Koptjevskaja-Tamm 2004). This semantic behavior parallels that of attributive adjectives, relative clauses and other noun modifiers (Nikolaeva & Spencer 2013).

Many of the world’s languages, such as English, do not overtly express the opposition between relational and non-relational nouns in possessive NPs.<sup>6</sup> Interestingly, the opposition nevertheless has cognitive reality in speakers of such languages, as demonstrated by Lichtenberk et al. (2011). They show that the interpretation of possessive NPs by speakers of English strongly depends on whether or not the possessed noun is inherently relational. Phrases with inherently relational nouns, such as ‘my child’, tend to yield the interpretation inherent to the meaning of the noun, i.e. the possessor’s own child, not someone else’s child. Phrases with inherently non-relational nouns, however, are open to various interpretations, among which prototypical ownership, location and control, as illustrated above. These findings demonstrate that while inherently

---

<sup>6</sup> The opposition may nevertheless have a syntactic correlate. In English, for instance, only nouns denoting non-relational concepts can normally function as subjects of a possessive predicate: compare *that book is John’s* to (?)*that leg is John’s*. The latter clause is typically acceptable only when the noun is possessed in an alienable manner, e.g. when referring to a detached leg or to a leg that inherently belongs to someone or something else. As pointed out by one of the referees, inalienable interpretations of the clause are possible in principle (e.g. in the context of a photograph depicting only a person’s legs), but highly unusual. Other examples are given in Partee (1997: 464), Van Valin and LaPolla (1997: 19) and Tsujiojka (2002: 115–117).



relational nouns yield one highly salient relationship to their possessor (i.e. the relationship inherent to their meaning, cf. footnote 4), no such salient relationship exists when the noun is inherently non-relational, and a variety of relationships are freely available. This in turn provides clear support for the distinction between inherently relational and inherently non-relational nouns described above, as well as its cross-linguistic relevance, even in languages without a formal alienability split.

In this paper, I focus solely on languages that, unlike English, make a formal distinction between alienable and inalienable possession in possessive NPs. This distinction is fairly common everywhere except in Eurasia and South-East Asia (Nichols & Bickel 2013b) and is generally assumed to develop via two diachronic routes, as briefly discussed in the introduction to this paper. The first route involves the innovative phonological reduction and morphological bonding of the possessor pronoun with alienable nouns, but not with inalienable nouns. In Nyulnyul (Australian, Australia), for instance, prefixes on body part nouns, such as *nga-* (1<sup>st</sup> person singular), *ya-* (1<sup>st</sup> person dual inclusive), *kurr-* (2<sup>nd</sup> person plural) and *irr-* (3<sup>rd</sup> person plural), derive historically from free pronouns, *ngay*, *yay*, *kurr* and *irr* respectively (see McGregor 1996: 272–278 for further discussion).

The second route involves the creation of a new construction for alienable possession, which typically happens on the basis of a demonstrative (Eksell Harning 1980: 19; Schuh 1983: 182–184), an invariant element expressing location or goal (Heine 1997: 144) or a lexical item meaning ‘possession’ or ‘thing’ (Eksell Harning 1980: 19, 24). Crucially, the new construction is not extended to inalienable possession, which continues to use the old, typically bound possessive morphology. This development is well documented for a number of Afro-Asiatic languages, such as Maltese, in which the possessive pronoun marking alienable possession consists of the inalienable possessive suffix combined with the preposition *ta’* ‘of’ (Koptjevskaja-Tamm 1996).

Both diachronic pathways contribute to the following cross-linguistic pattern: inalienable possession never shows a greater degree of linguistic distance than alienable possession (Haiman 1983: 793–795, 1985: 130–136). Two types of explanations have been offered for this pattern, to which I will return in Section 3.5.3. The important point here is that the notion of linguistic distance (and the related notion of cohesion, see Haspelmath 2008a) collapses three types of alienability contrasts, which are actually the result of distinct processes of grammaticalization: (i) overt possessive coding vs. lack of possessive coding, (ii) juxtaposition vs. bound expression and (iii) separably analyzable vs. portmanteau expression. Each contrast is briefly exemplified below.

The first contrast relates to the fact that the possessive relationship in the NP tends to be overtly specified only in alienable possession. Coding of this relationship usually takes the form of case affixes or adpositional marking. An example is provided for Nyangumarda (Australian, Australia) below, where the possessor takes the genitive case suffix *-mili* in alienable possession, as in (2a), while the possessor remains unmarked in inalienable possession, as in (2b):

- (2) a. *ngaju-mili japun*  
 1SG-GEN joey  
 'my joey (infant marsupial)'  
 b. *ngaju mùso*  
 1SG hand  
 'my hand' (Sharp 2004: 306, 313)

The second contrast captures the fact that the possessor has an independent expression in alienable possession, but a bound expression in inalienable possession. This is the case in Puyuma (Austronesian, Taiwan), as shown in (3a) and (3b) respectively:

- (3) a. *nantu basak*  
 3NOM.POSS bag  
 'their bags'  
 b. *temama-taw*  
 father-3.POSS  
 'their father' (Teng 2008: 64, 98)

The third contrast describes that the possessor often fuses with the inalienable noun stem, but not with the alienable noun stem. The fused (or portmanteau) expression of inalienable possession typically concerns high-frequency combinations of relational nouns and their possessor, such as 1<sup>st</sup> person possessors of kinship terms. This is demonstrated for Ungarinyin (Australian, Australia) in (4b) below. Example (4a) shows that the alienable possessive markers is a free form:

- (4) a. *dámbun ŋinaŋga*  
 country 1SG.POSS  
 'my country'

b. *idja*

1SG.mother

'my mother' (cf. *ɲara* 'mother') (Rumsey 1982: 51, 50)

Crucially, only the two contrasts exemplified in (3) and (4) illustrate oppositions in the form of the possessive person marker, which are the likely outcome of the increasing phonological reduction and morphological bonding of the person marker with the possessum.<sup>7</sup> This process is generally conceived of in terms of the cline in (5) (see e.g. Hopper & Traugott 1993: 7; Siewierska 2004: 262; Corbett 2006: 13):

(5) word > clitic > affix > fused marker (> Ø)

The contrast exemplified in (2), however, illustrates an opposition in function: while the person marker in (2a) uniquely identifies the relationship in the NP as a possessive one, as it is marked with the genitive case suffix *-mili*, the person marker in (2b) lacks this ability, since it remains unmarked. In other words, only the person marker in (2a) – but not the one in (2b) – uniquely identifies the relationship in the NP as a possessive one, and thus conveys possessor role information.<sup>8</sup> Loss of the ability to convey role information over time often goes paired with a decrease in other semantic distinctions, such as person, number, gender, inclusivity and honorificity (Bakker & Siewierska 2009). Although these changes are likely to interact with the formal ones depicted in (5), they are logically independent of each other. In fact, this is already demonstrated by Nyangumarda in (2), where the alienability contrast is implemented solely via the presence/absence of a genitive case marker, rather than via the formal realization of the person marker. A similar example, but then for bound person marking, comes from Udihe (Altaic, Russia). In Udihe, alienably possessed nouns (6a) and inalienably possessed nouns (6b) take the same set of person suffixes. However, like the independent person forms in Nyangumarda, these suffixes only overtly identify the possessive nature of the relationship in alienable possession, where they combine with the possessive marker *-ɲi*; in

<sup>7</sup> A related, but logically independent contrast is that of (segmental) length of coding: inalienable possessive coding is always shorter than alienable possessive coding (Haiman 1983: 795, Haspelmath 2008: 20–21).

<sup>8</sup> As pointed out by an anonymous referee, this does not mean that the inalienable person marker may not be interpreted as a possessor when co-occurring with a noun, as shown by its translation 'my'. The important point here is that while the marker in (2a) carries the possessor role information itself, the marker in (2b) does not.

inalienable possession the suffixes do not combine with a dedicated possessive marker, and thus do not carry the possessor role information themselves. Following terminology by Nichols & Bickel (2013b), the choice between such ‘simple’ versus ‘compound’ suffixes marks inalienable versus alienable possession:

- (6) a. *ja:-ŋi-ni*  
       cow-POSS-3SG  
       ‘his/her cow’  
    b. *neŋu-ni*  
       younger.sibling-3SG  
       ‘his/her younger sibling’ (Nikolaeva & Tolskaya 2001: 112, 111)

The aim of the present paper is to carefully tease apart the functional and formal changes involved in the grammaticalization of possessive person markers, unlike for instance Haiman (1983, 1985) and Haspelmath (2008a). Crucially, the ability of the person marker to convey possessor role information is treated here separately as a functional aspect of grammaticalization.<sup>9</sup> This development is characterized by the marker’s gradual loss in referential potential, and follows its own grammaticalization cline, as presented and discussed in the next subsection.

### 3.2.2 A four-part typology of referential/agreement markers

The functional dimension of the grammaticalization of independent pronouns into bound markers of agreement is captured by a four-part typology of person markers, drawing on recent work by Hengeveld (2012). The typology makes a systematic distinction between so-called *referential markers* and *agreement markers*, irrespective of their formal realization. The former is a neutral term for markers that function like pronouns: referential expressions of the (alienable or inalienable) possessor. The latter term is used for markers that are merely the result of an agreement rule: they simply copy the relevant features from the

---

<sup>9</sup> Interestingly, while the ability of free person forms (and nouns) to provide possessor role information is widely known as ‘the genitive case’ (when expressed inflectionally), no accepted terminology exists for the same function of bound person forms, although some suggestions have been made in the recent literature, including that of ‘agreement differentiation’ (Bickel & Nichols 2007: 197, 223–224) and ‘index-set’ (Haspelmath 2013a: 203).

possessor NP to the possessed noun, and make no independent referential contribution. Such markers are referred to in this paper as having no ‘referential potential’.

The distinction between referential markers and agreement markers has been the subject of much debate, and different authors have put the dividing line at different points (e.g. Jelinek 1984; Mithun 1991b; Bresnan & Mchombo 1987; Siewierska 1999, 2001, 2004: 120–127; Corbett 2003: 184–192; LeSourd 2006; Bickel & Nichols 2007: 232–234; Schultze-Berndt 2011, Croft 2013, Haspelmath 2013a). The typology applied in this paper differs essentially from those put forward in previous work in two respects. First, it does not focus solely on bound person forms (e.g. Siewierska 2004: 127), nor does it determine the referentiality of a marker based on its formal expression type (e.g. Nichols 1992: 79–80). Instead, it applies to markers of any morpho-phonological realization, i.e. free forms, clitics, affixes and fused forms. As such, the typology systematically separates function and form. Second, person forms ‘in between’ free pronouns and bound agreement markers are given a unified treatment as referential markers or as markers of agreement, depending on the distribution of (grammatical) feature information in the language and the construction in question. This approach departs from traditional analyses where such markers receive a uniform treatment as referential markers, agreement markers, or both. The approach taken here nevertheless is supported by recent work (e.g. Fedden et al. 2013; Iemmolo & Witzlack-Makarevich 2013), demonstrating that person markers show highly language-specific and construction-specific behavior, and can therefore hardly be considered a single type. Each of the two features of the typology are further discussed and exemplified below.

The four types of person markers in Hengeveld’s typology can be ordered along a likely path of grammaticalization, in which each marker reflects a diachronic stage in the development of pronoun towards agreement marker. The corresponding cline is given in (7) below:

- (7) unique referential marker  
       > appositional referential marker  
           > contextual agreement marker  
               > syntactic agreement marker

The position of a marker on this cline is based on (i) the possible co-occurrence of a corresponding possessor noun or pronoun inside the same possessive NP, and (ii) when the (pro)noun occurs optionally, the distribution of feature information between the (pro)noun and the person marker. The first diagnostic

distinguishes between so-called *unique referential markers* – the functional equivalents of deictic or anaphoric pronouns – and *syntactic agreement markers* – the functional equivalents of canonical agreement markers. The second diagnostic applies to the markers that occupy a diachronically intermediate position on the cline, distinguishing between so-called *appositional referential markers* and *contextual agreement markers*. The typology is summarized in Table 1.<sup>10</sup>

Table 1: A four-part typology of referential/agreement markers

Type of referential/agreement marker	unique referential marker	appositional referential marker	contextual agreement marker	syntactic agreement marker
Co-occurrence with possessor	banned		optional	obligatory
Marker is richer in terms of (grammatical) feature information than possessor NP	n.a.	yes	no	n.a.

Given the fact that unique referential markers are the functional equivalents of deictic or anaphoric pronouns, and pronouns give rise to agreement markers (rather than the other way around, see e.g. Lehmann 1982a, Siewierska 2004: 262, and the work by Givón and Ariel discussed below), possessive person markers are expected to start out as unique referential markers. As the name suggests, markers of this type are the sole instantiations of the possessor in the possessive NP. They behave as pronouns in that they may only co-occur with another possessor when the latter is located outside the possessive phrase boundaries, e.g. as a left or right dislocated topic (as in *John, his house burnt down completely.*) or simply as the antecedent in preceding discourse (as in *John is moving to Canada. His house will be sold.*). Other common labels for unique referential markers are ‘anaphoric agreement markers’ (Bresnan & Mchombo 1987: 741; Lehmann 1982a: 219; Siewierska 1999: 226) and

<sup>10</sup> Other types of markers are not logically possible, such as, for instance, ‘appositional agreement markers’, which are excluded since the possessor (pro)noun cannot be in an appositional relation with a non-referential expression, or ‘syntactic referential markers’, which are excluded because the obligatory expression of a possessor (pro)noun shows that the marker has lost the ability to refer on its own (cf. the discussion below).

'pronominal agreement markers' (Bickel & Nichols 2007: 233). A language with such markers is Macushi (Cariban; Brazil, Guyana, Venezuela). Inalienable possession is marked with a person prefix on the possessum, which is in complementary distribution with a possessor NP, whether nominal (8a) or pronominal (8b). This shows that the prefix functions as the (inalienable) possessor itself:

- (8) a. *kaikusi no'pî*  
       jaguar wife  
       'the jaguar's wife'  
    b. (\**mîkîrî*) *i-no'pî*  
       3           3.POSS-wife  
       'its wife' (Abbott 1991: 62, 85, 87)

Importantly, referential potential is determined for nominal and pronominal possessors individually in this paper. This is warranted by the fact that the complementary distribution between the person marker and the possessor NP may be split, depending on the (pro)nominal character of the latter (cf. Siewierska 1999: 228–229, 2004: 151–154).<sup>11</sup> In one type of split, attested among the languages in my sample, the marker co-occurs with a lexical NP, but not with an independent pronoun. In languages with this type of split, the marker is of the unique referential type, but only in pronominal contexts. In NPs with a nominal possessor, the marker functions as an appositional referential marker or a contextual agreement marker.<sup>12,13</sup> The result of taking into account such construction-specific splits in person marking is a more fine-grained analysis of referential potential than in existing typologies, such as Siewierska's (1999, 2004),

---

<sup>11</sup> Another split attested in the sample is one between proper nouns and common nouns: in Ungarinyin (Australia, Worroran; Rumsey 1982: 70–71) the possessive person marker may only co-occur with a proper noun, not with a common noun (which takes a genitive case marker). Similar splits also occur in other grammatical domains, of which main clauses have received most attention so far (e.g. Siewierska 2004: 149–162, Bickel 2010, Witzlack-Makarevich 2011: 175–181, and Van Lier & Van Rijn 2013 for an investigation of nominalizations).

<sup>12</sup> It is not logically possible for such a marker to be a syntactic agreement marker, because it can be used without a possessor NP, namely in pronominal contexts.

<sup>13</sup> The inverse pattern, i.e. co-occurrence with an independent pronoun but not with a lexical NP, is also attested among the world's languages, although not very frequently (Siewierska 2004: 152–154). The isolate Urarina (Peru) is the sole language demonstrating this pattern in my sample (Olawsky 2006: 339–340).

where such splits are not accounted for. For the sake of cross-linguistic comparability, I only provide examples of NPs with nominal possessors in this section.

Over time, the possessor (pro)noun may enter the possessive NP, as restrictions on its co-occurrence with the person marker are loosened. In the final stages of grammaticalization, the (pro)nominal occurs obligatorily, showing that the person marker has lost the ability to refer to the possessor on its own. Following Hengeveld (2012: 474, as well as Lehmann 1982a: 219 and Siewierska 2004: 126), I call such person markers ‘syntactic agreement markers’. Interestingly, I know of no language that employs such markers in possessive NPs. The few attested cases of syntactic agreement are found for subjects in main clauses, and are largely restricted to western Europe (Siewierska 1999: 238, 2004: 268–273).<sup>14</sup>

As mentioned previously, many of the world’s languages employ possessive person markers that are in between unique referential markers and markers of syntactic agreement. This third type of person marker co-occurs optionally with a corresponding (pro)noun. According to Givón (1976, 1983), the development of verbal markers of this type is the outcome of the overuse of topicalization constructions (‘The man, he came’), by which the dislocated topic becomes the subject and the (unique referential) marker becomes cliticized and eventually morphologically bound to the verb (‘The man he-came’). As pointed out by Siewierska (2004: 264), this is also a likely scenario for possessors, as they are typically human (or at least animate), definite and therefore highly topical. By contrast, Ariel (2000) argues that the development of such markers starts with the cliticization of first and second person possessors, due to their high accessibility in the memory store of the hearer. Further morpho-phonological reduction induces the need for a (pro)nominal expression of the possessor inside the possessive NP. A third potential source of this type of person marker are so-called ‘possessor raising/external possession constructions’ (cf. Haspelmath & König 1998), as proposed by, for instance, McGregor (1996).

A language in which a nominal and a prefixal expression of the possessor co-occur optionally is Nyulnyul (Australian, Australia). In (9a) the possessor is

---

<sup>14</sup> A language that is arguably in the process of developing possessive markers of this type is Skou (Skou, Indonesia), where the presence of a co-occurring possessor pronoun is optional but preferred (Donohue 2004: 269). Interestingly, as mentioned previously, the language already employs syntactic agreement markers for subjects in main clauses.



expressed by both a noun and a person prefix, while in (9b) the nominal is dropped and only the prefix remains.<sup>15</sup>

- (9) a. *bin wamb ni-mal*  
       this man 3MIN.POSS-hand  
       ‘this man’s hand’
- b. *ni-mal*  
       3MIN.POSS-hand  
       ‘his hand’ (McGregor 2012: 422, 117)

Person markers that co-occur optionally with a corresponding (pro)noun, as exemplified here for Nyulnyul, have received three types of analyses in the literature. Under the first analysis, the person marker is univocally regarded as a referential instantiation of the possessor. This means that when the (pro)noun is expressed in the NP, as in (9a), the marker and the (pro)noun are in an appositional relation, such that the possessor is referred to twice. This analysis is adopted in most functional frameworks (e.g. De Groot & Limburg 1986; Hengeveld & Mackenzie 2008; but also Kibrik 2011 and Schultze-Berndt 2011). Under the second analysis, the marker always expresses agreement. When the (pro)noun is absent, as in (9b), the marker is considered to agree with an underspecified possessor called ‘pro’. This analysis is adopted in much generative work (e.g. Jelinek 1984; Baker 1996) and partially in Nichols (1992: 59, 78–80). The third analysis considers the marker a case of agreement when the (pro)noun is expressed, as in (9a), but referential when it is dropped, as in (9b). This analysis is adopted in Siewierska’s (1999, 2004: 120–127) typology of person marking, who simply refers to such markers as ‘ambiguous agreement markers’.

In this paper, I adopt an alternative analysis to person markers that co-occur with optional possessors, as in Nyulnyul. Following Hengeveld (2012), this approach combines the first and second analyses discussed above, but without considering these to be present in parallel, as in the third analysis by Siewierska. Rather, the referential potential of possessive person markers is determined on a language-specific and a construction-specific basis. This means that the marker is referential in some languages, but expresses agreement in others. Moreover, even within the same language, some markers may be

---

<sup>15</sup> The construction may be additionally accompanied an emphatic or oblique pronoun, which is optional in inalienable possession (but obligatory in alienable possession; McGregor 2012: 422, p.c.). Since the focus here is on the possessive prefix, this pronoun is not given in the examples below.

referential in nature, while others express agreement. Markers of the first type refer independently, and may occur in apposition with a (pro)nominal expression of the possessor. They are correspondingly referred to as ‘appositional referential markers’. Markers of the second type agree with a possessor (pro)noun, which is overtly expressed in the NP, or is present only in the discourse context. They are correspondingly referred to as ‘contextual agreement markers’.

Such a language-specific and construction-specific approach to the referential potential of person marking is warranted by the fact that the absence of the possessor (pro)noun may have two logically independent reasons. On the one hand, the (pro)noun may remain unexpressed because the marker itself provides sufficient reference to the possessor. This is the case in languages with unique referential markers, such as Macushi (8), where the referential nature of the person marker prevents the use of a (pro)noun inside the NP altogether.

On the other hand, languages may leave the possessor (pro)noun unexpressed because it is retrievable from the discourse context. The extent to which arguments of verbs are actually expressed has been typologized under the name of *referential density* (Bickel 2003). Interestingly, this typological parameter also applies to possessive NPs. A case in point is Tidore (West Papuan, Indonesia), where possessors (as well as verbal arguments) tend to remain unexpressed when the context provides sufficient information. As a result, possessive NPs (as well as entire clauses) often lack any overt specification of the possessor, as illustrated in (10):

- (10) *Dadi towaro papa la ngone fo-wako*  
 So take.leave father so 1PL.INCL 1PL.INCL.A-return  
 ‘So take leave of (your) father so that we go home.’ (van Staden 2000: 404)

From facts like these we may conclude, as mentioned above, that a distinction has to be made between possessor (pro)nouns that are not expressed because the possessive person marker is referential by itself, and those that are not expressed because the language concerned has a low referential density. In the latter case the possessive person marker is an agreement marker, despite the absence of the possessor (pro)noun, and agreement occurs with the possessor in the discourse context.

In order to determine whether a possessive person marker is a unique referential marker or a contextual agreement marker a simple test may be applied. I assume here an analysis in which agreement is a matter of copying (grammatical) feature information from the possessor NP to the possessum

(following Corbett's (2003, 2006) notion of canonical agreement, see also Lehmann 1982a: 203 and Bickel & Nichols 2007: 229). In such an analysis, only those features can be copied that can be retrieved from the possessor (pro)noun. This means that if a possessive person marker is richer in terms of the semantic information expressed in grammatical features than the possessor (pro)noun itself, it cannot be a contextual agreement marker but must be an appositional referential marker. The examples from Nyulnyul, partially repeated from (9) above, and the isolate Burushaski (Pakistan) below illustrate this. In these examples, the optionality of the possessor NP is indicated with parentheses, and the expression of possessor role information by the person marker is abbreviated as 'POSS' in glossing:

- (11) a. (*bin wamb*) *ni-mal*  
 this man 3MIN.POSS-hand  
 'this man's hand'
- b. (*wamb-in*) *i-n-dam-Ø* *yiil jan*  
 man-ERG 3.NOM-CM-hit-3MIN.ACC dog 1MIN.OBL  
 'The man hit my dog.' (McGregor 2012: 585)
- (12) a. (*hir-e*) *i-yas*  
 man-GEN 3SG.HUM.M-sister  
 'the man's sister'
- b. (*hir*) *i-phús-im-i*  
 man.ABS 3SG.HUM.M-bound-AOR-3SG.HUM.M.A  
 'he tied the man up' (Grune 1998, p.c.; Berger 1998: 117)

Example (11a) shows that the possessive person marker expresses person and role information (and number, but as I will explain below, number could not be taken into account in this study), while only the person information can be copied from the possessor nominal, which lacks role information. The possessive person marker is thus richer in terms of the grammatical features expressed than the possessor nominal, and for that reason must be referential in nature. Example (11b) shows that the person marker in (11a) indeed expresses possessor role information, as subjects require a different person marker.<sup>16</sup> The marker in

<sup>16</sup> In this paper, I consider a set of person markers unique to the possessive NP when it differs in at least one form from another set. Partial correspondences between sets of person markers are cross-linguistically common, but not easily classified in a systematic way (see Siewierska 1998 for an attempt). Therefore, these have not been taken into account in this study.

(11a) is thus unique for possessive constructions, and as such expresses its possessor role.

Example (12a) shows that the possessive person marker expresses person, number and (human masculine) gender, features that are also encoded by the possessor nominal, which furthermore encodes role information via a genitive case affix. In this case an agreement analysis is warranted, as the relevant features to be copied to the person marker are indeed available from the possessor noun. The fact that the person marker in (12a) does not provide possessor role information is demonstrated in example (12b), which shows that the marker is used to mark non-actors in general.

Note that another possibility is for the person marker and the possessor (pro)noun to express the same amount of feature information. In this case, an agreement analysis of the person marker is also warranted, since the marker can copy all of the relevant information from the (pro)noun. An example from Koasati (Muskogean, United States) is given in (13):

- (13) a. (*jhocilím*) *im-laykí*  
           star       3-dung  
           ‘a meteor’ (lit. ‘star’s dung’)
- b. (*nitá*) *im-walí:ka-t*  
           bear    3-run-PST  
           ‘(He) ran away from the bear.’ (Kimball 1991: 433, 131)

Example (13a) shows that the possessive person marker expresses the same person information as provided by the possessor nominal. As shown in (13b), the marker in (13a) does not provide possessor role information, since it also marks a range of other roles in main clauses, among which malefactive, as in the case of *nitá* ‘bear’ in (13b). This means that the relevant information, in this case just person information, can be readily copied from the possessor nominal, and the marker is correspondingly analyzed as a contextual agreement marker.

To sum up, contextual agreement marking and appositional referential marking are two different person marking strategies, which can be distinguished from one another by considering the distribution of (grammatical) feature information between the possessor (pro)noun on the one hand and the

possessive person marker on the other.<sup>17</sup> While referential markers expand on the feature information of the possessor NP, agreement markers can only express those features provided by the possessor NP, as they are mere copies of the (pro)noun and thus have no independent semantic contribution to make to the utterance. Importantly, I only focus on the distribution of possessor role information in this paper, not on that of other features, such as number and gender. This has two main reasons. The first is that many of the world's languages lack such features. With respect to number, Haspelmath (2013b) for instance demonstrates that over 30% the 291 languages in his sample completely lack plural marking for (a set of) nouns. Moreover, as demonstrated by Rijkhoff (2002: 38–41, 104–117), many languages, including Nyulnyul and Koasati exemplified above, employ so-called 'set nouns': nouns that in their unmarked form do not denote a single item, like regular count nouns as in English, but rather a set of items, which may consist of one item (a singleton set) or multiple items that together form a collective (a collective set). Instead of number, such nouns may express a feature Rijkhoff refers to as 'nominal aspect', which concerns marking the kind of set involved. Importantly, more than half of the 52 languages in Rijkhoff's sample employs (a group of) set nouns, suggesting that the absence of grammatical number is a cross-linguistically frequent phenomenon. The complete absence of grammatical gender is presumably even more common: while more than half of 257 languages studied by Corbett (2013) does not have a nominal gender system, almost 70% of the 378 languages in Siewierska's (2013a) sample does not distinguish gender in independent person forms.

The second reason for not taking number, gender or other features into account is that their distribution between nouns and person markers cannot be determined in a straightforward way. As pointed out above, set nouns do not express number but nominal aspect, and are therefore not well comparable to person forms in terms of number information. Gender is a problematic feature in that the gender of a noun is usually determined by forms outside of the NP itself, and thus cannot be determined independently.

---

<sup>17</sup> Another diagnostic to determine whether indexes are referential or not that has been proposed in the literature (Bresnan & Mchombo 1987; Evans 1999, 2002; Austin & Bresnan 1996) can only be applied at the clausal level and is not well suitable for possessive NPs. It relies on differences in specificity and definiteness, i.e. the ability of the person marker to have a generic reading, to occur with nouns taking quantifiers or expressions such as 'nobody' or 'nothing', and their presence with nominals that are questioned. Since possessors are typically definite and specific, this diagnostic is hard to apply within the possessive NP (for further criticism, see Mithun 2003).

Importantly, the functions carried out by the bound person markers in examples (11) to (13) above could just as well have been carried out by free forms. This is demonstrated in example (14) from Ungarinyin and example (15) from Tiwi (both Australian, Australia):

- (14) a. (*Gádbuŋu*) *anàŋga* *dámbun*  
 Gadbuŋu 3SG.M.POSS country  
 ‘Gadbungu’s country’  
 b. *ada* *ŋima* *ɖjiri-guɖe*  
 sit I.will.do 3SG.M-COM  
 ‘I will sit down with him.’ (Rumsey 1982: 70, 74)
- (15) a. (*Purukupaɭi*) *ŋara* *mɔɭani*  
 Purukupaɭi 3SG.M son  
 ‘Purukupaɭi’s son’  
 b. *ŋara* *jiikaɭimi* *jikwani*  
 3SG.M he.made fire  
 ‘He made fire.’ (Osborne 1974: 74, 61)

The examples in (14) and (15) are functionally completely parallel to (11) and (13). In (14a) it is shown that the free possessive person marker carries possessor role information, while the possessor nominal does not. Example (14b) shows that the possessive person marker indeed carries role information, since another set of free person forms is used outside possessive constructions, e.g. for comitatives in main clauses. The possessive person marker thus expands on the role information provided by the possessor nominal, and therefore must be referential in nature, just like the bound person marker in Nyulnyul (11a). In (15a), by contrast, neither the person marker nor the possessor nominal carries possessor role information. As shown in (15b), the marker is not unique to possessive constructions, since it also marks actors in main clauses, among other roles. This means that the marker does not expand on the role information provided by the possessor nominal, and in this case an agreement analysis is warranted, just as with the bound person markers in Koasati (13a).

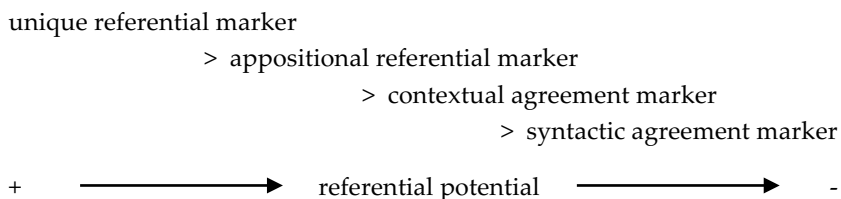
In sum, this paper applies a four-part typology of person markers, which makes a basic distinction between referential markers – those that instantiate the possessor – and agreement markers – those that copy information from a possessor (pro)noun to the possessed noun. Over time, a person marker is expected to develop from a referential marker to a marker of agreement, which is a development characterized by (i) its growing reliance on the expression of a

possessor (pro)noun inside the NP, and (ii) a gradual loss in (grammatical) feature information compared to the co-occurring possessor (pro)noun.<sup>18</sup> This process is expected to prefer markers of inalienable possession, as well as markers of a reduced morpho-phonological form (i.e. clitics rather than separate words and affixes rather than clitics). These predictions are tested in this study, and further discussed in the next section.

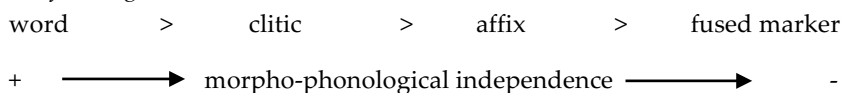
### 3.3 Hypotheses

The functional and formal changes involved in the diachronic development of possessive person markers can be conceived of as two distinct grammaticalization clines. The functional cline captures the marker's gradual loss of referentiality, following the four-part typology presented in the previous subsection. The formal cline captures the marker's gradual loss of morpho-phonological independence, or, formulated the other way around, its increase in morpho-phonological bonding with the possessum. Each cline is presented in (16) below:

(16) *The functional grammaticalization cline:*



*The formal grammaticalization cline:*



Three hypotheses are tested in this study. First, I investigate the relationship between alienable possession and inalienable possession in terms of function (hypothesis (i)) and form (hypothesis (ii)). I predict that if a language has a grammaticalized alienability split, person markers of alienable possession never

<sup>18</sup> The second diagnostic is not applied to fused forms, since they cannot be analyzed separately from the noun stem to which they attach, and are therefore unanalyzable in terms of the expression of possessor role information.

proceed further down the two clines than person markers of inalienable possession. This diachronic prediction has a clear synchronic counterpart: alienable possessive person markers show the same or a higher degree of referential potential and morpho-phonological independence as inalienable possessive person markers, but never a lower degree. The two corresponding hypotheses are formulated below:

- (i) A person marker used in alienable possession never shows a lower degree of referentiality than a person marker used in inalienable possession.
- (ii) A person marker used in alienable possession never shows a lower degree of formal independence than a person marker used in inalienable possession.

Hypothesis (i) predicts that, for instance, a language with contextual agreement markers in alienable possession employs contextual or syntactic agreement markers in inalienable possession, but never unique or appositional referential markers. Similarly, hypothesis (ii) predicts that, for instance, a language with person affixes in alienable possession may employ affixal or fusional markers in inalienable possession, but never free or cliticized ones. Both hypotheses follow previous findings in the literature showing that inalienable possessive marking is historically older, and thus more grammaticalized, than alienable possessive marking (such as Haiman 1983, 1985 and Haspelmath 2008a discussed above). The novelty of the hypotheses resides in the fact that the referential potential of alienable and inalienable means of person marking (hypothesis (i)) is investigated independently of its formal realization (hypothesis (ii)), as enabled by Hengeveld's four-part typology of referential/agreement markers presented in the previous subsection.

The third hypothesis tested in this study investigates the relationship between the two clines in (16), independently of the alienability distinction. I predict that as a marker moves down the functional cline, it cannot move up the formal cline. In other words, a marker that loses referential potential should never gain formal independence. The synchronic correlate of this diachronic prediction is that markers of a lower degree of referentiality (those on the right-hand side of the functional cline) never have a more independent expression form than markers of a higher degree of referentiality (those on the left-hand side of the functional cline) in individual languages. The corresponding hypothesis is formulated below:



- (iii) In individual languages, a possessive person marker with a lower degree of referentiality never has a more independent expression form than a possessive person marker with a higher degree of referentiality.

According to this hypothesis, there should be no languages, for instance, with appositional referential markers in the form of words, but unique referential markers in the form of clitics, or with contextual agreement markers in the form of clitics, but appositional referential markers in the form of affixes. As mentioned previously, this hypothesis is more specific about the relationship between function and form than the parallel path hypothesis in two respects. First, it expects the clines to be related in a relative sense, rather than in an absolute sense. For instance, it does not exclude the presence of languages where unique referential markers and appositional referential markers both have the form of clitics, or where appositional referential markers and contextual agreement markers both have the form of affixes. Hence, a move down the functional cline is not expected to necessarily go paired with a move down the formal cline. This prediction follows previous work on referential/agreement marking (e.g. Siewierska 2004: 262; Corbett 2006: 265), suggesting that person markers may very well occupy different positions on the two clines in different languages. Second, in addition to predicting a cross-linguistic tendency, the hypothesis in (iii) is expected to apply language-internally, and is tested by studying languages with minimally two sets of possessive person markers. This may be an alienable set and an inalienable set, or two alienable or inalienable sets of person markers, if a language employs more than one such set. The hypothesis thus extends beyond the alienability split and the corresponding hypotheses in (i) and (ii).

In the next section, I first present the language sample used to test each hypothesis (3.4.1) and then show which types of possessive NPs and person marking strategies are taken into account (3.4.2). Each hypothesis will be addressed in turn in Section 3.5.

## 3.4 Methodology

### 3.4.1 The language sample

The sample used in this study is created in such a way that it reflects the highest possible degree of genetic and geographical diversity. The genetic criterion is met by adopting the sampling method of Rijkhoff et al. (1993; Rijkhoff & Bakker

1998), which makes use of so-called ‘Diversity Values’ (henceforth ‘DV’). These values determine the number of languages to be selected from each language (sub)family based on the internal complexity of that (sub)family, given a particular preferred sample size. More internally complex (sub)families have a higher DV than less internally complex (sub)families, and are therefore represented by a higher number of languages in the sample. Isolates are, by definition, part of any sample. The method is applied to Ruhlen’s (1991) classification of the world’s languages.<sup>19</sup> The genetic criterion is combined with a geographical one: where possible, the languages selected on the basis of the genetic criterion are spoken in non-contiguous areas.

Within the restrictions of the genetic and geographical criteria, the sample satisfies a third, typological criterion. Given the hypotheses tested in this study, I made sure that the sample exclusively contains languages with a grammaticalized alienability split that use person marking in both alienable and inalienable possession. Note that the alienability split need not be made on the basis of two separate sets of (free or bound) person markers, as in Puyuma (3) and Ungarinyin (4) exemplified in Section 3.2.1. The sample also contains languages where the same set of person markers is used in both alienable and inalienable possession, and the distinction between both types of possession is made via the presence of an additional possessive morpheme in alienable possession, as in Bambara (2) and Udihe (6), and/or a difference in the referential potential of the person marker.

In just over half of the sample languages only NPs with a pronominal possessor could be investigated. In NPs with a nominal possessor, they only use person markers for inalienable possession, do not use possessive person markers at all (as in Macushi (8)), or lack an alienability split altogether.<sup>20</sup> Languages without any form of possessive person marking are not taken into account. Without language-specific diachronic information, it cannot be determined whether the absence of person marking marks the beginning or the end of

---

<sup>19</sup> Ruhlen’s classification is often criticized for positing large, little supported, language families, such as the Amerindian family (e.g. Kaufman 1990; Nichols 1990; Campbell & Poser 2008). The classification is nevertheless very suitable for creating relatively small language samples, since a minimal sample composed with the DV technique consists of one language per family. Moreover, the method requires the selection of a large number of languages from internally complex language families, which compensates for the use of a rougher classification like Ruhlen’s.

<sup>20</sup> Other logical possibilities are not attested, such as languages in which only alienable possession is person-marked (cf. Siewierska 2004: 138), or languages that employ an alienability split only with nominal possessors (cf. Dahl & Koptjevskaja-Tamm 1998: 47).

grammaticalization, making languages of this kind unreliable candidates for testing the hypotheses in this study.

For this study, I started with an initial sample of 62 languages.<sup>21</sup> Twenty (sub)families (including six isolates) do not contain any languages that have an alienability split, either with nominal or pronominal possessors. Accordingly, these (sub)families are excluded from the investigation. As noted in Section 3.2.1, the lack of an alienability split is especially common among Eurasian and South-East Asian languages. The three languages Etruscan, Hurrian and Meroitic selected by the sampling method were also removed from the sample due to insufficient data. The result is a 39-language sample, which is presented in Table 2 below. Note that, despite the exclusion of certain language (sub)families, all six of Dryer's (1992) macro-areas (Africa, Eurasia, South-East Asia & Oceanic, Australia & New Guinea, North America and South America) are well represented in the sample.

The language (sub)families excluded from the sample appear in shading. The counts between brackets specify per (sub)family how many languages were selected on the basis of the typological stratification given above (before the slash) out of the languages originally required by Rijkhoff et al.'s sampling technique (after the slash). For instance, only three out of the seven Austric language groups contain a language that implements an alienability split and uses person marking as a means of coding alienable and inalienable possession. Therefore, only three Austric languages are part of the sample. Language isolates are abbreviated as 'LI'; Burushaski is the only isolate included in this study.

Table 2: The 39-language sample

<b>Language family (Ruhlen 1991)</b>	<b>Subfamilies</b>	<b>Language selected</b>	<b>Language family (WALS)</b>
Afro-Asiatic	Chadic (1/1)	Lele	Afro-Asiatic
	Berber (1/1)	Tamashek	Afro-Asiatic

<sup>21</sup> The number of languages to be selected from each (sub)family in a sample counting 60 languages is specified in Rijkhoff et al. (1993: 186). However, Rijkhoff et al. (1993) make use of the first edition of Ruhlen's classification (Ruhlen 1987); in the second edition (Ruhlen 1991), which I use in this study, two additional first-order language families are distinguished: Korean-Japanese-Ainu and Kartvelian. Each requires representation by a single language. This brings the initial sample size to 62 languages.

Table 2: The 39-language sample

Language family (Ruhlen 1991)	Subfamilies		Language selected	Language family (WALS)
	Semitic (1/1)		Maltese	Afro-Asiatic
Altaic (1/1)			Udihe	Altaic
Amerind (8/9)	Northern (2/2)	Penutian (1/1)	Koasati	Muskogean
	Hokan (1/1)		Diegueño	Hokan
	Andean (1/1)		Urarina	isolate
	Eq.-Tucanoan (2/2)	Equatorial (1/1)	Paumarí	Arauan
		M.-Tucanoan (1/1)	Hup	Nadahup
	Ge-Pano-Carib (2/2)	M.-Carib (1/1)	Tiriyo	Cariban
	Ge-Pano (1/1)		Bororo	Bororoan
	Centr. Amerind (0/1)		Kiowa	Kiowa-Tanoan
	Chib.-Paezan (1/1)		Sanuma	Yanomam
Australian (4/4)	Nyulnyulan (1/1)		Nyulnyul	Nyulnyulan
	Worrorran (1/1)		Ungarinyin	Worrorran
	Pama-Nyungan (1/1)		Nyangu-marda	Pama-Nyungan
	Other (1/1)		Mangarayi	Mangarrayi-Maran
Austic (3/7)	Austro-Tai (2/5)	Daic (0/1)	–	–
		Austro-Malayo-Pol. (1/1)	Biak	Austronesian
		Atayalic (0/1)	–	–
		Paiwanic (1/1)	Puyuma	Austronesian
		Tsouic (0/1)	–	–

Table 2: The 39-language sample

Language family (Ruhlen 1991)	Subfamilies				Language selected	Language family (WALS)
	Austroasiatic (1/1)				Kharia	Austro-Asiatic
	Miao-Yao (0/1)				–	–
Basque (LI)					–	isolate
Burushaski (LI)					Burushaski	isolate
Caucasian (0/1)					–	–
Chukchi-Kamchatkan (0/1)					–	–
Elamo-Dravidian (0/1)					–	–
Eskimo-Aleut (1/1)					West Greenlandic	Eskimo-Aleut
Etruscan (LI)					–	–
Hurrian (LI)					–	–
Indo-Hittite (2/2)	Indo-European (1/1)				Icelandic	Indo-European
	Anatolian (1/1)				Hittite	–
Indo-Pacific (4/7)	Trans New Guinea (1/1)				Inanwatan	Marind
	Sko (1/1)				Skou	Skou
	E. Papuan (1/1)				Nasioi	East Bougainville
	W. Papuan (1/1)				Hatam	West Papuan
	Other subfamilies (0/3)				–	–
Kartvelian (0/1)					–	–
Ket (LI)					–	isolate
Khoisan (0/1)					Ju   'hoan	Kxa
Korean-Japanese-Ainu (0/1)					–	–
Meroitic (LI)					–	–
Na-Dene (1/1)					Slave	Na-Dene
Nahali (LI)					–	isolate
Niger-Kordofanian (5/5)	Niger-Congo (4/4)	N.-C. Proper (3/3)	Central N.-C. (2/2)	North Centr. N.-C. (1/1)	Dogon	Dogon

		South Centr. N.-C. (1/1)	Ewe	Niger-Congo
		West Atlantic (1/1)	Nigerian Fula	Niger-Congo
		Mande (1/1)	Bambara	Mande
		Kordofanian (1/1)	Krongo	Kadu
Nilo-Saharan (3/3)	East Sudanic (1/1)		Lango	Eastern Sudanic
	Centr. Sudanic (1/1)		Ngiti	Central Sudanic
	Kunama (1/1)		Kunama	isolate
	Nivkh (LI)		–	isolate
	Pidgins and Creoles (0/1)		–	–
Sino-Tibetan (1/2)	Sinitic (1/1)		Mandarin Chinese	Sino-Tibetan
	Tibeto-Karen (0/1)		–	–
	Sumerian (LI)		–	isolate
	Uralic-Yukaghir (0/1)		–	–

### 3.4.2 Delimitation of the constructions investigated

As pointed out in the previous subsection, this study exclusively focuses on languages with person marking at both ends of an alienability split. I examine NPs with both nominal and pronominal possessors, and person markers of any person (and number/gender). Note that in NPs with unique referential markers, the marker itself functions as the possessor pronoun. I do not consider possessive NPs that are morphosyntactically complex, which excludes NPs containing additional attributive modifiers or possessums embedded in possessor chains.

As is widely known, languages vary extensively in the types of nouns that participate in alienable and inalienable possession. This means that there is no one-to-one relationship between the inherently (non-)relational nature of concepts, and the nouns treated (in)alienably across individual languages. I accommodate this discrepancy by focusing solely on two classes of nouns, already mentioned in Section 3.2.1: kinship terms and/or body part terms (relational) versus concrete, inanimate and countable objects (non-relational).

Not only do these nouns most clearly qualify as respectively inherently relational and inherently non-relational on the basis of their semantic properties (items such as ‘leg’ and ‘father’ inherently require a possessor, while items such as ‘book’ or ‘basket’ do not), they also participate in any class of inalienable and alienable nouns across individual languages (as demonstrated above, if a language makes a formal alienability split, a set of kinship terms and/or body part terms always participates in the inalienable class). Moreover, kinship terms and body part terms are primary candidates for person marking: if a language uses person markers in NPs with a nominal possessor, these markers are at least used on kinship and/or body part terms (Stassen 1997: 39). This makes them the most suitable representatives of inalienable (i.e. relational) nouns for the purposes of this study. Note that other types of nouns that often participate in alienable possession, such as ‘house’ and ‘dog’, and inalienable possession, such as parts of wholes and spatial terms, do not contradict the findings of this study. They are incidentally used to illustrate alienable or inalienable possession when examples with nouns from the two classes above are not available.

Note that in a number of sample languages, only a subset of kinship and/or body part terms are treated inalienably. In Puyuma (Austronesian, Taiwan), for instance, only older generation kinship terms, such as ‘mother’, ‘father’ and ‘grandparent’, participate in the inalienable class, while younger generation kinship terms, as well as body parts, are found only in the alienable class (Teng 2008: 97). Similarly, in Nyulnyul only a small proportion of body part nouns (as well as some personal representations, such as ‘shadow’ and ‘name’, and some items of clothing) take the inalienable prefix-set (McGregor 1996: 257–264, 2012: 117–122). This includes the noun *-mal* ‘arm/hand’ exemplified in (9a) in Section 3.2.2. In order to make investigation of such languages possible, I focus solely on those items that participate in the inalienable class in the language in question, since these are generally also the items that are conceived of as the most prototypically relational – and thus as inherently and inseparably related to the possessor – in that language (see e.g. Chappell & McGregor 1996).<sup>22</sup> In Nyulnyul, for instance, only those body part nouns are inalienable that are conceived of as most important, visible and central to the possessor, which includes external body parts like *-mal* ‘arm/hand’, while those considered more peripheral to the possessor, such as internal organs and

---

<sup>22</sup> This observation is supported by Bally (1996), who points out that only those items are generally treated inalienably in a given language that are conceived of as belonging to a human being’s personal sphere, which means that they are viewed as “associated with a person in an habitual, intimate or organic way” (Bally 1996: 33).

bodily products, are formally treated as alienable nouns (McGregor 1996: 257–264). This means that only body part nouns from the former class are taken as representatives of inalienable nouns in Nyulnyul. Conversely, the same approach is adopted for concrete (inanimate), countable objects: when treated as inalienable rather than as alienable nouns in a given language, I focused solely on those nouns that receive the alienable coding – and that are thus conceived of as possessed in a temporary, non-inherent manner – in that language. With respect to Nyulnyul, this means that personal representations, and the items of clothing part of the inalienable class are not taken as representatives of alienable nouns in this language.

Finally, note that about one third of the sample languages uses more than one person marking strategy in alienable possession, inalienable possession or both. This paper takes into account the referential potential and formal expression type of each of these strategies, in order to display the full range of variation found in the realm of possessive person marking. Most often, the different strategies are in complementary distribution, their use governed by a number of factors. One such factor is the division of possessums into lexical subclasses. Diegueño (Hokan; Mexico, United States; Miller 2001: 146), for instance, employs two sets of positional referential markers: one for kinship nouns, and one for all other types of inalienable nouns, including body parts. Another source of distinct strategies of possessive person marking are properties of the possessor NP, such as person, number and its (pro)nominal character. In Lango (Eastern Sudanic, Uganda), for instance, alienable and inalienable possessors trigger a distinct set of suffixes on the possessed noun only when they are pronominal and singular. The alienability distinction is collapsed for plural, pronominal possessors, which are indexed with a single set of suffixes on the possessum (Noonan 1992: 78). Another example comes from Lele (Afro-Asiatic, Chad), in which markers for singular inalienable possessors are suffixed directly onto the possessum, while markers for plural inalienable possessors attach to the genitive marker *dí/dú* (Frajzyngier 2001: 70–71).

A few other languages mark the same possessum twice. This is the case for 1<sup>st</sup> and 2<sup>nd</sup> person dual and plural possessors in Kharia (Austro-Asiatic, India; Peterson 2011: 163), which trigger an enclitic on the inalienable possessum – as shown in example (17a) below – and optionally another enclitic – as shown in example (17b) below. Each possibility is counted as a separate marking strategy with its own referential potential.



- (17) a. *aba=ŋ*  
 father=1SG  
 'my father'
- b. *jiyom=na=iŋ*  
 life=1.POSS=1SG  
 'my life' (Peterson 2011: 166)

With this outline of the languages and constructions examined in the study, I come to the discussion of my results in the next section. I first discuss the relationship between alienable and inalienable possessive person marking in terms of the (cross-)reference/agreement distinction, as predicted by hypothesis (i) (Section 3.5.1). In addition to a discussion of the data (Section 3.5.1.1) a separate subsection is devoted to an interesting finding: the complete absence of possessive person markers expressing syntactic agreement (Section 3.5.1.2). The relationship between alienable and inalienable possessive person marking in terms of their formal expression, as predicted by hypothesis (ii), is discussed in Section 3.5.2. In Section 3.5.3, I provide an explanation for the findings discussed in 3.5.1 and 3.5.2 in terms of the inherent (non-)relationality of possessums in the two types of possession. The relationship between the function and form of possessive person markers, as predicted by hypothesis (iii), is addressed in a final subsection (3.5.4).

## 3.5 Results

### 3.5.1 Hypothesis (i): the referential potential of (in)alienable possessive person marking

#### 3.5.1.1 Data

Table 3 below presents the data relevant for hypothesis (i), which predicts a relationship between alienable and inalienable possessive person markers in terms of their referential potential. It gives the distribution of alienable (AL) and inalienable (INAL) possessive person markers in terms of the four possible types of referential potential in Hengeveld's typology in the 39 languages of the sample. The order of the four types (in the columns) reflects a gradual loss in referential potential, as captured by the corresponding grammaticalization cline in (16) above.

There are sixteen logically possible combinations of alienable and inalienable marking strategies in terms of referential potential (in the rows marked 'a' to 'p'). Those that would be counterexamples to hypothesis (i) are marked with an asterisk and appear in dark shading. Importantly, Table 3 counts combinations of marking strategies (N=70), not languages (N=39). Each row consists of a particular combination of marking strategies in a particular language. Since many of the sample languages employ more than one combination of alienable and/or inalienable marking, they may occupy multiple rows in the table. Furthermore, since referential potential may be split depending on the (pro)nominal character of the possessor, the counts of each combination are specified separately for NPs with a co-occurring possessor noun (in the fifth column) and those with a co-occurring possessor pronoun (in the penultimate column). Note that for unique referential markers, the column labeled 'possessor pronoun' refers to the status of the marker itself, not to the co-occurrence of an independent pronoun, which is by definition impossible for markers of this type. The co-occurrence of a lexical possessor is also impossible for such markers, as indicated by the cells marked 'n.a.' (not applicable). Finally, eight sample languages are not taken into account in this table (Dogon, Fula, Hittite, Inanwatan, Kharia, Kunama, Lango and Tamashek). The referential potential of their person markers could not be determined, as information on the co-occurrence of a possessor pronoun – either in alienable possession, inalienable possession or both – is lacking. Table 1 in Appendix II lists the sample languages that instantiate each combination.

The table shows that none of the combinations marked with an asterisk is attested, except for one, discussed below. Hence, possessive person markers used in alienable possession are almost never of a referential/agreement type further to the right than those used in inalienable possession. Instead, they prefer to be of the same referential/agreement type (combinations a, h and m) or of a type further to the left (combinations c and i). This finding provides clear support for hypothesis (i): in languages with an opposition between alienable and inalienable possession, alienable possessive person markers strongly tend to be of the same or a higher degree of referential potential as inalienable possessive person markers, but never of a lower degree of referential potential. This generalization comprises three sub-relations, which can be formulated as the following implicational universals. First, if unique referential markers are used in inalienable possession, they are also used in alienable possession (compare combination a to \*e, \*f and \*g). Second, if inalienable possession takes appositional referential markers, alienable possession takes such markers as well, or markers of the unique referential type, but never (contextual/syntactic)

Table 3: The distribution of (in)alienable person markers in terms of referential potential

	unique referential marker	appositional referential marker	contextual agreement marker	syntactic agreement marker	Frequencies		
					Possessor noun	Possessor pronoun	Total (N=70)
a	AL, INAL				n.a.	31	31
b	AL	INAL			n.a.	-	-
c	AL		INAL		n.a.	3	3
d	AL			INAL	n.a.	-	-
*e	INAL	AL			n.a.	-	-
*f	INAL		AL		n.a.	-	-
*g	INAL			AL	n.a.	-	-
h		AL, INAL			14	6	20
i		AL	INAL		6	3	9
j		AL		INAL	-	-	-
*k		INAL	AL		1	1	2
*l		INAL		AL	-	-	-
m			AL, INAL		2	3	5
n			AL	INAL	-	-	-
*o			INAL	AL	-	-	-
p				AL, INAL	-	-	-

agreement markers (compare combinations b and h in relation to \*k and \*l). Third, the use of inalienable contextual agreement markers implies the use of alienable markers of any referential potential except syntactic agreement (compare combinations c, i and m to \*o).<sup>23</sup>

The only language that poses two counterexamples to hypothesis (i) is Mangarayi (Australian, Australia). It uses a single set of person markers in both alienable and inalienable possession, which expresses agreement in the former, but is referential in nature in the latter, as shown by the asterisk-marked pattern

<sup>23</sup> Each of these correlations is tested for statistical significance on the basis of a Fisher's Exact test ( $p < 0.05$ ). A Contingency Coefficient (CC) was also computed, which indicates the strength of a statistically significant correlation irrespective of sample size. It ranges from 0 to 1, with 0.10 indicating a weak correlation and  $0.45 \leq$  indicating a strong correlation (Everitt 1977). The first relation is statistically significant ( $p < 0.0001$ ) and very strong ( $CC = 0.676$ ). The second and third relations are not significant, due to the skewed distribution of marking combinations in the table.

'k'. This is the case for both nominal and pronominal possessors. An example of alienable (18a) and inalienable (18b) possession with a nominal possessor is given below:

- (18) a. (*ŋa-ɓuɓuŋ-gu*)                      *ø-baŋam-ŋawu*  
           GEN.M-old.person-GEN.M    ABS.NT-camp-3SG.NF.POSS  
           'the old man's camp'
- b. (*landi*)    *jurgjurg-ŋawu*  
           tree        leaf-3SG.NF.POSS  
           'the leaves of a tree' (Merlan 1982: 66, 74)

The marker is analyzed as referential in nature in inalienable possession (18b), since there it carries information – more specifically possessor role information – that cannot be copied from the possessor NP. In alienable possession (18a), by contrast, the NP takes a genitive case affix, which means that the possessor role information can be readily copied onto the possessive person marker. The marker is correspondingly analyzed as expressing agreement. Note, however, that the fact that the role information in (18a) **can** be copied does not necessarily imply that it also **has** to be copied; in principle, it may still be the case that the marker provides the possessor role information independently, rather than copying this information from the possessor NP. In other words, constructions where the marker and the optional possessor NP share the possessor role information, as in (18a), are less clear-cut cases of agreement than those where the marker provides less information than the NP – and are therefore perhaps better considered to be unanalyzable in terms of (cross-)reference/agreement. As such, the pattern in (18) does not pose a strong counterexample to the hypothesis in (i). A true counterexample to the hypothesis would be a language without genitive case/adpositional marking, but with a unique set of person markers only in inalienable possession, not in alienable possession. None of the languages investigated in this study fit this description.<sup>24</sup>

The data in Table 3 show that the most frequently attested combinations of referential/agreement markers are those in which the alienable and inalienable

---

<sup>24</sup> This observation yields an interesting generalization open to further study: if person markers in inalienable possession express possessor role information, person markers in alienable possession express such information as well. A possible explanation for this generalization is that the presence of a possessor already inherently follows from the semantics of the inalienable noun, which is therefore in far less need of specifying the possessive relationship than non-relational nouns. Section 3.5.3 discusses this explanation in more detail.

possessive person marker share the same referential potential (56 out of the 70 attested combinations – a, h or m – i.e. 80%). Combination a – the use of a unique referential marker for both alienable and inalienable possession – is the most frequently attested (31 out of 70 combinations, i.e. 44%). An example is provided for Puyuma (Austronesian, Taiwan) below, which demonstrates the complementary distribution between the independent possessor pronoun and the alienable (19a) and inalienable (19b) possessive person markers:

- (19) a. *tu=ruma'*            (*\*taytaw*)  
           3.POSS=house    3SG  
           'his/her/their house'
- b. *i*            *tina-taw*            (*\*taytaw*)  
           SG.NOM    mother-3.POSS    3SG  
           'his/her/their mother' (Teng 2008: 94, p.c.)

Another common pattern is the use of possessive person markers of the appositional referential type in both alienable and inalienable possession. This is combination h, which has 20 instances in the sample (i.e. 29%). This pattern is most frequently attested in NPs with a nominal possessor (14 out of 20 combinations, i.e. 70%). A case in point is Ungarinyin (Australian, Australia), which uses independent person forms in alienable possession, but a set of person prefixes in inalienable possession. Both person forms function as appositional referential markers, as they carry possessor role information that the possessor nominal does not carry. This is demonstrated in respectively (20a), repeating (14a) in Section 3.2.2, and (20b) below:

- (20) a. (*Gádbuŋu*)    *anàŋga*            *dámbun*  
           Gadbuŋu    3SG.M.POSS    country  
           'Gadbungu's country'
- b. (*ari*)    *a-marlarr*  
           man    3SG.M-forehead  
           'the man's forehead' (Rumsey 1982: 70, 43; Spronck p.c.)

Combinations in which the two person markers are of a distinct referential potential are considerably less frequent (14 out of the 70 combinations – c, i and k – i.e. 20%). An example of combination c is provided below for the language isolate Urarina (Peru), where the independent person markers coding alienable possession occur in complementary distribution with an additional, independent pronoun (21a), while the person clitics coding inalienable

possession occur optionally with the same pronoun (21b). Hence, co-occurrence restrictions are loosened only for inalienable possession, which directly reflects its more grammaticalized status in comparison to alienable possession:

- (21) a. (\**ii*)    *ii*    *kuriki*  
           2SG    2SG    money  
           ‘your money’  
       b. (*ii*)    *dz=aria*  
           2SG    2SG=family  
           ‘your family’ (Olawsky 2006: 562, 340, p.c.)

A language with combination *i*, for both nominal and pronominal possessors, is Bororo (Bororoan, Brazil). As shown below, the language employs appositional referential markers in the form of clitics in alienable possession (22a), but contextual agreement markers in the form of prefixes in inalienable possession (22b):

- (22) a. (*Barae*)        *eno=moto*  
           (Brazilian)    3PL.POSS=land  
           ‘Brazil’ (lit. ‘Brazilians land’)  
       b. (*Kuruiedi*)    *u-mana*  
           Kuruiedi        3SG-older.brother  
           ‘Kuruiedi’s older brother’ (Crowell 1979: 197, 215)

An explanation for the fact that combinations of (in)alienable possessive person markers with the same referential potential are more common than those with a distinct referential potential is that alienability splits are often already implemented via the formal expression of the person forms. That is, the alienable and inalienable possessive person forms are already of a distinct formal expression type. This is the case in each of the examples provided in (19) to (22). A language in which the alienable and inalienable person marker share the same expression type as well as the same referential potential is Koasati (Muskogean, United States). Koasati uses two sets of contextual agreement markers: one for alienable possession (as well as for indirect objects, benefactives, malefactive and subjects of a class of stative verbs), and another for inalienable possession (as well as for direct objects and subjects of another class of stative verbs) (Kimball 1991: 112). This is shown in respectively (23a), which repeats (13a) in Section 3.2.2, and (23b):

- (23) a. (*jhocilím*) *im-laykí*  
 star 3-dung  
 'a meteor' (lit. 'a star's dung')
- b. (*albatá*) *ø-a:nipó*  
 alligator 3-prepared.meat  
 'alligator meat' (Kimball 1991: 433, 438)

The distribution of marking combinations in the table demonstrates that, in keeping with hypothesis (i), alienable possession shows a stronger preference for (unique/appositional) referential marking (63 out of 70 combinations, i.e. 90%) than inalienable possession (53 out of 70 combinations, i.e. 76%). Hence, referential markers are more likely to be found in alienable possession, than in inalienable possession. However, these counts also show that both alienable possession and inalienable possession – when considered in their own right – prefer the use of referential markers over that of agreement markers. In the languages investigated, possessive person markers thus tend to be referential in nature rather than expressing agreement, irrespective of the type of possession that they mark. A possible explanation for this tendency may be sought in the low contextual salience of possessors in comparison to verbal arguments, especially subjects and objects. Following Keenan and Comrie's (1977) accessibility hierarchy, and subsequent work by Givón (1983) and Ariel (1988, 1990), possessors are less contextually salient – and therefore less accessible in the mind of the hearer – than (subject and object) arguments of verbs. This may be attributed to the embedded position of possessors in the clause. Importantly, this may trigger the need for a more salient means of referent tracking, i.e. a marker that refers to the possessor and that overtly specifies its possessor role.

Finally, Table 3 shows that while many languages employ possessive person markers of the unique referential type (34 out of 69 combinations, i.e. 49%), none of the sample languages employs markers of the syntactic agreement type. As mentioned in Section 3.2.2, syntactic agreement markers are not attested in possessive NPs, and appear to be restricted to subjects in main clauses. The absence of such markers is striking given the common use of possessive person markers among the world's languages. Possible motivations for the complete absence of possessive syntactic agreement markers are briefly discussed in the next subsection. Interestingly, some of these motivations conversely explain the common use of unique referential markers among the sample languages, which is also briefly discussed in the section below.

### 3.5.1.2 The absence of syntactic agreement markers in possessive NPs

An interesting finding given in the previous subsection is that none of the sample languages, nor any language I know of, employs possessive NPs with syntactic agreement markers. This means that, in contrast to verbal person forms, possessive person forms are never obligatorily accompanied by a corresponding NP; in the presence of a possessive person form, a (pro)nominal expression of the possessor always occurs optionally, or is banned from the NP altogether. Two general explanations may be formulated for this finding, which partly resemble those provided by Siewierska (1999, 2004: 268–273) to account for the rarity of syntactic agreement markers in the clausal domain.

First, syntactic agreement markers may not develop since the possessor NP is dropped so often that it will not become obligatory. As discussed in Section 3.2.2, reasons for dropping the possessor NP may be two-fold. On the one hand, the person marker may be referential by itself. This is the case in many of the languages investigated in this study, since possessive markers strongly tend to be referential in nature, as demonstrated in the previous subsection. On the other hand, the possessor NP may remain unexpressed because it is retrievable from the discourse context. Languages with a low referential density, such as Tidore (11) exemplified in Section 3.2.2, show extreme examples of such context-induced omission of the NP. Both factors dispense the need for expressing the possessor as a free NP, thereby preventing the person marker from further grammaticalization.

A further, related reason for leaving the possessor NP unexpressed is the general informativeness of the person marking paradigm, i.e. the extent to which the marker expresses role information, as well as other features, such as person, number, gender, inclusivity and honorificity. As pointed out by Siewierska (2004: 269–272), the loss of such information – that is the rise of homophonous forms in the paradigm – may induce the presence of a free NP, for reasons of disambiguation or to convey information not expressed in the person forms. Siewierska shows that this is a likely scenario for the rise of syntactic agreement markers in the clausal domain. Conversely, the retention of semantic information may provide a likely motivation for the absence of syntactic agreement markers in possessive NPs: when the possessive person marker already provides sufficient information to identify the referent, an overt NP is relatively redundant, and may therefore remain unexpressed. This explanation applies to many of the languages in the sample, as they use paradigms of possessive person marking that are relatively informative, showing little homophony. The expression of the possessor NP may therefore be left to special



contexts only, such as focus, emphasis or contrast. In fact, most of the sample languages employ possessive person markers that are richer in terms of grammatical feature information – specifically possessor role information – than the co-occurring NP, which is illustrative of their referential nature. As noted above, this may discharge the need for a (pro)nominal expression of the possessor even further.

A second explanation for the absence of possessive syntactic agreement markers is the fact that the person marker may be lost before it reaches this final stage of grammaticalization. On the one hand, a loss of person marking may be due to formal changes: new person markers may be used once the old ones start to fuse with the possessum stem. This process often goes paired with the merging of paradigmatic distinctions, as discussed above, and/or a loss in referential potential, as further discussed in Section 3.5.4.

On the other hand, the person marker may simply fall into disuse. As pointed out by Siewierska (1999: 243, 2004: 276–278), this is often the result of language contact, which triggers the use of new, free forms in favor of the old, mostly bound, forms. This process, also known as ‘diffusion’, is a form of indirect borrowing: only structural patterns are borrowed, here the phenomenon of using independent person forms, rather than actual morphemes. A case in point is the language isolate Urarina (Peru). Under the influence of Spanish, younger Urarina speakers tend to abolish the proclitics used on inalienable possessums (24a) in favor of the independent person markers used also with alienable possessums (24b). A result of this process is the gradual neutralization of Urarina’s alienability opposition.

- (24) a. *ka=tihja*  
           1SG=foot  
           ‘my foot’  
       b. *kanu tihja*  
           1SG foot  
           ‘my foot’ (Olawsky 2006: 353)

In sum, there are a number of different scenarios that may prevent the development of syntactic agreement markers in possessive NPs. The absence of such markers shows that the optional expression of a possessor NP in the presence of a person marker is diachronically stable. Following Van Gijn (2011), this is due to the fact that such markers, reflecting the stages directly prior to syntactic agreement, are functionally optimal: they identify the possessor referent, but without obligatorily expressing information twice. This distribution

of functions is maximally economical, and therefore likely to be diachronically stable. As demonstrated in this section, this explanation may account not only for the cross-linguistic rarity of syntactic agreement markers in clauses, but also for the absence of such markers in possessive NPs.

Finally, note that three of the reasons provided above for the absence of possessive syntactic agreement markers may conversely account for the common use of possessive unique referential markers, as pointed out in the previous subsection. First, unique referential markers already refer to the possessor on their own, which may dispense the need for an additional referential expression of the possessor inside the possessive NP. Second, markers of this type may be highly informative in terms of the grammatical features that they express, and therefore do not require a lexical or pronominal possessor for disambiguation or to convey information not expressed in the person forms. As pointed out by Siewierska (2004: 268), the relative informativeness of person forms may be connected to their formal expression; as will be demonstrated in Section 3.5.4, unique referential markers strongly tend to be separate words, which makes them inherently less likely to lose information as a result of fusion with the possessum. A third reason relates to the frequency with which unique referential markers are expressed in discourse: in languages where pronominal expressions of the possessor are frequently dropped, for instance because they have a low referential density, such markers are less likely to further grammaticalize into markers of appositional cross-reference. This has been described by, for instance, Bisang (2008, 2009) as the lack of ‘overt complexity’ in Southeast Asian languages.

This subsection concludes the discussion of alienable and inalienable possessive person markers in terms of their referential potential. In the next subsection, I turn to the relationship between alienable and inalienable possessive person markers terms of their formal expression, as predicted by hypothesis (ii).

### **3.5.2 Hypothesis (ii): the formal expression of (in)alienable possessive person marking**

In this section, I present the data relevant to hypothesis (ii), which predicts a relationship between alienable and inalienable means of person marking and their formal expression type. Table 4 below presents the different types of marking combinations attested among the 39 languages of the sample. The order of the four expression types (in the columns) captures the gradual reduction in

morpho-phonological form predicted by the corresponding cline in (16). Otherwise, the table has the same structure as Table 3 above. Note that the total number of combinations in Table 4 (N=82) is higher than in Table 3 (N=70), since Table 4 includes the marking strategies excluded from Table 3 (which are unanalyzable in terms of referential potential, but analyzable in terms of formal expression type). Table 2 in Appendix II lists the sample languages that instantiate each combination of marking strategies.

The data in Table 4 confirm hypothesis (ii) without any counterexamples: person markers of alienable possession are never less formally independent, or formulated the other way around, never show a greater degree of formal bonding with the possessum than markers of inalienable possession. Four sub-relations follow from this generalization. First, if the inalienable person marker is a free form, the alienable marker is too (compare combination a to \*e, \*f and \*g). Second, inalienable markers in the form of clitics always combine with alienable markers in the form of words, never an affix or a fused form (compare combination b to \*k and \*l). The combination of alienable and inalienable possessive clitics is a logical possibility, but is not attested among the languages investigated. Third, when the inalienable marker is an affix, the alienable marker is a word, a clitic or an affix, but never a fused form (compare combinations c, i and m to \*o). The fourth and final relation involves the use of fused forms for inalienable possession, which combine with alienable markers in the form of words in the languages investigated. Combinations involving cliticized, affixal or fused alienable markers (j, n and p) are logically possible, but again not attested in the sample.<sup>25</sup>

The frequencies in the table show that combinations involving clitics and/or fused person forms are relatively uncommon (12 out of 82 instances, i.e. 15%). The rarity of person clitics is presumably due to the fact that clitics are recognized on the basis of different criteria or are not recognized in the available descriptions at all. As mentioned in Section 3.2.1, fused person forms arise only for certain (high-frequency) combinations of possessors and possessums, such as 1<sup>st</sup> person possessors of kinship terms.

---

<sup>25</sup> Only the first relation is statistically significant ( $p=0.004$ ) and moderately strong ( $CC=0.291$ ). As with hypothesis (i), the other relations are not statistically significant, due to the skewed distribution of combinations in the table.

Table 4: The distribution of (in)alienable person markers in terms of formal expression type

	word	clitic	affix	fused marker	Frequencies		
					Possessor noun	Possessor pronoun	Total (N=82)
a	AL, INAL				1	12	13
b	AL	INAL			-	4	4
c	AL		INAL		11	23	34
d	AL			INAL	1	3	4
*e	INAL	AL			-	-	-
*f	INAL		AL		-	-	-
*g	INAL			AL	-	-	-
h		AL, INAL			-	-	-
i		AL	INAL		2	2	4
j		AL		INAL	-	-	-
*k		INAL	AL		-	-	-
*l		INAL		AL	-	-	-
m			AL, INAL		9	14	23
n			AL	INAL	-	-	-
*o			INAL	AL	-	-	-
p				AL, INAL	-	-	-

Comparing the distribution of combinations in the table shows that there is a slight preference for the alienable and inalienable possessive person marker to have a distinct formal expression: 46 out of the 82 attested combinations (i.e. 56%) are of type b, c, d or i, while 36 of the 80 combinations (i.e. 44%) are of type a or m. Combination c is the most frequently attested among the languages investigated (34 out of 82 combinations, i.e. 42%). It involves the use of person words in alienable possession, but of person affixes in inalienable possession. A number of relevant examples have been provided in previous sections: Puyuma (3), Ungarinyin (20) and Urarina (21).

A second frequent pattern is the coding of both alienable and inalienable possession by means of person affixes, i.e. combination m, which has 23 instances (i.e. 28%). In languages with this pattern, the alienability split is instead implemented via a difference in referential potential and/or the fact that the affixes form two separate sets. The latter was illustrated for Mangarayi (18) and Koasati (23) in Section 3.5.1.1. The former is illustrated for Udihe (Altaic, Russia) below. As previously shown in (6) in Section 3.2.1, Udihe employs one set of

person suffixes in alienable and inalienable possession, which expresses possessor role information only in alienable possession. The person marker thus expands on the role information provided by the possessor NP, but only in alienable possession, as shown in (25a), where it is correspondingly analyzed as an appositional referential marker. In inalienable possession, the suffix is analyzed as a contextual agreement marker, as shown in (25b):

- (25) a. (*nuani*) *ja:-ŋi-ni*  
           3SG       cow-POSS-3SG  
           ‘his/her cow’  
       b. (*nuani*) *neŋu-ni*  
           3SG       younger.sibling-3SG  
           ‘his/her younger sibling’ (Nikolaeva & Tolskaya 2001: 112, 111)

Interestingly, none of the languages investigated realizes person markers differently depending on the nominal or pronominal character of the possessor. There is no language, for instance, that uses person affixes with nominal possessors, but independent person forms with (or as) pronominal possessors, or that uses person clitics with nominal possessors, but person affixes with (or as) pronominal possessors. The absence of such splits yields the following generalization: if a language uses a set of person markers with a nominal possessor, it also always uses the same set of person markers with (or as) a pronominal possessor. This generalization is directly related to the absence of syntactic agreement markers: in the presence of a possessive person marker, a possessor NP is always optional, and the person marker thus occurs in both nominal and pronominal contexts.

In general, Table 4 reveals that person markers used in alienable possession tend to be realized as words (55 out of 82 instances, i.e. 67%), while inalienable person markers tend to be realized as affixes (61 out of 82 combinations, i.e. 74%). The preference for alienable marking with free person forms is strongest among pronominal possessors (42 out of the 58 combinations involving possessor pronouns – a to d – i.e. 72%). This is due to the fact that free person forms strongly tend to be unique referential markers, i.e. markers in complementary distribution with other referential expressions of the possessor, as illustrated for Puyuma (19) in Section 3.5.1.1. The preference of inalienable marking with affixal person forms is strongest among NPs with nominal possessors: only 2 non-affixal inalienable markers are attested for such possessors in the sample (one word, in combination a, and one fused form, in combination d).

Finally, I compare the formal expression of person markers in NPs with nominal possessors to that in NPs with pronominal possessors, irrespective of the alienability opposition. The figures in Table 4 show that possessive person markers strongly prefer an affixal expression form when they co-occur with a nominal possessor (22 out of the 24 combinations a possessor noun – c, i, and m – i.e. 92%). Conversely, they prefer to be realized as words in NPs with a pronominal possessor (42 out of the 58 combinations involving possessor pronouns – a to d – i.e. 72%). As noted above, this preference is due to the fact that independent person markers tend to function as unique referential markers, and thus instantiate the possessor pronoun themselves. The co-occurrence of an independent person marker with an additional, independent pronoun in the same possessive NP is rare. Together, these observations suggest that the co-occurrence of a possessor (pro)noun is much more likely when the marker has the form of affix than when it has form of a word. The corresponding data for this generalization are presented in Section 3.5.4, which investigates the relationship between referential potential and formal expression type in the languages of the sample. The next subsection first provides an explanation for the generalizations in (i) and (ii) that pertains to the semantic opposition between relational and non-relational nouns, as discussed in Section 3.2.1.

### 3.5.3 A semantic explanation of the findings

Together, the findings provided in the previous two subsections support the claim that inalienable possession never shows a greater degree of linguistic distance than alienable possession (cf. Section 3.2.1). Haiman and Haspelmath each provide a distinct motivation for this asymmetry, which will be discussed in turn below. I will subsequently argue for an alternative explanation of the findings, building on the opposition between inherently relational and inherently non-relational nouns.

The first explanation, proposed by Haiman (1983: 793–795, 1985: 130–136), provides an iconic motivation of alienability splits. According to Haiman (see also Croft 2008), the conceptual distance between possessor and possessum is greater in alienable possession than in inalienable possession, which is iconically reflected in the greater linguistic distance between the corresponding items.

An alternative explanation, relying on economy, is provided by Haspelmath (2008a, see also Nichols 1988: 579, 1992: 121). He argues that kinship terms and body part terms tend to occur with a possessor NP significantly more often than other (non-relational) nouns. This makes the possessive relationship

increasingly predictable to the hearer, allowing the speaker to use less expressive (more cohesive) coding for inalienable possession.

In this paper, I propose an alternative explanation, which combines the two accounts given above. Like Haspelmath (but unlike Haiman), I recognize that patterns of possessive marking are to a high degree economically motivated, as resulting from the predictability of inalienable possessors. However, like Haiman (but unlike Haspelmath), I argue that this predictability follows directly from the inherent relationality of nouns such as kinship terms and body part terms, rather than from the relative frequency with which such nouns occur possessed in discourse. Since such nouns are inherently relational, their relationship to a possessor follows automatically from their semantics. As a result, inalienable possessive relations are highly predictable and in far less need of expressive means of coding than alienable possessive relations, both in terms of function – the use of referential markers rather than of agreement markers – and in terms of form – the use of independent person markers rather than of bound person markers.<sup>26</sup>

The major advantage of a semantically based explanation of the attested patterns is that it applies irrespectively of language-specific (and construction-specific) differences in the degree to which possessor NPs are expressed. Recall from the discussion of Tidore in Section 3.2.2 that the extent to which languages use overt possessors varies greatly, due to differences in referential density. In fact, it may well be that inalienable possessors are more likely to be omitted than alienable possessors, since they tend to be directly related to the discourse context. This is demonstrated by Dahl & Koptjevskaja-Tamm (1998: 43–44), who point out that kinship terms tend to have possessors of 1<sup>st</sup> or 2<sup>nd</sup> person, while possessors of body part terms typically denote referents that are explicitly mentioned in the same or an immediately preceding clause (as in external possession constructions, such as *I hit him in the head*, see also Taylor 1996). As a result, possessors of such items are redundant from a communicative point of view.

Crucially, the type of frequency asymmetry argued for by Haspelmath is highly sensitive to differences in the extent to which individual languages express possessor NPs. In fact, this is demonstrated by his own corpus data, which show that in Spanish inalienable nouns much more frequently occur

---

<sup>26</sup> As noted previously, kinship and body part terms need not all be treated as inalienable nouns in individual languages. However, if an item is perceived as inalienably possessed in a given language, it is also formally treated as such, and thus adheres to the explanation formulated here.

without a possessor than with a possessor, due to the common use of body part constructions like *levanta la mano* ‘raise your (lit. the) hand’ (Haspelmath 2008a: 20). The fact that a possessor is present in the discourse context is nevertheless inherent to the semantics of the inalienable possessum, and thus applies cross-linguistically. Moreover, the opposition between relational and non-relational items is not restricted to the possessive NP, but pertains also to other grammatical domains. Clear examples of other relational items are verbs and adpositions, while other non-relational items are nouns accompanied by attributive adjectives or relative clauses. The inherent relationality of the former types of items is expected to account for asymmetries in morphosyntactic coding similar to those discussed in this paper.

### 3.5.4 Hypothesis (iii): the relationship between function and form

The third and final hypothesis tested in this paper predicts a relationship between the referential potential (function) and the expression type (form) of possessive person markers, both across and within individual languages. Table 5 below presents the cross-linguistic distribution of the three attested types of referential/agreement markers (in the columns) in terms of their formal expression type (in the rows). The table excludes syntactic agreement markers, which are not attested, as well as fused person forms, which are unanalyzable in terms of referential potential (cf. footnote 18) and relatively rare (cf. Table 4 in Section 3.5.2). Moreover, it excludes six sample languages (Dogon, Fula, Hittite, Kunama, Lango and Tamashek) as they do not employ any possessive person markers of which the referential potential could be determined. Kharia and Inanwatan, excluded from Table 3 in Section 3.5.1.1, do participate in this table, as the referential potential of (some of) their possessive markers could be determined. In total, the table thus contains information from 33 out of the 39 sample languages. An overview of the marking strategies in each sample language is provided in Table 3 in Appendix II.

The table shows that the distribution of the different types of person markers among the languages in the sample is far from uniform. The vast majority of markers is of the (unique or appositional) referential type (respectively 47 and 43 instances), while far less markers are of the (contextual) agreement type (only 25 instances). These data support the observation, made in Section 3.5.1.1, that possessive person markers tend to be referential in nature, rather than cases of agreement. With respect to formal expression type, the table



Table 5: The referential potential and formal expression of (in)alienable person markers

		possessor type	Referential potential (function)			Total
			unique referential marker	appositional referential marker	contextual agreement marker	
Expression type (form)	word	noun	n.a.	8	-	8
		pronoun	31	2	-	33
<b>subtotal</b>			<b>31</b>	<b>10</b>	<b>-</b>	<b>41</b>
	clitic	noun	n.a.	2	-	2
		pronoun	1	1	3	5
<b>subtotal</b>			<b>1</b>	<b>3</b>	<b>3</b>	<b>7</b>
	affix	noun	n.a.	18	13	31
		pronoun	15	12	11	38
<b>subtotal</b>			<b>15</b>	<b>30</b>	<b>24</b>	<b>69</b>
<b>Total</b>			<b>47</b>	<b>43</b>	<b>27</b>	<b>117</b>

reveals that the markers in the sample prefer to be expressed as affixes (67 instances) rather than as separate words (41 instances). Person markers realized as clitics are very poorly attested (only 7 instances), which, as noted previously, may perhaps be attributed to their identification by means of varying criteria.

Despite the skewed distribution of person markers in the table, the data show a clear asymmetry in the way referential markers and agreement markers are expressed. Unique referential markers are the most likely to take the form of separate words (31 out of 47 instances, i.e. 66%), followed by appositional referential markers (10 out of 43 instances, i.e. 23%). There are no contextual agreement markers in the form of separate words, although these are attested outside the sample, for instance in the extinct isolate Atakapa (United States; Swanton 1929) and in Tiwi (Australian, Australia; Osborne 1974: 74–75).

Instead, contextual agreement markers strongly tend to be realized as affixes (24 out of 27 instances, i.e. 90%), followed by appositional referential markers (30 out of 43 instances, i.e. 70%). Unique referential markers are in turn the least likely to be expressed as affixes (15 out of 47 instances, i.e. 32%). These figures clearly demonstrate that the likelihood of a possessive person marker to have an expression form on the left-hand side of the formal cline decreases with every step to the right of the functional cline. In other words, functionally more grammaticalized markers also tend to be formally more grammaticalized, while functionally less grammaticalized markers also tend to be formally less

grammaticalized. From a diachronic perspective, this suggests that there is a strong tendency for functional and formal changes to proceed in the same direction: as a marker loses referential potential, it also tends to lose – but not gain – formal independence, e.g. it develops from a free form to an affix on the possessum, but not the other way around. This in turn could be taken as evidence that function and form are parallel paths.

However, at the same time, the data clearly show that the parallel path hypothesis cannot be made absolute: in many of the languages investigated, function and form do not synchronically coincide, suggesting that a move down the functional cline need not go paired with a move down the formal cline, and vice versa. This is demonstrated in Table 5, where it is shown that possessive person markers expressing contextual agreement may take the form of clitics (3 instances), while possessive person markers expressing (unique or appositional) cross-reference may take the form of affixes (respectively 15 and 30 instances). In fact, such so-called ‘mismatches’ (Lehmann 1982a: 236) in function and form are rather common: as mentioned above, 32% of all attested unique referential markers is affixal, and an affixal expression form is even preferred among markers of the appositional referential type (70%). Ample examples of affixal referential markers have been provided in the previous subsections. An example of a contextual agreement marker in the form of a clitic, used with both nominal (26a) and pronominal possessors (26b), is provided below for Kharia (Austro-Asiatic, India):

- (26) a. (*Raʔa=yaʔ*)    *ayo=ɖom*  
           Rata=GEN    mother=3.POSS  
           ‘Rata’s mother’  
       b. (*am=aʔ*)    *ɖhangar=nom*  
           2SG=GEN    servant=2.POSS  
           ‘your servant’ (Peterson 2011: 87, 164)

Thus, the data show that possessive person markers may, and in fact often do, demonstrate variable behavior in terms of referential potential (function) and expression type (form). This clearly demonstrates that function and form cannot be directly related in a one-to-one relationship, thereby arguing against an absolute interpretation of the parallel path hypothesis. Moreover, it demonstrates that functional and formal changes cannot be collapsed into a single cline, as commonly done in the traditional grammaticalization literature (cf. the cline in (1) given in the introduction).

Instead of an absolute relationship between function and form, the data reveal that the relationship between both dimensions of grammaticalization is of a relative nature. This can be demonstrated by comparing the referential potential and expression type of possessive person markers in individual languages. The result of this comparison is given in Table 6 below. The table presents the distribution of the 27 logically possible combinations of possessive person markers of a distinct referential potential (in the rows) in terms of their formal expression (in the columns), as well as the languages employing each combination. Like Table 5, Table 6 excludes syntactic agreement markers and fused person forms. The following abbreviations are used: 'UR' for unique referential marker, 'AR' for appositional referential marker and 'CA' for contextual agreement marker.

Table 6: The distribution of markers of a distinct referential potential in terms of formal expression type

	word	clitic	affix	Languages		
				Possessor noun	Possessor pronoun	Total (N=17)
i	UR, AR			n.a.	-	-
ii	UR	AR		n.a.	-	-
iii	UR		AR	n.a.	Tiriyo	1
*iv	AR	UR		n.a.	-	-
*v	AR		UR	n.a.	-	-
vi		UR, AR		n.a.	-	-
vii		UR	AR	n.a.	-	-
*viii		AR	UR	n.a.	-	-
ix			UR, AR	n.a.	-	-
x	UR, CA			n.a.	-	-
xi	UR	CA		n.a.	Urarina	1
xii	UR		CA	n.a.	Tiriyo	1
*xiii	CA	UR		n.a.	-	-
*xiv	CA		UR	n.a.	-	-
xv		UR, CA		n.a.	-	-
xvi		UR	CA	n.a.	-	-
*xvii		CA	UR	n.a.	-	-
xviii			UR, CA	n.a.	-	-
xix	AR, CA			-	-	-
xx	AR	CA		-	-	-

Table 6: The distribution of markers of a distinct referential potential in terms of formal expression type

	word	clitic	affix	Languages		
				Possessor noun	Possessor pronoun	Total (N=17)
xxi	AR		CA	Tiriyo, Ungarinyin, Biak	-	3
*xxii	CA	AR		-	-	-
*xxiii	CA		AR	-	-	-
xxiv		AR, CA		-	-	-
xxv		AR	CA	Bororo	Bororo	2
*xxvi		CA	AR	-	-	-
xxvii			AR, CA	Udihe, Paumari, Ungarinyin, Mangarayi, Biak	Udihe, Tiriyo, Mangarayi	9

Without any counterexamples, the table reveals that if a language employs two sets of possessive person markers of a distinct referential potential, the person marker of a lower degree of referential potential never has a more independent expression form than the marker of a higher degree of referential potential. Instead, lower referential markers are always less or equally formally independent as higher referential markers. For instance, there is one language in which appositional referential markers have the form of clitics and contextual agreement markers have the form of affixes (Bororo; pattern xxv), and six languages in which both types of markers have the form of affixes (Udihe, Paumari, Ungarinyin, Mangarayi, Biak and Tiriyo; pattern xxvii). However, there is no language in which appositional referential markers have the form of affixes, while contextual agreement markers have the form of affixes (pattern xxvi). These data provide clear support for hypothesis (iii). From a diachronic perspective, they suggest that as a marker moves down the functional cline, it cannot move up the formal cline. Thus, a loss in referential potential never goes paired with a gain in formal independence. However, a loss in referential potential need not go paired with a loss in formal independence either, as was already demonstrated above. Together, these findings suggest that function and form develop in the same direction, but need not develop at the same pace.

On a final note, Table 6 demonstrates that not every combination predicted by hypothesis (iii) is also attested among the sample languages. This is mainly due to the fact that in many of the languages investigated, possessive person markers are not of a distinct, but of the same referential potential. This is demonstrated in Table 7, which presents the distribution of combinations of possessive person markers sharing the same referential potential (in the rows) in terms of their expression type (in the columns). Note, however, that none of the combinations in this table contradict the hypothesis in (iii).

Table 7: The distribution of markers of the same referential potential in terms of formal expression type

	word	clitic	affix	Frequencies		
				Possessor noun	Possessor pronoun	Total (N=77)
i	UR, UR			n.a.	14	14
ii	UR	UR		n.a.	2	2
iii	UR		UR	n.a.	17	17
iv		UR, UR		n.a.	-	-
v		UR	UR	n.a.	1	1
vi			UR, UR	n.a.	9	9
vii	AR, AR			1	-	1
viii	AR	AR		-	-	-
ix	AR		AR	9	2	11
x		AR, AR		-	-	-
xi		AR	AR	1	-	1
xii			AR, AR	9	6	15
xiii	CA, CA			-	-	-
xiv	CA	CA		-	-	-
xv	CA		CA	-	-	-
xvi		CA, CA		-	1	1
xvii		CA	CA	-	-	-
xviii			CA, CA	2	3	5

In sum, the data discussed in this section suggests that, although (a loss in) referential potential strongly correlates with (a loss in) formal independence, this relationship cannot be made absolute, since highly referential markers may be, and in fact often are, formally bound, while lower referential markers may be formally independent. Instead, the relationship between function and form is

shown to be relative in nature: lower referential markers are never more formally independent than higher referential markers, in individual languages. Together, these findings suggest that functional and formal changes proceed in the same direction, but need not proceed at the same pace.

### 3.6 Conclusions

The diachronic development of free pronouns to markers of agreement is a widely debated topic in the literature (e.g. Givón 1976; Corbett 2003; Siewierska 1999, 2004: 261–273). Focusing on person markers in possessive NPs, the paper shows that this development is characterized, on the one hand, by a loss in referential potential, and, on the other hand, by a reduction in morpho-phonological form. Each dimension follows its own grammaticalization cline, which is investigated separately for alienable and inalienable possessive person markers in a worldwide sample of 39 languages.

The referential potential of individual person markers is captured by a four-part typology (Hengeveld 2012), which provides a refinement to existing typologies of referential/agreement markers in two ways. First, it is able to distinguish referential markers and agreement markers independently of their formal expression type. Second, it straightforwardly determines the referential potential of a cross-linguistically common type of possessive person marker: those that optionally co-occur with a possessor noun or pronoun inside the same possessive NP. Such person markers are given a unified analysis as referential markers or as markers of agreement, depending on the distribution of grammatical feature information in the language and the construction in question. This approach is supported by recent work (e.g. Fedden et al. 2013; Iemmolo & Witzlack-Makarevich 2013), demonstrating that person markers show highly language-specific and construction-specific behavior, and can therefore hardly be regarded as a single type. The result is a much more fine-grained analysis of person marking in terms of referentiality than in existing typologies of (cross-)reference/agreement.

The paper tests three hypotheses, each of which is fully borne out by the data. The results for the first two hypotheses demonstrate a robust relationship between alienable and inalienable possessive person markers in terms of referential potential (hypothesis (i)) and in terms of formal expression type (hypothesis (ii)). It is demonstrated that, across individual languages, person marking of inalienable possession is never more referential and never more formally independent than person marking of alienable possession. In fact,

inalienable possessive marking tends to be less referential and formally more reduced than alienable means of coding. These findings are explained in terms of the inherent relationality of inalienable nouns as opposed to the inherent non-relationality of alienable nouns. Since the presence of a possessive relationship follows automatically from the semantics of the inalienable noun, it is in far less need of a referential and morpho-phonologically independent expression of the possessor than alienable nouns. The diachronic implication of these results is that inalienable possessive person marking is more grammaticalized, both in terms of function and in terms of form, than alienable possessive person marking. These results support previous findings in the literature (Haiman 1983: 793–795; Haspelmath 2008a: 18–22; Nichols 1992: 121–122).

The third hypothesis tested in this study demonstrates a strong relationship between the referential potential and the formal expression type of possessive person markers, independently of alienability splits. It is shown that, in individual languages, markers of a lower degree of referentiality never show a greater degree of formal independence than markers of a higher degree of referentiality. Differently put, higher referential markers never show a greater degree of bonding with the possessum than lower referential markers in individual languages. This finding provides clear support for the unidirectionality of grammaticalization processes (see e.g. Haspelmath 1999; Hopper & Traugott 1993: 126–128), as it suggests that a loss in referential potential never goes paired with a gain in formal independence. However, at the same time, it is shown that a loss in referential potential need not go paired with a loss in formal independence either: while referential markers may, and in fact often do, have the form of affixes, agreement markers may have the form of clitics. Thus, the relationship between function and form is relative, rather than absolute. From a diachronic perspective, this suggests that function and form develop in the same direction, but not at the same pace. Notably, this is a much more specific relationship between function and form than is often presumed in the literature, most prominently by proponents of the parallel path hypothesis.

In sum, this study demonstrates that only when different dimensions of grammaticalization are systematically separated in synchronic typological work can their precise diachronic interplay be revealed. This not only applies to person markers in possessive NPs, but also to those in other grammatical domains, most prominently (main) clauses and adpositional phrases, for which the same conflation of function and form is often presupposed. The identification and comparison of (splits in) the referential potential and the formal expression of person markers in each of these domains is an interesting

future enterprise. The typology of referential/agreement marking applied in this paper provides a suitable framework for pursuing this aim.