**Abstract:** This study explores the coding asymmetry between independent and dependent possessive person forms (as in English *mine/my*) from a cross-linguistic perspective. On the basis of a typological survey of 70 geographically and genealogically diverse languages, this paper identifies three universal tendencies: the length universal, the constituent order universal and the alienability universal. First, the length universal claims that independent possessive person forms are either longer or as long as the dependent possessive person forms. Second, the constituent order universal claims that the internal constituent order of the person form and the substantivizer correlates with the constituent order of the possessor and the possessum. In addition, in languages where both a composite possessive marker and a composite substantivizer are used, the possessive marker is always closer to the person root. Finally, the alienability universal claims that the possessive person forms in the alienable possessive constructions are more likely to be used as independent possessive person forms than those in the inalienable possessive constructions. These universal tendencies are instances of form-frequency correspondence, which is shown by corpus evidence from three languages.

**Keywords:** independent person forms, length, constituent order, alienability, universal, frequency

---

* This paper is supported by the European Research Council (ERC Advanced Grant 670985, Grammatical Universals).
1. **Introduction**

Many languages distinguish between dependent possessive person forms (like English *my*) and independent possessive person forms (like English *mine*). Independent possessive person forms are typically used in anaphoric as well as predicative contexts. For this paper, a systematic typological study of independent possessive person forms was conducted using a sample of 70 languages. The study confines itself to possessive person forms and does not focus on full nominal possessors.

Despite the previous rich typological literature on adnominal possessive constructions (e.g. Ultan 1978; Nichols 1988; Chappell & McGregor 1996; Heine 1997; Stolz et al. 2008; Aikhenvald 2013; van Rijn 2016a; Haspelmath 2017; among many others), the coding asymmetry between independent and dependent possessive person forms has been subject to much less inquiry. The only cross-linguistic study of the difference between the two types of possessive person forms concerns pidgin and creole languages (Haspelmath et al. 2013; Michaelis 2019).

By and large, I collected data by manually searching for descriptions of possessive constructions in reference grammars. In order to obtain a representative language sample, I tried to maximize the geographical and genealogical diversity of the languages. Languages were chosen from as many language families as possible, and the final sample consists of 70 languages from 61 language families. In addition, the sample covers all six macro-areas of Glottolog. More specifically, I chose 9 African languages, 16 Eurasian languages, 12 North American languages, 12 South American languages, 18 languages from the area of Papunesia, and 3 languages from Australia. A list of the languages is given in the Appendix and Figure 1 shows the worldwide distribution.

**Figure 1. The sample of 70 languages**

---

1. An anonymous reviewer pointed out that languages might treat these two functions in different ways. However, most grammatical descriptions do not provide further comments on this issue.
2. One of the reviewers point out that this paper may relate the issues to more general ones, namely the independent noun phrases in general. However, due to space limitations it is not possible to resolve all issues at one time.
3. I started out trying to choose one language from each family, but it is very hard to apply this principle thoroughly due to the availability of data. In addition, languages north of latitude 40 N are somewhat underrepresented because they are usually overrepresented in samples (thanks to the reviewer for pointing this out).
4. Information about language families and macro-areas was extracted from Glottolog.
5. The term ’Papunesia’ is also taken from Glottolog and covers Papuan and Austronesian languages.
6. All maps in this paper are generated by the lingtypology package in R.
The aim of this paper is twofold. The first purpose is to identify cross-linguistic tendencies pertinent to independent possessive person forms. The second purpose is to explain the universal tendencies with reference to the form-frequency correspondence hypothesis, which is further supported by some corpus evidence.

Three universal tendencies are identified in this paper: the length universal, the constituent order universal and the alienability universal.7

(i) The length universal: independent possessive person forms are either longer than or identical in length to the corresponding dependent ones.

(ii) The constituent order universal: the internal constituent order of the person root and the function-indicating morpheme in the independent possessive person forms corresponds to the order of the possessor and the possessum. Furthermore, in languages where both a composite possessive marker and a composite substantivizer are used, the possessive marker is always closer to the person root.

(iii) The alienability universal: if a language has an alienability split and the inalienable dependent possessive person forms are used to form the independent possessive person form, then the alienable dependent possessive person forms can also be used to form the independent possessive person form.

7 Here I do not distinguish universals and universal tendencies. I call them universals in the following sections for two reasons. Firstly, no exception is found to these universals; Secondly, the term universal is shorter and more practical.
These claims can initially be illustrated by the following examples.

Regarding the length universal, it is clearly shown that in comparison with the corresponding dependent possessive person forms, the independent ones are either longer, as illustrated in (1)-(3), or the same in length, as illustrated in (4)-(5).

These examples also serve as evidence for the constituent order universal. In many languages, independent possessive person forms consist of a person root\(^8\) and a substantivizer. A substantivizer is a function-indicating morpheme that allows the possessive person form to stand alone without an overt head noun. It is clear that the internal constituent order of the person root and the substantivizer in the independent possessive person form is on a par with the constituent order of the possessor and the possessum. Taking example (1) as an illustration, the possessor \textit{titte} (their) is preposed to the possessum \textit{čohojo-pul-gele} (‘knife’) in (1a), and in the independent person form in (1b), the possessive person root \textit{titte} also precedes the substantivizer -l’e.

(1) Kolyma Yukaghir (Yukaghir, Eurasia)
\begin{itemize}
  \item[a.] \textit{titte} čomōj čohojo-pul-gele  
  3PL.POSS big knife-PL-ACC 
  ‘their big knife’ (Maslova 2003: 235)
  \item[b.] \textit{titte}-l’e  
  3PL.POSS-SUBST 
  ‘theirs’ (Maslova 2003: 236)
\end{itemize}

(2) Huehuetla Tepehua (Totonacan, North America)
\begin{itemize}
  \item[a.] \textit{ki-maka7}  
  1SG-hand 
  ‘my hand’ (Kung 2007: 354)
  \item[b.] \textit{ki(n)-7anu7}  
  1SG-SUBST 
  ‘mine’ (Kung 2007: 399)
\end{itemize}

(3) Kilen (Tungusic, Eurasia)
\begin{itemize}
  \item[a.] \textit{min-i} \textit{əniə-mi}  
  1SG-POSS mother-1SG
\end{itemize}

\(^8\) Person roots are the forms that refer to person (first person, second person or third person). Here I use the term person root as a cover term for free person forms and bound person forms (or ‘person indexes’, following Haspelmath 2013).
‘my mother’ (Zhang 2013: 94)

b. ₍i adi pit’xə min-əŋkə.
   This several book 1SG-SUBST
   ‘These books are mine.’ (Zhang 2013: 105)

(4) Hmong Njua (Hmong-Mien, Eurasia)

a. vi1l pəŋ3l lɛ33 ti44
   1SG POSS CLF bowl
   ‘my bowl’ (Wang 1985: 50)

b. vi1l pəŋ3l
   1SG POSS
   ‘mine’ (Wang 1985: 78)

(5) Huallaga Quechua (Quechuan, South America)

a. qam-pa surti-ki
   2SG-POSS fate-2SG
   ‘your fate’ (Weber 1989: 255)

b. qam-pa
   2SG-POSS
   ‘yours’ (Weber 1989: 255)

Some languages have a composite possessive marker and a composite substantivizer, as illustrated by Desano. For these languages, the possessive marker is always closer to the person root than the substantivizer. As is shown in (6b), the possessive marker ya is closer to the root ~igu.

(6) Desano (Tucanoan, South America)

a. ~igu=ya ~budu-ku
   3SG:M=POSS Tobacco-CLF:tree
   ‘his tobacco plant’ (Silva 2012: 198-199)

b. ~igu=ya=go
   3SG:M=POSS=SUBST
   ‘his (wife)’ (Silva 2012: 199)

The last generalization concerns alienability splits and the choice of independent possessive person forms, as illustrated by the example of Bororo in (7). We see that the independent possessive person form (as shown in 7c) has the same form as the person form in alienable possessive constructions (as shown in 7b), while the person form in
inalienable possessive constructions is different (as shown in 7a).

(7) Bororo (Bororoan, South America)

a. *i-ke*
   1SG.INAL-food \(\text{inalienable}\)
   ‘my food’

b. *ino* \(\text{tori}\)
   1SG.AL stone \(\text{alienable}\)
   ‘my stone’

c. *ino*
   1SG.IND \(\text{independent}\)
   ‘mine’ (Crowell 1979: 215-217)

The following sections provide more evidence for these claims and the rest of the paper is organized as follows. In section 2, the relevant comparative concepts are defined and the main coding types are summarized. The three universal tendencies are discussed in more detail in sections 3-5. In section 6, I argue that some of these universal tendencies can be explained by form-frequency correspondence and coding efficiency.

2. Comparative concepts and first observations

2.1 Main comparative concepts

Typological studies are based on comparative concepts instead of descriptive categories (Haspelmath 2010) and defining comparative concepts is crucial at the outset of a cross-linguistic survey. The first comparative concept that needs to be defined is the possessive relation. As is customary in typology (e.g. Koptjevskaja-Tamm 2002; van Rijn 2016; Haspelmath 2017), by possessive relations I mean the following three relation types: kinship relations (e.g. *my son*), part-whole relations (e.g. *your legs*) and ownership (e.g. *her necklace*). The linguistic forms used to refer to such relations are called adnominal possessive constructions (or adpossessive constructions). Such constructions are nominal constructions that express possessive relations and consist of a possessor modifier and at least a notional possessed noun in the same nominal phrase.

Person forms (Sievierska 2004) are linguistic forms that encode primarily the grammatical category of person. In the same vein, possessive person forms are
linguistic forms that encode person-form possessors. In languages, there are different types of person forms. A dependent possessive person form (dep.PR) is a person form that can be used in adpossessive constructions where the possessed noun is overt (e.g. her house). An independent possessive person form (indep.PR) is a person form that can be used alone without an overt possessed noun to express the possessive meaning (e.g. hers).

All person forms contain person roots, namely the forms that refer to person (first, second, or third person). In addition, an overt marker may also occur in possessive person forms. Typical markers of this kind are possessive markers and substantivizers. A possessive marker, also called genitive marker, is a marker used to indicate the possessive relation on the possessor. For instance, in Mandarin Chinese wǒ de (‘my’). de is a possessive marker that indicates the role of the possessor. A substantivizer is a form that is used to allow the person form to stand alone as a nominal without an overt noun (e.g. the -s in hers). In practice, I extract the form for the substantivizer by removing the dependent person forms from the independent forms.

In different languages possessive markers and substantivizers might have different status. Some of these markers may have a more concrete meaning, but some of them do not. Those with a concrete meaning can be illustrated by the possessive marker khòòng in Lao. In addition to the possessive meaning, it also has the meaning of ‘thing’ in Lao. Those without a concrete meaning can be illustrated by the possessive marker de in Mandarin Chinese. Regardless of whether the possessive markers or substantivizers have a concrete meaning or not, they are taken as part of the possessive person forms as long as they cannot occur on their own.

Moreover, different languages may have different diachronic pathways for possessive person forms (Michaelis 2019). For instance, in some European languages, the dependent forms look like reduced form from the independent ones (e.g. Spanish mi ‘my’ and mio ‘mine’; German mein ‘my’ and meiner ‘mine’; French notre ‘our’ and nôtre ‘ours’). However, it is not always clear which pathways other languages take. Furthermore, even if the dependent forms are reduced from the independent forms, it is still clear that there is an additional morpheme in the independent forms. For this reason, these additional morphemes in the independent forms are considered substantivizers.

---

* Many thanks to one of the anonymous reviewers for pointing this out.
regardless of their diachronic pathways.

Another important point regarding person forms is the contrast between free and bound forms. Free person forms are person forms that can be used on their own, and bound person forms (i.e. indexes) are person forms that cannot be used on their own (following the definition in Haspelmath et.al 2013). A typical free person form is illustrated by the example from Hmong Njua in (8), and a typical bound person form is illustrated by the example from Chimalapa Zoque with the form ʔəm- in (9).

(8) Hmong Njua (Hmong-Mien, Eurasia)

\[
\begin{array}{lll}
\text{maŋ}^{55} & \text{paŋ}^{31} & \text{ka}^{35} \text{tə}^{11} \\
2PL & POS & corn \\
\end{array}
\]

‘your corn’ (Wang 1985:88)

(9) Chimalapa Zoque (Mixe-Zoque, North America)

\[\text{ʔəm-ʔəkwih}\]

2-shirt

‘your shirt’ (Johnson 2000: 264)

Both free person forms and bound person forms are taken into account, but in most cases the distinction between them does not play an important role. This distinction is only relevant when it is connected to the distinction between dependent and independent possessive person forms.

In addition, alienability is also an important notion that is relevant in the formal distinction of possessive person forms. The distinction between alienable and inalienable possession has been widely observed in the world’s languages. By and large, inalienable possession refers to a possessive relation where possessums cannot be separated from their owner, and all other possessive relations are alienable possession (Heine 1997: 85). On the other hand, an alienability split is a formal split (Nichols 1988, van Rijn 2016a, Haspelmath 2017) in the coding of adnominal possessive constructions that is represented by the different coding patterns of inalienable and alienable possession. Alienable nouns often constitute an open set in languages whereas inalienable nouns tend to be represented by a closed set of nouns (Nichols 1988). While alienability is widely attested in languages, there is a great variation in the membership of inalienable nouns across languages. Despite the variation in both form and meaning, Nichols (1988) identified the cross-linguistic tendency that kinship and body-part nouns are more likely to be treated as inalienable nouns.
This study focuses on person forms, and I therefore only take into account such cases where an alienability split involves different possessive person forms. The dependent possessive person forms in the inalienable constructions are called inalienable dependent possessive person forms or inalienable possessors (inalien.dep.PR). On the other hand, the dependent possessive person forms in the alienable constructions are called alienable dependent possessive person forms or alienable possessors (alien.dep.PR). This distinction is crucial to understanding the alienability universal.

2.2 Main types of independent possessive person forms

Languages differ in their way of forming independent possessive person forms. In this study, I distinguish seven different coding types, which will be illustrated in the following.¹⁰

The first type is the POST-SUBSTANTIVIZER type, which is represented by languages where the independent possessive person form is formed by adding a postposed substantivizer to the corresponding dependent possessive person form, without involving an alienability split. This type can be represented by English *your-s* and also illustrated by an example from Trió. As is shown in (10), the independent possessive form *i-pakoro* (10b) is formed by suffixing a substantivizer *-ja-no_ro* to the dependent possessive form *i-* (10a).

(10) Trió (Cariban, South America)

a. *i-pakoro*
   3SG-house
   ‘his house’ (Carlin 2004:95)

b. *i-ja-no_ro*
   3SG-SUBST
   ‘his’ (Carlin 2004:158)

There are further complexities in some languages. For instance, in Teribe different person forms take different substantivizers, as shown in the paradigm of person forms in (11). Although the substantivizers are different, it is clear that all of the substantivizers are postposed. Thus, this is still considered to be the POST-
The second type is the POST-SUBSTANTIVIZER.alienable type, which is represented by languages where a substantivizer is attached to the alienable dependent possessive person form. As is illustrated by the example from Kaluli in (12), the inalienable dependent possessive person form is *n*- (12a), and the alienable dependent possessive person forms is *ni* (12b). The independent form *ni-no* (12c) is formed by suffixing *-no* to the alienable dependent possessive form *ni* (12b) instead of the inalienable one.

(12) Kaluli (Bosavi, Papunesia)

a. *n-o:*

  1SG.POSS.INAL-mother

  ‘my mother’ (Grosh & Grosh 2004: 15)

b. *ni so:lo:-wo: tambo*

  1SG.POSS.AL family-TOP all

  ‘my entire family’ (Grosh & Grosh 2004: 54)

c. *kabi we ni-no:.*

  axe this 1SG.POSS.AL-SUBST

  ‘This axe is mine.’ (Grosh & Grosh 2004: 41)

Another example is taken from Mauwake, where multiple strategies are used to form independent possessive person forms. In Mauwake, inalienable possession (13a) is
coded differently from alienable possession (13b). However, only the alienable possessor is used to form the independent possessor. In addition, three strategies are equally possible with no difference in meaning (Berghäll 2015: 102): a genitive form (13c), genitive form plus a demonstrative (13d), and the genitive form or plain form plus a dative form (13e).

(13) Mauwake (Nuclear Trans New Guinea, Papunesia)

a. w-iawi
   3.INAL-father inalienable
   ‘his/their father’ (Berghäll 2015: 63)

b. wiena owowa=pa
   3.AL.POSS village=LOC alienable
   ‘their village’ (Berghäll 2015: 246)

c. Ikiwosa yena.
   head 1SG.POSS
   ‘The head is mine.’

d. Fikera pun wienda nain=ke.
   Kunai.grass too 3PL.POSS that=FOC
   ‘The kunai grass is theirs, too.’ (Berghäll 2015: 102)

e. Maa nain yo/yena efurik.
   thing that 1SG/1SG.POSS 1SG.DAT
   ‘That thing is mine.’ (Berghäll 2015: 102)

The third type is the POST-SUBSTANTIVIZER.inalienable type, which is represented by languages where a substantivizer is attached to the inalienable dependent possessive person form. This type is very rare, and is illustrated by Canela-Kraho in (14). In Canela-Kraho, pronominal possessors in inalienable possessive constructions are person indexes, as in (14a). In contrast, alienable possessors are marked by attaching an possessive suffix -jô. It is clear that the independent form i-tekjê (14c) is formed by attaching an substantivizer tekjê to the inalienable possessor.

(14) Canela-Kraho (Nuclear Macro-Je, South America)

a. i-pur
   1SG-field
   ‘my field’ (Popjes & Popjes 1986: 134)
b. *i-jō* \_\_\_ wapo

1SG-POSS knife

‘my knife’ (Popjes & Popjes 1986: 169)

c. *i-tekjē*

1SG-SUBST

‘mine’ (Popjes & Popjes 1986: 134)

The fourth type is the **PRE-SUBSTANTIVIZER** type, which is represented by languages where the independent possessive person form is formed by adding a preposed substantivizer to the corresponding dependent possessive person form, without involving an alienability split. This type is illustrated by the example of Anywa in (15). It is clear that the independent possessive person form *már-če* (15c) is formed by preposing a substantivizer *már* to the dependent possessive person form *-če* (15a).

(15) Anywa (Nilotic, Africa)

a. *tōoGG-* če

spear-3SG

‘his spear’ (Reh 1996: 143)

b. *kāagā nānēen mār-če.*

Kaga see SUBST-3SG

‘Kaga checked on his.’ (Reh 1996: 166)

The fifth type is **PRE-SUBSTANTIVIZER, alienable** type, in which independent possessive person form are formed by preposing a substantivizer to the alienable dependent possessive person form, as illustrated by Awjila Berber in (16). In Awjila Berber, kinship terms are inalienable nouns that have a special set of possessive suffixes. In contrast, other nouns are alienable nouns. As illustrated in (16), the independent possessive person form of first person singular *wa=*nn-*ūk* (16c) is formed by preposing a substantivizer *wa* to the alienable possessive person form =nn-*ūk* (16a), instead of the inalienable form -Ø (16b).

(16) Awjila Berber (Afro-Asiatic, Africa)

a. *aẓiṭ=onn-ūk*

donkey=POSS-1SG.AL

‘my donkey’ (van Putten 2013: 133)

b. *abba-Ø*

father-1SGINAL
‘my father’ (van Putten 2013: 118)

c. wa=nn-ïk

SUBST=POSS-1SG.AL

‘mine’ (van Putten 2013: 120)

The sixth type is the IDENTICAL type, which is represented by languages where independent possessive person forms are identical to the corresponding dependent forms. This is illustrated by the example from Berik in (17).

(17) Berik (Tor-Orya, Papunesia)

a. gwola imna
dog 2SG.POSS

‘your dog’ (Westrum 1988: 161)

b. imna
2SG.POSS

‘yours’ (Westrum 1988: 161)

The seventh type is the IDENTICAL.alienable type. In this type, the possessive marker in the dependent possessive person form can be omitted when the possessed noun is an inalienable noun. Although the juxtaposition strategy can be used for inalienable constructions, the independent possessive person forms are always identical to the ones used in alienable constructions, as is illustrated by the example of Mandarin Chinese in (18).

(18) Mandarin Chinese (Sino-Tibetan, Eurasia)

a. wǒ didì
1SG younger.brother

‘my younger brother’

b. wǒ de shū
1SG POSS book

‘my book’

c. wǒ de
1SG POSS

‘mine (my younger brother/my book)’

Preliminary information of the main types of independent person forms in my sample is summarized in Table 1. It is clear that some patterns predominate over other patterns. The most common type is the POST-SUBSTANTIVIZER, and the second most
common type is IDENTICAL.alienable. Other types only have scattered instances.

Table 1 Main Types of Independent possessive person forms

<table>
<thead>
<tr>
<th>Type</th>
<th>Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. POST-SUBSTANTIVIZER</td>
<td>28</td>
</tr>
<tr>
<td>2. POST-SUBSTANTIVIZER.alienable</td>
<td>8</td>
</tr>
<tr>
<td>3. POST-SUBSTANTIVIZER.inalienable</td>
<td>2</td>
</tr>
<tr>
<td>4. PRE-SUBSTANTIVIZER</td>
<td>3</td>
</tr>
<tr>
<td>5. PRE-SUBSTANTIVIZER.alienable</td>
<td>2</td>
</tr>
<tr>
<td>6. IDENTICAL</td>
<td>7</td>
</tr>
<tr>
<td>7. IDENTICAL.alienable</td>
<td>20</td>
</tr>
</tbody>
</table>

Total 70

The complete information pertinent to the independent possessive person forms is shown in the supplementary material. Multiple sources of substantivizers and possessive markers are also shown in the supplementary material. It is interesting that multiple sources have converged to shape similar patterns in which the independent possessive person forms are either longer or have the same length as equally long as the dependent ones.

3. The length universal

Based on the comparative concepts and basic observations in the last section, the following generalization can be formulated:

(19) **Universal 1**: Independent possessive person forms are either longer than or identical in length to the corresponding dependent ones.

The length universal predicts that there is one type of language that is logically possible but should not occur, namely languages where independent forms are shorter than the corresponding dependent ones. The length universal has already been illustrated by the examples in the foregoing. In this section, I will provide more examples to illustrate this universal. In (20)-(22), the independent possessive person

---

* ‘Length’ means the phonological segments involved in the forms, and superassemblental elements are not taken into account.
forms are longer, and in (23)-(25), the independent possessive person forms have the same length as the dependent ones.

(20) Betta Kurumba (Dravidian, Eurasia)
   a. yan ki:ri-əl wan na:ku kayli.
      1SG.POSS house-LOC about four hen
      ‘There are some 4 hens in my house.’ (Coelho 2003: 144)
   b. ɨdə pa:mbə-ŋgu kaynni yan-ɖə.
      this snake-like rope 1SG.POSS-SUBST
      ‘This snake-like rope is mine.’ (Coelho 2003: 220)

(21) Cashibo-Cacataibo (Pano-Tacanan, South America)
   a. ‘ён пити
      1SG.POSS food
      ‘my food’ (Zariquiey Biondi 2011: 226)
   b. ‘ё=нан
      1SG=SUBST
      ‘mine’ (Zariquiey Biondi 2011: 324)

(22) Zialo (Mande, Africa)
   a. wò-làà
      2PL-village
      ‘your village’ (Babaev 2010: 158)
   b. wò-dɛ́y
      2PL-SUBST
      ‘yours’ (Babaev 2010: 65)

(23) Urim (Nuclear Torricelli)
   a. nimpa karek wor wekg a-kitn-en
      dog chicken good two POSS-2SG-ATR
      ‘your two good animals, dog and chicken’ (Hemmilä & Luoma 1987:127)
   b. əl³-kitn-en
      POSS-2SG-ATR
      ‘yours’ (Hemmilä & Luoma 1987:46)

(24) Waris (Border, Papunesia)

* l is added to show emphatic meaning.
a. *deuv-pa ka-na-mba sambla sahoklal.*
   house-TOP 1SG-POSS-TOP two bad
   ‘My two houses (are) no good.’ (Brown 1990: 38)

b. *ka-na dihel-v.*
   1SG-POSS exist-PRS
   ‘I have one.’ (= ‘Mine exists.’) (Brown 1990: 36)

(25) Lao (Tai-Kadai, Eurasia)

   a. *majbanhat khòòng khɔɔj njaau lyyn khòòng cau.*
      ruler POSS 1SG long bigger POSS 2SG
      ‘My ruler is longer than yours.’ (Morev et.al 1979: 89)

   b. *mèèl (khòòng) caw4*
      mother POSS 2SG
      ‘your mother’ (Enfield 2007: 94)

In Lao, as illustrated in (25), the independent possessive person form is the same
as the alienable dependent person form, and this type is documented as ‘same as
alienable PR’ in the Supplementary material, and documented as ‘longer/same’ in Table
2.

The evidence for the length universal is summarized in Table 2. In my 70-language
sample, 42 languages always show longer independent possessive person forms; 21
languages show either longer or same length; and 7 languages always show independent
forms of the same length. Among the 21 languages of the ‘longer/same’ group, only 1
language (i.e. Mauwake) has multiple available strategies for independent possessive
person forms (as was illustrated in example 13). The remaining 20 languages show
either longer forms or the same length because of an alienability split. In these
languages, alienable possessive person forms are always longer than the inalienable
ones, and the independent possessive person forms are the same as the alienable
possessive person forms, thus they are also longer than the inalienable ones.
Consequently, the independent forms are longer compared with the inalienable
possessors, and are the same in length when compared with the alienable possessors.

Table 2. Evidence for the length universal
<table>
<thead>
<tr>
<th>Value</th>
<th>Languages</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>longer</td>
<td>42</td>
<td>60%</td>
</tr>
<tr>
<td>longer/same</td>
<td>21</td>
<td>30%</td>
</tr>
<tr>
<td>same</td>
<td>7</td>
<td>10%</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure 2 shows the areal distribution of different patterns. In the map, languages with longer independent possessive person forms are represented by blue dots; languages with either longer or same length possessive person forms are represented by orange dots; languages where independent and dependent possessive person forms are identical in length are represented by green dots.

**Figure 2 The map showing the length of independent possessive person forms**

4. **The constituent order universal**

A closer look at the internal constituent order of independent possessive person forms allows us to formulate the following cross-linguistic generalizations:
(26) **Universal 2** The constituent order universal:

**Universal 2a.** In independent possessive person forms, the internal constituent order of the person root and the substantivizer/possessive marker is the same as the order of the corresponding dependent possessive person form and the possessum.

**Universal 2b.** When both a composite possessive marker and a composite substantivizer are used in the independent possessive person form, the possessive maker is always closer to the person root than the substantivizer.

In this context, not all of the 70 languages in my sample are relevant here for the following reasons. Firstly, only composite forms\(^{14}\) play a role for Universal 2, and as a result languages with suppletive forms are ignored. Secondly, the constituent order is supposed to be the dominant constituent order. If there is no dominant constituent order in adnominal possessive constructions in a language, then the language is not relevant to these universals.

In addition, when there are multiple dependent possessive person forms, only the one that has a corresponding independent possessive person form is taken into account. For instance, as is shown in the example of Piro in (27), both the attached form (27a) and the separate person form (27b)\(^{15}\) can be used adnominally, but only the free form has a corresponding independent possessive person form (27c).

(27) Piro (Kiowa-Tanoan, North America)

a. *n-fi*ma-ne

1SG-fish-PSSD

‘my fish’ (Hanson 2010: 30)

b. *hita* sawli-te

1SG machete-PSSD

‘my machete’ (Hanson 2010: 49)

c. *hita*-ni-ni wa *fima.*

1SG-SUBST-IMP.DECL REF fish

‘(The) fish is mine.’ (Hanson 2010: 52)

Huallaga Quechua also has two different ways to encode a possessive relation:

---

\(^{14}\) Composite forms can be analyzed as consisting of several smaller units. For instance, *your-s* in English is composite, consisting of *your* and *s.*

\(^{15}\) Here I haven’t found a minimal pair, but both ‘machete’ and ‘fish’ are alienable nouns, and they show similar structure. According to Hanson (2010: 49): “On nouns, independent possessors are in full complementary distribution with the pronominal prefixes. This is the case for both alienable and inalienable nouns.”
head marking in (28a) and double marking in (28b) with an additional independent person form. Only the free person form in (28b) has a corresponding independent form, seen in (28c), that is identical to the dependent one.

(28) Huallaga Quechua (Quechuan, South America)

a. \textit{wasi-ki}

\hspace{1cm} \text{house-2SG}

\hspace{1cm} ‘your house’ (Weber 1989: 54)

b. \textit{qam-pa} \textit{surti-ki}

\hspace{1cm} \text{2SG-POSS fate-2SG}

\hspace{1cm} ‘your fate’ (Weber 1989: 255)

c. \textit{qam-pa}

\hspace{1cm} \text{2SG-POSS}

\hspace{1cm} ‘yours’ (Weber 1989: 255)

\textit{Wari’} is one of those languages where both the possessum and the substantivizer are preposed to the possessor, as is illustrated in (29).

(29) Wari’ (Chapacuran, South America)

a. \textit{wina-m}

\hspace{1cm} \text{head-2SG}

\hspace{1cm} ‘your head’ (Everett & Kern 1997: 148)

b. \textit{mene-m}

\hspace{1cm} \text{SUBST-2SG}

\hspace{1cm} ‘yours’ (Everett & Kern 1997: 148)

The foregoing examples represent languages where a substantivizer is added to the possessive person form when the possessum noun does not occur. It may seem quite natural to have the substantivizer on the same side as the possessed noun, since the substantivizer may be analyzed as occupying the slot of the possessed noun in the independent form. Hence, it is not hard to understand why the substantivizer occurs on the same side as the possessum. However, that possessive markers should also be located on the same side as the possessed noun is not so straightforward. Nevertheless, languages still tend to be coded this way.

In some languages, the dependent and independent possessive person forms have the same form, but still the order of the person root and the possessive marker corresponds to the order of the possessive person form and the possessum, as illustrated
in (30) and (31).

(30) Hmong Njua (Hmong-Mien, Eurasia)

a. \( pi^{33} pa\dot{n}^{31} na^{55} \)
\[ \text{1PL POSS corn} \]
‘our corn’ (Wang 1985: 75)

b. \( pi^{33} pa\dot{n}^{31} \)
\[ \text{1PL POSS} \]
‘ours’ (Wang 1985: 68)

(31) South Wa (Austroasiatic, Eurasia)

a. \( ma^{wu} te\dot{i}^{e} n\dot{o}^{h} \)
\[ \text{money POSS 3SG} \]
‘his money’ (Zhou & Yan 1984: 52)

b. \( te\dot{i}^{e} n\dot{o}^{h} \)
\[ \text{POSS 3SG} \]
‘his’ (Zhou & Yan 1984: 52)

A potential exception comes from Tauya,\(^{16}\) as illustrated in (32). In Tauya, the preferred position for possessive person forms is before the possessed noun, but the possessive marker follows the possessive person form, as shown in (32a-b). However, this should not be counted as a counter-example for two reasons. Firstly, although the possessed noun prefers the preposed position, it is also possible to be postposed. Secondly, when there is a case marker, the preferred word order is possessor-possessed noun, as shown in (32c), which is in line with the universal tendency.

(32) Tauya (Nuclear Trans New Guinea, Papuasia)

a. \( wate \text{ } ne-pi \)
\[ \text{house 3SG-POSS} \]
‘his/her house’ (MacDonald 1990: 131)

b. \( ʔe-ra \text{ } ya-pi-ʔa. \)
\[ \text{DEM-TOP 1SG-POSS-indicative} \]
‘That’s my house.’ (MacDonald 1990: 132)

c. \( ya-pi \text{ } wate-ʔai \)
\[ \text{1SG-POSS house-adessive.case} \]

---

\(^{16}\) This example was pointed out by an anonymous reviewer.
‘my house’ (MacDonald 1990: 132)

Due to the diversity of diachronic paths that languages went through, the reasons for this phenomenon are different from language to language, but the cross-linguistic trend is clearly observed.

In the following, I will give some examples for Universal 2b. As shown in (33), in Wintu the substantivizer -o is located further away from the person base ne than the possessive marker -t.

(33) Wintu (Wintuan, North America)
   a. ne-t
      1SG-POSS
      ‘my’
   b. ne-t-o
      1SG-POSS-SUBST
      ‘mine’ (Pitkin 1984:219)

Korean also shows a similar picture. As illustrated in (34), the possessive marker -uy is closer to the person root than the substantivizer -geot.

(34) Korean (Koreanic, Eurasia)
   a. na-uy
      1SG-POSS
      ‘my’
   b. na-uy-geot
      1SG-POSS-SUBST
      ‘mine’ (personal knowledge)

Another example comes from Awjila Berber. As is shown in (35), in Awjila Berber both the possessive marker ənn and the substantivizer wa are preposed to the person root –ùk. The substantivizer is located at the outer layer, which is further away from the person root than the possessive marker ənn.

(35) Awjila Berber (Afro-Asiatic, Africa)
   a. aẓit=ənn-ùk
      donkey=POSS-1SG
      ‘my donkey’ (van Putten 2013: 133)
   b. wa=nn-ùk
      SUBST=POSS-1SG
‘mine’ (van Putten 2013: 120)

The evidence for Universal 2 is summarized in Table 3. As already mentioned earlier, languages with flexible word order and languages where the substantivizer/possessive marker is not composite are left aside as irrelevant. Apart from these 6 irrelevant languages, the rest of the languages support my claim. The evidence is summarized in Table 3.

The **ALL POSTPOSED** type is represented by languages where the possessive marker, the substantivizer and the possessed noun are all postposed to the person root. The **ALL PREPOSED** type is represented by languages where the possessive marker, the substantivizer and the possessed noun are all preposed to the person root. The **POSTSUB+POST.PM** type is represented by languages where both the substantivizer and the possessed noun are postposed to the person root, while the possessive marker does not play a role since it is not a composite form in these languages. Likewise, the **PRESUB+POST.PM** type is represented by languages where both the substantivizer and the possessed noun are preposed to the person root. In addition, the areal distribution of the languages is given in Figure 3.

<table>
<thead>
<tr>
<th>Value</th>
<th>Languages</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL POSTPOSED</td>
<td>37</td>
<td>53%</td>
</tr>
<tr>
<td>ALL PREPOSED</td>
<td>5</td>
<td>7%</td>
</tr>
<tr>
<td>POSTSUB+POST.PM</td>
<td>19</td>
<td>27%</td>
</tr>
<tr>
<td>PRESUB+PRE.PM</td>
<td>3</td>
<td>4%</td>
</tr>
<tr>
<td>irrelevant</td>
<td>6</td>
<td>9%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>70</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Figure 3. The areal distribution languages as evidence for Universal 2**
5. The alienability universal

This section discusses the alienability universal, so that only languages with alienability splits in possessors\(^{17}\) are included. In my sample, 33 languages are relevant for this universal. In addition, 8 languages show alienability splits but do not show differences in possessive person forms. In these languages, alienability splits are realized by other formal strategies, such as the obligatoriness of occurring in possessive constructions, where inalienable nouns are obligatorily possessed, but alienable nouns are not.

An alienability split is determined by the nature of the possessum, typically resulting in coding asymmetries of adnominal possessive constructions. On the basis of the comparative concepts of alienable and inalienable dependent possessive person forms defined in the second section, the following generalization pertinent to alienability can be formulated:

(36) **Universal 3:** If a language shows an alienability split and the inalienable dependent possessive person forms (inalien.dep.PR) are used to form the independent possessive person form (indep.PR), then the alienable dependent possessive person forms (alien.dep.PR) can also be used to form the independent possessive person form.

\(^{17}\) That is, languages where different dependent possessive person forms are found for inalienable and alienable possession.
The alienability universal can be visualized in the following simplistic semantic map for possessive person forms:

**Independent PR — alienable dependent PR — inalienable dependent PR**

Conceivably, alienable and inalienable possessors may have two different independent person forms.\(^{18}\) Suppose the form for alienable PR is PR1 in a certain language, and the inalienable PR is PR2 in this language, the following contrast can be made:

\[
\begin{align*}
(37) & \quad \text{a. My}_1 \text{ watch is pretty, yours}_1 \text{ is ugly. \quad alienable} \\
& \quad \text{b. My}_2 \text{ leg is long, yours}_2 \text{ is short. \quad Inalienable}
\end{align*}
\]

However, in real life such a contrast is not very common.\(^{19}\) For instance, in Anindhilyakwa, the inalienable form is different from the alienable form, but the independent form is the same as the alienable form, even when it is used to refer to inalienable nouns.

\[
(38) \quad \text{Anindhilyakwa (Gunwinyguan, Australia)}
\]

\[
\begin{align*}
& \quad \text{a.} \quad \text{nvngenv}-m-\text{alhvka} \\
& \quad \quad \text{1SG-INAL-foot} \\
& \quad \quad \quad \quad \text{‘my footprints’ (van Egmond 2012: 113)} \\
& \quad \text{b.} \quad \text{nganyangwa} \quad \text{dhv-mamawura} \\
& \quad \quad \text{1SG.POSS} \quad \text{F-CLF.VEG.sun} \\
& \quad \quad \quad \quad \text{‘my watch’ (van Egmond 2012: 155)} \\
& \quad \text{c.} \quad \text{nganyangwa} \quad \text{mamvdhakba.} \\
& \quad \quad \text{1SG.POSS} \quad \text{VEG.tail} \\
& \quad \quad \quad \quad \text{‘The tail is mine.’ (van Egmond 2012: 285)}
\end{align*}
\]

A similar phenomenon is also found in Mandarin Chinese, as shown in example (39). In inalienable constructions, the possessor can be juxtaposed with the possessed noun, but in alienable constructions the possessive marker must be added when the possessed noun does not occur. However, such instances as example (39) are not very frequent, possibly because in Mandarin Chinese kinship terms that are treated as inalienable nouns are high in animacy, which leads to their prominence in speech.

\(^{18}\) Many thanks to the anonymous reviewer for pointing this out.

\(^{19}\) Here I have to admit that mini-pairs of this type are very hard to find, which lead to an obvious limitation of this study. In many languages of my sample, only one set of forms is found as independent possessive person forms. Only some of the grammars are explicit regarding this issue.
Consequently, it is less likely that kinship terms will be omitted in the first place. In addition, politeness plays an important role in Chinese, and it is therefore even less common to omit the head noun when the possessed noun is the eldership (e.g. ‘father’, ‘mother’, ‘older brother’, ‘grandfather’…).

(39) Mandarin Chinese (personal knowledge)

\[
\text{wǒ mèimei hěn cōngmíng, nǐ=de hěn bèn.} \\
1\text{SG younger.sister very clever 2SG=POSS very stupid} \\
\text{‘My younger sister is very clever, yours is very stupid.’}
\]

Another example comes from Bororo, where three sets of possessive person forms are used for inalienable possession, alienable possession and possession of domesticated nouns, as shown in (40). When the head noun does not occur and the possessive person forms stand on their own, the alienable possessive person forms are used even if the noun is of the inalienable class (Crowell 1979: 217). In addition, when the nouns that belong to the domesticated animal class do not occur, the possessive person forms of this set are also used alone.

(40) Possessive person forms in Bororo (Crowell 1979: 215-216)

<table>
<thead>
<tr>
<th>Inalienable (e.g. ke ‘food’)</th>
<th>Alienable (e.g. tori ‘stone’)</th>
<th>Domesticated animal (e.g. kogariga ‘chicken’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG i-ke</td>
<td>ino tori</td>
<td>inagu kogariga</td>
</tr>
<tr>
<td>2SG a-ke</td>
<td>ako tori</td>
<td>akagu kogariga</td>
</tr>
<tr>
<td>3SG u-ke</td>
<td>o tori</td>
<td>aku kogariga</td>
</tr>
<tr>
<td>Coreferential ti-ge</td>
<td>to-dori</td>
<td>tagu kogariga</td>
</tr>
<tr>
<td>Reciprocal pu-ge</td>
<td>pu-dori</td>
<td>pugagu kogariga</td>
</tr>
<tr>
<td>1PL.INCL pa-ge</td>
<td>pago tori</td>
<td>pagagu kogariga</td>
</tr>
<tr>
<td>1PL.EXCL xe-ge</td>
<td>xeno tori</td>
<td>xenagu kogariga</td>
</tr>
<tr>
<td>2PL ta-ge</td>
<td>tago tori</td>
<td>tagagu kogariga</td>
</tr>
<tr>
<td>3PL e-ke</td>
<td>eno tori</td>
<td>enagu kogariga</td>
</tr>
</tbody>
</table>

Perhaps for inalienable possessive constructions it is less frequent to leave out the possessed nouns, and it is therefore quite unlikely for languages to develop dedicated independent forms for them. In addition, some languages use bound person forms for inalienable possession, and bound forms by definition cannot stand on their own. Thus,

\* This claim is only based on the intuition of a couple of native speakers, and I have no corpus data to support it.
it is easy to understand why languages that have dedicated inalienable independent forms are extremely rare. In my typological survey, I have not found any language that has this characteristic.

In turning to the evidence supporting the Universal 3, a number of logically possible patterns are listed in Table 4. What is not predicted by the semantic map (Independent PR — alienable dependent PR — inalienable dependent PR) is type (c), where independent possessive person forms are the same as the inalienable dependent possessive person forms, but different from the alienable dependent possessive person forms. Type (c) is not attested in my sample. In other words, I have not found any exception to this universal.

<table>
<thead>
<tr>
<th>Table 4 Possible coding patterns of possessors regarding alienability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indep.PR</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>a. X</td>
</tr>
<tr>
<td>b. X</td>
</tr>
<tr>
<td>c. X</td>
</tr>
</tbody>
</table>

Type (a) is illustrated by the example from Tapieté, where three different forms are used for the inalienable dependent possessor (41a), alienable dependent possessor (41b), and independent possessor (41c). Tapieté is the only language in my sample where three person index series seem to show entirely different forms when the person of the possessor is third person singular.

(41) Tapieté (Tupian, South America)

a. *h-o’o*

3SG.INAL-flesh **inalienable**

‘his/her flesh’ (González 2005: 120)

b. *yi-i’a*

3SG.AL-plate **alienable**

‘his/her plate’ (González 2005: 112)

c. *ñ-a’ampowa*
3SG-SUBST independent
‘his/hers’ (González 2005:243)

Except for the third person singular in Tapieté, the remaining languages in type (a) show the pattern with the alienable possessive person form and the independent possessive person form sharing the same root, as is illustrated by examples from Ute and Choctaw.

In Ute, although the three forms are different, the examples clearly show that the alienable dependent possessive person form and the independent possessive person form have the same root form *uwa*, and that the inalienable dependent possessive person form behaves differently.

(42) Ute (Uto-Aztecan, North America)
   a. *kwasi*-a-*u*
      tail-POSS-3SG inalienable
      ‘his/her tail’ (Givón 2011: 100)
   b. *‘uwa-yas* kani
      3SG-POSS house alienable
      ‘his/her house’ (Givón 2011: 205)
   c. *‘uwa-ya-su*
      3SG-POSS-SUBST independent
      ‘his/hers’ (Givón 2011: 403)

Another example of type (a) is Choctaw. Although the three forms in this language are different, it is still possible to see a common root *a* in all three forms.

(43) Choctaw (Muskogean, North America)
   a. *sa-shki’*
      1SG.INAL-mother inalienable
      ‘my mother’ (Broadwell 2006: 49)
   b. *a-katos*
      1SG.AL-cat alienable
      ‘my cat’ (Broadwell 2006: 49)
   c. *á-mmi’*
      1SG-SUBST independent
      ‘mine’ (Broadwell 2006: 95)
Type (b) is illustrated by examples from Moskona and Lao. In Moskona, the inalienable dependent form is a person index *di* (44a), while the alienable dependent form is a possessive pronoun *dadin* (44b). The independent form (44c) is the same as the alienable form.21

(44) Moskona (East Bird’s head, Papunesia)

a. *di-ositym*
   
   1SG-nose inalienable
   
   ‘my nose’ (Gravelle 2010:85)

b. *dadin1SG. mowos*
   
   1SG.POSS village alienable
   
   ‘my village’ (Gravelle 2010: 156)

c. *dadin*
   
   1SG.POSS independent
   
   ‘mine’ (Gravelle 2010: 189)

In Lao, the alienable form (45b) and the independent form (45c) are the same, while the inalienable form is different in that it is not marked by the possessive marker *khòòng*.

(45) Lao (Tai-Kadai, Eurasia)

a. *mèél caw4*
   
   mother 2SG inalienable
   
   ‘your mother’ (Enfield 2007: 94)

b. *majbanhat khòòng caw4*
   
   ruler POSS 2SG alienable
   
   ‘your ruler’ (Morev et.al 1979: 89)

c. *khòòng caw4*
   
   POSS 2SG independent
   
   ‘yours’ (Morev et.al 1979: 89)

To summarize, as we can see above these coding types have different ways of showing a coding asymmetry in alienability. Type (a) is represented by languages where the three forms are different, even though similar roots of person form can still be captured. Type (b) is represented by languages where independent possessive person

---

21 It is not clear whether the set of possessive pronouns can also be used for inalienable possession or not. In fact, even for the alienable possession it is quite rare to omit the head nouns in Moskona (Gravelle 2010:189).
forms are the same as alienable dependent possessive person forms, but different from the inalienable dependent possessive person forms.

In reference to the four patterns above, the data for alienability splits regarding the above-mentioned four patterns is summarized in Table 5, and the areal distribution is given in Figure 4.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Languages</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [indep.PR], [alien.dep.PR], [inalien.dep.PR]</td>
<td>13</td>
<td>39%</td>
</tr>
<tr>
<td>b. [indep.PR, alien.dep.PR], inalien.dep.PR</td>
<td>20</td>
<td>61%</td>
</tr>
<tr>
<td>c. [indep.PR, inalien.dep.PR], alien.dep.PR</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure 4. The map of alienability split types

---

22 In some languages not all person forms show alienability difference. For instance, in Zialo, there is only partial difference between alienable possessive person forms and inalienable possessive person forms in the first person exclusive singular and third person singular. When there is a difference, the independent form chooses the alienable form instead of inalienable form. This is the least typical case in our sample. However, since the distinction can be captured, I still include these languages.
6. Towards functional explanations

6.1 The length universal as an instance of form-frequency correspondence

With respect to the length universal, I propose that frequency plays an important role in the coding asymmetry between independent and dependent possessive person forms.\(^{23}\) As is argued by Haspelmath (2008), more frequent forms tend to be shorter because of an efficiency principle (see also Zipf 1935, Haspelmath et al. 2014).

Dependent possessive person forms are more frequent than the corresponding independent person forms in terms of token frequency, presumably because they show different distributional constraints. Noun phrases with dependent possessive person forms can occur in all situations in which non-possessed nouns can occur. In contrast, independent possessive person forms are used only when the listener can infer the possessed noun from the context.

This claim is thus that independent possessive person forms are rarer than dependent ones in all languages. In the following, I will cite data from 3 languages. Table 6 shows the different frequencies of independent and dependent possessive person forms in the corpus of contemporary American English.\(^{24}\) We see that the independent possessive person forms are much less frequent than the dependent ones.

### Table 6. Frequency of possessive person forms in English

<table>
<thead>
<tr>
<th>Form</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>dependent possessive person form</td>
<td></td>
</tr>
<tr>
<td>my</td>
<td>1,239,555</td>
</tr>
<tr>
<td>your</td>
<td>883,613</td>
</tr>
<tr>
<td>her</td>
<td>1,845,549</td>
</tr>
<tr>
<td>our</td>
<td>702,300</td>
</tr>
<tr>
<td>their</td>
<td>1,417,175</td>
</tr>
<tr>
<td>independent possessive person form</td>
<td></td>
</tr>
<tr>
<td>mine</td>
<td>32,654</td>
</tr>
<tr>
<td>yours</td>
<td>13,190</td>
</tr>
<tr>
<td>hers</td>
<td>10,412</td>
</tr>
<tr>
<td>ours</td>
<td>8,271</td>
</tr>
<tr>
<td>theirs</td>
<td>5,344</td>
</tr>
</tbody>
</table>

\(^{23}\) There might also be other reasons contributing to the length universal. One reviewer points out that independent forms represent a whole noun phrase rather than being attached to a noun, resulting in a more prominent role in the sentence. This in turn makes it more likely to be prosodically prominent and shields it from being reduced.

\(^{24}\) [http://corpus.byu.edu/coca/](http://corpus.byu.edu/coca/)
In Mandarin Chinese and Korean, we see the same frequency asymmetry between independent and dependent possessive person forms. Table 7 shows the frequency asymmetry in Mandarin and Table 8 shows the frequency asymmetry in Korean.

The data for Mandarin Chinese comes from the contemporary spoken Chinese in the CCL corpus. I have chosen a subset of the CCL corpus as my sample and extracted the data manually. As can be seen from Table 7, in Mandarin, dependent forms are more frequent than the independent ones, and the length of dependent and independent possessive person forms are identical. This result supports the form-frequency correspondence hypothesis.

### Table 7. Frequency of possessive person forms in Mandarin

<table>
<thead>
<tr>
<th>Form</th>
<th>Person</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>dependent possessive person form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wǒ de</td>
<td>1SG</td>
<td>964</td>
</tr>
<tr>
<td>nǐ de</td>
<td>2SG</td>
<td>642</td>
</tr>
<tr>
<td>tā de</td>
<td>3SG</td>
<td>1,128</td>
</tr>
<tr>
<td>wōmen de</td>
<td>1PL</td>
<td>380</td>
</tr>
<tr>
<td>nīmen de</td>
<td>2PL</td>
<td>46</td>
</tr>
<tr>
<td>tāmen de</td>
<td>3PL</td>
<td>190</td>
</tr>
<tr>
<td>independent possessive person form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wǒ de</td>
<td>1SG</td>
<td>12</td>
</tr>
<tr>
<td>nǐ de</td>
<td>2SG</td>
<td>17</td>
</tr>
<tr>
<td>tā de</td>
<td>3SG</td>
<td>6</td>
</tr>
<tr>
<td>wōmen de</td>
<td>1PL</td>
<td>2</td>
</tr>
<tr>
<td>nīmen de</td>
<td>2PL</td>
<td>0</td>
</tr>
<tr>
<td>tāmen de</td>
<td>3PL</td>
<td>4</td>
</tr>
</tbody>
</table>

The data for Korean comes from the corpus of the National Institute of Korean Language. As is shown in Table 8, the dependent possessive person forms are far more frequent than the independent ones (both the full forms and the reduced independent forms). In addition, the independent possessive person forms are longer

---

25 http://ccl.pku.edu.cn:8080/ccl_corpus/
26 Many thanks to my Korean colleague Youjeong Oh for helping me extracting the data from the Korean corpus.
27 https://ithub.korean.go.kr/user/corpus/corpusSearchManager.do
than the corresponding dependent ones in Korean. This result also supports the form-frequency correspondence hypothesis.

Table 8. Frequency of possessive person forms in Korean

<table>
<thead>
<tr>
<th>Form</th>
<th>Person</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent possessive person form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>na-uy</td>
<td>1SG</td>
<td>11,393</td>
</tr>
<tr>
<td>jeo-uy</td>
<td>1SG (honorific)</td>
<td>1,082</td>
</tr>
<tr>
<td>neo-uy</td>
<td>2SG</td>
<td>1,218</td>
</tr>
<tr>
<td>geu-uy</td>
<td>3SG.M</td>
<td>35,775</td>
</tr>
<tr>
<td>geunyeo-uy</td>
<td>3SG.F</td>
<td>8,661</td>
</tr>
<tr>
<td>wuli-uy</td>
<td>1PL</td>
<td>13,645</td>
</tr>
<tr>
<td>dangsin-deul-uy</td>
<td>2PL</td>
<td>120</td>
</tr>
<tr>
<td>gue-deul-uy</td>
<td>3PL</td>
<td>8,472</td>
</tr>
<tr>
<td>Independent possessive person form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>na-uy-geot</td>
<td>1SG</td>
<td>23</td>
</tr>
<tr>
<td>je-uy-geot</td>
<td>1SG (honorific)</td>
<td>3</td>
</tr>
<tr>
<td>neo-uy-geot</td>
<td>2SG</td>
<td>3</td>
</tr>
<tr>
<td>dangsin-uy-geot</td>
<td>2SG (honorific)</td>
<td>1</td>
</tr>
<tr>
<td>geu-uy-geot</td>
<td>3SG.M</td>
<td>1</td>
</tr>
<tr>
<td>geunyeo-uy</td>
<td>3SG.F</td>
<td>1</td>
</tr>
<tr>
<td>neo-uy-geot</td>
<td>1PL</td>
<td>14</td>
</tr>
<tr>
<td>dangsin-deul-uy-geot</td>
<td>2 PL (honorific)</td>
<td>0</td>
</tr>
<tr>
<td>gue-deul-uy-geot</td>
<td>3PL</td>
<td>1</td>
</tr>
<tr>
<td>Reduced independent possessive person form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nae-geot</td>
<td>1SG</td>
<td>140</td>
</tr>
<tr>
<td>je-geot</td>
<td>1SG (honorific)</td>
<td>3</td>
</tr>
<tr>
<td>uri-geot</td>
<td>1PL</td>
<td>30</td>
</tr>
<tr>
<td>ne-geot</td>
<td>2SG</td>
<td>38</td>
</tr>
<tr>
<td>dangsin-geot</td>
<td>2SG (honorific)</td>
<td>3</td>
</tr>
<tr>
<td>dangsin-deul-geot</td>
<td>2PL (honorific)</td>
<td>0</td>
</tr>
</tbody>
</table>

This explanation in terms of form-frequency correspondence and coding efficiency.

\* In this table, the reduced forms do not cover all person forms. This is because those third person forms are very informal and are not found in the corpus.
may be contrasted to an alternative explanation in terms of iconicity (Haiman 1983). In line with the iconicity hypothesis, it can be argued that the independent possessive forms are usually longer because they involve two referents instead of one. For instance, the form my in English is only concerned with the possessor, while the form mine in English is concerned with both the possessor and the possessed thing. However, problems arise with this explanation in the case of languages where the dependent possessive person forms can be used alone without any further marking. In such cases, these dependent forms have the same forms as the independent ones. For instance, wǒ de (‘my’) in Mandarin refers to the possessor when used with a head noun, but it also refers to the possessed thing (‘something of mine’) when the head noun is not overt.

In addition, one could say that because the abstract syntactic structure contains an additional layer, the independent possessive person form is longer than the dependent one, since the former stands alone as an NP, but the latter is only the determiner in the NP. However, this is only one way of analyzing the syntactic structure of independent possessor NPs, but it is unclear what syntactic evidence there is. On the other hand, since the world’s languages display quite diverse surface structures, such hypotheses about deep structures are not easy to test, especially when so-called zero-forms are involved. In contrast, the empirical approach in this study does not presuppose such hypotheses about the deep structure of possessive person forms and instead shows that the form-frequency correspondence provides a better basis for an explanation.

More light can be shed on the nature of this issue with reference to Croft’s theory. Croft (1991) identifies three types of pragmatic functions (reference, modification and predication) and proposes that the cross-linguistic marking patterns of nouns, verbs and adjectives are induced by the prototypicality of the association between pragmatic functions and lexical semantics (i.e. objects, properties and actions), as shown in Table 9. When the association is not a typical one, a function-indicating morpheme (or function-indicator) occurs. For instance, destroy is an action and its typical pragmatic function is predication. When used in the modification function, a function-indicating morpheme -ed occurs, forming destroyed.

---

* One reviewer pointed out that Haiman (1983) did not discuss the contrast between independent and dependent possessors. This is true. However, his work is extremely influential, and questions that in line with his thought were raised when I presented my paper. For this reason, I still discuss this issue here.

* Many thanks to one of the reviewers for pointing this out.
Table 9. English examples of marked and unmarked correlations (cf. Croft 1991)

<table>
<thead>
<tr>
<th>Objects</th>
<th>Reference</th>
<th>Modification</th>
<th>Predication</th>
</tr>
</thead>
<tbody>
<tr>
<td>vehicle</td>
<td>vehicle's vehicular</td>
<td>be a/the vehicle</td>
<td></td>
</tr>
<tr>
<td>(unmarked noun)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Properties</td>
<td>whiteness</td>
<td>white</td>
<td>be white</td>
</tr>
<tr>
<td>(unmarked adj)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actions</td>
<td>destruction</td>
<td>destroyed</td>
<td>destroy</td>
</tr>
<tr>
<td>(unmarked verb)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Following this line of thought, we can now turn to the associations between person forms and pragmatic functions, as shown in Table 10. The usual association between a person form and pragmatic function is reference; hence, the plain person form is used for this function. The default referential use is self-reference in which the person form refers to the person themselves. When the person form is used as a modifier, the pragmatic function changes from reference to modification. Accordingly, a function-indicator is used in some languages to indicate this unusual pragmatic function. When the person form is used to refer to the possessed thing, the pragmatic function change again from modification to reference. In this case, it is called other-reference because the referent is oriented towards the possessed thing instead of the person themselves. This kind of association is still rarer than the association between modification and person forms, because at least the person still refers to themselves when used as a modifier. For instance, in English, you is the unmarked form for reference. In the form your, r is a function-indicator to indicate that the function is changed to modification. In the form yours, s is a function-indicator to indicate that the function is changed to reference again.

Table 10. The association between person forms and pragmatic functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Reference (self)</th>
<th>Modification</th>
<th>Reference (other)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>usual</td>
<td>unusual</td>
<td>rare</td>
</tr>
<tr>
<td>Person forms</td>
<td>plain form</td>
<td>dependent possessive</td>
<td>independent possessive</td>
</tr>
<tr>
<td>Example</td>
<td>you</td>
<td>your</td>
<td>yours</td>
</tr>
</tbody>
</table>

Along these lines, self-reference is a usual pragmatic function of person forms,
while the function of referring to other referents is rare. In light of the form-frequency correspondence hypothesis, the frequent association between a person form and self-reference should be encoded with either less or the same amount of linguistic material in comparison with a rarer association between the person form and the possessed thing.

In this way, we can argue that the key reason for the length universal lies not in that the forms are concerned with more referents, but in the fact that they have a rarer association with the pragmatic function of ‘other-reference’. This rarer association may result in a longer form due to the presence of the function-indicator.

### 6.2 The grammaticalization cline

In this section, I will discuss the functional explanations for Universal 2 with reference to the diachronic source of substantivizers. More specifically, the constituent order universal can be explained in terms of a grammaticalization cline of the possessive markers and substantivizers. This idea is inspired by Lehmann (2015: 72-79), who discusses the grammaticalization of possessive markers or relators with evidence from Japanese, Chinese and Persian. Under his hypothesis, possessive markers in many languages (e.g. in Japanese) are grammaticalized from nouns that are anaphorically related to the head noun (Lehmann 2015: 75). Based on this hypothesis and my sample of 70 languages, I propose the following grammatical cline:

(46) Anaphoric noun > substantivizer > possessive marker> fused/reduced flag > zero

In the grammatical cline shown in (46), from left to right the linguistic form goes through a loss of referentiality and at the same time a process of bleaching in meaning. This grammatical cline is only tentative because I do not have any diachronic evidence to support it. However, the variety of the independent possessive person forms in different languages at a synchronic level in my sample can be used as indirect evidence for this grammatical cline.

This phenomenon can be seen with a few more words about the glossings. In the examples above, for comparative purposes I have glossed all of the morphemes that allow the dependent possessive person forms to stand alone as ‘substantivizers’. However, some real-life language-specific examples are far more complicated and diverse than the global picture described earlier in the paper. In some languages, the substantivizer still has a concrete meaning such as ‘thing’ or ‘possession’. As a matter of fact, many languages use an anaphoric noun as a substantivizer. Although it is not
always clear in the grammatical descriptions to what extent the meaning of the substantivizer is bleached, the glossing in the grammatical descriptions can still serve as clues that point to the diachronic source of the substantivizer. For instance, in Lao, the substantivizer *khòòng* is glossed as an attributive marker, although it is described as having the meaning of ‘thing’ elsewhere in the grammar. In contrast, in Toro Tegu Dogon, the substantivizer is glossed as ‘possession’, as shown in (47).

(47) Toro Tegu Dogon (Dogon, Africa)

a. ú nénù
   2SG dog
   ‘your dog’ (Heath 2010: 122)

b. ú sí
   2SG possession
   ‘yours’ (Literally: your possession) (Heath 2010: 247)

All of the possible diachronic sources of the possessive markers and substantivizers are documented in the supplementary material, which can be considered as indirect evidence for the grammatical cline proposed above. Providing the grammatical cline holds true, how does it serve as the explanation for the constituent order universal? The answer is relatively straightforward. Since the substantivizer and the possessive marker have both developed from the anaphoric noun that is anaphorically related to the possessum, it would be quite natural for the anaphoric form to occupy the same slot as the antecedent. Consequently, the substantivizer and the possessive marker are on the same side as the possessum.

However, there may be more diachronic pathways for different languages. As suggested by Michaelis (2019), there are at least two types of diachronic pathways that give rise to the result of longer independent forms: lengthening scenario or shortening scenario. In the lengthening scenario, the independent forms are developed by adding an additional morpheme to the dependent forms. In the shortening scenario, the dependent forms are developed by a process of shortening (e.g. Old English *min* to Modern English *my*). The explanation with the grammatical cline proposed above is compatible with the lengthening scenario, but it is not clear how this could also be compatible with the shortening scenario. At this stage, it is hard to puzzle this issue out without delving into the exact diachronic pathway of each language. This topic may be considered as a subject for further research.
6.3 Bound and free person forms

The contrast between bound and free person forms in many languages contributes to both the length universal and the alienability universal.

Firstly, dependent possessive forms may be bound forms, especially in head-marking languages, and independent possessive person forms are by definition exclusively free forms. When there are multiple sets of person forms in a language, free forms are normally longer than bound forms. As a result, independent possessive person forms are longer than dependent ones.

Secondly, many languages with an alienability split show head-marking patterns, where the possessor is coded as a person index attached to the noun. In such languages, the alienable possessor is usually coded by adding a possessive marker, while the independent possessor is coded by adding a substantivizer. All three forms are therefore different, as is the case with coding Type (a): ‘[indep.PR], [alien.dep.PR], [inalien.dep.PR]’. An alternative scenario is that the alienable possession displays dependent-marking, which is typically longer than the bound form in the inalienable construction. Thus, the alienable forms are more likely to be free forms since they are always longer than the inalienable forms. Consequently, the alienable forms are used as the independent forms in many languages as the inalienable forms are always bound forms in these languages and thus are not available, which gives rise to Type (b) of ‘[indep.PR, alien.dep.PR], [inalien.dep.PR]’.

7. Conclusions

In this article, I have proposed three universal tendencies with a sample of 70 languages worldwide. I have claimed that independent possessive person forms are either longer than or as long as the corresponding dependent possessive person forms. I have also observed that the constituent order of an independent possessive person form is the same as the constituent order of the corresponding dependent person form and the possessed noun. In addition, in languages where a composite possessive marker and a composite substantivizer are used, the possessive marker is always closer to the person root. Finally, I have claimed that possessive person forms in alienable constructions are more likely to be used as independent possessive person forms than the corresponding inalienable ones. These universals are called the length universal, the
constituent order universal and the alienability universal.

These universals are independent of the marking types of adnominal possessive constructions. We see from the languages discussed in this paper that these universals are relevant and hold true whether the adnominal possessive construction in a language is dependent marking, head marking or double marking.

On top of that, I have tried to explain these universals based on functional motivations and diachronic factors. The length universal can be explained as a case of form-frequency correspondence and coding efficiency, as illustrated by the corpus evidence from English, Mandarin Chinese and Korean. That independent person forms are often longer is because they are much rarer than the corresponding dependent ones. They are rare in two senses. Firstly, they are less frequently used (i.e. lower token frequency). Secondly, it is rare to use a person form to refer to the possessed thing. The constituent order universal can be explained by means of the hypothesis of a possible grammatical cline from anaphoric noun to function-indicating morpheme and finally to zero, and the alienability universal can be explained by the contrast between the bound and free person forms.

**Abbreviations:** 1 first person; 2 second person; 3 third person; ACC accusative; AL alienable; ART article; ATR attributive marker; COP copular; CLF5 the fifth nominal class; DAT dative; DEF definite; FOC contrastive focus; IMP.DECL impersonal declarative; INAL inalienable; IND independent possessor; LOC locative; M masculine; POSS possessive marker; PL plural; PR possessor; PRS present tense; PSSD possessed; REF referential article; SG singular; SUBST substantivizer; TOP topic; VEG the vegetable nominal class.

**References:**

Aikhenvald, Aleksandra Y. 2013. Possession and ownership: A cross-linguistic perspective.


Hanson, Rebecca. 2010. *A grammar of Yine (Piro)*. Bundoora: La Trobe University dissertation.


Hemmilä, Ritva and Pirkko Luoma. 1987. Urim grammar. SIL.


## Appendix  The language sample

<table>
<thead>
<tr>
<th>Language</th>
<th>Family</th>
<th>Macro-area</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambulas</td>
<td>Ndu</td>
<td>Papunesia</td>
<td>Wilson 1980</td>
</tr>
<tr>
<td>Amkoe</td>
<td>Khoisan</td>
<td>Africa</td>
<td>Collins &amp; Gruber 2014</td>
</tr>
<tr>
<td>Anindilyakwa</td>
<td>Gunwinyguan</td>
<td>Australia</td>
<td>van Egmond 2012</td>
</tr>
<tr>
<td>Anywa</td>
<td>Nilotic</td>
<td>Africa</td>
<td>Reh 1996</td>
</tr>
<tr>
<td>Assiniboine</td>
<td>Siouan</td>
<td>North America</td>
<td>Cumberland 2005</td>
</tr>
<tr>
<td>Awjila Berber</td>
<td>Afro-Asiatic</td>
<td>Africa</td>
<td>van Putten 2013</td>
</tr>
<tr>
<td>Bao'an Tu</td>
<td>Mongolic</td>
<td>Eurasia</td>
<td>Fried 2010</td>
</tr>
<tr>
<td>Bardi</td>
<td>Nyulnyulan</td>
<td>Australia</td>
<td>Bowern 2012</td>
</tr>
<tr>
<td>Bena</td>
<td>Atlantic-Congo</td>
<td>Africa</td>
<td>Morrison 2011</td>
</tr>
<tr>
<td>Berik</td>
<td>Tor-Orya</td>
<td>Papunesia</td>
<td>Westrum 1988</td>
</tr>
<tr>
<td>Betta Kurumba</td>
<td>Dravidian</td>
<td>Eurasia</td>
<td>Coelho 2003</td>
</tr>
<tr>
<td>Bororo</td>
<td>Bororoan</td>
<td>South America</td>
<td>Crowell 1979</td>
</tr>
<tr>
<td>Canela-Krahô</td>
<td>Nuclear-Macro Je</td>
<td>South America</td>
<td>Popjes &amp; Popjes 1986</td>
</tr>
<tr>
<td>Cashibo-Cacataibo</td>
<td>Pano-Tacan (Panoan)</td>
<td>South America</td>
<td>Zariquiey Biondi 2011</td>
</tr>
<tr>
<td>Cavineha</td>
<td>Pano-Tacan</td>
<td>South America</td>
<td>Guillaume 2008</td>
</tr>
<tr>
<td>Choctaw</td>
<td>Muskogean</td>
<td>North America</td>
<td>Broadwell 2006</td>
</tr>
<tr>
<td>Desano</td>
<td>Tucanoan</td>
<td>South America</td>
<td>Silva 2012</td>
</tr>
<tr>
<td>English</td>
<td>Indo-European</td>
<td>Eurasia</td>
<td>personal knowledge</td>
</tr>
<tr>
<td>Ese Eja</td>
<td>Pano-Tacan</td>
<td>South America</td>
<td>Vuillermet 2012</td>
</tr>
<tr>
<td>Gaahmg</td>
<td>Eastern Jebel</td>
<td>Africa</td>
<td>Stirtz 2012</td>
</tr>
<tr>
<td>Gilaki</td>
<td>Indo-European</td>
<td>Eurasia</td>
<td>Rastorgueva et al. 2012</td>
</tr>
<tr>
<td>Halkomelem</td>
<td>Salishan</td>
<td>North America</td>
<td>Suttles 2004</td>
</tr>
<tr>
<td>Hınıq</td>
<td>Nakh-Daghestanian</td>
<td>Eurasia</td>
<td>Forker 2013</td>
</tr>
<tr>
<td>Hmong Njua</td>
<td>Hmong-Mien</td>
<td>Eurasia</td>
<td>Wang 1985</td>
</tr>
<tr>
<td>Huallaga Huánuco Quechua</td>
<td>Quechuan</td>
<td>South America</td>
<td>Weber 1989</td>
</tr>
<tr>
<td>Huehuetla Tepehua</td>
<td>Totonacan</td>
<td>North America</td>
<td>Kung 2007</td>
</tr>
<tr>
<td>Ingush</td>
<td>Nakh-Daghestanian</td>
<td>Eurasia</td>
<td>Nichols 2011</td>
</tr>
<tr>
<td>Isthmus-Mecayapan Nahuatl</td>
<td>Uto-Aztec</td>
<td>North America</td>
<td>Wolgemuth 1981; Jorden 1963; Kaiser et.al 2013</td>
</tr>
<tr>
<td>Japanese</td>
<td>Japonic</td>
<td>Eurasia</td>
<td>2013</td>
</tr>
<tr>
<td>Kaluli</td>
<td>Bosavi</td>
<td>Papunesia</td>
<td>Grosh &amp; Grosh 2004</td>
</tr>
<tr>
<td>Karajá</td>
<td>Nuclear-Macro-Je</td>
<td>South America</td>
<td>Ribeiro 2012</td>
</tr>
<tr>
<td>No.</td>
<td>Language</td>
<td>Subgroup</td>
<td>Regional</td>
</tr>
<tr>
<td>-----</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>32</td>
<td>Kilen</td>
<td>Tungusic</td>
<td>Eurasia</td>
</tr>
<tr>
<td>33</td>
<td>Komnzo</td>
<td>Upper-Morehead</td>
<td>Papunesia</td>
</tr>
<tr>
<td>34</td>
<td>Korean</td>
<td>Koreanic</td>
<td>Eurasia</td>
</tr>
<tr>
<td>35</td>
<td>Lao</td>
<td>Tai-Kadai</td>
<td>Eurasia</td>
</tr>
<tr>
<td>36</td>
<td>Mandarin Chinese</td>
<td>Sino-Tibetan</td>
<td>Eurasia</td>
</tr>
<tr>
<td>37</td>
<td>Mauwake</td>
<td>Nuclear-trans-new-Guinea</td>
<td>Papunesia</td>
</tr>
<tr>
<td>38</td>
<td>Menya</td>
<td>Angan</td>
<td>Papunesia</td>
</tr>
<tr>
<td>39</td>
<td>Moskona</td>
<td>East Bird's head</td>
<td>Papunesia</td>
</tr>
<tr>
<td>40</td>
<td>Navajo</td>
<td>Athabaskan-Eyak-Tlingit</td>
<td>North America</td>
</tr>
<tr>
<td>41</td>
<td>Northern Tujia</td>
<td>Sino-Tibetan</td>
<td>Eurasia</td>
</tr>
<tr>
<td>42</td>
<td>Okspamin</td>
<td>Nuclear-trans-new-Guinea</td>
<td>Papunesia</td>
</tr>
<tr>
<td>43</td>
<td>Piro</td>
<td>Kiowa-Tanoan</td>
<td>North America</td>
</tr>
<tr>
<td>44</td>
<td>South Wa</td>
<td>Austroasiatic</td>
<td>Eurasia</td>
</tr>
<tr>
<td>45</td>
<td>Southern Pomo</td>
<td>Pomoan</td>
<td>North America</td>
</tr>
<tr>
<td>46</td>
<td>Southern Yukaghir</td>
<td>Yukaghir</td>
<td>Eurasia</td>
</tr>
<tr>
<td>47</td>
<td>Supyire</td>
<td>Atlantic-Congo</td>
<td>Africa</td>
</tr>
<tr>
<td>48</td>
<td>Sursurunga</td>
<td>Austronesian</td>
<td>Papunesia</td>
</tr>
<tr>
<td>49</td>
<td>Tapieté</td>
<td>Tupian</td>
<td>South America</td>
</tr>
<tr>
<td>50</td>
<td>Tauya</td>
<td>Nuclear-trans-new-Guinea</td>
<td>Papunesia</td>
</tr>
<tr>
<td>51</td>
<td>Tehit</td>
<td>West Bird's head</td>
<td>Papunesia</td>
</tr>
<tr>
<td>52</td>
<td>Teiwa</td>
<td>Timor-Alor-Pantar</td>
<td>Papunesia</td>
</tr>
<tr>
<td>53</td>
<td>Teribe</td>
<td>Chibchan</td>
<td>North America</td>
</tr>
<tr>
<td>54</td>
<td>Tidore</td>
<td>North Halmahera</td>
<td>Papunesia</td>
</tr>
<tr>
<td>55</td>
<td>Toro Tegu Dogon</td>
<td>Dogon</td>
<td>Africa</td>
</tr>
<tr>
<td>56</td>
<td>Trió</td>
<td>Cariban</td>
<td>South America</td>
</tr>
<tr>
<td>57</td>
<td>Tí'ixa</td>
<td>Khoe-Kwadi</td>
<td>Africa</td>
</tr>
<tr>
<td>58</td>
<td>Urim</td>
<td>Nuclear-Torricelli</td>
<td>Papunesia</td>
</tr>
<tr>
<td>59</td>
<td>Urubú-Kaapor</td>
<td>Tupian</td>
<td>South America</td>
</tr>
<tr>
<td>60</td>
<td>Ute</td>
<td>Uto-Aztecan</td>
<td>North America</td>
</tr>
<tr>
<td>61</td>
<td>Vaeakau-Taumako</td>
<td>Austronesian</td>
<td>Papunesia</td>
</tr>
<tr>
<td>62</td>
<td>Wari'</td>
<td>Chapacuran</td>
<td>South America</td>
</tr>
<tr>
<td>63</td>
<td>Waris</td>
<td>Border</td>
<td>Papunesia</td>
</tr>
<tr>
<td>64</td>
<td>West Yugur</td>
<td>Turkic</td>
<td>Eurasia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Wintu</td>
<td>Penutian</td>
<td>North America</td>
<td>Pitkin 1984</td>
</tr>
<tr>
<td>Yareba</td>
<td>Yareban</td>
<td>Papuasia</td>
<td>Weimer &amp; Weimer 1975</td>
</tr>
<tr>
<td>Yidp</td>
<td>Pama-Nyungan</td>
<td>Australia</td>
<td>Dixon 1977</td>
</tr>
<tr>
<td>Yimas</td>
<td>Lower-Sepik-Ramu</td>
<td>Papuasia</td>
<td>Foley 1991</td>
</tr>
<tr>
<td>Zialo</td>
<td>Mande</td>
<td>Africa</td>
<td>Babaev 2010</td>
</tr>
<tr>
<td>Zoque Chimalapa</td>
<td>Mixe-Zoque</td>
<td>North America</td>
<td>Johnson 2000</td>
</tr>
</tbody>
</table>

The list of reference grammars


Fehn, Anne-Maria. 2014. A grammar of Ts'ixa (Kalahari Khoe). Universität zu Köln.
Hanson, Rebecca. 2010. A grammar of Yine (Piro). La Trobe University.
Hemmilä, Ritva and Pirkko Luoma. 1987. Urim grammar. SIL.


Westrum, Peter. 1988. Berik grammar sketch. Irian XVI.


Zhang, Paiyu. 2013. The Kilen language of Manchuria: grammar of a moribund Tungusic language. Hong Kong: University of Hong Kong.
