Explanation in typology

Diachronic sources, functional motivations and the nature of the evidence

Edited by

Karsten Schmidtke-Bode

Natalia Levshina

Susanne Michaelis

Ilja Seržant

Conceptual Foundations of Language Science

language science press

Conceptual Foundations of Language Science

No scientific work proceeds without conceptual foundations. In language science, our concepts about language determine our assumptions, direct our attention, and guide our hypotheses and our reasoning. Only with clarity about conceptual foundations can we pose coherent research questions, design critical experiments, and collect crucial data. This series publishes short and accessible books that explore well-defined topics in the conceptual foundations of language science. The series provides a venue for conceptual arguments and explorations that do not require the traditional book-length treatment, yet that demand more space than a typical journal article allows.

In this series:

1. N. J. Enfield. *Natural causes of language*.

ISSN: 2363-877X

Explanation in typology

Diachronic sources, functional motivations and the nature of the evidence

Edited by

Karsten Schmidtke-Bode

Natalia Levshina

Susanne Michaelis

Ilja Seržant

language
science
press

Karsten Schmidtke-Bode , Natalia Levshina , Susanne Michaelis & Ilja Seržant (eds.). 2018. *Explanation in typology*: *Diachronic sources, functional motivations and the nature of the evidence* (Conceptual Foundations of Language Science). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/000

ISBN: no digital ISBN
no print ISBNs!
ISSN: 2363-877X
no DOI
ID not assigned!

Cover and concept of design: Ulrike Harbort
Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

Language Science Press
Unter den Linden 6
10099 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

no logo

Contents

Introduction

Karsten Schmidtke-Bode
Leipzig University

The present volume addresses a foundational issue in linguistic typology and language science more generally. It concerns the kinds of explanation that typologists provide for the cross-linguistic generalizations they uncover, i.e. for so-called universals of language. The universals at issue here are usually probabilistic statements about the distribution of specific structures, such as the classic Greenbergian generalizations about word order and morphological markedness patterns. Some examples are given in (1)–(4) below:

(1) With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional. (**Greenberg1963**)

(2) A language never has more gender categories in nonsingular numbers than in the singular. (**Greenberg1963**)

(3) If a language uses an overt inflection for the singular, then it also uses an overt inflection for the plural. (**Croft2003**: 89, based on **Greenberg1966**: 28)

(4) In their historical evolution, languages are more likely to maintain and develop non-ergative case-marking systems (treating S and A alike) than ergative case-marking systems (splitting S and A). (**BickelEtAl2015**: 5)

As can be seen from these examples, cross-linguistic generalizations of this kind may be formulated in terms of preferred types in synchronic samples or in terms of higher transition probabilities for these types in diachronic change (cf. also **Greenberg1978**; **Maslova2000**; **Cysouw2011**; **Bickel2013** for discussion

of the latter approach). But this is, strictly speaking, independent of the question we are primarily concerned with here, namely how to best account for such generalizations once they have been established.

The most widespread typological approach to explanation is grounded in functional properties of the preferred structural types: For example, typical correlations in the ordering of different types of phrases (e.g. object–verb and NP–postposition) have been argued to allow efficient online processing (e.g. **Hawkins1994**; 2004). Markedness patterns in morphology (e.g. the distribution of zero expression in case, number or person systems) have been attributed to economy, i.e. the desire to leave the most frequent and hence most predictable constellations unexpressed, or rather to a competition between economy and the motivation to code all semantic distinctions explicitly (e.g. **Haiman1983**; **Comrie1989**; **Aissen2003**; **Croft2003**; **Haspelmath2008**; among many others). The general idea behind this approach is thus that speech communities around the world are subject to the same kinds of cognitive and communicative pressures, and that the languages they speak tend to develop structures that respond to these pressures accordingly, or, as **Bickel2014** puts it, "in such a way as to fit into the natural and social eco-system of speakers: that they are easy to process, that they map easily to patterns in nonlinguistic cognition, and that they match the social and communicative needs of speakers."

There is a clear parallel to evolutionary biology here, in that languages are said to *converge* on similar structural solutions under the same functional pressures, just like unrelated species tend to develop similar morphological shapes in order to be optimally adapted to the specific environment they co-inhabit (**Deacon1997**; **Caldwell2008**; **EvansLevinson2009**; **Givón2010**). When applied to language, this line of explanation at least implicitly invokes what is known as "attractor states", i.e. patterns of structural organization that languages are drawn into in their course of development.[1] For this reason, one could also speak of a **result-oriented** approach to explanation.

There is, however, another way of looking at the same patterns, one that redirects attention from the functional properties to the diachronic origins of the linguistic structures in question. On this view, many universal tendencies of order and coding are seen as by-products, as it were, of recurrent processes of morphosyntactic change, notably grammaticalization, but without being adap-

[1] The term attractor state (or basin of attraction) is adopted from the theory of complex dynamic systems (cf., e.g., **Cooper1999**; **HoweLewis2005**; **Holland2006**), which has become increasingly popular as a way of viewing linguistic systems as well (cf. **BecknerEtAl2009** and **Port2009** for general overviews, and **Haig2018** (among others) for its application to typological data).

tive in the above sense: There is no principled convergence on similar structural traits because these traits might be beneficial from the perspective of processing, iconicity or economical communicative behaviour. Instead, the current synchronic distributions are argued to be long-term reflections of individual diachronic trajectories, in particular the diachronic sources from which the structures in question originate. **Givón1984** and **Aristar1991**, for example, suggested that certain word-order correlations may simply be a consequence of a given ordering pair (e.g. Gen–N & Rel–N, or V–O & Aux–V) being directly related diachronically: Auxiliaries normally grammaticalize from main verbs that take other verbs as complements, and since these complements follow the verb in VO languages, they also follow the auxiliary in the resulting Aux–V construction; the mirror-image pattern holds for OV languages (cf. also **Lehmann1986**: 12–13). If this line of reasoning extends to most other word-order pairs, there is no need to motivate the synchronic correlations in functional-adaptive terms, e.g. by saying that the correlations arise *in order to* facilitate efficient sentence processing.

In the domain of morphology, **Garrett1990** argued that patterns in case marking, specifically of differential ergative marking, are exhaustively explained by the properties of the source of the ergative marker: When ergative case arises from the reanalysis of instrumental case, the original characteristics of the latter, such as a restriction to inanimate referents, are directly bequeathed to the former. The result is a pattern in which animate A-arguments are left unmarked, but since this is a direct "persistence effect" (**Hopper1991**) of the history of the ergative marker, there is again no need for an additional functional-adaptive explanation in terms of other principles, such as a drive for economical coding patterns. Rather than being result-oriented, then, this way of explaining universals can be characterized as **source-oriented**.

Such source-oriented explanations thus move away from attractor states of grammatical organization and often emphasize the importance of "attractor trajectories" instead (**BybeeBeckner2015**: 185): In some domains of grammar, the patterns of reanalysis and ensuing grammaticalization are so strikingly similar across the world's languages that it is not surprising that they yield similar outcomes, such as strong correlations between V–O & Aux–V or V–O & P–NP ordering. In other cases, it is argued that many individual, and partly very different, diachronies are capable of producing a uniform result, but without any consistent functional force driving these trajectories. **Cristofaro2017**, for instance, claims that this is the case for plural markers: An initial system without number marking can develop an overt plural morpheme from many different sources – usually by contextual reanalysis – and thus ultimately come to contrast a zero singular

with an overt plural, but these developments are neither triggered nor further orchestrated by a need for economical coding: They do not happen to keep the (generally more frequent) singular unmarked and the (generally less frequent) plural overtly signalled.

In other words, whether the individual diachronic trajectories are highly similar or rather diverse, the premise of the source-oriented approach is that they can scale up to produce a predominant structural pattern in synchronic samples. Hence they obviate the need for highly general functional principles tying these patterns together.

While the source-oriented approach was still a more marginal position in previous volumes on explaining language universals (e.g. **Hawkins1988a**; **Good2008**), it has gained considerable ground over the last decade, notably in a series of articles by Cristofaro (e.g. **Cristofaro2012**; 2014; 2017) but also in other publications (e.g. **Creissels2008**; **GildeaZúñiga2016**). Moreover, while the basic thrust of the two explanatory approaches is straightforward, clarification is needed on a number of – equally fundamental – details. After all, both approaches are functionalist in nature, as they rely on domain-general mechanisms (cf. **Bybee2010**) to explain the emergence of language structure and linguistic universals; and in both approaches, these mechanisms constrain how languages "evolve into the variation states to which implicational and distributional universals refer" (**Hawkins1988b**). But as **Plank2007** notes, "what is supposed to be the essence and force of diachronic constraints would merit livelier discussion." It is the goal of the present book to offer precisely a discussion of this kind.

The volume begins with a programmatic paper by **Martin Haspelmath** on what it means to explain a universal in diachronic terms. He aims to clarify how diachrony is involved in result-oriented and source-oriented accounts, respectively, and thus lays out a general conceptual framework for the explanation of universals. At the same time, Haspelmath opens the floor for debating the strengths and weaknesses of the two explanatory accounts at issue here. His own position is that, in many cases, current source-oriented explanations are ill-equipped to truly explain the phenomena they intend to account for, and hence cannot replace result-oriented motivations. Haspelmath's arguments for this position, as well as his terminological proposals, provide a frame of reference to which all other contributions respond in one way or another.

The lead article is followed by two endorsements of source-oriented explanations, articulated by **Sonia Cristofaro** and **Jeremy Collins**, respectively. They both describe the approach in widely accessible terms, allowing also readers outside of linguistic typology to appreciate the general argument as well as the

specific examples discussed. The phenomena themselves involve domains that are particularly well-known for being explained in functional-adaptive terms, namely differential argument marking, number marking and word-order correlations, and these are all argued to be best captured by persistence effects from their respective diachronic origins.

We then proceed to papers that allow for progressively more room for functional-adaptive motivations and, importantly, for methodological discussions on how to obtain evidence for such pressures. Accordingly, all of these papers adduce novel empirical data and discuss them in light of the present debate.

Matthew Dryer's paper is an immediate follow-up on Collins' discussion of word-order correlations. On the one hand, Dryer argues that the various correlates of adposition–noun ordering (e.g. OV and NP–P, and Gen–N and NP–P) are, indeed, best accounted for in source-oriented terms. In particular, only this approach proves capable of explaining the occurrence (and the individual semantic types) of both prepositions and postpositions in SVO languages. On the other hand, however, Dryer contends that there are some significant correlations for which a source-based account either fails to offer an explanation or else makes the opposite prediction of the patterns we find synchronically. Dryer concludes, therefore, that neither a purely source-based nor a purely result-based explanation is sufficient to deal with word-order correlations.

In a similar fashion as Dryer's paper, **Holger Diessel**'s article demonstrates that different aspects of the same grammatical domain – in this case adverbial clause combinations – are amenable to different types of explanation. Whereas some typological characteristics of the structure and position of adverbial clauses can successfully be accounted for by their source constructions (and thereby supplant earlier processing-based explanations), others are still best rendered in functional-adaptive (and hence result-oriented) terms.

Karsten Schmidtke-Bode offers a review of Hawkins' (2004, 2014) research programme of "processing typology", examining the plausibility of Hawkins' functional-adaptive ideas in diachronic perspective. On a theoretical level, it is argued that a predilection for efficient information processing is operative mostly at the diffusion stage of language change, regardless of the source from which the respective constructions originate. On a methodological level, the paper proposes that the cross-linguistic predictions of Hawkins' programme can be tested more rigorously than hitherto by combining static and dynamic statistical models of large typological data sets; this is demonstrated in a case study on the distribution of article morphemes in VO- and OV-languages, respectively.

An important methodological point is also made by **Ilja A. Seržant**, who claims

that functional-adaptive pressures may not actually surface in standard typological analyses but do become visible in qualitative data from transition phases. Based on diachronic data from Russian, he shows how the development of differential object marking was crucially influenced by considerations of ambiguity avoidance (and hence a classic functional-adaptive motivation), over and above the constraints inherited from the source construction. In the absence of such longitudinal data, transition phases can be identified on the basis of synchronic variability, and Seržant shows that a wide variety of languages currently exhibit variation in differential object marking that mirrors the diachronic findings from Russian, and that is not predictable from the source meaning of the marker in question.

Susanne Maria Michaelis adds another source of data to the debate at hand. She argues that creole languages provide a unique window onto the relationship between synchronic grammatical patterns and their diachronic trajectories, as the latter are often relatively recent and also accelerated when compared to normal rates of grammatical change. The developments are, consequently, more directly accessible and less opaque than in many other cases. By inspecting creole data on possessive forms in attributive and referential function (e.g. *your* versus *yours*), Michaelis finds evidence for the development of the same kinds of coding asymmetries that this domain offers in non-contact languages around the world. She proposes that the data are indicative of result-oriented forces that drive diverse diachronic pathways towards the same synchronic outcome. This stance contrasts most explicitly with Cristofaro's, who interprets such situations in exactly the opposite way (i.e. as providing evidence *against* a unifying functional explanation).

Natalia Levshina, finally, adopts an entirely different methodological approach to illuminate the present discussion: In her paper, she showcases the paradigm of artificial language learning, which can be employed to inspect whether users of such newly acquired languages develop performance biases that are in keeping with hypothesized functional principles, such as an increasingly efficient distribution of morphological marking. Her case study clearly demonstrates such biases and discusses where they may ultimately come from, i.e. how they fit into the new conceptual framework of constraints offered by Haspelmath's position paper.

The volume is rounded off by a brief **epilogue** in which **Karsten Schmidtke-Bode** and **Eitan Grossman** summarize and further contextualize the arguments put forward by the contributors.

Overall, the purpose of the present book is to provide a state-of-the-art overview

of the general tension between source- and result-oriented explanations in linguistic typology, and specifically of the kinds of arguments and data sources that are (or can be) brought to bear on the issue. It should be made clear from the outset that the two types of explanation are framed as antagonistic here even though in most cases, an element of both will be needed in order to fully account for a given grammatical domain. As we emphasize in the epilogue, the diachronic source of a grammatical construction certainly constrains its further development, but the major issue at stake here is the extent to which result-oriented, functional-adaptive motivations enter these developments as well. By the end of the day, universals of language structure will thus differ in the *degree* to which they are shaped by such adaptive pressures.

Acknowledgements

The present volume originated in the context of the project *Form-frequency correspondences in grammar* at Leipzig University. The support of the European Research Council (ERC Advanced Grant 670985, Grammatical Universals) is gratefully acknowledged. An oral precursor to this volume was a workshop on the topic at the 49th Annual Conference of the Societas Linguistica Europaea in Naples in 2016, co-organized by the editors of this book. We would like to thank the participants and the audience of that workshop for insightful contributions and discussion. We would also like to thank Eitan Grossman and Mark Dingemanse for extensive feedback on all papers in the present volume. Finally, we are grateful to Sebastian Nordhoff and his team at Language Science Press as well as the participants of Language Science Press's community proofreading.

Chapter 1

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 (**HaspelmathEtAl2014**). 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 predictable 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 (**KochOesterreicher2012** [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, **LupyanDale2010** analyzed data from the *World atlas of language structures* (**HaspelmathEtAl2005**) 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, **BentzWinter2013** 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 (**Meyer2005**; 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 **Haspelmath2018** [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 predictable meanings.

Such functional-adaptive explanations have a diachronic component (**Bybee1988**): 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 dependent/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 2018 [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 **Haspelmath2018** [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. c. Is this bike yours?

In a recent study, **Ye2017**[1] has found that in the world's languages independent possessive person-forms like English *mine*, French *le mien* 'mine', and Mandarin *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 *wo de* 'my'. Coding length here refers to the number of segments in the signal, or possibly to the amount of biomechanical effort (see **NapoliEtAl2014** 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 *wo de* [I GEN].

Table **??** 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

[1] **Ye2017** analyzes a sample of 69 genealogically and areally unrelated languages.

person-form, such as in Italian *la mia* (with kinship terms like sorella 'sister')[2]. 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	*wo de shu* I GEN book 'my book'	*wo de* I GEN 'mine'	
additional marker	Lezgian	*zi ktab* I.GEN book 'my book'	*zi-di* I.GEN-SUBST 'mine'	**Haspelmath1993**
additional stem	Kanuri	*fewá-ndé* COW-1PL.POSS 'our cows'	*kaá-nde* INDEP-1PL 'ours'	Cyffer (1998:31f.)
additional article	Italian	*mia sorella* 'my sister'	*la mia* mine	Schwarze (1988:44, 286f.)
longer form	Coptic	*p-ek-ran* ART-2SG-name 'your name'	*p-ô-k* ART-INDEP-2SG 'yours'	**Haspelmath2015**

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 shorter[3].

Now the claim is that these coding asymmetries reflect asymmetries of frequency of use. More frequent meanings (here: dependent possessives) are more

[2] If 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'.

[3] See also **Croft1991**, 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 (1991:51), e.g. marked predicative nominals vs. unmarked nouns, or marked predicative adjectives vs. unmarked attributive adjectives.

predictable 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 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 Chinese[4]) 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 counter-parts. Table **??** shows data from British English.

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

Dependent	**Token frequency**	**Independent**	**Token frequency**
my	145,250	*mine*	6,067
your	132,598	*yours*	4,059
our	92,314	*ours*	1,658
their	251,410	*theirs*	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 **Ye2017**). 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.

[4]For frequency counts in Korean and Mandarin Chinese, see **Ye2017**.lingler & ructure dataset.
t. ellemmatiken, bin aber leider nicht fündig geworden. Jena schon alle verfügbaren Chinesisch-Gramm Online.
iel (eds.). 2017. ctable meaning are coded with longer forms.ndruck, wie die Aysmmetrien so auss

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 predictable meanings need more coding and more predictable 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 (**Bybee2007**), or the rarer member of the grammatical opposition is lengthened[5] (**Haspelmath2008**). 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 shortenend 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 **??** 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 from *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 suffx *-ig,* which occurs in other derived adjectives (like *selb-ig* 'same', *bärt-ig* 'bearded', *ehrgeiz-ig* 'ambitious'). As we see in Tables 3a and 3b, the array of source constructions and diachronic pathways which give rise to longer independent possessive person-forms is very diverse.

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 re-

[5]Here, 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.

Language	Strategy	Dependent form	Independent form
English	phonological reduction of dependent form	*my*	*mine*

Table ??a: Shortened dependent form

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

Table ??b: Lengthened independent form

cruiting 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 **KortmannLunkenheimer2011**) go back 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 **Allen2002**, and for the notion of exaptation, see **Lass1990**, 2017, Norde & Van de **Velde2016** 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*, **MichaelisEtAl2013**, 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 (**HaspelmathMichaelis2017**; **MichaelisHaspelmath2018**). 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?

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 and APiCS **Consortium2013** 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 **??**.

The following Table **??**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.

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 **??** 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 **??**). These languages do not contradict

[6] See already **SeurenWekker1986** for the notion of transparency in the creolization process. Findest Du das zu redundant zu dem schon Gesagten?tegy angeht, aber es gibt halt nur diese eine strategy.ation.of the definite

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

Table 3: Dependent and independent possessive person-forms in some creole languages

Creole language	Dependent form	Independent form
Bislama	*blong yu*	*blong yu*
(Meyerhoff2013)	[POSS 2SG] 'your'	[POSS 2SG] 'yours'
Kinubi	*tá-i*	*tá-i*
(Luffin2013)	[POSS-1SG] 'my'	[POSS-1SG] 'mine'
Batavia Creole	*minya*	*minya sua*
(Maurer2013)	[1SG.POSS] 'my'	[1SG.POSS POSS] 'mine'
Martinican Creole	*-mwen*	*ta mwen*
(ColotLudwig2013)	[1SG.POSS] 'my'	[POSS 1SG.POSS] 'mine'
Pichi	*yù*	*yù yon*
(Yakpo2013)	[2SG.POSS], [2sg] 'you'	[2SG.POSS own] 'yours'
Palenquero	*mi*	*ri mi*
(Schwegler2013)	[1SG.POSS] 'my'	[of 1SG.POSS] 'mine'

Table 4: 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

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' (**ManfrediPetrollino2013**), 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 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 (**Hagemeijer2013**), 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 (**Fattier2013**)
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' (**Fattier2013**). 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 construc-

tion. 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 feature 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') (**Finney2013**).

There is a fourth source of independent forms involving a general (adjectival) nominalizer, such as 'one'. In Berbice Dutch, there is a general nominalizer *-jɛ* which is added to the personal pronoun *ɛkɛ* [1SG.POSS]/[1SG] 'my' ('I'), resulting in *ɛkɛ-jɛ* [1SG.POSS-NMLZ] 'mine' (see Table **??**b). This nominalizer goes back to Eastern Ijo, the substrate language of Berbice Dutch, where it has singular non-human reference, whereas in Berbice Dutch it has grammaticalized into a generic nominalizer (**Kouwenberg2013**).

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-**HolzschuhKlingler2013**, 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 ex-

plained by a functional-adaptive constraint of coding efficiency: More frequently expressed meanings (dependent possessives) need less costly signal encoding because they are highly predictable, whereas less frequently expressed meanings need more robust signal encoding because they are less predictable (**Haspelmath2018** [this volume]; see **NorcliffeJaeger2016** and **JaegerBuz2018** 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 **Croft1991**). 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

and thousands of years of speakers performing communicatively efficient speech acts in marking the less predictable 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 predictable meanings by more morphosyntactic material, which accumulated over generations and eventually contributes to the shaping of syntactic categories (see **NorcliffeJaeger2016**:171[8]).

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 **Cristofaro2017**, and **Cristofaro2018** [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 **Haspelmath2018** [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 Haitain 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

[8] "Communicative efficiency therefore holds explanatory potential not just for patterns of real-time language use, but also for the shape of grammars." (**NorcliffeJaeger2016**:171).

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 predictable 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 predictable 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 **??**). 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* (**Steinkrüger2013**), 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 predictable 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 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 predictable meanings tend to be coded with longer forms than, or equally long forms as, the more predictable 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 predictable 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 predictable 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 predictable meanings come to be coded with more material than more frequent and therefore more

predictable 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 an external reviewer 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

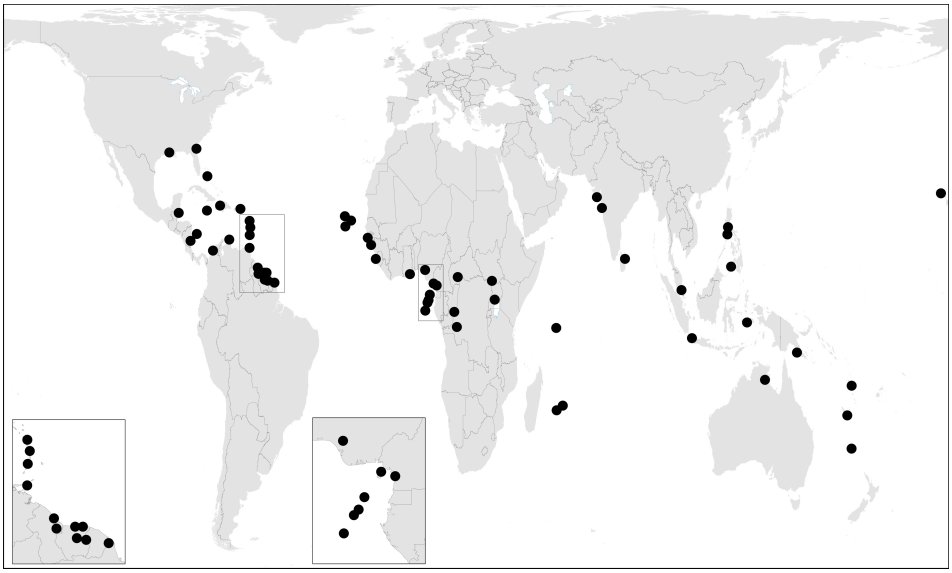

Figure 1: Distribution of the 59 creole languages in APiCS (for more information see apics-online.info) (CC BY-SA 4.0, Hans-Jörg Bibiko, MPI-SHH/Jena)

```
%%move bib entries to  localbibliography.bib
@article{Allen2002,
author = {Allen, Cynthia L.},
journal = {\textit{Language Sciences}},
```

pages = {189–211},
sortname = {Allen, Cynthia L.},
title = {The development of ‘strengthened’ possessive pronouns in {English}},
volume = {24.},
year = {2002}
}

@misc{Bentz2013,
author = {Bentz, Christian and Bodo Winter},
note = {Languages with more second language learners},
year = {2013}
}

tend to lose nominal case. \textit{Language Dynamics and Change} 3. 1–
27.

Bybee, Joan L. 2007. Introduction. In: Joan L. Bybee (ed.), \textit{Frequency of u
22. Oxford: Oxford University Press.

@incollection{Bybee2006,
address = {Cambridge},
author = {Bybee, Joan L.},
booktitle = {\textit{Linguistic universals}},
editor = {Ricardo Mairal and Juana Gil},
pages = {179–194},
publisher = {Cambridge University Press},
sortname = {Bybee, Joan L.},
title = {Language change and universals},
year = {2006}
}

@incollection{Bybee1988,
address = {Oxford},
author = {Bybee, Joan L.},
booktitle = {\textit{Explaining language universals}},

```
editor = {John A. Hawkins},
pages = {350–379},
publisher = {Blackwell},
sortname = {Bybee, Joan L.},
title = {The diachronic dimension in explanation},
year = {1988}
}


Colot, Serge \& Ludwig, Ralph. 2013. Martinican Creole structure dataset. In Sus

@incollection{Cristofaro2017,
address = {Berlin},
author = {Cristofaro, Sonia},
booktitle = {\textit{Dependencies in language}},
editor = {Nick J. Enfield},
pages = {9–23},
publisher = {Language Science Press},
title = {Implicational universals and dependencies},
year = {2017}
}


Cristofaro, Sonia. \#2018 [this volume]. Taking diachronic evidence seriously: R
\#\#. Berlin: Language Science Press.

@book{Croft1991,
address = {Chicago},
author = {Croft, William},
publisher = {The University of Chicago Press},
title = {\textit{Syntactic categories and grammatical relations: The cognitive o
year = {1991}
}


@book{Cyffer1998,
address = {Cologne},
author = {Cyffer, Norbert},
```

publisher = {Köppe},
title = {\textit{A sketch of {Kanuri}}},
year = {1998}
}

Fattier, Dominique. 2013. Haitian structure data set. In Susanne Maria Michaelis,

Finney, Malcolm Awadajin. 2013. Krio structure dataset. In Susanne Maria Michaelis

Hagemeijer, Tjerk. 2013. Santome structure dataset. In Susanne Maria Michaelis, Ph

@book{Haspelmath1993,
address = {(Mouton Grammar Library 9). Berlin},
author = {Haspelmath, Martin},
publisher = {Mouton de Gruyter},
title = {\textit{A grammar of {Lezgian}}},
year = {1993}
}

@incollection{Haspelmath2008,
address = {Oxford},
author = {Haspelmath, Martin},
booktitle = {\textit{Language universals and language change}},
editor = {Jeff Good},
pages = {185--214},
publisher = {Oxford University Press},
title = {Creating economical morphosyntactic patterns in language change},
year = {2008}
}

@incollection{Haspelmath2015,
address = {Berlin},
author = {Haspelmath, Martin},
booktitle = {\textit{{Egyptian}-Coptic linguistics in typological perspective}},
editor = {Eitan Grossman, Martin Haspelmath and Tonio Sebastian Richter},

pages = {103–143},
publisher = {De Gruyter Mouton},
title = {A grammatical overview of {Egyptian} and Coptic},
year = {2015}
}

Haspelmath, Martin. \#2018. \#Can cross-linguistic regularities be explained by
\#\#. Berlin: Language Science Press.

Haspelmath, Martin \& APiCS Consortium. 2013. Independent pronominal possessors.
3. http://apics-online.info/parameters/39.chapter.html.

@article{Haspelmath2014,
author = {Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog and
journal = {\textit{Journal of Linguistics}},
number = {3},
pages = {587–625},
sortname = {Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog a
title = {Coding causal-noncausal verb alternations: {{A}} form-frequency corresp
volume = {50},
year = {2014}
}

@incollection{Haspelmath2017,
address = {Amsterdam},
author = {Haspelmath, Martin and Susanne Maria Michaelis},
booktitle = {\textit{{{Europe}an} Perspectives {VI}: Selected Papers from the 8t
editor = {Isabelle Buchstaller and Beat Siebenhaar},
pages = {3–22},
publisher = {Benjamins},
title = {Analytic and synthetic: {{T}}ypological change in varieties of {Europea
year = {2017}
}

@misc{Jaeger2018,

author = {Jaeger, T. Florian and Esteban Buz},
note = {Signal reduction and linguistic encoding. \textit{Handbook of psycholingui
year = {2018}
}

@incollection{Koch2012,
address = {Frankfurt/Main},
author = {Koch, Peter and Wulf Oesterreicher},
booktitle = {\textit{Communicative spaces}. \textit{Variation, contact, and change
editor = {Claudia Lange, Beatrix Weber and Göran Wolf},
pages = {441--473},
publisher = {Lang},
title = {Language of immediacy —~language of distance. {{O}}rality and Literacy fr
year = {2012}
}

Kouwenberg, Silvia. 2013. Berbice Dutch. In Susanne Maria Michaelis, Philippe Maur

@book{Kortmann2013,
address = {\textit{The electronic world atlas of varieties of English.} Leipzig},
editor = {Kortmann, Bernd \ and Kerstin Lunkenheimer},
note = {//ewave-atlas.org.},
publisher = {Max Planck Institute for Evolutionary Anthropology. http:},
sortname = {Kortmann, Bernd \ and Kerstin Lunkenheimer},
title = {\biberror{no title}},
year = {2013}
}

@article{Lass1990,
author = {Lass, Roger},
journal = {\textit{Journal of Linguistics}},
number = {1},
pages = {79–102},
title = {How to do things with junk: {{{E}}}xaptation in language evolution},
volume = {26},

year = {1990}
}

@article{Lass2017,
author = {Lass, Roger},
journal = {\textit{Exaptation and language change}. Amsterdam: John Benjamins, 2
pages = {117–126},
title = {Review of Muriel Norde \& Freek {Van} de Velde, eds},
volume = {1.},
year = {2017}
}

Luffin, Xavier. 2013. Kinubi structure dataset. In Susanne Maria Michaelis, Phil

@misc{Lupyan2010,
author = {Lupyan, Gary and Rick Dale},
note = {Language structure is partly determined by social structure. \textit{PLo
year = {2010}
}

Manfredi, Stefano \& Sara Petrollino. 2013. Juba Arabic structure dataset. In Su

Maurer, Philippe. 2013. Angolar structure dataset. In Susanne Maria Michaelis, P

@article{Meyer2008,
author = {Meyer, Julien},
journal = {\textit{Journal of the International Phonetic Association}},
number = {1},
pages = {69--94},
title = {Typology and acoustic strategies of whistled languages: {{P}}honetic co
volume = {38},
year = {2008}
}

```
@book{Meyer2005,
address = {Lyon},
author = {Meyer, Julien},
publisher = {Université Lumière Lyon 2},
title = {Description typologique et intelligibilité des langues sifflées, approche
year = {2005}
}


\begin{styleHTMLPreformatted}
Meyerhoff, Miriam. 2013. Bislama structure dataset. In Susanne Maria Michaelis, Ph
\end{styleHTMLPreformatted}

\begin{styleHTMLPreformatted}
@book{Michaelis2018,
address = {In Andrej Malchukov \& Walter Bisang, \textit{Handbook of areal pattern
author = {Michaelis, Susanne Maria  and  Martin Haspelmath},
publisher = {De Gruyter Mouton},
title = {Grammaticalization in creole languages: {{A}}ccelerated functionalization
year = {2018}
}

\end{styleHTMLPreformatted}

@book{Michaelis2013,
address = {Oxford},
booktitle = {\textit{The atlas of pidgin and creole language structures}},
editor = {Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath \ and  Magn
publisher = {Oxford University Press},
sortname = {Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath \ and  Ma
title = {\textit{The atlas of pidgin and creole language structures}},
year = {2013}
}


@article{Napoli2014,
author = {Napoli, Donna Jo, Nathan Sanders  and  Rebecca Wright},
journal = {\textit{Language}},
```

```
number = {2},
pages = {424–456},
title = {On the linguistic effects of articulatory ease, with a focus on sign la
volume = {90},
year = {2014}
}


Neumann-Holzschuh, Ingrid  \& Thomas A. \citealt{Klingler2013}. Louisiana Creole

@article{Norcliffe2016,
author = {Norcliffe, Elisabeth  and  T. Florian Jaeger},
journal = {\textit{Language and Cognition}},
number = {2},
pages = {167–205},
sortname = {Norcliffe, Elisabeth  and  T. Florian Jaeger},
title = {Predicting head-marking variability in {Yucatec Maya} relative clause p
volume = {8},
year = {2016}
}


@book{Norde2016,
address = {Amsterdam},
booktitle = {\textit{Exaptation and language change}},
editor = {Norde, Muriel \ and  Freek Van de Felde},
publisher = {Benjamins},
sortname = {Norde, Muriel \ and  Freek Van de Felde},
title = {\textit{Exaptation and language change}},
year = {2016}
}


@book{Schwarze1988,
address = {Tübingen},
author = {Schwarze, Christoph},
publisher = {Niemeyer},
title = {\textit{{Grammatik} der italienischen {Sprache}}},
```

year = {1988}
}

Schwegler, Armin. 2013. Palenquero structure dataset. In Susanne Maria Michaelis, P

Seržant, Ilja. \#2018. The discriminatory function of case in differential object m
\#\#. Berlin: Language Science Press.

Steinkrüger, Patrick O. 2013. Zamboanga Chabacano. In Susanne Maria Michaelis, Phi

Yakpo, Kofi. 2013. Pichi structure dataset. In Susanne Maria Michaelis, Philippe Ma

@book{Ye2017,
address = {Independent and dependent possessive person forms},
author = {Ye, Jingting},
publisher = {\textstylegmail{A cross-linguistic study}. Talk at the Diversity lingu
title = {\biberror{no title}},
year = {2017}
}

Explanation in typology