

Chapter 8

Support from creole languages for functional adaptation in grammar: Dependent and independent possessive person-forms

Susanne Maria Michaelis

Leipzig University & Max Planck Institute for the Science of Human History
(Jena)

It seems to be a robust empirical observation that independent possessive person-forms (such as English *mine*, *yours*, *hers*) are always longer than (or as long as) the corresponding adnominal possessive person-forms (such as English *my*, *your*, *her*). Since adnominal forms are also much more frequent in discourse than independent forms, this universal coding asymmetry can be subsumed under the grammatical form-frequency correspondence hypothesis (Haspelmath et al. 2014). In other words, the fact that independent possessive forms are longer can be seen as a functional response to the need to highlight rarer, less predicatable forms.

In this paper, I present evidence from creole languages and show that irrespectively of their young age and extremely accelerated grammaticalization processes, these high-contact languages confirm the coding asymmetry. Moreover, creole languages, just as non-creole languages, show a diverse array of diachronic pathways all leading eventually to longer independent possessive person-forms. Such a case of multi-convergence of structures through very different diachronic processes strongly suggests that the current patterns cannot be explained exclusively on the basis of the sources and the kinds of changes that commonly give rise to independent (and adnominal) possessive forms, but that there is an overarching functional efficiency principle underlying these coding asymmetries.



1 Introduction

Languages are functionally adapted to their users' needs in a variety of ways. This can be seen in a range of different domains, such as (i) text genres, (ii) social structure and (iii) the ecological environment. The genre of informal, spontaneous face-to-face communication is reflected in grammatical features of loosely connected discourse with mainly coordinated or juxtaposed sentences, many hesitation phenomena, overlapping utterances, and piecemeal structuring of information in accordance with online processing needs, whereas text genres intended for formal, planned, out-of-context, written communication show densely integrated information, multiple syntactic embedding strategies and therefore longer sentences, and greater syntagmatic variation (Koch & Oesterreicher 2012[1985] [1985]). Secondly, languages are adapted to the social structuring of their users, for instance to the percentage of second language speakers in a speech community: In a well-known study, Lupyan & Dale (2010) analyzed data from the *World atlas of language structures* (Haspelmath et al. 2005) and found that the greater the number of second language speakers in a speech community, the simpler are aspects of the morphology of the languages spoken by these communities. In a similar vein, Bentz & Winter (2013) found that languages with many second language speakers tend to have fewer morphological cases. And third, it has been shown that speakers adapt their languages to their ecological environments, for example by using whistled speech in distant communication to overcome the background noise of rural environments (Meyer 2005; 2008).

In the present chapter, I will look at yet another instance of functionally adapted linguistic structures: efficiency-based universal coding asymmetries in grammar, also called form-frequency correspondences (see Haspelmath 2019 [this volume]). More specifically, I will discuss one specific universal coding asymmetry resulting from asymmetric frequency of use patterns in discourse: the difference between dependent and independent possessive person-forms. Independent person-forms such as *mine*, *yours*, *hers*, and *ours* are coded with forms that are longer than or equally long as dependent possessive person-forms such as *my*, *your*, *her*, and *our*. I claim that the reason for this is a general efficiency principle: Less frequent and therefore more surprising meanings need more costly coding than more frequent and therefore more predicatable meanings.

Such functional-adaptive explanations have a diachronic component (Bybee 1988): Since the current system is often rigidly conventional, the adaptive forces must have been active in earlier diachronic change. But how can we understand such a development? Functionally adapted coding asymmetries, as seen in depen-

dent/independent possessive person-forms, are the outcome of hundreds, sometimes thousands of years of language change processes. These processes reflect countless speech acts between interlocutors adding up incrementally and resulting in the crystallization of functionally adapted grammatical structures over time. As grammatical change progresses at an extremely slow pace compared to other cultural evolutionary processes, the step-by-step changes which bring about functionally adapted grammatical structures are often opaque or difficult to trace, even in languages with a well-documented written history (see Seržant 2019 [this volume]). To circumnavigate this difficulty, I will focus on creole languages, which are born out of extremely accelerated change processes in the context of the European colonial expansion, roughly during the 16th to 20th centuries. These high-contact languages have evolved their complex grammatical structures within only a few hundred years. In this way they are a good test case for functional-adaptive change processes because creoles demonstrate in a kind of fast motion what happens to grammatical structures under functional pressures, which in less contact-influenced languages would have taken hundreds (or thousands) of years to evolve. In this way, creoles open a unique window on grammatical change processes which in these languages can be traced gradually from their transparent source constructions to various further grammaticalized stages, processes which are supposed to be operative in all languages at all times, but which take much more time to proceed in languages less heavily influenced by contact.

I make two main points in this paper:

(i) Evidence from creole languages indeed confirms the coding asymmetry: Independent person-forms are coded with forms that are always longer than, or as long as, the dependent person-forms, but never shorter.

(ii) Creole languages, just as non-creole languages, show a diverse array of diachronic pathways all leading eventually to longer independent possessive person-forms. Such a case of multi-convergence of structures through very different diachronic processes strongly suggests that there is an overarching functional efficiency principle underlying these coding asymmetries (see Haspelmath 2019 [this volume]).

After introducing the coding asymmetry in possessive person-forms in §2, in §3 I discuss various types of source constructions and diachronic pathways which lead to longer independent possessive person-forms. Then in §4, I present a range of cases from creole languages and their various diachronic pathways. In §5, I consider but ultimately reject some alternative explanations against the background of the functional efficiency-based explanation adopted in this article.

2 Coding asymmetry: Dependent vs. independent possessive person-forms

Dependent possessive person-forms always occur together with an overt noun within a nominal phrase, as in *your house*, whereas independent possessive person-forms occur without an overt noun, as in *mine*. In the latter case, the referent of the noun is understood from the context because of an anaphoric relationship, as in (1a) and (1b), or because of a predicative use, as in (1c).

- (1) English
 - a. Your house is bigger than mine. (= ‘than my house’)
 - b. Their dog is in a kennel, but ours sleeps under my bed. (= ‘our dog’)
 - c. Is this bike yours? (= ‘your bike’)

In a recent study, Ye (2017)¹ has found that in the world’s languages independent possessive person-forms like English *mine*, French *le mien* ‘mine’, and Mandarin Chinese *wo de* ‘mine’ are coded with forms that are longer than or equally long as the corresponding dependent possessive person-forms, such as English *my*, French *mon* ‘my’, or at least not shorter, as illustrated by Mandarin Chinese *wo de* ‘my’. Coding length here refers to the number of segments in the signal, or possibly to the amount of biomechanical effort (see Napoli et al. 2014 with regard to sign languages). Most importantly, examples of counter-asymmetric coding are not attested, i.e. there are no languages where the dependent possessive person-forms are longer than independent possessive person-forms, e.g. **mine house* vs. *my* ‘mine’. Note that (in)dependent possessive person-form can be manifested through a range of language-specific structures, also embracing complex forms, such as combinations of articles or adpositions with pronouns, as in French *le mien* and Mandarin Chinese *wo de* [I GEN].

Table 1 shows a number of different types of correspondences between dependent and independent person-forms in the world’s languages: Firstly, many languages code the two types of person-forms identically and thus with equally long forms, as for instance in Mandarin Chinese. In other languages, the independent person-form has an additional marker compared to the dependent form. This can be a substantivizer, as in Lezgian (*-di*), or an additional stem, as in Kanuri (*kaá-*). In some languages the definite article is used to form the independent person-form, such as in Italian *la mia* (with kinship terms like *sorella* ‘sister’).²

¹Ye (2017) analyzes a sample of 69 genealogically and areally unrelated languages.

²If nouns like *casa* ‘house’ or *libro* ‘book’ were considered, Italian would be classified just like Chinese (identical pattern) because there would be no coding difference: *la mia casa* ‘my house’ vs. *la mia* ‘mine’, *il mio libro* ‘my book’ vs. *il mio* ‘mine’.

Yet another synchronic pattern in independent person-forms consists in having extra material on the dependent form, as in Coptic *p-ô-k* [ART-INDEP-2SG] ‘yours’ (vs. *p-ek-ran* [ART-2SG-name] ‘your name’).

Table 1: Some types of correspondences of dependent and independent person-forms

Pattern type	Language	Dependent person-form	Independent person-form	Source
identical	Mandarin Chinese	<i>wo de shu</i> I GEN book ‘my book’	<i>wo de</i> I GEN ‘mine’	
additional marker	Lezgian	<i>zi ktab</i> I.GEN book ‘my book’	<i>zi-di</i> I.GEN-SUBST ‘mine’	Haspelmath (1993: 110)
additional stem	Kanuri	<i>fewá-ndé</i> COW-1PL.POSS ‘our cows’	<i>kaá-nde</i> INDEP-1PL ‘ours’	Cyffer (1998: 31f.)
additional article	Italian	<i>mia sorella</i> ‘my sister’	<i>la mia</i> ‘mine’	Schwarze (1988: 44,286f.)
longer form	Coptic	<i>p-ek-ran</i> ART-2SG-name ‘your name’	<i>p-ô-k</i> ART-INDEP-2SG ‘yours’	Haspelmath (2015: 277)

Apparently the only possible generalization which can be drawn from the typological variation is that the independent person-form is always longer than, or as long as, the dependent person-form, but never shorter³.

Now the claim is that these coding asymmetries reflect asymmetries of frequency of use. More frequent meanings (here: dependent possessives) are more predicatable and therefore speakers or signers can reduce the amount of the linguistic signal in taking into account how much of the signal hearers and receivers (in sign languages) need in order to successfully reconstruct the intended meaning. By contrast, less frequent meanings (here: independent possessives) are in

³See also Croft (1991), who very similarly predicts “function-indicating morphosyntax” in all the atypical combinations of lexical semantic class and pragmatic functions, whereas typical combinations lack function-indicating markers (Croft 1991: 51), e.g. marked predicative nominals vs. unmarked nouns, or marked predicative adjectives vs. unmarked attributive adjectives.

need of a greater amount of signal coding for the hearer to be able to infer the meaning.

Indeed, frequency counts of three large text corpora of three different languages (English, Korean, and Mandarin Chinese⁴) confirm the hypothesis that dependent and independent person-forms are unequally spread over discourse in such a way that dependent possessive person-forms are generally more frequent than their independent counterparts. Table 2 shows data from British English.

Table 2: (In)dependent possessive person-forms in the British National Corpus

Dependent	Token frequency	Independent	Token frequency
<i>my</i>	145,250	<i>mine</i>	6,067
<i>your</i>	132,598	<i>yours</i>	4,059
<i>our</i>	92,314	<i>ours</i>	1,658
<i>their</i>	251,410	<i>theirs</i>	976

Interestingly, frequency counts from Mandarin Chinese, a language without a coding asymmetry in possessive person-forms, give the same results as counts for English and Korean, which have the coding asymmetry in possessive person-forms (see Ye 2017). Therefore, the prediction is that we find similar frequency distributions of dependent and independent possessive person-forms in all languages, independently of whether the universal coding asymmetry is grammaticalized or not.

3 Types of source constructions and diachronic pathways

As noted earlier, synchronic universal coding asymmetries have a diachronic correlate because the adaptive forces must have been active in earlier stages of the language and have kept shaping grammatical structures according to the functionally motivated efficiency principle: less predicatable meanings need more coding and more predicatable meanings need less coding.

There is a wide variety of sources and diachronic pathways by which independent possessive person-forms come to be longer than the dependent forms. Generally, one can distinguish two scenarios: either the more frequent member of the grammatical opposition is shortened (Bybee 2007), or the rarer member of

⁴For frequency counts in Korean and Mandarin Chinese, see Ye 2017.

the grammatical opposition is lengthened⁵ (Haspelmath 2008). In the shortening scenario, speakers assess what hearers can predict and adjust their articulations accordingly, resulting in shortening of the signal of the more frequent form of a grammatical opposition. In this way, Old English *min* ‘my’ was eventually shortened to Modern English *my*, likewise Old Spanish *mío* was shortened to Modern Spanish *mi*. The Coptic contrast between *pôk* ‘yours’ and *pek* ‘your’ that we saw in Table 1 is likewise attributable to shortening of the earlier full person-form *pôk* to *pek-*. The shortened form became a dependent person-form whereas the old form *pôk* became restricted to the independent function (Eitan Grossman p.c.).

The lengthening scenario can be described as follows: When hearers are in danger of making wrong predictions, speakers tend to help them by using forms which – compared to the rarer member of the opposition – have been lengthened with some extra material. One example comes from German, where the independent form *der mein-ig-e* [DEF 1SG.POSS-INDEP-MASC.SG.NOM] ‘mine’ is based on the dependent form *mein* ‘my’ plus an additional suffix *-ig*, which occurs in other derived adjectives (like *selb-ig* ‘same’, *bärt-ig* ‘bearded’, *ehrgeiz-ig* ‘ambitious’). As we see in Tables 3 and 4, the array of source constructions and diachronic pathways which give rise to longer independent possessive person-forms is very diverse.

Table 3: Shortened dependent form

Language	Strategy	Dependent form	Independent form
English	phonological reduction of dependent form	<i>my</i>	<i>mine</i>

The different strategies range from the use of a dummy noun (‘my thing’, ‘my property’), intensified person forms (‘my own’), the use of adpositions (‘of my’) and definite articles (‘the my’) to general nominalizer (‘my one’). One special strategy to arrive at longer independent possessive person-forms consists in recruiting already existing pronominal (lengthened) forms which have been used for other grammatical functions. One example comes from Middle English varieties, where the independent possessive forms *her-n*, *our-n*, *their-n* (still surviving in English dialects today, see Kortmann & Lunkenheimer 2013) go back

⁵Here, the term ‘lengthening’ mainly refers to processes by which a given linguistic form is expanded or augmented by new lexical or morphosyntactic material. But – in principle – lengthening may also pertain to phonological/phonetic processes, such as vowel lengthening or gemination.

Table 4: Lengthened independent form

Language	Strategy	Dependent form	Independent form
German	affixal lengthening	<i>mein</i> [1SG.POSS]	<i>der mein-ige</i> [DEF 1SG.POSS-INDEP]
Arabic	dummy noun: 'property'	<i>-ii</i> [1SG.POSS]	<i>milk-ii</i> [property-1SG.POSS]
Greek	intensified person form 'own'	<i>mu</i> [1SG.POSS]	<i>dhikó mu</i> [INTENS 1SG.POSS]
Diu Indo-Portuguese	use of adposition 'of, for'	<i>mi</i> [1SG.POSS]	<i>də mi</i> [of 1SG.POSS]
Albanian	use of definite article	<i>im</i> [1SG.POSS]	<i>im-i</i> [1SG.POSS-DEF]
Berbice Dutch	general nominalizer	<i>εke</i> [1SG.POSS], [1SG]	<i>εke-je</i> [1SG.POSS-NMLZ]
English (dialectal)	exaptation	<i>her</i> [3SG.F.POSS]	<i>her-n</i> [3SG.POSS-INDEP]

to erstwhile feminine dative case-marked pronominal forms with the suffix *-n* (*hire-n* [3SG.FEM.DAT] 'to her'). In Middle English, such dative forms got re-used, or "exapted", to function as independent possessive forms, also under the additional analogical pressure from the *my/mine* and *thy/thine* oppositions (see Allen 2002, and for the notion of exaptation, see Lass 1990; 2017; Norde & Van de Velde 2016 and the discussion below).

Irrespectively of the shortening or the lengthening scenario, ALL these developments result in coding asymmetries which work in the SAME direction: The less frequent member (here the independent possessive person-form) is coded with a form that is always coded as least as long as the more frequent member of the pair, but never shorter.

Now how do creole languages fit into this picture? In the next section, I will consider possessive person-forms in various creole languages from around the world (based on the *Atlas of pidgin and creole language structures*, Michaelis et al. 2013, apics-online.info) to check whether the universal trend identified by typological work can be supported by these high-contact languages.

4 Diverse pathways in creoles

Before looking at possessive person-forms in creole languages, I would like to highlight one characteristic feature of these languages which is crucial for the argument put forward in this paper: Creole languages show an unusual amount of freshly grammaticalized material due to an accelerated pace of grammatical change processes (Haspelmath & Michaelis 2017; Michaelis & Haspelmath forthcoming). Examples come from tense-aspect-mood markers, such as the Negerhollands future tense marker *lo* < *loo* 'go' < Dutch *lopen* 'run', or the Jamaican anterior marker *wehn* < English *been*. Creoles also show newly grammaticalized case markers, such as the dative marker *pe* in Diu Indo-Portuguese (< Portuguese *para*), the accusative marker *ku* in Papiá Kristang (< Portuguese *com* 'with'), or voice markers, such as the reciprocal marker *kanmarad* in Seychelles Creole (< French *camarade*). The explanation for these widespread newly grammaticalized markers appears to be as follows: Speakers communicating in high-contact situations which involve many second language speakers tend to rely on extra transparency of their utterances in order to successfully get their messages across.⁶ These instances of extra transparency give rise to newly grammaticalized structures by refunctionalizing erstwhile content words or otherwise less grammaticalized constructions, as seen in the examples cited above.

Turning to possessive forms, let us now consider the following three guiding questions:

- Do creoles confirm the universal coding asymmetry discussed in this paper?
- Does the need for extra transparency translate into freshly grammaticalized constructions also in the domain of possessive person-forms?
- Which kinds of source constructions give rise to the various possessive person-forms?

⁶See already Seuren & Wekker (1986) for the notion of transparency in the creolization process.

The answer to the first question is a straightforward yes: The creole evidence, which comes from 59 creoles world-wide with different lexifier and substrate languages (see Haspelmath & APiCS Consortium 2013 and Figure 1 in the Appendix), confirm the universal coding asymmetry: Independent possessive person-forms are coded with forms that are longer than or equally long as dependent possessive person-forms. Some examples are given in Table 5.

Table 5: Dependent and independent possessive person-forms in some creole languages

Creole language	Dependent form	Independent form
Bislama (Meyerhoff 2013)	<i>blong yu</i> [POSS 2SG] ‘your’	<i>blong yu</i> [POSS 2SG] ‘yours’
Kinubi (Luffin 2013)	<i>tá-i</i> [POSS-1SG] ‘my’	<i>tá-i</i> [POSS-1SG] ‘mine’
Batavia Creole (Maurer 2013)	<i>minya</i> [1SG.POSS] ‘my’	<i>minya sua</i> [1SG.POSS POSS] ‘mine’
Martinican Creole (Colot & Ludwig 2013)	<i>-mwén</i> [1SG.POSS] ‘my’	<i>ta mwén</i> [POSS 1SG.POSS] ‘mine’
Pichi (Yakpo 2013)	<i>yù</i> [2SG.POSS], [2sg] ‘you’	<i>yù yon</i> [2SG.POSS own] ‘yours’
Palenquero (Schwegler 2013)	<i>mi</i> [1SG.POSS] ‘my’	<i>ri mi</i> [of 1SG.POSS] ‘mine’

The following Table 6 presents a quantitative overview of the different construction types found in creole languages of APiCS. Here, only languages with an exclusive value assignment are considered (48 out of 59 creole languages).

Likewise, the answer to the second question raised above is positive: The majority of the possessive person-forms are indeed freshly grammaticalized and therefore still transparent enough to be traced quite closely with respect to the different diachronic processes that have brought about their coding asymmetry.

8 Support from creole languages for functional adaptation in grammar

Table 6: Distribution of different construction types over 48 creoles in independent possessive person-forms (APiCS Feature 39)

Coding pattern	Feature value	Number of creole languages in APiCS
Symmetry	Identical to dependent pronominal possessor	20
Asymmetry	Special adposition plus pronoun	9
	Other word plus dependent pronominal possessor	13
	Special form for independent pronominal possessor	6
	Total	48

Coding asymmetries explicitly allow for the two forms of an opposition to be equally long (either overtly or zero-coded)⁷, as is the case in Mandarin Chinese *wo de* ‘my’, ‘mine’ cited above. As Table 6 shows, there are quite a number of creole languages which show this coding pattern, i.e. no length difference in the coding of both forms, as for instance in Tok Pisin *bilong mi* [POSS 1SG] ‘my’, ‘mine’ or the related language Bislama (see Table 5). These languages do not contradict the universal coding asymmetry, as they do not show the opposite coding pattern, i.e. longer dependent forms against shorter independent forms.

Let us now turn to creole languages for which we can attest a coding asymmetry in possessive forms. As for the source constructions, I will first look at cases of shortening that parallel the English development from *mine* to *my*. One example comes from Juba Arabic, where the original form *bita-i* [POSS-1SG] ‘my/mine’ gets shortened and at some point reanalyzed as the dependent possessive *tá-i* ‘my’, as in *ída tái* [hand 1SG.POSS] ‘my hand’ (Manfredi & Petrollino 2013), whereas the older non-shortened form *bita-i* continues to be used as the independent possessive form meaning ‘mine’.

However, the vast majority of asymmetric correspondence types in creole languages – as in non-creole languages – follow the second scenario described in §3: the coding asymmetry comes about by some process of expanding the less

⁷See also Croft (1991: 58f.), who calls such cases NEUTRAL evidence.

frequent member of the grammatical opposition. One widespread source is the use of an adposition going back to ‘of’ or ‘for’ in one of the European lexifier languages French, Portuguese, English etc. An example comes from Portuguese-based Santome (Hagemeyer 2013), where the dependent possessive person-form *mu* ‘my’, which is expanded by the genitive preposition *ji* (< Portuguese *de* ‘of’), gives rise to the independent possessive form *ji mu* ‘mine’. Jamaican *fi-mi* ‘mine’ is another instance of the lengthening of the dependent form *mi* ‘1SG.POSS’ (and also 1SG ‘I’) by the preposition *fi* ‘for’ (< English *for*).

A second source construction for independent possessive person-forms in creole languages involves the use of a dummy noun, such as ‘part’ or ‘thing’ (as mentioned above), as in Haitian Creole *pa m nan* [part 1SG.POSS DEF] ‘mine’ (lit. ‘my part’, *pa* < French *part* ‘part’) as opposed to dependent forms, such as *-m* (*nan*) [1SG.POSS (DEF)] ‘my’ in *se m* [sister POSS.1SG] ‘my sister’. The polysemous morpheme *pa*, which in some contexts still has the original lexical meaning ‘part’, has grammaticalized into a possessive form which can also be used in contexts where the possessor is stressed, as in (2).

- (2) Haitian Creole (Fattier 2013)
 Liv pa m nan bèl.
 book POSS POSS.1SG DEF beautiful
 ‘MY book is beautiful.’

However, the non-stressed noun phrase would be *liv m* [book POSS.1SG] ‘my book’ (Fattier 2013). Here, we clearly see that the postposed morpheme *pa* in *pa m* does not denote a part of something, but has grammaticalized into a possessive marker, as the literal meaning ‘book my part’ is not available for this construction. The same holds for the independent possessive form *pa m nan* ‘mine’: the meaning is not ‘my this part’, but *pa* has become part of the newly grammaticalized independent possessive form ‘mine’.

A third source construction for independent possessive forms features an intensifier which is added to the dependent possessive, as in Krio *mi yon* [1SG.POSS INTENS.OWN] ‘mine’ (the dependent possessive form being *mi* ‘my’) (Finney 2013).

There is a fourth source of independent forms involving a general (adjectival) nominalizer, such as ‘one’. In Berbice Dutch, there is a general nominalizer *-je* which is added to the personal pronoun *εke* [1SG.POSS]/[1SG] ‘my’ (‘I’), resulting in *εke-je* [1SG.POSS-NMLZ] ‘mine’ (see Table 4). This nominalizer goes back to Eastern Ijo, the substrate language of Berbice Dutch, where it has singular nonhuman reference, whereas in Berbice Dutch it has grammaticalized into a generic nominalizer (Kouwenberg 2013).

A fifth source can be illustrated with an example from Reunion Creole, where the determiner/demonstrative *sa* is one of the lengthening elements (besides the genitive preposition *d*) in the independent possessive person-form *sa d mwen* [DEM of 1SG] ‘mine’, compared to the dependent form *mon* [1SG.POSS] ‘my’.

In some creole languages the source construction is not known, as in Louisiana Creole. Here, the marker *kenn* is used as a morpheme to code the independent possessive person-forms, as in *mo-kenn* [1SG.POSS-POSS] ‘mine’. This morpheme could perhaps be traced back to a 2SG.FEM independent person-form in French *tienne* ‘yours’, which has developed into /kien/, which would then have analogically spread to the whole paradigm, as in *mo-kenn* [1SG.POSS-POSS] ‘mine’, *to-kenn* [2SG.POSS-POSS] ‘yours’, *li-kenn* [1SG.POSS-POSS] ‘his’ (Neumann-Holzschuh & Klingler 2013, Neumann-Holzschuh p.c.). The unusual feature in this scenario is the idea that it is the second-person form which analogically spreads to all other persons, and not the more frequent 1SG or 3SG forms. Whether this is the right reconstruction of the origin of *kenn* is not clear.

Generalizing over all instances of newly grammaticalized independent possessive forms in creole languages, we can state that irrespectively of the diverse source constructions, it is the independent possessive person-form that, in ALL instances, is longer than, or as long as, the dependent person-form, but never shorter.

5 Possible alternative explanations

We have seen that the cross-creole data support the universal coding asymmetry in possessive person-forms, and that this synchronic asymmetry can be explained by a functional-adaptive constraint of coding efficiency: More frequently expressed meanings (dependent possessives) need less costly signal encoding because they are highly predicatable, whereas less frequently expressed meanings need more robust signal encoding because they are less predicatable (Haspelmath 2019 [this volume]; see Norcliffe & Jaeger 2016 and Jaeger & Buz 2018 for supporting psycholinguistic evidence in other domains of morphosyntax). Before concluding this paper, I will consider several alternative explanations, but reject them all as less convincing.

5.1 Semantics, iconicity, and syntax

Some functional linguists might argue for an alternative, semantically based or iconicity-based explanation here, namely that the independent possessive form is

semantically more complex in that it combines possession and referentiality, and so additional material has to be adduced in order to express this more complex concept, or to compensate for the absence of an overt nominal.

But I would reject such a proposal because it is not obvious that independent possessors are semantically more complex. Rather, we can think of the situation as follows: Possessors refer to objects and persons, but at the same time, when used in possessive constructions, they also express properties, like adjectives. In the most frequent use, possessive forms (again like adjectives) have a modification function, as in *my house* (the “unmarked” use in terms of Croft 1991). But when possessive forms are used in the less frequent referential function, as in *mine*, specific marking is needed to highlight this unusual noun-like usage. Semantically, there is not really any difference in complexity of both kinds of person-forms: dependent possessive forms combine person and property with regard to possession in a MODIFICATION function, whereas independent person-forms combine person and property with regard to possession in a REFERENCE function. There is thus only a difference in the propositional function in which the semantic concepts are expressed (modification vs. reference), but there is no ADDITIONAL semantic complexity in independent possessive person-forms.

Likewise, some linguists might argue that the motivation for the coding asymmetry is purely syntactic, as the two possessive forms occupy different syntactic slots. As the modifier, such as French *mon*, cannot occur as the head of a NP, it has to be transformed into a noun by what Croft (1991: 58f.) calls “function-indicating markers”, thus yielding *le mien* ‘mine’ in French. The use of the definite article represents one of the lengthening processes in independent possessive person-forms that I described above. But I would interpret the mere use of function-indicating markers as the frozen grammaticalized results of hundreds or thousands of years of speakers performing communicatively efficient speech acts by marking the less predicatable meanings with more elaborate linguistic matter. In this respect, there is no contradiction between today’s syntax and yesterday’s (and earlier) speakers’ preferences to highlight less predicatable meanings by more morphosyntactic material, which accumulated over generations and eventually contributes to the shaping of syntactic categories (see Norcliffe & Jaeger 2016: 171⁸).

⁸“Communicative efficiency therefore holds explanatory potential not just for patterns of real-time language use, but also for the shape of grammars” (Norcliffe & Jaeger 2016: 171).

5.2 Diachronic change as a possible explanatory factor

Yet a different type of explanatory account might propose that the diachronic origins of the relevant patterns give rise to the observed cross-linguistic distributions (see Cristofaro 2017, and Cristofaro 2019 [this volume]). The claim would be that the kinds of sources and diachronic pathways that bring about the observed patterns are tightly constrained (mutational constraints, see Haspelmath 2019 [this volume]) and, crucially, that the coding asymmetry is a direct but incidental result of how independent possessive person-forms emerge from their respective sources.

The strongest argument against such a possible claim, and for an interpretation of the data in terms of a functional-adaptive, result-oriented approach, is the fact that we see convergence of multiple sources and pathways toward a UNIFORM outcome. In particular, the asymmetric coding can come about through shortening or through lengthening. If there were no overarching functional constraint, we would expect many more counter-examples in the data, i.e. cases where the dependent possessive person-forms are longer than the independent ones, such as dependent **mine book* vs. independent **my* ‘mine’, or German dependent **mein-iges Buch* ‘my book’ vs. independent **mein* ‘mine’, or Jamaican dependent **fi-mi buk* ‘my book’ vs. independent **mi* ‘mine’. But this is not what we find.

The creole data make clear that there is a surprisingly large array of source constructions which enter the pool of possible dependent and independent possessives. Many of these source constructions had different communicative functions when they were first grammaticalized. The use of a dummy noun ‘part’, for instance, which is the source of current Haitian Creole independent possessive *pa m nan* ‘mine’, may have started out as a predicative focus construction, such as ‘this is MY part’. This focussing function is still present in constructions like in example (2). But at some point, the morpheme *pa* got refunctionalized into the phrase *pa m nan*, which eventually got grammaticalized into the independent possessive person-form ‘mine’. How did this happen? I assume that speakers must have somehow felt that they needed a more elaborate, more fully marked form to convey to hearers that a less predicatable meaning (independent possessive) was expressed. Therefore they chose (elements of) an already existing construction, here the focus construction, and through a kind of inflationary overuse grammaticalized it into the independent possessive form *pa m nan*, where the morpheme *pa* does not have the meaning ‘part’ anymore. It is only at this moment that speakers created a grammatical opposition between a dependent and an independent possessive form.

Another source of a longer independent possessive person-form is the use of a preposition ‘of/for’ together with a possessive person form ‘my/I’, yielding complex forms, such as ‘of my’ or ‘for me’, as seen in the Jamaican independent possessive form *fi-mi* ‘mine’ (vs. dependent possessive *mi* [1SG.POSS] ‘my’/[1SG] ‘I’, already cited above). Forms like *fi-mi* may go back to a kind of predicative construction, such as ‘this is for me/this is of my’. But here again, at some point in time, the creators of Jamaican refunctionalized the chunk *fi-mi* to fit the need to highlight the more unusual, less predicatable independent possessive meaning ‘mine’.

In this context, another fact makes a source-oriented account less convincing. Quite a few creole languages show lengthened forms, such as *fi-mi*, not only in the independent, but also in the dependent possessive person-form, as for instance in Zamboanga Chabacano *dimíyo* (‘of.1SG’) ‘my/mine’ or in Tok Pisin *bilong mi* (of.1SG) ‘my/mine’. This is the situation where there is no length difference in both forms, as illustrated for Mandarin Chinese in §2 (identical pattern in Table 1). If a hypothesized predicative construction were the source of the independent possessive person-forms, it certainly cannot be the source for the dependent form. Therefore, here we must allow for some kind of analogical extension to the dependent forms. Interestingly, it is only in the dependent possessive function that *dimíyo* can be shortened to *mí* (Steinkrüger 2013), thus again giving rise to a new coding asymmetry in the predicted direction: the independent possessive form *dimíyo* ‘mine’ is longer than the dependent possessive form *mí* (similar to English *mine/my* and Juba Arabic *bitai/tái*).

Coming back to both lengthening scenarios of independent possessive forms described above: The crucial point here is the fact that the change process from a focus or predicative construction to an independent possessive form should not be seen as a self-propelling grammaticalization process, but as a result of speakers’ unconscious choices to communicate efficiently by highlighting the less predicatable meaning, thus ultimately bringing about functionally adapted linguistic structures. In other words: If speakers did not sense the communicative need to mark independent possessives with more linguistic material, they would not drag parts of a focus or predicative construction into an emergent independent possessive person-form in the first place.

Therefore, speaking of SYNCHRONIC “lengthening” strategies in independent possessive forms, as I have done in the previous sections, could be misinterpreted. What generations of speakers really do while communicating is recruiting ALREADY EXISTING structures (lexical or grammatical) to fit new grammatical functions (parts of old focus constructions and old predicative constructions

are used to express new independent possessive forms). Linguists subscribing to the source-oriented approach would probably completely agree with this statement. But, as I laid out in the preceding paragraph, there is a second part to this story, where mere persistence accounts fail to explain the data: While recruiting existing structures for new grammatical functions, speakers unconsciously comply with the efficiency principle. As a result of the cumulative individual speech acts, we observe ever changing functionally adapted structures, which overwhelmingly point into the SAME direction: rarer, less predicatable meanings tend to be coded with longer forms than, or equally long forms as, the more predicatable meaning, but never with shorter forms.

Moreover, the examples of Haitian Creole *pa* and Jamaican *fi-mi* make clear that a functional-adaptive approach in terms of coding efficiency has no problem with the fact that the function or motivation of the source construction, here a focus or predicative construction, is different from the function at the synchronic level, here the independent possessive meaning. However, what is important is the fact that speakers always refunctionalize existing lexical or grammatical material in a predicatable way. In many cases, the newer grammatical functions that are expressed with already grammaticalized material follow quite narrow grammaticalization paths. In other more extreme cases, speakers exapt existing grammatical material to make it fit to their communicative needs, i.e. highlighting less predicatable meanings. This is the case with the erstwhile Middle English dative case form *hern* that was exapted into the independent possessive form (see §3). The mere existence of such exaptations in grammatical change supports the idea that the source constructions can be irrelevant for the synchronic grammatical patterns. But what is indeed effective in every utterance and gives rise to universal coding asymmetries is the overarching functional efficiency principle in signal coding: Spend as little energy as necessary to reach the intended goals, from which it follows that less frequent and therefore less predicatable meanings come to be coded with more material than more frequent and therefore more predicatable meanings.

Thus, creole languages help sharpen our understanding of functional-adaptive forces unfolding in situations of unusually accelerated language change.

Acknowledgments

I am grateful to Martin Haspelmath, to the co-editors of the present volume and to Eitan Grossman and Mark Dingemanse for their comments on an earlier draft of this paper. Furthermore, the support of the European Research Council (ERC

Advanced Grant 670985, Grammatical Universals) is gratefully acknowledged. This paper is closely related to a joint workshop talk with Martin Haspelmath at the SLE meeting in Naples, September 2016.

Appendix

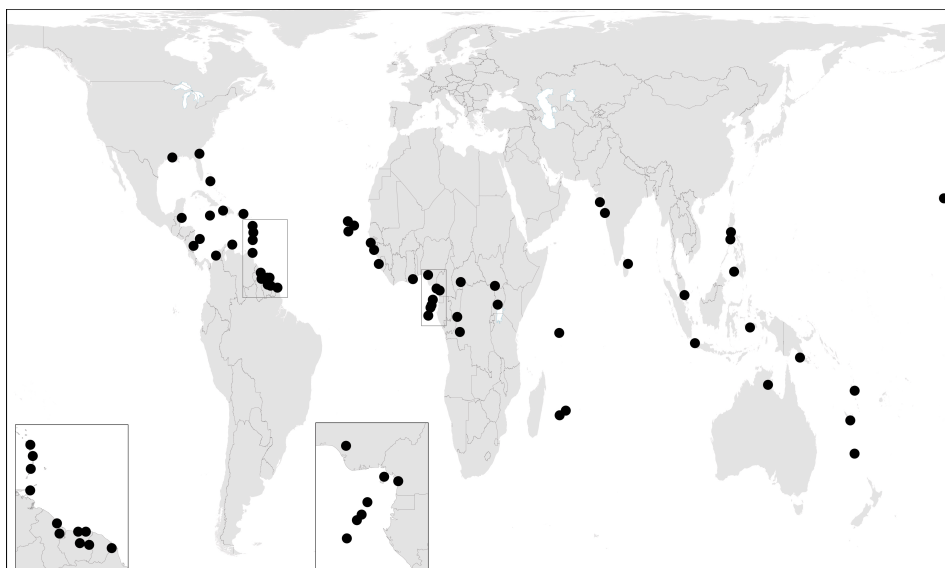


Figure 1: Distribution of the 59 creole languages in APiCS (for more information see apics-online.info) (CC BY-SA 4.0, Hans-Jörg Bibiko, MPI-SHH Jena)

References

- Allen, Cynthia L. 2002. The development of ‘strengthened’ possessive pronouns in English. *Language Sciences* 24(3). 189–211. DOI:10.1016/S0388-0001(01)00028-6
- Bentz, Christian & Bodo Winter. 2013. Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change* 3(1). 1–27. DOI:10.1163/22105832-13030105
- Bybee, Joan L. 1988. The diachronic dimension in explanation. In John A. Hawkins (ed.), *Explaining language universals*, 350–379. Oxford: Blackwell.

- Bybee, Joan L. 2007. Introduction. In Joan L. Bybee (ed.), *Frequency of use and the organization of language*, 5–22. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780195301571.001.0001
- Colot, Serge & Ralph Ludwig. 2013. Martinican creole structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/49>, accessed 2017-7-2.
- Cristofaro, Sonia. 2017. Implicational universals and dependencies. In Nick J. Enfield (ed.), *Dependencies in language: On the causal ontology of linguistic systems*, 9–22. Berlin: Language Science Press. DOI:10.5281/zenodo.573777
- Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 25–46. Berlin: Language Science Press. DOI:10.5281/zenodo.2583806
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- Cyffer, Norbert. 1998. *A sketch of Kanuri*. Cologne: Rüdiger Köppe.
- Fattier, Dominique. 2013. Haitian structure data set. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/49>, accessed 2017-7-2.
- Finney, Malcolm Awadajin. 2013. Krio structure data set. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/15>, accessed 2017-7-1.
- Hagemeyer, Tjerk. 2013. Santome structure data set. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/35>, accessed 2017-7-1.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin, New York: Mouton de Gruyter. DOI:10.1515/9783110884210

- Haspelmath, Martin. 2008. Creating economical morphosyntactic patterns in language change. In Jeff Good (ed.), *Language universals and language change*, 185–214. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199298495.003.0008
- Haspelmath, Martin. 2015. A grammatical overview of Egyptian and Coptic. In Eitan Grossman, Martin Haspelmath & Tonio Sebastian Richter (eds.), *Egyptian-Coptic linguistics in typological perspective*, 103–143. Berlin, New York: De Gruyter Mouton. DOI:10.1515/9783110346510.103
- Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by constraints on change? In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 1–23. Berlin: Language Science Press. DOI:10.5281/zenodo.2583804
- Haspelmath, Martin & APiCS Consortium. 2013. Alignment of case marking of personal pronouns. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The atlas of pidgin and creole language structures*, 232–235. Oxford: Oxford University Press.
- Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog & Elif Bamyacı. 2014. Coding causal–noncausal verb alternations: A form–frequency correspondence explanation. *Journal of Linguistics* 50(3). 587–625. DOI:10.1017/S0022226714000255
- Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.
- Haspelmath, Martin & Susanne Maria Michaelis. 2017. Analytic and synthetic: Typological change in varieties of European languages. In Isabelle Buchstaller & Beat Siebenhaar (eds.), *Language variation – European perspectives VI: Selected papers from the 8th International Conference on Language Variation in Europe (ICLaVE 8), Leipzig 2015*, 3–22. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/silv.19.01has
- Jaeger, T. Florian & Esteban Buz. 2018. Signal reduction and linguistic encoding. In Eva M. Fernández & Helen Smith Cairns (eds.), *The handbook of psycholinguistics*, 38–81. Oxford: Wiley-Blackwell. DOI:10.1002/9781118829516.ch3
- Koch, Peter & Wulf Oesterreicher. 2012[1985]. Language of immediacy – language of distance: Orality and literacy from the perspective of language theory and linguistic history. In Claudia Lange, Beatrix Weber Weber & Göran Wolf (eds.), *Communicative spaces: Variation, contact, and change. Papers in honour of Ursula Schaefer*, 441–473. Frankfurt/Main: Peter Lang. DOI:10.15496/publikation-20415

- Kortmann, Bernd & Kerstin Lunkenheimer (eds.). 2013. *The electronic world atlas of varieties of English*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://ewave-atlas.org/>.
- Kouwenberg, Silvia. 2013. Berbice Dutch. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages. Volume 1: English-based and Dutch-based languages*, 275–284. Oxford: Oxford University Press.
- Lass, Roger. 1990. How to do things with junk: Exaptation in language evolution. *Journal of Linguistics* 26(1). 79–102. DOI:10.1017/S0022226700014432
- Lass, Roger. 2017. Review of *Exaptation and language change*, eds. Muriel Norde & Freek Van de Velde, Amsterdam: John Benjamins, 2016. *Diachronica* 34(1). 117–126. DOI:10.1075/dia.34.1.06las
- Luffin, Xavier. 2013. Kinubi structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/63>, accessed 2017-7-2.
- Lupyan, Gary & Rick Dale. 2010. Language structure is partly determined by social structure. *PLoS ONE* 5(1). e8559. DOI:10.1371/journal.pone.0008559
- Manfredi, Stefano & Sara Petrollino. 2013. Juba Arabic structure data set. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/64>, accessed 2017-7-2.
- Maurer, Philippe. 2013. Angolar structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/36>, accessed 2017-7-2.
- Meyer, Julien. 2005. *Description typologique et intelligibilité des langues sifflées, approche linguistique et bioacoustique*. Lyon: Université Lumière Lyon 2 (Doctoral dissertation).
- Meyer, Julien. 2008. Typology and acoustic strategies of whistled languages: Phonetic comparison and perceptual cues of whistled vowels. *Journal of the International Phonetic Association* 38(1). 69–94. DOI:10.1017/S0025100308003277
- Meyerhoff, Miriam. 2013. Bislama structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

- tionary Anthropology. <http://apics-online.info/contributions/23>, accessed 2017-7-2.
- Michaelis, Susanne Maria & Martin Haspelmath. Forthcoming. Grammaticalization in creole languages: Accelerated functionalization and semantic imitation. In Andreij L. Malchukov & Walter Bisang (eds.), *Handbook of areal patterns of grammaticalization and cross-linguistic variation in grammaticalization scenarios*. Berlin: De Gruyter Mouton.
- Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). 2013. *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.
- Napoli, Donna Jo, Nathan Sanders & Rebecca Wright. 2014. On the linguistic effects of articulatory ease, with a focus on sign languages. *Language* 90(2). 424–456. DOI:10.1353/lan.2014.0043
- Neumann-Holzschuh, Ingrid & Thomas A. Klingler. 2013. Louisiana Creole structure data set. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/53>, accessed 2017-7-2.
- Norcliffe, Elizabeth & T. Florian Jaeger. 2016. Predicting head-marking variability in Yucatec Maya relative clause production. *Language and Cognition* 8. 167–205. DOI:10.1017/langcog.2014.39
- Norde, Muriel & Freek Van de Velde (eds.). 2016. *Exaptation and language change*. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/cilt.336
- Schwarze, Christoph. 1988. *Grammatik der italienischen Sprache*. Tübingen: Niemeyer.
- Schwegler, Armin. 2013. Palenquero structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/48>, accessed 2017-7-2.
- Seržant, Ilja A. 2019. Weak universal forces: The discriminatory function of case in differential object marking systems. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 149–178. Berlin: Language Science Press. DOI:10.5281/zenodo.2583816
- Seuren, Pieter A. M. & Herman Wekker. 1986. Semantic transparency as a factor in creole genesis. In Pieter Muysken & Norval Smith (eds.), *Substrata versus*

universals in creole genesis, 57–70. Amsterdam, Philadelphia: John Benjamins.
DOI:10.1075/cll.1.05seu

- Steinkrüger, Patrick O. 2013. Zamboanga Chabacano. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages. Volume 2: Portuguese-based, Spanish-based and French-based languages*, 156–162. Oxford: Oxford University Press.
- Yakpo, Kofi. 2013. Pichi structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/19>, accessed 2017-7-2.
- Ye, Jingting. 2017. *Independent and dependent possessive person forms: A cross-linguistic study*. Paper presented at the Diversity linguistics seminar, Leipzig University.

