

Chapter 6

Attractor states and diachronic change in Hawkins’s “Processing Typology”

Karsten Schmidtke-Bode

Leipzig University and Friedrich Schiller University Jena

This paper provides an assessment of John Hawkins’s (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 Hawkins (2014), concerning the typology and diachrony of article morphemes. This will allow us to carve out both strengths and weaknesses of Hawkins’s 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’s approach to cross-linguistic variation, laid out most comprehensively in Hawkins (1994; 2004; 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 typology”, see Hawkins 2007). In a nutshell, Hawkins argues that efficient information 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's 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's "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 (Langacker 1987; Kemmer & Barlow 2000). 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. Haiman 1983; Croft 2003; Haspelmath 2008).

At the same time, however, Hawkins's 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 (see Diessel 2016); 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, see Bybee 2010); 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. Cristofaro 2017; Collins 2019 [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's 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 Hawkins (2014), 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 (Croft 2000; 2006; Aitchison 2013) 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's publications contain a number of indications as to how his efficiency principles influence innovation, diffusion and propagation processes. I will tackle 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 Hawkins (2014: 10) 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. Croft 2000; Cristofaro 2017; Cristofaro 2019 [this volume]), it remains mysterious, indeed, how Hawkins's 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: Keller's (1994) "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. Cristofaro (2017) claims that there is no empirical evidence at all for this scenario in linguistics, but this assessment is overly pessimistic: Rosenbach (2008), 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" (Rosenbach

¹Although the present section is specifically about Hawkins's work, it actually applies to "functional-adaptive constraints" (Haspelmath 2019 [this volume]) more generally, not least because Hawkins's processing typology draws on and incorporates similarly-minded principles from many other functionalist typologists (e.g. Greenberg, Comrie, Keenan, Givón, Haiman, Croft, Haspelmath, etc.).

2008: 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 (see also Haspelmath 1999; Nettle 1999; Enfield 2014 for similar positions).² On this view, then, Hawkins's 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 (see also Seiler 2006).

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 & Comrie's (1977) 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 (see, e.g., Sinnemäki 2014 for a quantitative study).

Many other cases of such differential marking are collected in Haspelmath (2008) 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

²Note that recent mathematical models of language change (e.g. Blythe & Croft 2012) 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. Trudgill 2004) do not produce the empirical patterns of propagation that have been established in historical linguistics and sociolinguistics. However, Blythe & Croft (2012) 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.

³The term DIFFUSION is best-known in the context of "lexical diffusion" (Wang 1969), 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*), see De Smet 2012). Diffusion is thus synonymous with the term ACTUALIZATION (Timberlake 1977, Andersen 2001 and many others, most recently De Smet 2012).

(e.g. longer processing domains correlating with resumptive pronouns) and the relative unexpectedness (‘surprisal’) of a given configuration (see Norcliffe et al. 2015 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 Haig (2018) to explain “why differential object indexation is an attractor state” (Haig 2018: 781) in the grammaticalization of object pronouns.

In Hawkins’s 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 (Wasow et al. 2011). As Fox & Thompson (2007) observe, a sentence like

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

would sound “quite awkward” (Wasow et al. 2011: 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. Jaeger (2010) 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 (Kurumada & Jaeger 2015; Norcliffe & Jaeger 2016; Kurumada & Grimm 2017). 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’s efficiency predictions. And as, for example, Seržant (2019 [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 (see Gentner & Smith 2012; Blevins & Blevins 2009). 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 (Broadwell 2006: 140); and Creissels (2018) 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, Dalrymple & Nikolaeva (2011) 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*

⁴For 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” (Creissels 2018: 81).

ne vais pas. 'I'm not going.' > *Je ne sais pas*. 'I don't know.'). In other words, if analogy works in many other instances of grammaticalization, why not in those cases that involve differential marking? Siding with Hawkins (and Haspelmath 2019 [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 (2019 [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 (see also Seržant 2019 [this volume]).

Interestingly, while Joan Bybee is now often cited as a representative of source-oriented typology, her relevant publications 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" (Bybee 2008: 108; see also Bybee 1988). 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?*, see Ogura 1993), 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*, see van Gelderen 2014: 172). In a similar vein, I would thus argue that the diffusion or actualization stage is highly relevant for the kinds of effect that lead to efficient typological marking patterns.

In conclusion, I consider Hawkins's 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

⁵See also Smith (2001) 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" (Smith 2001: 214), just as I argued more generally above.

Let us finally turn to the realm of INNOVATION, i.e. Hawkins's suggestions as 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" (Hawkins 2014: 88) 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 (see Dunn et al. 2011 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" (Bybee & Beckner 2015) – open to all languages with similar source constructions due to the same mechanisms of reanalysis –, Hawkins's 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 Hawkins (2014).

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's 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 1, 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

⁶Although 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., Croft (2001), Beckner & Bybee (2009) or Bates & MacWhinney (1989), Hawkins's 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.

intervening material (e.g. in AP-N sequences like *the very delicious meal*). In an 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 Hawkins 2014: 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, Hawkins’ (2014) 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; Bauer 1993: 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-Daghestanian, Lezxic; Haspelmath 1993: 343)
 Ada-z balk’an aku-na.
 he-DAT horse see-AOR
 ‘He saw the horse.’

In support of the hypotheses in (2), Hawkins cites some of Dryer's (2005) *WALS* data, which show, indeed, that definite-article words are relatively more frequent in VO-languages (more on the data below).

Hawkins's 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 (see also Diessel 2011 for a similar critique). Hawkins's 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 (see, e.g., Gahl & Garnsey 2004; Levy 2008; Auer 2009). Furthermore, Hawkins's diachronic thinking adds a new dimension to classic research in grammaticalization. As Good (2008: 7) 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's 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 (see also Hawkins 1986; 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 Hawkins (2014).

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's 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 (Haspelmath 2011), 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 Hawkins (2014), to examine the typological distribution of article morphemes. Dryer’s *WALS* data, in their most recent version, are set out in Table 1.

Table 1: Distribution of articles in different word-order types (Dryer 2013a,c)

	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	

For the purposes of testing our revised version of Hawkins’s 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, see Dryer 2013a for details) – thus incorporating cases of incipient grammaticalization; (iii) we include languages with indefinite articles only.⁷ The conflated form of the data thus looks like in Table 2.

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 Hawkins’s (2014) 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 (see, e.g., Bickel 2011). To this end, I am seeking converging evidence from two complementary quantitative approaches to the data, namely mixed-effects logistic regression (see also Cysouw 2010; Jaeger et al. 2011) and

⁷Some 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 (see supplementary material SM3.2).

⁸For similar raw data, see also Dryer (2009), who endorses Hawkins’s processing explanation.

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

	VO	OV	Totals
ART morph	243	142	385
No ART morph	70	111	181
Totals	313	253	566

Bickel’s (2011; 2013) Family Bias Method (which is particularly suitable to testing hypotheses formulated in diachronic terms). In the supplementary materials to this paper⁹, 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 ($\beta = -0.73$, $p < 0.001$).¹¹ Although the model is not particularly good overall, probably missing important further predictors ($R^2_c = 0.14$, $C = 0.72$), Hawkins’s 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 3 shows that this effect is stable (i.e. never reversed) across all six macro areas. In sum, the global typological picture is consistent with Hawkins’s processing account, even when tested against a more comprehensive data set and with more rigorous modes of examination.

⁹See <http://www.kschmidtkebode.de/publications> or <http://doi.org/10.5281/zenodo.2577480>.

¹⁰All statistical analyses were performed in R 3.3.1 (R Development Core Team 2016). I am grateful to Taras Zakharko and Balthasar Bickel for making their Family Bias algorithm freely available (Zakharko & Bickel 2011ff.).

¹¹All regression analyses I performed are based on generalized linear mixed-effects models that include genealogical and macro-areal dependencies as random effects (see 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.

6 Attractor states and diachronic change in Hawkins’s “Processing Typology”

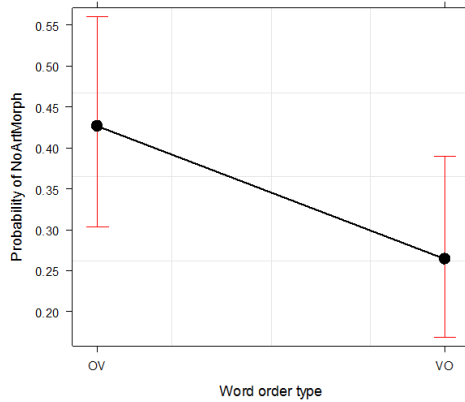


Figure 2: Effect of word order type on the probability of (not) having articles in a mixed-effects model (see SM3 for details)

Table 3: (Rounded) family biases for articles in different word-order types ($N_{\text{total}} = 217$ 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

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 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 (Dryer 2013b) in all VO languages in the same sample as above ($N_{\text{total}} = 278$ languages).

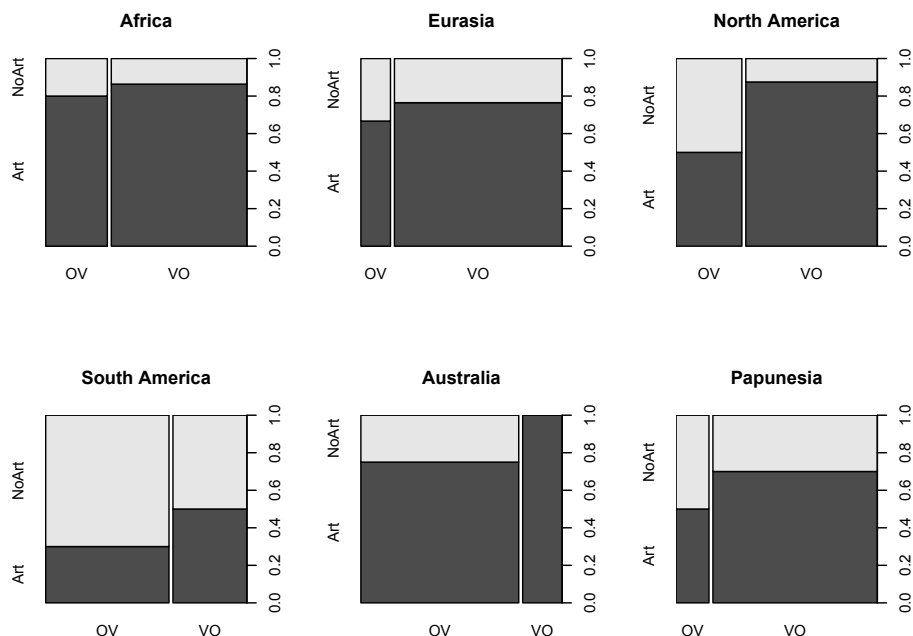


Figure 3: Family biases by macro area (see SM2 for details)

Across several different statistical models (and operationalizations of the hypothesis, see SM4), I did not find support for Hawkins’s efficiency hypothesis. In one analysis, for example, I probed whether 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

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

¹²Moreover, if we look at VO languages which are beginning to use a demonstrative as a definite article (N = 26 in Dryer 2013a), Hawkins’s 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, 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.

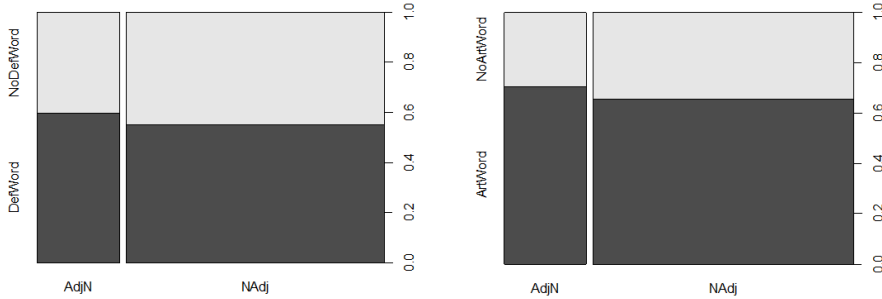


Figure 4: Occurrence of articles in VO languages depending on the position of adjectives (left plot: definiteness words only; right plot: all article-like words; for the corresponding mixed-effects models, see SM4)

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

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

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

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, Ross (2001) discusses an interesting case of an intense contact situation in which the Austronesian language Takia adapted its VO syntax

¹³Bickel (2013) 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.

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 (see 4a), which would be fully in line with Hawkins's prediction. 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 (see 4b-c).

- (4) a. Proto-Western Oceanic (Ross 2001: 142)
a tam^Wata a-ña
DET man that-3SG
'that man'
- b. Takia (Austronesian, Oceanic; Ross 2001: 140)
Waskia tamol an
Waskian man that
'that Waskia man'
- c. Waskia (Nuclear Trans New Guinea, Madang; Ross 2001: 140)
Waskia kadi mu
Waskia man that
'that Waskia man'

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

¹⁴Diachronic research has actually put forward a number of plausible candidates for such factors. A prominent one since at least Vennemann (1975) is the loss of a case system and the concomitant rigidification of constituent order, which favours the development of articles to express information-structural distinctions that were previously coded by a more flexible word order (see also Hawkins 2004; Hewson & Bubenik 2006; Fischer 2010; Carlier & Lamiroy 2014). Another possible factor is the loss of an aspectual system (Abraham 1997; Leiss 2000; 2007). However, especially the former type of explanation is often viewed critically (e.g. Selig 1992 on Romance; McColl Millar 2000 on English; Leiss 2000 on Germanic), and there is currently no proposal as to how various factors may conspire to explain the synchronic distribution of articles (see also Lüdtke 1991). For some further information and preliminary typological analyses of these factors, interested readers are kindly referred to SM5.

We conclude, then, that Hawkins (2014) 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 (2019 [this volume]) or a "weak cognitive bias" with "significant population-level consequences" (Thompson et al. 2016: 4530).

4 Concluding remarks

The present contribution has taken a closer look at John Hawkins's "processing typology", a research programme that fully subscribes to functional-adaptive motivations for grammatical structure. In the first part of the paper, I discussed where such motivations are possibly operative in diachronic change. In my view, a case can be made for Hawkins's 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's 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 under specific structural conditions. In the brief case study presented here, we saw that Hawkins's NP processing hypothesis provides a neat match to the global typological data, even when these are analyzed in more rigorous and hence more appropriate ways than in Hawkins (2014). 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 Culbertson et al. 2012; see also Levshina 2019 [this volume]). As a matter of fact, a particularly strong aspect of Hawkins's work (especially in Hawkins 2004; 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.

References

- Abraham, Werner. 1997. The interdependence of case, aspect, and referentiality in the history of German: The case of the verbal genitive. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 29–61. Cambridge: Cambridge University Press.
- Aitchison, Jean. 2013. *Language change: Progress or decay?* 4th edn. Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511809866.018

- Andersen, Henning (ed.). 2001. *Actualization: Linguistic change in progress*. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/cilt.219
- Auer, Peter. 2009. On-line syntax: Thoughts on the temporality of spoken language. *Language Sciences* 31. 1–13. DOI:10.1016/j.langsci.2007.10.004
- Bates, Elizabeth & Brian MacWhinney. 1989. Functionalism and the competition model. In Brian MacWhinney & Elizabeth Bates (eds.), *The cross-linguistic study of sentence processing*, 3–73. New York: Cambridge University Press.
- Bauer, Winifred. 1993. *Maori*. Abingdon; New York: Routledge.
- Beckner, Clay & Joan L. Bybee. 2009. A usage-based account of constituency and reanalysis. *Language Learning* 59(1). 27–46. DOI:10.1111/j.1467-9922.2009.00534.x
- Bickel, Balthasar. 2011. Statistical modeling of language universals. *Linguistic Typology* 15. 401–413. DOI:10.1515/lity.2011.027
- Bickel, Balthasar. 2013. Distributional biases in language families. In Balthasar Bickel, Lenore A. Grenoble, David A. Peterson & Alan Timberlake (eds.), *Language typology and historical contingency*, 415–444. Amsterdam, Philadelphia: John Benjamins. DOI:10.5167/uzh-86870
- Blevins, James P. & Juliette Blevins (eds.). 2009. *Analogy in grammar: Form and acquisition*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199547548.001.0001
- Blythe, Richard A. & William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88(2). 269–304. DOI:10.2307/23251832
- Broadwell, George Aaron. 2006. *A Choctaw reference grammar*. Lincoln; London: University of Nebraska Press.
- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/tsl.9
- Bybee, Joan L. 1988. The diachronic dimension in explanation. In John A. Hawkins (ed.), *Explaining language universals*, 350–379. Oxford: Blackwell.
- Bybee, Joan L. 2008. Formal universals as emergent phenomena: The origins of structure preservation. In Jeff Good (ed.), *Linguistic universals and language change*, 108–121. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199298495.003.0005
- Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511750526.011
- Bybee, Joan L. & Clay Beckner. 2015. Emergence at the cross-linguistic level: Attractor dynamics in language change. In Brian MacWhinney & William O'Grady (eds.), *The handbook of language emergence*, 183–200. Oxford: Blackwell. DOI:10.1002/9781118346136.ch8

- Carlier, Anne & Béatrice Lamiroy. 2014. The grammaticalization of the prepositional partitive in Romance. In Silvia Luraghi & Tuomas Huumo (eds.), *Partitives and related categories*, 477–519. Berlin: Mouton de Gruyter. DOI:10.1515/9783110346060.477
- Collins, Jeremy. 2019. Some language universals are historical accidents. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 47–61. Berlin: Language Science Press. DOI:10.5281/zenodo.2583808
- Creissels, Denis. 2018. The Obligatory Coding Principle in diachronic perspective. In Sonia Cristofaro & Fernando Zúñiga (eds.), *Typological hierarchies in diachrony and diachrony*, 59–109. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/tsl.121.02cre
- Cristofaro, Sonia. 2017. Implicational universals and dependencies. In Nick J. Enfield (ed.), *Dependencies in language: On the causal ontology of linguistic systems*, 9–22. Berlin: Language Science Press. DOI:10.5281/zenodo.573777
- Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 25–46. Berlin: Language Science Press. DOI:10.5281/zenodo.2583806
- Croft, William. 2000. *Explaining language change: An evolutionary approach*. London: Harlow. DOI:10.1075/jhp.6.1.09rau
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780198299554.001.0001
- Croft, William. 2003. *Typology and universals*. 2nd edn. Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511840579
- Croft, William. 2006. Evolutionary models and functional-typological theories of language change. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 68–91. Oxford: Blackwell. DOI:10.1002/9780470757048.ch4
- Culbertson, Jennifer, Paul Smolensky & Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122(3). 306–329. DOI:10.1016/j.cognition.2011.10.017

- Cysouw, Michael. 2010. Dealing with diversity: Towards an explanation of NP-internal word order frequencies. *Linguistic Typology* 14(2–3). 253–286. DOI:10.1515/lity.2010.010
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511993473
- De Smet, Hendrik. 2012. The course of actualization. *Language* 88(3). 601–633. DOI:10.1353/lan.2012.0056
- Diessel, Holger. 2011. Review article of *Language, usage and cognition* by Joan L. Bybee. *Language* 87(4). 830–844. DOI:10.2307/41348862
- Diessel, Holger. 2016. *Word order correlations: Grammaticalization, nominalization, and analogy*. Paper presented at the Workshop on Mechanisms of Grammatical Change, Wissenschaftskolleg zu Berlin.
- Dryer, Matthew S. 2005. Definite articles. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 154–157. Oxford: Oxford University Press.
- Dryer, Matthew S. 2009. The branching direction theory of word order correlations revisited. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 185–208. Dordrecht: Springer. DOI:10.1007/978-1-4020-8825-4_10
- Dryer, Matthew S. 2013a. Definite articles. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/37>, accessed 2016-12-12.
- Dryer, Matthew S. 2013b. Order of adjective and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/87>, accessed 2016-12-18.
- Dryer, Matthew S. 2013c. Order of object and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/83>, accessed 2018-12-12.
- Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson & Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473. 79–82. DOI:10.1038/nature09923
- Enfield, N. J. 2014. *Natural causes of language: Frames, biases, and cultural transmission*. Berlin: Language Science Press. DOI:10.17169/langsci.b48.77
- Fischer, Susann. 2010. *Word-order change as a source of grammaticalisation*. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/la.157

- Fox, Barbara A. & Sandra A. Thompson. 2007. Relative clauses in English conversation: Relativizers, frequency, and the notion of construction. *Studies in Language* 31(2). 293–326. DOI:10.1075/sl.31.2.03fox
- Gahl, Susanne & Susan M. Garnsey. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80(4). 748–775. DOI:10.1353/lan.2004.0185
- Gentner, Dedre & Linda Smith. 2012. Analogical reasoning. In Vilayanur S. Ramachandran (ed.), *Encyclopedia of human behavior*, 2nd edn., 130–136. Amsterdam: Elsevier. DOI:10.1007/978-3-319-47829-6_1076-1
- Good, Jeff. 2008. Introduction. In Jeff Good (ed.), *Linguistic universals and language change*, 1–19. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199298495.003.0001
- Haig, Geoffrey. 2018. The grammaticalization of object pronouns: Why differential object indexing is an attractor state. *Linguistics* 56(4). 781–818. DOI:10.1515/ling-2018-0011
- Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819. DOI:10.2307/413373
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin, New York: Mouton de Gruyter. DOI:10.1515/9783110884210
- Haspelmath, Martin. 1999. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18(2). 180–205. DOI:10.1515/zfsw.1999.18.2.180
- Haspelmath, Martin. 2008. Creating economical morphosyntactic patterns in language change. In Jeff Good (ed.), *Language universals and language change*, 185–214. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199298495.003.0008
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80. DOI:10.1515/flin.2011.002
- Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by constraints on change? In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 1–23. Berlin: Language Science Press. DOI:10.5281/zenodo.2583804
- Hawkins, John A. 1986. *A comparative typology of English and German: Unifying the contrasts*. London, Sydney: Croom Helm. DOI:10.4324/9781315687964
- Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press. DOI:10.1017/CBO9780511554285

6 *Attractor states and diachronic change in Hawkins's "Processing Typology"*

- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199252695.001.0001
- Hawkins, John A. 2007. Processing typology and why psychologists need to know about it. *New Ideas in Psychology* 25. 87–107. DOI:10.1016/j.newideapsych.2007.02.003
- Hawkins, John A. 2012. The drift of English towards invariable word order from a typological and Germanic perspective. In Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford handbook of the history of English*, 622–632. Oxford: Oxford University Press. DOI:10.1093/oxfordhb/9780199922765.013.0053
- Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199664993.001.0001
- Hewson, John & Vit Bubenik. 2006. *From case to adposition: The development of configurational syntax in Indo-European languages*. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/cilt.280
- Jaeger, T. Florian. 2010. Redundancy and reduction: How speakers manage syntactic information density. *Cognitive Psychology* 61(1). 23–62. DOI:10.1016/j.cogpsych.2010.02.002
- Jaeger, T. Florian, Peter Graff, William Croft & Daniel Pontillo. 2011. Mixed effect models for genetic and areal dependencies in linguistic typology. *Linguistic Typology* 15(2). 281–320. DOI:10.1515/lity.2011.021
- Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.
- Keller, Rudi. 1994. *Language change: The invisible hand in language*. London: Routledge.
- Kemmer, Suzanne & Michael Barlow. 2000. Introduction: A usage-based conception of language. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, vii–xxviii. Stanford, CA: Center for the Study of Language & Information.
- Kurumada, Chigusa & Scott Grimm. 2017. Communicative efficiency in language production and learning: Optional plural marking. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, 2500–2506. <http://www.sas.rochester.edu/lin/sgrimm/publications/Kurumada%20Grimm2017.pdf>, accessed 2017-7-3.
- Kurumada, Chigusa & T. Florian Jaeger. 2015. Communicative efficiency in language production: Optional case-marking in Japanese. *Journal of Memory and Language* 83. 152–178. DOI:10.1016/j.jml.2015.03.003
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Vol. I: Theoretical prerequisites*. Stanford: Stanford University Press.

- Leiss, Elisabeth. 2000. *Artikel und Aspekt: Die grammatischen Muster von Definitheit*. Berlin: Mouton de Gruyter. DOI:10.1515/9783110825961
- Leiss, Elisabeth. 2007. Covert patterns of definiteness/indefiniteness and aspectuality in Old Icelandic, Gothic, and Old High German. In Elisabeth Stark, Elisabeth Leiss & Werner Abraham (eds.), *Nominal determination: Typology, context constraints and historical emergence*, 73–102. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/slcs.89.06lei
- Levshina, Natalia. 2019. Linguistic Frankenstein, or How to test universal constraints without real languages. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 203–221. Berlin: Language Science Press. DOI:10.5281/zenodo.2583820
- Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 106. 1126–1177. DOI:10.1016/j.cognition.2007.05.006
- Lüdtke, Helmut. 1991. Überlegungen zur Entstehung des bestimmten Artikels im Romanischen. *Linguistica* 31. 81–97. DOI:10.4312/linguistica.31.1.81-97
- McColl Millar, Robert. 2000. Some suggestions for explaining the origin and development of the definite article in English. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 275–310. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/slcs.53.14mil
- Nettle, Daniel. 1999. *Linguistic diversity*. Oxford: Oxford University Press.
- Norcliffe, Elizabeth, Alice C. Harris & T. Florian Jaeger. 2015. Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience* 30(9). 1009–1032. DOI:10.1080/23273798.2015.1080373
- Norcliffe, Elizabeth & T. Florian Jaeger. 2016. Predicting head-marking variability in Yucatec Maya relative clause production. *Language and Cognition* 8. 167–205. DOI:10.1017/langcog.2014.39
- Ogura, Mieko. 1993. The development of periphrastic *do* in English: A case of lexical diffusion in syntax. *Diachronica* X. 51–85. DOI:10.1075/dia.10.1.04ogu
- R Development Core Team. 2016. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.r-project.org>.
- Rosenbach, Anette. 2008. Language change as cultural evolution: Evolutionary approaches to language change. In Regine Eckardt, Gerhard Jäger & Tonjes Veenstra (eds.), *Variation, selection, development: Probing the evolutionary model of language change*, 23–74. Berlin, New York: Mouton de Gruyter. DOI:10.1515/9783110205398.1.23

- Ross, Malcolm. 2001. Contact-induced change in Oceanic languages in North-West Melanesia. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Areal diffusion and genetic inheritance*, 134–166. Oxford: Oxford University Press.
- Seiler, Guido. 2006. The role of functional factors in language change: An evolutionary approach. In Ole Nedergaard Thomsen (ed.), *Competing models of linguistic change, evolution and beyond*, 163–182. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/cilt.279.11sei
- Selig, Maria. 1992. *Die Entwicklung der Nominaldeterminanten im Spätlatein: Romanischer Sprachwandel und lateinische Schriftlichkeit*. Tübingen: Narr.
- Seržant, Ilja A. 2019. Weak universal forces: The discriminatory function of case in differential object marking systems. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 149–178. Berlin: Language Science Press. DOI:10.5281/zenodo.2583816
- Sinnemäki, Kaius. 2014. A typological perspective on differential object marking. *Linguistics* 52(2). 281–313. DOI:10.1515/ling-2013-0063
- Smith, John Charles. 2001. Markedness, functionality and perseveration in the actualization of a morphosyntactic change. In Henning Andersen (ed.), *Actualization: Linguistic change in progress*, 203–224. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/cilt.219.10smi
- Thompson, Bill, Simon Kirby & Kenny Smith. 2016. Culture shapes the evolution of cognition. *PNAS* 113(16). 4530–4535. DOI:10.1073/pnas.1523631113
- Timberlake, Alan. 1977. Reanalysis and actualization in syntactic change. In Charles N. Li (ed.), *Mechanisms of syntactic change*, 141–177. Austin: University of Texas Press.
- Trudgill, Peter. 2004. *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh: Edinburgh University Press.
- van Gelderen, Elly. 2014. *A history of the English language*. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/z.183
- Vennemann, Theo. 1975. An explanation of drift. In Charles Li (ed.), *Word order and word order change*, 269–305. Austin: University of Texas Press.
- Wang, William S-Y. 1969. Competing changes as a cause of residue. *Language* 45(1). 9–25. DOI:10.2307/411748
- Wasow, Thomas, T. Florian Jaeger & David M. Orr. 2011. Lexical variation in relativizer frequency. In Horst J. Simon & Heike Wiese (eds.), *Expecting the unexpected: Exceptions in grammar*, 175–195. Berlin, New York: Mouton de Gruyter. DOI:10.1515/9783110219098.175

Karsten Schmidtke-Bode

Zakharko, Taras & Balthasar Bickel. 2011ff. *Familybias: Family bias estimation. R package*. <https://github.com/IVS-UZH>, accessed 2016-11-17.