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 (see 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 adaptive 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

[1] The term attractor state (or basin of attraction) is adopted from the theory of complex dynamic systems (e.g. **Cooper1999**; **HoweLewis2005**; **Holland2006**), which has become increasingly popular as a way of viewing linguistic systems as well (see **BecknerEtAl2009** and **Port2009** for general overviews, and **Haig2018** or **Nichols2018** for very recent applications to typological data).

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 (see 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 (**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 corre-

lations, 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. Diessel focuses specifically on the structure and development of preposed adverbial clauses and argues that some of their typological characteristics, notably the properties of their subordinating morphemes, receive a satisfactory explanation in terms of the respective source construction(s), thereby supplanting earlier processing-based explanations. On the other hand, he proposes that the position of adverbial constructions (in general) is clearly subject to a number of functional-adaptive pressures, and that these may already have affected the diachronic sources from which the current preposed adverbial clauses have grammaticalized.

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 certain functional-adaptive pressures may not actually surface in standard typological analysis because they are weak forces, clearly at work but also easily overridden by other, language-specific factors. Because of their weak nature, they may not be directly visible any-more in a synchronic type, but they can be detected in qualitative data from transition phases. Based on diachronic data from Russian, Seržant 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 syn-chronic 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

Attractor states and diachronic change in Hawkins' "Processing Typology"

Karsten Schmidtke-Bode
Leipzig University

This paper provides an assessment of John Hawkins' (2004; 2014) programme of explaining cross-linguistic regularities in terms of functional-adaptive principles of efficient information processing. In the first part of the paper, I systematize how such principles may possibly affect the diachronic development of languages, and I argue that evidence for efficient coding can be obtained primarily from the actualization process, rather than the innovation stage that is at the focus of purely source-based approaches to explaining universals. In the second part of the paper, I present a small case study on a specific prediction made in **Hawkins2014**, concerning the typology and diachrony of article morphemes. This will allow us to carve out both strengths and weaknesses of Hawkins' programme in its current manifestation.

1 Introduction

In debating the role of source- and result-oriented explanations in typology, a research programme that merits discussion is John Hawkins' approach to cross-linguistic variation, laid out most comprehensively in **Hawkins1994**, (2004) and (2014). The overarching hypothesis of these works is that many cross-linguistic generalizations about grammatical structure can be explained as adaptations to efficient information processing ("processing ty-pology", cf. **Hawkins2007**). In a nutshell, Hawkins argues that efficient in-formation processing can be achieved by (i) "minimizing domains" in which certain semantic and syntactic relations are processed, (ii) "minimizing forms" whenever their information content is recoverable from the context or long-term statistical knowledge, (iii) arranging elements in such a way that the ultimate message can be transmitted as rapidly

and accurately as possible, i.e. without delays, false predictions, backtracking, etc. These efficiency principles are thus attractors that are assumed to affect linguistic choices in usage events and ultimately also the conventionalized shapes of grammars.

Hawkins' programme is one of the most systematic attempts to ground typological data in psycholinguistic research and to link it to the arena of language use; in this spirit, it is similar, for example, to the work of Bybee (1985; 2010) and Croft (2001; 2003). Moreover, Hawkins' "performance-grammar-correspondence hypothesis", according to which grammatical rules are basically crystallized usage preferences, echoes one of the key tenets of the usage-based theory of language (cf. **Langacker1987**; **KemmerBarlow2000**). And some specific efficiency principles, such as the "minimization of forms" in proportion to their degree of predictability, even have exact parallels in other functional-typological works (e.g. **Haiman1983**; **Croft2003**; **Haspelmath2008**).

At the same time, however, Hawkins' work is not always received uncritically within usage-based linguistics. Among other things, it is couched in a formal phrase-structure architecture that appears to presuppose the existence of many grammatical categories (cf. **Diessel2016**); some of its principles have been criticized for not being truly domain-general but perhaps specific to language (such as a pressure for short constituent recognition domains, cf. **Bybee2010**); and crucially in the present context, Hawkins has also been criticized for neglecting or underestimating the diachronic dimension behind the phenomena he attempts to explain (e.g. **Cristofaro2017**; **Collins2018** [this volume]). But to the extent that Hawkins does make reference to historical developments, the nature and plausibility of his diachronic claims are worth investigating in more detail, which is precisely what the present contribution aims to do.

To this end, the first part of the paper develops a systematization, in the usage-based framework, of how Hawkins' functional-adaptive principles possibly affect the diachronic development of languages. I argue that there is solid evidence for efficient information processing in the moulding of grammar, suggesting that there is a place for result-oriented processes, beside source determination, in accounting for typological distributions. In the second part of the paper, I exemplarily focus on a diachronic prediction made in **Hawkins2014**, according to which languages of different word-order types show markedly different propensities for grammaticalizing definite articles (the prediction will be formulated more precisely as we go along). This miniature case study will not only serve as a testing ground for this specific efficiency-based hypothesis, but also allow us to identify some general merits and potential problems of processing typology.

2 The diachronic dimension in processing typology

In the usage-based approach, language change is conceived as a multi-step process (cf. **Croft2000**, 2006; **Aitchison2013**) that starts by breaking a convention in the form of a linguistic innovation ("altered replication"), followed by the spread of that innovation through both the linguistic system ("diffusion") and the speech community ("propagation"). Hawkins' publications contain a number of indications as to how his efficiency principles influence innovation, diffusion and propagation processes. I will tackle of each of them briefly, in reverse order, as this reflects an increasing degree of explicitness of the respective proposals.[1]

As for propagation processes, Hawkins is usually reticent with regard to the forces that implement efficient structures, even though the central diachronic mechanism in his programme is that of "selection": efficient variants are said to be selected relatively more frequently than their inefficient counterparts, until they may ultimately oust the inefficient ones completely. It is in this way, Hawkins argues, that preferred patterns in performance can conventionalize into grammatical rules, although he concedes in **Hawkins2014** that it is presently poorly understood how exactly this "translation from performance to grammar" works. Now, if one subscribes to the view that propagation is entirely driven by sociolinguistic forces like prestige, solidarity and the resulting accommodation (e.g. **Croft2000**; **Cristofaro2017**, 2018 [this volume]), it remains mysterious, indeed, how Hawkins' very idea of selection processes can fit in.

On the other hand, there are well-known accounts of language change in which propagation is not exclusively a social phenomenon: **Keller1994**'s (**Keller1994**) "Invisible Hand" theory, for example, leaves room for functional considerations in the selection process. Some of Keller's classic examples of invisible-hand processes, such as the emergence of a traffic jam or a short-cutting footpath do not, in fact, involve social motives: People follow a certain course of action because they primarily consider its functional advantages, regardless of the sociolinguistic profile of the person whose behaviour they adopt. **Cristofaro2017** claims that there is no empirical evidence at all for this scenario in linguistics, but this assessment is overly pessimistic: **Rosenbach2008**, in a detailed examination of evolutionary accounts of language change, concludes that "the evidence available does not speak for the *exclusive* role of social factors in the selection process" (**Rosenbach2008**:

[1] Although the present section is specifically about Hawkins' work, it actually applies to "functional-adaptive constraints" (**Haspelmath2018** [this volume]) more generally, not least because Hawkins' processing typology draws on and incorporates similarly-minded principles from many other functionalist typologists (e.g. Greenberg, Comrie, Keenan, Givón, Haiman, Croft, Haspelmath, etc.).

44; emphasis in original). Therefore, I currently see no reason to dismiss a priori a theory in which both social and functional selection pressures can be operative in propagation (cf. also **Haspelmath1999**; **Nettle1999**; **Enfield2014** for similar positions).[2] On this view, then, Hawkins' efficiency principles are relevant to, and hence at least partially drive, the propagation process, although empirical evidence that clearly disentangles functional and social selection processes is, of course, very hard to come by (cf. also **Seiler2006**).

The empirical picture is clearer, in my view, when it comes to DIFFUSION or ACTUALIZATION processes, i.e. the spread of an innovation through the linguistic system.[3] Although Hawkins himself does not speak of diffusion or actualization, the process is actually highly germane to his research, as many of the phenomena he discusses in support of his efficiency theory are cases of limited diffusion. In relativization, for example, a well-known pattern is for resumptive pronouns, once they have been innovated, not to spread across the entire range of relativization sites, but to be restricted to certain sections of Keenan & **Comrie1977**'s (**Comrie1977**) accessibility hierarchy (as in Hausa, Hebrew, Welsh and many other languages). Similarly, when object case markers develop and spread within the linguistic system, they typically end up being confined to animate, definite or pronominal objects, rather than being extended across the board (cf., e.g., **Sinnemäki2014** for a quantitative study). Many other cases of such differential marking are collected in **Haspelmath2008** and subsumed by Hawkins (2004; 2014) under his "Minimize Forms" principle: The marker in question is applied to those environments that require more processing effort, and is left out economically elsewhere. Processing effort, in turn, may be related to various factors, notably constraints on working memory (e.g. longer process-

[2] Note that recent mathematical models of language change (e.g. **BlytheCroft2012**) clearly show that selection as such is a crucial element of propagation processes, in as far as alternative models of propagation that do not rely on a weighting of linguistic variants (e.g. **Trudgill2004**) do not produce the empirical patterns of propagation that have been established in historical linguistics and sociolinguistics. However, **BlytheCroft2012** also concede that their model cannot distinguish between social and functional factors in selection, i.e. it leaves open which of these is more vital in the propagation process or how they possibly interact.

[3] The term DIFFUSION is best-known in the context of "lexical diffusion" (**Wang1969**), which refers to the successive spread of a phonetic or morphosyntactic innovation to different lexical items (e.g. the Progressive construction to more and more lexical verbs, or final consonant devoicing to all relevant words). In the present paper, I am using the term diffusion in a broader sense, comprising also the application of an innovated grammatical marker or construction to a new morphosyntactic environment (e.g. the extension of *all but* in its historically younger sense 'almost' from adjectival uses (*This was all but remarkable*) to verbal environments (*He all but fell down*), cf. De **Smet2012**). Diffusion is thus synonymous with the term ACTUALIZATION (cf. **Timberlake1977**, **Andersen2001** and many others, most recently De **Smet2012**).

ing domains correlating with resumptive pronouns) and the relative unexpectedness ('surprisal') of a given configuration (cf. **NorcliffeEtAl2015** on memory- and expectation-based processing in cross-linguistic perspective). For example, discourse participants and animate entities are more likely to be subjects than objects, hence it is precisely these kinds of objects that are more surprising. Paired with the Hawkinsian assumption of efficiency on the part of the speaker, it is also only these objects that need to be marked overtly. A similar surprisal-based account is provided by **Haig2018** to explain "why differential object indexing is an attractor state" (**Haig2018**) in the grammaticalization of object pronouns.

In Hawkins' programme, then, all of these cases are amenable to an explanation in terms of efficient information processing. I believe that this account is presently superior to purely source-oriented typologies such as Cristofaro's, for the following reasons.

Firstly, there is solid evidence for efficiency where the occurrence of a particular marker is optional. This can be observed, for example, with variable relativizers that, other things being equal, show up less frequently when a relative clause is statistically expected given the previous co-text, and vice versa (cf. **WasowEtAl2011**). As **FoxThompson2007** observe, a sentence like

(1) *This was the ugliest set of shoes* [*I ever saw in my life*].

would sound "quite awkward" (**WasowEtAl2011**: 181) if the relative clause were introduced by *that*; according to Wasow et al., this is precisely because a relative clause is expected in this context, which is in turn why relative *that* tends to be omitted efficiently. **Jaeger2010** shows that similar predictability effects account for a large portion of the variability of the English complementizer *that*.

Importantly, the same kinds of effect also show up in psycholinguistic experimentation, and in languages other than English. For example, recent studies have shown that optional case marking in Japanese, optional indexation in Yucatec Maya relative clauses or optional plural marking in an artificial language exhibit an efficient distribution in the participants' linguistic behaviour, other things being tightly controlled for (cf. **KurumadaJaeger2015**; **NorcliffeJaeger2016**; **KurumadaGrimm2017**). All of these synchronic effects are independent of the historical source of the respective marker. In other words, no matter how a particular relativizer emerges, it comes to be applied in ways that are consonant with Hawkins' efficiency predictions. And as, for example, Seržant (2018 [this volume]) shows, such optional marking can conventionalize into more fixed grammatical patterns over time.

Secondly, to the extent that the "Minimize Form" effects are typologically

sound (i.e. independent of geographical and genetic affiliations), they are in contrast with a powerful principle that we observe elsewhere in grammars, viz. the potent force of analogy (cf. **GentnerSmith2012**; **BlevinsBlevins2009**). Analogy is the driving force behind lexical diffusion, and where it runs to (near-)completion, the result is a productive grammatical rule in the traditional sense. Time and again, historical studies show just how sweeping analogical extension can be: By incremental diffusion processes, English has conventionalized a rule that every main clause requires an overt subject, and every lexical verb now needs *do*-support if it is to occur in an interrogative clause. In other languages, split alignment systems are gradually being eliminated in favour of unified marking: for example, younger speakers of Choctaw (Muskogean: USA) are in the process of re-shaping split intransitivity into a nominative-accusative system with consistent coding for the S argument (**Broadwell2006**); and **Creissels2018** argues more generally that there are strong analogical pressures on languages to regain consistent alignment patterns if these get disrupted by grammaticalization processes.[4]

In view of these analogical forces, one may wonder why systems of differential resumption, differential object marking or differential possessive marking exist quite pervasively, and in highly systematic ways. In Cristofaro's account, they are persistence effects, i.e. they are all due to the fact that the unmarked meanings are perceived as incompatible with the source construction. For example, when an object marker originates from a topic marker, it is expected to be restricted to object NPs whose properties are most closely associated with topicality (or topic-worthiness), such as pronominal, animate and definite entities, and not to apply elsewhere. In fact, **DalrympleNikolaeva2011** argue that such erstwhile topic markers are often extended to animate and/or definite objects, thus diffusing in principled ways to create DOM patterns that may plausibly be linked to the source construction. However, given the powers of analogy, why does diffusion stop there? If it is really the source construction pulling its weight here, one may wonder why it does not do so in many other cases.

Just consider what is perhaps the textbook example of a development that standardly overrides effects from the source construction, namely diffusion processes in grammaticalization. It is by analogical extension that the *going-to*-future has spread to inanimate subjects (*The icicle is going to break off.*), and that the French negative marker *pas* has been extended beyond contexts of directed motion (*Je ne vais pas* > *Je ne sais pas.*). In other words, if analogy works in many other

[4]For example, "many languages in which the grammaticalization of a new TAM form resulted in [tense-based split ergative alignment] have undergone a subsequent evolution that can be characterized as regularization under the pressure of analogy" (**Creissels2018**).

instances of grammaticalization, why not in those cases that involve differential marking? Siding with Hawkins (and **Haspelmath2018** [this volume]) here, I find it convincing that an attractor state of efficient coding shapes the development of grammatical systems, especially in light of the behavioural evidence cited above.

Thirdly, while I agree with Cristofaro (2018 [this volume]) that functional principles should be "visible" in the diachronic development of particular structures, I find her interpretation of this requirement too narrow: She demands that the alleged motivations be present at the innovation stage of a grammatical construction and hence directly influence its emergence. But as we have just seen, it is often during the actualization phase that functional-adaptive principles are operative, regardless of how or why a given marker originated in the first place (cf. also Seržant 2018 [this volume]).

Interestingly, while Joan Bybee is now often cited as a representative of source-oriented typology, her relevant publications (e.g. **Bybee1988**; 2008) reveal a broader perspective than Cristofaro's: "Identifying the causal mechanisms [that lead to typological generalizations] requires a detailed look at *all* the properties of a change – including its directionality, gradualness, spread through the community and through the lexicon" (**Bybee2008**). Crucially, it is precisely in lexical-diffusion processes that many of her well-known frequency effects apply: For example, Bybee's "conserving effect" of token frequency explains why highly entrenched main verbs like *speak, think* and *mean* resisted the innovative *do*-support in *wh*-questions for a long time (e.g. *What spekest thou?*, cf. **Ogura1993**), or why the change from *-th* to *-s* in the third-person of English verbs affected the most frequent verbs last (notably *hath* and *doth*, cf. van **Gelderen2014**: 172). In a similar vein, I would thus argue that the diffusion or actualization stage is highly relevant for the kinds of effects that lead to efficient typological marking patterns.

In conclusion, I consider Hawkins' account (and functional-adaptive motivations of similar kinds) capable of explaining why certain changes do *not* happen – particularly, why we find that analogical extensions are systematically brought to a halt even though they are so commonly carried through in other domains of grammar.[5]

Let us finally turn to the realm of INNOVATION, i.e. Hawkins' suggestions as

[5]See also **Smith2001** for a similar view: He investigates the diachronic development of agreement loss in Romance participles and argues that while parsing principles cannot be held responsible for the *rise* of participial agreement, they did play a role in its gradual disappearance. Specifically, Smith claims that agreement was retained longest in those environments where it was most beneficial for processing. Therefore, "functionality is here acting as a brake on actualization" (**Smith2001**), just as I argued more generally above.

to why, where and when certain grammatical structures emerge. A central concept here is that of correlated evolution: When a language changes in one part of the grammar, Hawkins often expects to see "ripple effects" (**Hawkins2014**) in domains that are linked to the changing subsystem by certain efficiency principles. For example, since Hawkins assumes that phrases of different types (VPs, PPs, NPs, etc.) show harmonic ordering patterns to allow efficient sentence processing, a change from OV to VO is predicted to engender innovations in PPs and NPs as well (cf. **DunnEtAl2011** and the papers in *Linguistic Typology* 15(2) for ample discussion of this issue). In the present context, perhaps the most interesting claim with regard to innovation is that efficiency principles can predict the occurrence of grammaticalization: While many grammaticalization paths are universal "attractor trajectories" (**BybeeBeckner2015**) – open to all languages with similar source constructions due to the same mechanisms of reanalysis –, Hawkins' efficiency principles predict under which structural conditions (e.g. in which language "types") particular events of grammaticalization are more or less likely to happen. In the remainder of this paper, I will briefly discuss a specific example of such a hypothesis developed by **Hawkins2014**.

3 A test case for processing typology

In his 2014 monograph, Hawkins examines the structure of noun phrases (NPs) from a processing perspective. Across the world's languages, NPs often contain elements in addition to the head noun that, in Hawkins' view, can function as processing cues to the recognition (or online "construction") of an NP, such as articles, classifiers and related morphemes.[6] Hawkins argues that such elements are more efficient in VO languages than in OV languages: As illustrated schematically in Figure **??**, an additional NP constructor C in a VO language can shorten the domain for the construction of the VP (V+NP), especially if N is delayed by intervening material (e.g. in AP-N sequences like *the very delicious meal*). In an

[6] Although Hawkins frames the idea of "online construction" in terms of syntactic trees, nodes and categories, the basic intuition behind it is functional in nature: Translated into the usage-based parlance of, e.g., **Croft2001**, **BecknerBybee2009** or **BatesMacWhinney1989**, Hawkins' idea is that a referential expression should be recognizable as such, based on reliable cues in the speech stream. Referential expressions (or NPs, for that matter) are arguably best cued by nouns and determiners, and the construction of an NP is thus facilitated by the early availability of such "constructing categories" within the string of units that ultimately belong to the NP. More generally, I believe that Hawkins is thus actually quite compatible with usage-based and construction-grammatical conceptions of processing, even though he uses terminology that is closely associated with generative syntax.

OV language, by contrast, additional NP constructors lengthen this dependency domain, no matter where they occur in the NP:

Figure 1: V-NP processing in VO- and OV-languages (adapted from **Hawkins2014**: 125)

From these considerations, one might derive the following prediction:

(2) While all languages have source constructions for articles (notably demonstratives for definite articles and the numeral 'one' for indefinite articles), the grammaticalization of these sources into more general NP markers should be a more productive historical process in VO languages than in OV languages. As a result, the synchronic typological distribution of articles is significantly different in the two language types.

As a matter of fact, **Hawkins2014**' (**Hawkins2014**) prediction is narrower in scope: He applies it only to definite articles, and only to independent definite articles (i.e. words and clitics rather than affixes). The following examples illustrate the language types that are expected to be frequent according to (2):

(3) a. VO with definite article
Maori (Austronesian, Oceanic; **Bauer1993**: 256)
I kite ia i te whare.
T/A see 3SG OBJ DET house
'She saw the house.'

b. OV without definite article
Lezgian (Nakh-Dagestanian, Lezgic; **Haspelmath1993**: 343)
Ada-z balk'an aku-na.
he-DAT horse see-AOR
'He saw the horse.'

In support of the hypotheses in (3), Hawkins cites some of **Dryer2005 WALS**'s (**Dryer2005 WALS**) data, which show, indeed, that free-standing definite articles are relatively more frequent in VO-languages (more on the data below).

Hawkins' approach, as illustrated by this specific example, has a number of assets: For instance, it emphasizes the importance of the linear dimension of language, which tightly constrains production and parsing processes but which tended to be neglected by (at least early) cognitive-linguistic and construction-based approaches to grammar (cf. also **Diessel2011** for a similar critique). Hawkins' work is clearly pioneering here, and in the recent usage-based literature, related

notions like contextual predictability, informativity and projective links have come to take a highly prominent place (cf., e.g., **GahlGarnsey2004**; **Levy2008**; **Auer2009**). Furthermore, Hawkins' diachronic thinking adds a new dimension to classic research in grammaticalization. As **Good2008** points out, work on grammaticalization typically offers "permissive explanations [...], that is, it focuses on particular grammaticalization paths without, in general, accounting for what factors will cause one language, but not another, to instantiate those paths." Hawkins' approach elevates this "permissive" nature of explanation to what Good (ibid.) calls a "probabilistic" one: It attempts to explain why certain grammaticalization processes are set in motion only (or preferably) in certain language types or at certain points in time (cf. also **Hawkins1986**, 2012 for representative work along these lines).

But just how convincing are such claims and the empirical support that Hawkins provides for them? In the present case, I have a number of reservations about the picture drawn in **Hawkins2014**.

To begin with, I do not quite see why the hypothesis is restricted to the development of definite articles, as indefinite articles should qualify equally well as NP constructors. Similarly, Hawkins' preoccupation with word-based processing (which is prominent throughout his 2014 book), to the neglect of affixes with identical functions, is not sufficiently motivated. In addition to the problem that free and bound markers are very hard to distinguish consistently for cross-linguistic comparison (**Haspelmath2011**), it remains unclear if there is a measurable psycholinguistic difference between word- and affix-processing. As long as there is no evidence for the view that free and bound definiteness markers are parsed in fundamentally different ways, we should rather take a more embracing approach to the data and ask whether VO- and OV-languages differ in their propensity to grammaticalize article morphemes from their respective source constructions.

With these considerations in mind, the first step of the empirical assessment is, just like in **Hawkins2014**, to examine the typological distribution of article morphemes. Dryer's *WALS* data, in their most recent version, are set out in Table 1:

For the purposes of testing our revised version of Hawkins' hypothesis, we need to discard the languages without a dominant order of V and O ("ndo"), and we basically conflate the figures in the first four rows of Table 1 and contrast them with those in the final row. In other words, (i) we consider both free and bound definiteness morphemes; (ii) we include those languages which are beginning to use a demonstrative like an article (row 3, cf. **Dryer2013a** for details) – thus incorporating cases of incipient grammaticalization; (iii) we include languages

Table 1: Distribution of articles in different word-order types (**Dryer2013**)

	VO	OV	ndo	Totals	
Distinct ART word	144	52	14	210	
DEF affix	49	33	6	88	ART
DEM used as ART	30	33	5	68	
Only INDEF ART	20	24	0	44	
No ART	70	111	14	195	NO ART
Totals	**313**	**253**	**39**	**605**	

with indefinite articles only.[7] The conflated form of the data thus looks like in Table 2:

Table 2: Distribution of articles in different word-order types (reorganized)

	VO	OV	Totals
ART morph	243	142	395
No ART morph	70	111	181
Totals	**313**	**253**	**566**

The distribution in Table 2 looks conspicuously skewed, but of course these are raw data that are not controlled for genetic and areal effects.[8] Therefore, what **Hawkins2014**' (**Hawkins2014**) analysis clearly needs to be augmented with (in this case as well as virtually all others in his book) is proper statistical modelling according to contemporary standards (cf., e.g., **Bickel2011**). To this end, I am seeking converging evidence from two complementary quantitative approaches to the data, namely mixed-effects logistic regression (cf. also **Cysouw2010**; **JaegerEtAl2011**) and Bickel's (2011; 2013) Family Bias Method (which is particularly suitable to testing hypotheses formulated in diachronic terms). In the supplementary ma-

[7] Some readers may object to this way of grouping the data. For example, one might reasonably argue that languages in which demonstratives are used with some article-like functions should *not* be said to have "proper" articles (yet). However, even when such languages are classified differently for statistical purposes, the results remain the same in many respects (cf. supplementary material SM3.2).

[8] For similar raw data, cf. also **Dryer2009**, who endorses Hawkins' processing explanation.

terials to this paper[9], I offer a more detailed, non-technical introduction to the Family Bias Method (SM1), as well as the statistical properties of all models (SM2–5). For reasons of space, I here confine myself to describing some major results of the analyses.[10]

Figure 2 shows that there is a significant effect of word order on the occurrence of articles in a mixed-effects regression model (β = -0.73, p < 0.001).[11] Although the model is not particularly good overall, probably missing important further predictors (R^2_c = 0.14, C = 0.72), Hawkins' hypothesized effect is clearly present, as the probability of *not* having articles (y-axis) increases significantly as we go from VO to OV (x-axis):

In the Family Bias estimations, too, it turns out that, among those families that do not just show a chance distribution of articles, VO families are about 2.6 times more likely to develop articles than OV families. This is illustrated in Table 3, and Figure **??** shows that this effect is stable (i.e. never reversed) across all six macro areas:

Table 3: (Rounded) family biases for articles in different word-order types (N_{total} = 217 genetic units, 99 of which are estimated to be "biased" (as opposed to internally diverse); Fisher exact test, p = 0.039)

	VO	OV	Totals
ART morph	50	19	69
No ART morph	15	15	30
Totals	**65**	**34**	**99**

In sum, the global typological picture is consistent with Hawkins' processing account, even when tested against a more comprehensive data set and with more rigorous modes of examination.

Recall, however, that a second prediction of this account is that articles are especially useful in those VO languages that have modifiers before the head noun

[9] Cf. http://www.kschmidtkebode.de/publications. #ZENODO

[10] All statistical analyses were performed in *R* 3.3.1 (*R* Development Core **Team2016**). I am grateful to Taras Zakharko and Balthasar Bickel for making their Family Bias algorithm freely available (**ZakharkoBickel2011+**).

[11] All regression analyses I performed are based on generalized linear mixed-effects models that include genealogical and macro-areal dependencies as random effects (cf. SM3). The model in Figure 2, for example, contains by-family and by-area random intercepts for the distribution of articles, while a by-area random slope for the word-order effect did not improve the model significantly and was hence excluded from the final model.

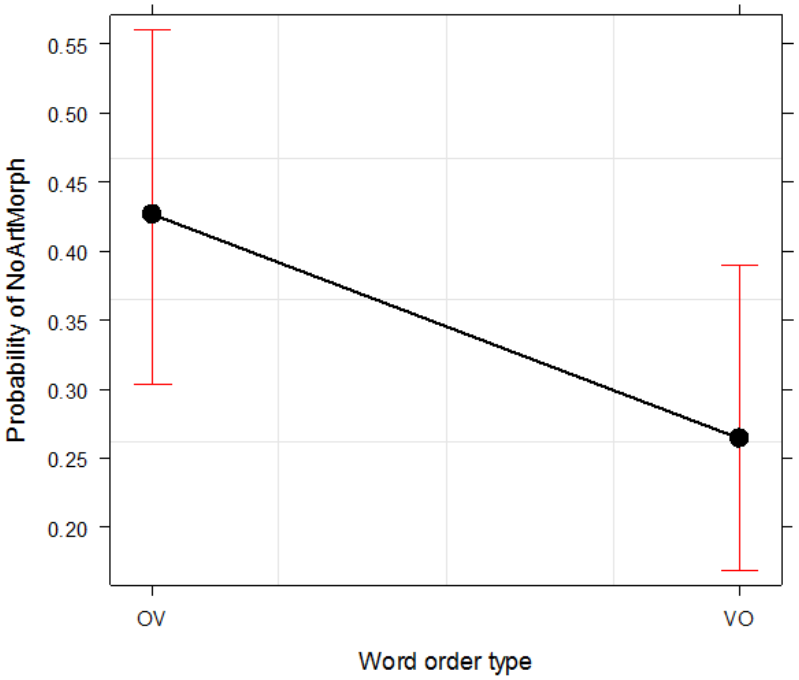

Figure 2: Effect of word order type on the probability of (not) having articles in a mixed-effects model (cf. SM3 for details)

Figure 3: Family biases by macro area (cf. SM2 for details)

in NPs (*a very delicious meal*). One would thus expect, for example, that the grammaticalization of articles is particularly productive in VO languages with ADJ-N order, and, from an efficiency perspective, less so in those with N-ADJ order. I tested this by examining the order of nouns and adjectives (**Dryer2013b**) in all VO languages in the same sample as above (N_{total} = 278 languages). Across several different statistical models (and operationalizations of the hypothesis, cf. SM4), I did not find support for Hawkins' efficiency hypothesis. In one analysis, for example, I probed whether free-standing article words are more likely in VO languages with ADJ-N order than in those with N-ADJ order. Figure 4 shows that this is neither the case for definite articles nor for articles in general.[12]

[12]Moreover, if we look at VO languages which are beginning to use a demonstrative as a definite article (N = 26 in **Dryer2013a**), Hawkins' account would lead us to expect that such incipient grammaticalization is particularly frequent in the constellation DEM-N and ADJ-N (and again less frequent if N precedes both the ADJ and the DEM). Now, of the 26 languages in question,

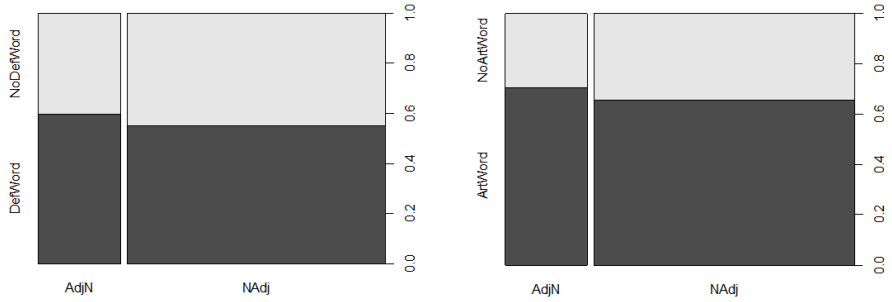

Figure 4: Occurrence of articles in VO languages depending on the position of adjectives (left plot: definiteness words only; right plot: all article-like words; for the corresponding mixed-effects models, cf. SM4)

Clearly, this picture does not speak for a critical processing pressure being at work. And the same conclusion actually carries over to OV languages: It is true that, to the extent that these languages show a reduced propensity for developing articles, they manage to keep NP processing domains slightly shorter; but there are several indications that this pressure cannot be particularly strong.

First, our Family Bias calculations show, in addition to the findings from above, that *none* of the large OV families in the sample actually exhibits a significant bias (towards or against articles) in the first place; they are all internally diverse, i.e. with no more than chance distributions of articles (Table 4).[13]

Second, from a more qualitative perspective, there is suggestive evidence that a potential efficiency motivation in OV languages is easily overridden by other factors. For example, **Ross2001** discusses an interesting case of an intense contact situation in which the Austronesian language Takia adapted its VO syntax to the OV structure of its Papuan contact language Waskia. Ross argues that, in the wake of this restructuring, Takia speakers must have shed the prenominal article word in NPs (cf. (4a)), which would be fully in line with Hawkins' predic-

22 are N-DEM and four are DEM-N. It is the latter type that is interesting here, and we find that two of these four languages are ADJ-N and the other two N-ADJ. Again, no clear pattern along Hawkinsian lines can be detected here.

[13] **Bickel2013** suggests that the (minimum) strength of a universal pressure can be calculated on the basis of the proportion of biased families k among all families of a particular kind n (here: OV families): $\hat{s} = (k+1)/(n+2)$. Based on the figures in Table 4, we obtain $\hat{s}_{(OV)} = (0+1)/(11+2) = 0.077$. This estimate is so small in magnitude that one is forced to conclude that there is no particular pressure at all on OV languages with regard to the development of articles.

Table 4: Distribution of biases (for or against) articles among large families in the sample (N_{total} = 29 genetic units)

	VO	OV	Totals
significantly biased	12	0	12
internally diverse	6	11	17
Totals	**18**	**11**	**29**

tion. At the same time, however, the degree of linguistic accommodation was so intense that Takia speakers also did something else: They grammaticalized a postnominal deictic element into a postnominal demonstrative with some article-like functions, reproducing exactly the article pattern in Waskia (cf. (4b-c)).

(4) a. Proto-Western Oceanic (**Ross2001**)
a tam^{w}ata a-ña
DET man that-3sg
'that man'

Proto-Western Oceanic

b. Takia (Austronesian, Oceanic; **Ross2001**: 140)
Waskia tamol **an**
Waskian man **that**
'that Waskia man'

Takia

c. Waskia (Nuclear Trans New Guinea, Madang; **Ross2001**: 140)
Waskia kadi **mu**
Waskia man **that**
'that Waskia man'

Waskia

In other words, Takia speakers chose precisely the diachronic route that Hawkins would predict to be disfavoured, which goes to show that the alleged processing pressure cannot have been very strong, after all. In this connection, one may also recall that our regression model from above, while bringing out a significant global effect from word order type, did not provide a particularly good fit to the data. The substantial amount of variation in the data that it cannot account for must thus be attributed to other, possibly stronger factors.[14]

[14] Diachronic research has actually put forward a number of plausible candidates for such factors. A prominent one since at least **Vennemann1975** is the loss of a case system and the concomitant rigidification of constituent order, which favours the development of articles to express

We conclude, then, that **Hawkins2014** correctly predicts a global difference between OV- and VO-languages in the development of articles. But the present analysis also revealed some challenges for this account. Therefore, it still needs to be established by future research whether the global correlation between word order type and the absence of articles really reflects a *causal* connection between these two phenomena, and whether this could be attributed to efficient information processing. If it turns out that Hawkins is correct, the findings in the present section suggest that we would be dealing with a 'weak universal pressure' in the sense of Seržant (2018 [this volume]) or a "weak cognitive bias" with "significant population-level consequences" (**ThompsonEtAl2016**: 4530).

4 Concluding remarks

The present contribution has taken a closer look at John Hawkins' "processing typology", a research programme that thoroughly subscribes to functional-adaptive motivations for grammatical structure. In the first part of the paper, I summarized where such motivations are possibly operative in diachronic change. In my view, a case can be made for Hawkins' efficiency considerations in the process of actualization, i.e. when a linguistic innovation comes to be extended to a principled, cross-linguistically similar subset of potential application sites (as in differential flagging and indexing, relativization, etc.). In this respect, I consider Hawkins' account as superior to purely source-oriented explanations of grammatical patterns. Of course, this does not deny that persistence accounts are relevant to typological patterns – they clearly are; but it argues against persistence as the sole or perhaps even the dominant explanatory principle for grammatical universals.

A more ambitious but also undoubtedly more problematic move is to link parsing and efficiency to certain innovation processes, such as when a particular grammaticalization channel is predicted to be set in motion only un-der specific structural conditions. In the brief case study presented here, we saw that Hawkins' NP processing hypothesis provides a neat match to the global typo-

information-structural distinctions that were previously coded by a more flexible word order (cf. also **Hawkins2004**; **HewsonBubenik2006**; **Fischer2010**; **CarlierLamiroy2014**). Another possible factor is the loss of an aspectual system (cf. **Abraham1997**; **Leiss2000**, 2007). However, especially the former type of explanation is often viewed critically (e.g. **Selig1992** on Romance; McColl **Millar2000** on English; **Leiss2000** on Germanic), and there is currently no proposal as to how various factors may conspire to explain the synchronic distribution of articles (cf. also **Lüdtke1991**). For some further information and preliminary typological analyses of these factors, interested readers are kindly referred to SM5.

logical data, even when these are analyzed in more rigorous and hence more appropriate ways than in **Hawkins2014**. On the other hand, the details of neither the typological picture nor individual diachronic studies produce evidence for a strong pressure on languages to develop into the predicted directions. Therefore, the hypothesis that speakers of OV languages are significantly less inclined than speakers of VO languages to grammaticalize additional NP constructors, remains plausible but currently rather weakly substantiated.

What we would need to see to make it more convincing is a triangulation of (i) typological data that are large enough to take several alternative predictor variables from the literature into account (e.g. case and aspect systems, the presence of other NP constructors such as classifiers), (ii) diachronic data from languages that have undergone (or are in the process of undergoing) changes in basic word order, (iii) behavioural evidence, such as psycholinguistic experimentation with artificial languages (e.g. along the lines of **CulbertsonEtAl2012**; cf. also **Levshina2018** [this volume]). As a matter of fact, a particularly strong aspect of Hawkins' work (especially in **Hawkins2004**; 2014) is that it generally attempts precisely this kind of methodological cross-fertilization; but for the domain at issue here, such an approach has yet to be fleshed out in sufficient detail.

Abbreviations

The paper follows the *Leipzig Glossing Rules*.
Additional abbreviation: T/A = tense/aspect marker

Acknowledgements

The research for this paper was carried out 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. I would like to thank John Hawkins, Mark Dingemanse, the co-editors of the present volume as well as the audiences of the Diversity Linguistics Conference (Leipzig, March 2017), the 3rd Usage-Based Linguistics Conference (Jerusalem, July 2017) and the Syntax of the World's Languages VIII Conference (Paris, September 2018) for very helpful feedback on previous versions of this paper. The usual disclaimers apply.

Chapter 2

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 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

[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 2 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.

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

[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

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 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 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 **??** and 4, 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 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 **KortmannLunkenheimer2011**) go back to

[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.

Table 3: Shortened dependent form

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

Table ??

Table 4: Lengthened independent 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]

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

[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

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 5.

Table 5: 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'

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).

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

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 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 exam-

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

ple 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 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 feature an intensifier which is added to the dependent possessive, as in Krio *mi yon* [1SG.POSS IN-

TENS.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 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 (**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 explained 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 psy-

cholinguistic 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 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

[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).

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 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* (**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

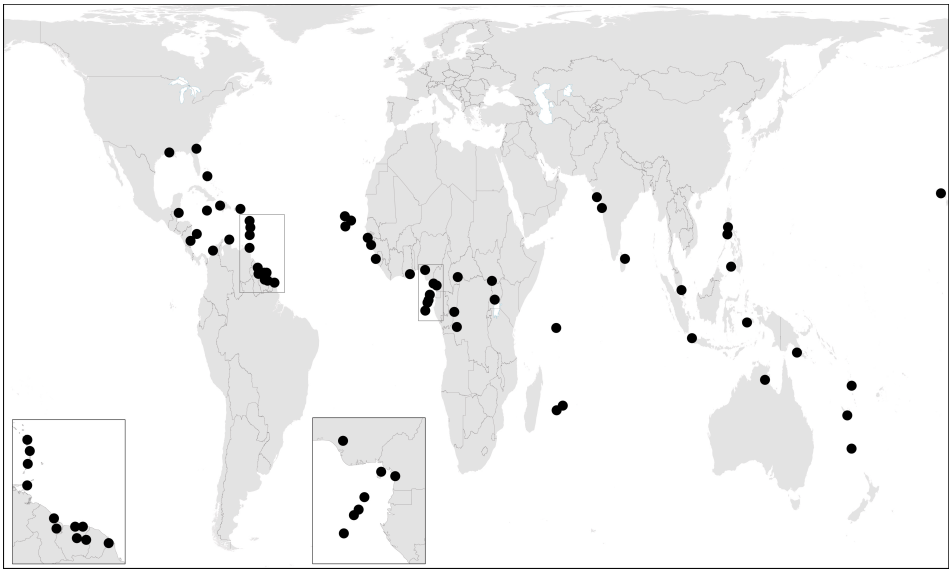

Figure 1: Distribution of the 59 creole languages in APiCS (for more information see apics-online.info) (CC BY-SA 4.0, Hans-Jörg Bibiko, MPI-SHH/Jena)

Chapter 3

Some language universals are historical accidents

Jeremy Collins
Radboud University Nijmegen

In this short paper, I elaborate on previous work by **Givón1971** and **Aristar1991** to argue that a substantial part of the well-known word order correlations is best explained by grammaticalisation processes. Functional-adaptive accounts in terms of processing or learning constraints are currently weakly substantiated, and they suffer from the fact that they do not adequately control for language-internal inheritance patterns. More generally, historical relatedness between different types of phrases constitutes an important confound in typological research, one that needs to be taken seriously before word order correlations are motivated by anything other than the diachronic patterns that link the word order pairs in question.

1 Introduction

There are surprisingly few properties that all languages share. Almost every attempt at articulating a genuine language universal tends to have at least one exception, as documented in **EvansLevinson2009**. However, there are non-trivial properties that are found in if not literally all languages, enough of them and across multiple language families and independent areas of the world, that they demand an explanation.

An example is the fact that languages have predictable word orders. If a language has the verb before the object, it tends to have prepositions rather than postpositions, as in English; if the verb is after the object, it is a good bet that the language will have postpositions rather than prepositions (**Greenberg1963**). The ordering of different elements such as a possessed noun and its possessor, or a noun and elaborate modifiers (complex adjective phrases, relative clauses), are to some extent free to vary among languages, but again tend to fall into correlating

types (**Dryer1992**; **Dryer2011**). Why should knowing the word order of one category in a language help predict the orderings of other categories? One prominent view holds that these patterns reflect an innate harmonic ordering principle of Universal Grammar, which is ultimately argued to solve the logical problem of language acquisition (**Pinker1994**; **Baker2001**; **Roberts2007**). This would amount to what Haspelmath (this volume) calls a "representational constraint" on the shape of grammars. Another possible explanation is that word-order correlations have evolved in the service of efficient language processing (e.g. **Hawkins1994**; **KirbyHurford1997**), i.e. for functional-adaptive reasons. We find this view in the functional-typological literature (e.g. **Dryer1992**; **EvansLevinson2009**) as well as in computer simulations in the literature on language evolution (**VanEverbroeck1999**).

However, I would argue that many of these patterns are not evidence of our psychological preferences, but are accidental consequences of language history. More specifically, they are accidental in the sense that they arise as a by-product of grammaticalisation processes. These processes do not seem to have word order correlations as a goal, nor is there good evidence for a "pull force" in that direction. Accordingly, grammaticalisation is an *alternative* to functional motivations here, and an understanding of this historical dimension is thus crucial to explaining word order correlations. In this short paper, I first elaborate this claim (§2) based on an earlier publication (**Collins2012**), before I outline its consequences for typological theory and practice (§3). In doing so, I am extending a line of argumentation by **Givón1971** and **Aristar1991**, but I relate the discussion specifically to the concerns of the present volume, and to Haspelmath's position paper in particular.

2 Word order correlations as a result of grammaticalisation

Grammaticalisation is the process by which new grammatical categories can be formed from other (often lexical) categories. For example, Mandarin Chinese has a class of words which might be called prepositions from a cross-linguistic point of view but which clearly have their historical roots in verbs. An example is 从 *cóng*, which in modern Mandarin is a preposition meaning 'from' but which in Classical Chinese was a verb meaning 'to follow'. It has lost its ability to be used as a full verb, requiring another verb such as 'come' in the sentence, just as English requires a verb in the sentence *I come from London.* Other Chinese prepositions such as 跟 *gēn* 'with' also have a verbal origin, and many preposition-like words such as 给 *gěi* 'for' and 在 *zài* 'in/at' even retain verbal meanings ('give' and

'to be present') and verbal syntax (such as being able to be used as the sole verb in the sentence and to take aspect marking). These patterns of inheritance directly explain why the two types of constituents (i.e. PP and VP) have the same word order: Prepositions and verbs were once the same category, and they simply have not changed their word orders since then. Since the verb precedes its NP object in classical and modern Chinese, its prepositional offspring in modern Chinese also precedes its NP complement. Interestingly, Chinese also has postpositions, such as *li* 'in', and these, too, are simply continuations of their lexical sources (cf. also **Dryer2018** [this volume]). Thus *li* is etymologically 'interior' or 'village', hence *fangzi li* 'in the house' might be glossed more literally as 'the house's inside'. Again, the ordering of the younger construction as noun (*fangzi*)–postposition (*li*) reflects the order of the older construction with genitive (*fangzi*)–noun (*li*). Very similar remarks apply to Niger-Congo languages like Dagaare in Ghana, which also shows typologically mixed adpositional phrases (**Bodomo1997**).

More generally, the pattern of adpositions inheriting the ordering of the noun or verb they derive from is replicated in different language families: We find it in many Oceanic languages (**LynchEtAl2002**: 51), where adpositions are transparently nouns and reflect whatever ordering of genitive–noun the language has (hence it can be either prepositional, as in Hawaiian, or postpositional, as in Motu); we also see it in Indo-European languages (e.g. English *across* < 13 ct. Anglo-French *an cros* 'on cross' (**BordetJamet2010**: 16)), in Japanese (e.g. *kara* 'from' < 'way', *si* restrictive particle < 'do' (**Frellesvig2010**: 132–135)), in Australian languages in which adpositions are morphologically still nouns (**Dixon2002**), in Tibetan and Burmese (De**Lancey1997**), and so on. **HeineKuteva2007** even remark that "we are not aware of any language that has not undergone such a process".

Grammaticalisation can also often explain the ordering of verb and object correlating with genitive and noun ordering (**Dryer2011**). Certain types of verb phrase derive historically from noun phrases made up of a nominalised verb and its patient argument in a possessive construction. An example is Ewe:

(1) Ewe (Atlantic-Congo, Gbe; **Claudi1994**: 220)

Ewe

Me-le é-kpɔ dzí.
1SG.-be.at 3SG.POSS/OBJ-see surface/on
'I am seeing him.' (lit. 'I am on his seeing.').

Ewe is normally SVO but employs the genitive–noun ordering here ('his seeing'), creating a construction which is SOV. Nominalisations of this kind are used cross-linguistically for expressing aspect (such as the continuous aspect

in Ewe), for subordinate clauses (expressing 'I was surprised that he saw me' as 'I was surprised at his seeing of me' in Javanese, cf. **Ogloblin2005**: 618) and for voice marking (in Austronesian languages, cf. **Himmelmann2005**: 174). These verb phrases can become the most frequently used and unmarked verb phrases in the languages, thus the basic verb–object order of a language can evolve from a genitive–noun construction, even if the nominal origins of the verb form are no longer transparent.

This development of (main-clause) verb phrases from nominalised verbs with a possessor object is again attested in very different language families, although it is more complicated to reconstruct. A typical example is the evolution of VOS ordering in Proto-Austronesian, which has been inherited by over a thousand Austronesian languages or evolved further into SVO or VSO (**Adelaar2005**). It is now generally accepted that verb phrases in Austronesian languages evolved from nominalising verbs, with a sentence such as 'The children are looking for the house' deriving from a Proto-Austronesian construction of the type 'The children are the searchers of the house'. **StarostaEtAl1982** as well as **Kaufman2009** present several pieces of evidence in favour of this diachronic hypothesis: For example, the voice markers on verbs derive from nominalising morphemes, cognates of which still exist in Tagalog and other languages, such as the locative voice marker *an* which is also used for deriving place names (*aklat-an* 'library' < *aklat* 'book'). Moreover, the direct object of the verb is marked with the genitive marker *ng* or put into the genitive case if a pronoun. Both nominalisation and the use of equational sentences of the form AB 'A is B' are extremely common in conservative Austronesian languages and presumably in Proto-Austronesian, allowing this frequently used construction to become a standard form of predication. Thus the verb–object ordering in Austronesian languages derives simply from the noun–genitive ordering of Proto-Austronesian, which is still retained in these languages. At a stroke this word order correlation is accounted for in roughly a sixth of the world's languages.

As **Sasse2009** notes in a comment on **Kaufman2009**, the situation in Austronesian is "not as 'exotic' as it seemed to be at first sight, especially not for a Semiticist or an Afroasiaticist". He notes that the Cushitic languages also replaced their finite verb forms with participles and are used with dative marking on the agent, in effect saying 'I have heard it' as 'To me was hearing' (**Sasse2009**); and that the dative pronouns eventually grammaticalised further to finite verbal morphology. This change also took place in the Iranian and Indo-Aryan languages, stretching over a large linguistic area.

Sasse also notes independent developments of agents marked with genitive

case in Mayan and Inuit languages, and **Gildea1997** made a similar reconstruction for the Cariban language family, of which the famous OVS language Hixkaryana is an example: It has genitive marking on the object, effectively expressing 'the enemy will destroy the city' as 'it will be the city's destruction by the enemy' (**Gildea1997**), explaining among other things why the subject is placed last, and why it has ergative marking. One can add to this list many languages in Asia, as described in **YapEtAl2011**, such as Tibeto-Burman languages that often use nominalised forms in main clauses (e.g. 'goat-killing exists' for 'he is killing a goat', cf. De**Lancey2011**: 349), and even Japanese, in which argument markers such as *ga* were originally genitive markers (**Shinzato2011**). Examples of Niger-Congo languages such as Ewe were given earlier and are discussed by **Claudi1994**, while Heine describes how many Nilo-Saharan and Chadic languages render desiderative sentences in the following way:

(2) Angas (Afro-Asiatic, Chadic; **Heine2009**: 31)
Musa rot dyip kə-shwe.
Musa want harvest POSS-corn
'Musa wants to harvest corn.' (lit. 'Musa wants the harvesting of the corn.')

Angas

The historical data thus show that these processes of grammatical change are not limited to individual languages or families but can instead be found much more widely, and independently of one another. They lead us to predict, then, that ultimately all correlations between the ordering of elements in verb phrases (V–NP), adpositional phrases (P–NP) and possessive noun phrases (GEN–NP) are due to direct historical connections between pairs of phrases (cf. also **Croft2003**: 77–78 for more discussion of such pairs). In the next section, I consider the implications of this assumption for both explanation and methodology in linguistic typology.

3 Consequences for typology

As historical evidence for the grammaticalisation account is accumulating, one may ask whether this makes alternative, functional-adaptive explanations invalid. Recall from above that on non-nativist approaches, word order correlations are often argued to make sentences easier or more efficient to parse in real time, as compared to sentences with mixed head–dependent ordering patterns (e.g. **Hawkins2004**). Is it possible that these factors play a role alongside grammaticalisation, such that, for example, processing demands filter out certain difficult-to-

process constructions, as **KirbyHurford1997** suggest (cf. also **Christiansen2000**)? Put somewhat differently, could it not be the case that grammaticalisation happens to produce orderings that are easy to parse?

There is currently not much evidence to substantiate this view. From a theoretical perspective, there is no indication that the processes involved in grammaticalisation are instigated by considerations of efficient parsing or learning. They happen through pragmatic inference in specific communicative contexts (**HopperTraugott2003**: Ch. 4), through widespread metaphorical mappings (cf. **Deutscher2005**: Ch. 4) and by means of chunking of repeated sequences (**Bybee2002**). Through these mechanisms, a new construction begins to emerge that gradually emancipates from its original lexical source. Since it is gradual, this process often creates a chain of intermediate cases, such as denominal adpositions in Tibetan, some of which still require genitive marking (e.g. *mdun* 'front') while others have shed this marking (e.g. *nang* 'inside'; cf. De**Lancey1997**: 58–59). In other words, grammaticalisation has its origin in common non-linguistic processes (cf. also **Bybee2010**: 6–8) and has predictable consequences, such as the gradual and sometimes only partial elimination of the morphology associated with the source. Importantly, a hallmark of grammaticalisation is syntagmatic "freezing" (**Croft2000**: 159; cf. also **Lehmann2015**: 168), so that the order of the elements in the new construction mirrors the order of elements in the source. The result is a "correlation" between the syntagmatic structure of the old and the new construction, but one that effectively rests on inertia rather than overarching processing principles that work towards a correlation.

From a methodological perspective, processing and learning accounts are an example of a broader trend of the "ad hoc search for functions that match the universals to be explained", as **Kirby1999** puts it. Attempts in the evolutionary literature to simulate processing or learning with computers in order to derive Greenberg's word order universals (e.g. Van **Everbroeck1999**; **KirbyChristiansen2003**), have a particularly "just-so" flavour: All that computer simulations can do is show that processing or learning preferences of individuals can cause these correlations to emerge over time, all other historical factors being equal, not that they are actually responsible. What we would thus need is independent historical evidence that processing concerns do, in fact, guide historical change. There are some attempts to show this, for example, in earlier English (e.g. **Fischer1992**; **ClarkEtAl2008**), when the language appeared to converge on the word order correlations after a period of freer word order. This could indeed be evidence for word order correlations emerging at least in part out of processing considerations; but there are other possibilities in this case which need to be inves-

tigated further, such as it being related to the rise of analytic verb forms and periphrastic *do*, to the loss of inflections or as a result of contact from French (cf. also **FischervanderWurff2006**: 187–188 for some of the controversies). The historical role of processing is unclear even in this case, and there is no conclusive cross-linguistic evidence for it either.

One possibility for establishing such causal relations cross-linguistically would be to look for cases of correlated evolution, i.e. situations in which a change in one word order can be shown to be followed by a change in another word order in the history of a language, or in its descendants. For example, if a language has verb-object order and prepositions but then changes to having object-verb order and postpositions, then this suggests that the two word orders are functionally linked (if this event takes place after any grammaticalisation linking these verbs and postpositions). The only solid statistical test of this so far has been a widely discussed study by **DunnEtAl2011**. Dunn and colleagues examined the ways in which four language families have developed (Bantu, Austronesian, Indo-European and Uto-Aztecan) and tested models of word order change using a Bayesian phylogenetic method for analysing correlated evolution. They found that some word orders do indeed change together: For example, the order of verb and object seems to change simultaneously with the order of adposition and noun in Indo-European. A model in which these two word orders are dependent is preferred over a model in which they are independent with a Bayes factor of above 5, a conventional threshold for significance. This seems to vindicate the idea that adpositions and verb-object order are functionally linked in Indo-European, and the pattern also holds up in Austronesian. It does not show up in the smaller and younger families Uto-Aztecan and Bantu, although that may be because of the low statistical power of this test when applied to small language families (cf. **CroftEtAl2011**). But a more important drawback is that there is no control for language contact. What could be happening is that some Indo-European languages in India have different word orders because of the languages that they are near, such as Dravidian languages, which also have object-verb order and postpositions. A similar point could be made about the Austronesian languages that undergo word order change, which are found in a single group of Western Oceanic languages on the coast of New Guinea, which is otherwise dominated by languages with object-verb order and postpositions.

In the context of the present discussion, an important result of **DunnEtAl2011**'s (**DunnEtAl2011**) paper is that word orders are very stable, staying the same over tens of thousands of years of evolutionary time (i.e. summing the time over multiple branches of the families). In this light, it is also instructive to note that

some typologically "mixed" or non-correlating languages show the same inert behaviour: Despite the fact that grammaticalisation has produced a mixture of prepositions and postpositions (e.g. in Chinese or Dagaare), the resulting systems have also survived for many generations, or even thousands of years, without showing any inclination to change. This, too, is a problem for processing-based theories, which sometimes explicitly predict that such inconsistencies should die out (e.g. **KirbyHurford1997**).

In the absence of convincing evidence for functional-adaptive motivations, I suggest that we accept that different types of syntactic constituents share their ordering patterns because they are historically related to each other, i.e. because they are linked by common ancestry. This also has important methodological consequences for typology. The kind of historical relatedness we observe here qualifies as a subtle, language-internal variant of Galton's problem (cf. **Cysouw2011** for an introduction), and it is thus actually a *confound* in typological samples. Just as other, more widely known, types of historical relatedness, such as a genealogical or areal interaction between two data points in a sample, need to be controlled for before one can test for a typological correlation, so does the language-internal historical relatedness between the grammatical patterns that make up that correlation. Put differently, languages in which possessor arguments are known to have developed from former object arguments and have simply adopted their order from this source, do not constitute an independent data point in support of the alleged word order correlation. For typological practice, this entails that we need large databases of attested grammaticalisation pathways, and that we need to examine more carefully the actual markers and their (likely) etymologies before we set out to test a functional-adaptive hypothesis. In principle, it would then be possible to inspect whether certain grammaticalisation pathways tend to be taken only in certain types of languages; for example, do postpositions only develop from nouns in a genitive construction ('table's head' > 'table on') if the language also places the verb after the object? It is easy enough to find exceptions to that, such as Dagaare (Atlantic-Congo), which has taken this route to postpositions despite being a VO language (**Bodomo1997**). But in a large database, we might still find interesting structural constraints, as well as geographical patterns, that could potentially speak for or against functional-adaptive motivations in addition to grammaticalisation.

For now, the major point is that the historical non-independence of data points can create correlations that are not causal. Such spurious correlations are well-known from non-linguistic research (cf., e.g., the spurious correlation between chocolate consumption and Nobel Prize winners; cf. also **RobertsWinters2013**

for further discussion), and my claim in this paper is that this is a serious methodological pitfall in the domain of word order correlations. Given the naturalness of grammaticalisation, and the above observation that word orders tend to be preserved and long retained after grammaticalisation, invoking functional-adaptive motivations to explain the correlations in question is not only redundant, but actually wrong-headed. It is as if one wanted to claim that there was a deeper ecological reason why chimpanzees and humans share 98.8% of their DNA, rather than just the primary historical reason, which is that they have a common ancestor.

Having said this, it should be pointed out that I am neither arguing against functional-adaptive explanations in general, nor am I denying the relevance of processing to understanding word order patterns as such, including some combinations of word order that tend to be preferred over others. For example, the fact that VO languages strongly tend to have postnominal relative clauses is plausibly related to processing constraints (**Hawkins2004**). Similarly, correlations between numeral–noun and adjective–noun ordering do not have a clear explanation in terms of grammaticalisation, but they do seem to be functionally linked and hence show interesting dependencies in experiments in artificial language learning (e.g. **CulbertsonEtAl2012**; cf. also **Dryer2018**) [this volume]). But with more and more diachronic evidence coming to light, historical links between many grammatical categories (VPs, auxiliaries, genitives, adpositions) can no longer be dismissed as marginal and as "lack[ing] generality" (**Hawkins1983**). Our default assumption, then, should be that the core word order correlations are first and foremost an accidental by-product of grammaticalisation.

Haspelmath2018 (this volume) actually acknowledges this type of explanation, at least for the ordering patterns of adpositional phrases, and labels it a "mutational constraint" – a situation in which historical sources and grammaticalisation pathways directly determine the synchronic outcomes and hence make functional-adaptive explanations superfluous. On the other hand, he rejects "common pathways" as too weak to have explanatory power in typology. But how common is "common", and when do we begin to speak of a mutational constraint? It is perfectly possible that common pathways (such as those documented in **HeineKuteva2002**; 2007), while not exhausting the possible sources and routes, are still frequent enough to produce a principled synchronic result. Therefore, I disagree with Haspelmath (**??**) that we need not be able to understand the diachronic patterns behind a universal tendency if there is a good functional-adaptive motivation available for it. In the case of word order correlations, and possibly other domains of grammar, it is the other way around: We first need to

understand the diachronic links between different types of phrases and then control for them when we attempt to establish whether there are universal correlations beyond historical dependencies at all. It may turn out that the real question is why it should ever be the case that the order of grammaticalised categories, such as adpositions, genitives or auxiliaries does *not* correlate with that of their source constructions.

4 Conclusion

Word order correlations are often invoked as evidence for universals of language acquisition or language processing. In this paper, I have argued that, before we can do so, it is important to understand the historical background of these patterns, which standard interpretations do not take into account. Given the naturalness and the non-teleological nature of grammaticalisation processes, it should be our default assumption that the order of grammaticalised categories retains the order of their respective source constructions. From this perspective, word order correlations are far from mysterious and, in many cases, do not require functional-adaptive motivations (such as specific processing principles) or innate constraints (such as a head-ordering parameter). Instead, the correlations arise during the creation of new constructions by extending old constructions. The grammaticalisation processes involved are well-understood and ubiquitous (cf. **Bybee2015**). And although we will never be able to have a full picture of the possible routes that lead to adpositions, auxiliaries, genitives, etc., the ones we know of seem common enough to produce the correlations in question. At the very least, they constitute language-internal dependencies, in Galton's spirit, that need to be controlled for in any typological investigation of word order correlations, in addition to areal dependencies that hold across languages. If they are not, one runs the risk of erroneously inferring causation from correlation, as the word order correlations would appear so strong that they require a deeper explanation, when in fact they are largely dependencies built into the sample.

5 Acknowledgements

I would like to thank the editors of the present volume, and Karsten Schmidtke-Bode in particular, for detailed discussion and extensive editorial help in compiling this paper, which is based largely on an earlier publication (**Collins2012**) and a more recent blog post (**Collins2016**).

Explanation in typology