University of Copenhagen, 2020 February 28 (Symposium to honour Peter Harder)

Grammaticalization and the efficiency theory of asymmetric coding

MARTIN HASPELMATH
MPI-SHH Jena and Leipzig University

1. The efficiency theory of asymmetric coding in grammar

– When asymmetric coding is cross-linguistically systematic, it is explained by frequency-induced predictability (Greenberg 1966; Croft 1990/2003; Diessel 2019).

– Grammatical coding is efficient in the sense that it minimizes speaker effort and maximizes clarity for the hearer: Longer forms are used when the meanings to be conveyed are less predictable.

Table 1: Examples of grammatical coding asymmetries

FREQUENT MEANING:	RARER MEANING:	
singular	plural	(*book – book-**s***)
nominative (A/S)	accusative (P)	(*he – hi-**m***)
allative	ablative	(*to – **from***)
positive	comparative	(*small – small-**er***)
present	future	(*go – **will** go*)
affirmative	negative	(*go – **don't** go*)
predicative verb	predicative adjective	((*they*) *play – (they) **were** small*)

– Grammatical coding is often symmetric, because of the competing motivation of explicitness. But when it is **universally asymmetric**, the explanation is always efficiency.

– There is no need to appeal to "markedness" (Haspelmath 2006) or "iconicity of complexity" (Haspelmath 2008).

2. Two dynamic views of how we can understand grammar

(A) understanding grammar by understanding grammatical change

mutational constraints: constraints on possible changes
(e.g. Bybee 2006)

(B) understanding grammar by understanding efficiency of language use

functional-adaptive constraints: constraints on possible results
(e.g. Hawkins 2014)

Plus: a non-dynamic way: through **biocognitive constraints**
(perhaps an innate "grammar toolkit"; Jackendoff 2002)

2.1. Mutational constraints

"the true universals of language are not synchronic patterns at all, but the mechanisms of change that create these patterns"
(Bybee 2006: 179; also Bybee 2003, the original version)

Bickel (2007: 240):
"It is a matter of current debate whether universal preferences result
(a) from preference principles that **guide** (or "select") the result of diachrony (Kirby 1999, Nettle 1999, Haspelmath 1999b),
(b) or from locally motivated preferred **pathways of change** (Bybee 2001, Blevins 2004, grammaticalization literature)"

Plank (2007): **achronic** laws vs. **diachronic** laws

S. A. Anderson (2016):
"there are no (or at least very few) substantive universals of language, and the regularities arise from **common paths of diachronic change** having their basis in factors outside of the defining properties of the set of cognitively accessible grammars"

Cristofaro (2019: 27):
"Cross-linguistically recurrent grammatical configurations do not appear to arise because of principles that favour those particular configurations in themselves... [This calls for] **a source-oriented approach to typological universals**, in which the patterns described by individual universals are accounted for in terms of **the actual diachronic processes that give rise to the pattern**, rather than the synchronic properties of the pattern in itself."

Bybee's big insight of the 1980s: grammatical markers of tense, aspect and modality develop in **recurrent** ways across languages (Bybee 1985; Bybee & Dahl 1989; Bybee 2006):

I. The perfective path

(i) "be," "have" + PP > RESULTATIVE
(ii) "come (from)"
(iii) "finish" > COMPLETIVE
} > ANTERIOR > PERFECTIVE/PAST

II. The present/imperfective path

(i) "be located at"
(ii) "movement while"
(iii) reduplication
} > PROGRESSIVE > PRESENT/IMPERFECTIVE

III. Future

(i) "want"
(ii) "movement towards"
(iii) "soon," "after"
} > INTENTION > FUTURE

The best-known constraint:
Grammaticalization is unidirectional/irreversible
– degrammaticalization does **not** exist (Lehmann 2015[1982]; Haspelmath 1999; 2004; but see Norde 2009)

A phonological constraint:
nasal vowels **only** develop from nasalization before nasal stops

VN > ṼN > Ṽ > V

This explains that nasal vowels occur only in languages with oral vowels and nasal stops and that nasal vowels are less frequent than oral vowels (Bybee 2006, citing Greenberg 1969).

But other "common paths of change" do **not** have corresponding mutational constraints:

– do perfective/past forms **only** develop from anteriors?
(no, they can come from earlier past forms, like the Germanic *-ed* past form)

– do future forms **only** develop from intention forms?
(no, they can also come from old presents, Haspelmath 1998b)

2.2. Functional-adaptive constraints

Even though all individual languages show many idiosyncrasies (and even downright dysfunctional patterns),
what is general across languages makes good functional sense, in most cases.

For example:

• **Vowel systems** tend to show dispersion that is optimal from an acoustic point of view (cf. Gordon 2016: 59)

TABLE 3.3. Most common vowel inventories of different sizes compared with those of the same size predicted to occur by Liljencrants and Lindblom (1972)

Most common	Liljencrants and Lindblom predicted
3 vowels	
i u a — 19 of 24	i u a
4 vowels	
i e o a — 9 of 34	i u ɛ a
i u e a — 7 of 34	
i u ə a — 6 of 34	
5 vowels	
i u e o a — 113 of 130	i u ɛ æ/a ɑ
6 vowels	
i u e ə o a — 30 of 84	i ʉ u ɛ ɔ a

• **Lexical systems** show a kind of distribution of meanings over words that is efficient. For example, words for 'snow' and 'ice' are distinct primarily where the temperatures are low, i.e. where snow and ice are frequent occurrences (Regier et al. 2016)

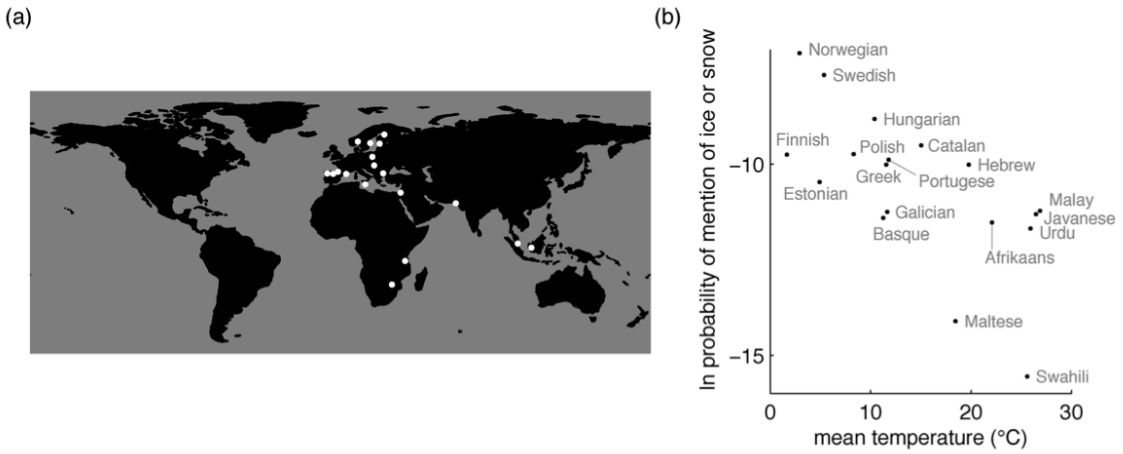

Fig 3. Results of cross-language Twitter analysis. (a) Locations associated with the 18 languages represented in the Twitter dataset we consider. (b) The natural log probability of mention of ice or snow in a given language as a function of the mean temperature where that language is spoken. Mixed effects logistic regression revealed that temperature is negatively associated with probability of mention of ice or snow, as a fixed effect ($\beta = -0.29, \chi^2(1) = 21696$, $p < 10^{-15}$), when including random intercepts for each language family.

• **Word orders** tend to favour efficient constituent recognition, as observed by Hawkins (1994; 2014)

(4) a. *The woman* $_{VP}$[*waited* $_{PP1}$[*for her son*] $_{PP2}$[*in the cold but not unpleasant wind*]].
1 2 3 4 5

---

b. *The woman* $_{VP}$[*waited* $_{PP2}$[*in the cold but not unpleasant wind*] $_{PP1}$[*for her son*]].
1 2 3 4 5 6 7 8 9

---

3. The ubiquity of coding asymmetries in grammar

Table 2: More examples of grammatical coding asymmetries

FREQUENT MEANING:	RARER MEANING:	
singular	plural	*book – book-**s***
nominative (A/S)	accusative (P)	*he – hi-**m***
allative	ablative	*to – **from***
positive	comparative	*small – small-**er***
present	future	*go – **will** go*
affirmative	negative	*go – **don't** go*
instrumental	comitative	Welsh *a – **gyda***
male	female	German *König – König-**in***
cardinal	ordinal	*seven – seven-**th***
present tense	past tense	*play – play-**ed***
active	passive	*plays – **is** play**ed***
basic	applicative	German *fahren – **be**-fahren*
disjoint	reflexive	*her – her**self***
predicative verb	predicative adjective	*(they) play – (they) **were** small*)
inanimate object	animate object	Spanish *veo la casa – veo **a** la mujer*

person-downstream	person-upstream	French *il me le présente – *me lui (**à lui**)*
introverted reflexive	extroverted reflexive	Russian *myli-**s'** /nenavideli **sebja***
inalienable possessor	alienable possessor	Maltese *id-i – il-ktieb **tiegħi***
3rd person indicative	2nd person	Spanish *canta-Ø – canta-**s***
2nd person imperative	3rd person imperative	Turkish *bak – bak-**sın***
noncausal	causal	French *bouillir – **faire** bouillir*
same-subject 'want'	different-subject 'want'	German *(ich will) gehen – **dass er** geht*

Asymmetric coding is also found in the lexicon, e.g.

horse	*vs.*	*hippopotamus*	
car	*vs.*	*cabriolet*	
church	*vs.*	*cathedral*	(frequent words are shorter: Zipf 1935; 1949)

4. Bybee vs. Jespersen: Mutational vs. functional-adaptive constraints

Bybee (2006: 179; 191):

> "[T]he true universals of language are not synchronic patterns at all, but **the mechanisms of change that create these patterns** ... These mechanisms create paths of change which are often similar cross-linguistically. As a by-product of these paths, synchronic states may also bear some resemblance to one another."

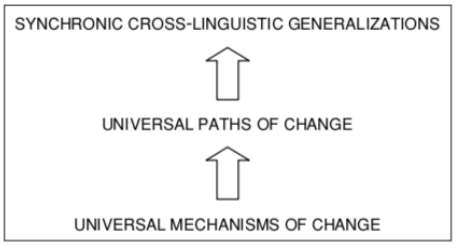

Figure 8.1 The relations among three tiers of cross-linguistic phenomena.

Bybee (2012): domain-general processes:
"sequential processing, neuromotor automatization, categorization and inference-making"

Jespersen (1941: 15-17):

> "In linguistic changes we see the constant interplay of two opposite tendencies, one of an individual, and the other of a social character, one towards **ease** and the other toward **distinctness**. The former is the **tendency to take things easy** and to follow the line of least resistance–to say it bluntly, an outcome of human indolence or laziness... The opposite tendency is **an effort to be clearly and precisely understood**, and to make as vivid and convincing an impression on the hearer as possible; each articulation is therefore made slowly and distinctly, and great exertion is made to choose the most lucid and forcible expression... In extreme cases this may lead to pompousness and over-emphasis."

- language structure tends to be a compromise between ease of production and ease of perception

ease of production	**ease of perception**	
Bequemlichkeit	Deutlichkeit	von der Gabelentz 1891
force of unification	force of diversification	Zipf 1949
economy	clarity	Malkiel 1968
signal simplicity	perceptual optimality	Langacker 1977
economy	iconicity	Haiman 1983
R principle	Q principle	Horn 1984
markedness	faithfulness	Optimality Theory

Now what is the role of grammaticalization?

Bybee (2006; 2012):
grammaticalization is one of the kinds of changes that occur because of the domain-general mechanisms

Jespersen (1941): grammaticalization plays almost no role

BUT: Haspelmath (1998a: 321), influenced by Lüdtke (1980; 1986)

> "there are two dimensions along which synchronic variation exists: a phonetic dimension (cf. 3a-f), and a syntactic-semantic dimension (cf. 3g-h). As Lüdtke (1980, 1986) has emphasized, both these dimensions are open on one side and have a natural limit on the other side, but they differ crucially in that the phonetic dimension is open-ended toward the pole of **ease of production**, whereas the syntactic-semantic dimension is open-ended toward the pole of **ease of perception**."

(3)

…			
a.	ˌkāsḍmaˈkɔm		
b.	ˈkansḍə ma ˈkɔmɲ		
c.	ˈkansḍə mal ˈkɔmɲ		
d.	ˈkansḍ ḍu maːl ˈkɔmɲ	g.	**Kann du mal komm?*
e.	ˈkanst duː maːl ˈkɔmən =	h.	*Kannst du mal kommen?*
f.	*ˈkanəst ˈduo ˈmahal ˈkɔmɔn	i.	*Könntest du vielleicht mal kommen?*
		j.	*Könnten Sie bitte mal herkommen?*
		h.	*Wären Sie bitte so freundlich, zu mir herüber zu kommen?*
		…	

ease of production ↓ ease of perception

This would mean that Jespersen-style **efficiency considerations** actually explain the unidirectionality of grammaticalization.

This is very different from Bybee's explanation in terms of domain-general mechanisms, where efficiency plays no role.

However: Haspelmath (1999a) proposes a different explanation of unidirectionality, in terms of extravagance

> cf. Jespersen's
> "great exertion is made to choose the most lucid and forcible expression... In extreme cases this may lead to **pompousness** and over-emphasis"

Haspelmath (1999a) attributes such "pompousness" not to the need for clarity, but to the **motivation of extravagance.** This is because commonly, new constructions evidently arise without any need, e.g. the English *gonna*-Future, or the German *von*-Genitive.

5. Can the constraints on grammaticalization explain universal results?

No: grammaticalization arises because of **extravagance**, and its properties do not explain efficient asymmetric coding

it is an **inflationary** process (Dahl 2001), and in no way result-oriented

On the contrary: inflationary processes are typically **disruptive**, and they may create strange synchronic patterns

e.g. English *you* 'polite address' > normal 2nd person pronoun

	though com-est	>	*you come*	*(old)*
	she com-es	>	*she come-es*	*(current)*

e.g. German

	ich kam	*ich bin gekommen*	
	'I came'	'I have come'	(old)
>	'I came' (rare)	'I came' (common)	(current)

But these strange patterns are highly unusual – we can (partially) understand them diachronically, but they do not reflect any tendencies.

But what we want to explain is the striking universal tendencies of uniform asymmetric coding!

> And there is no explanation why, for example, there are few freshly grammaticalized *present-tense auxiliaries*
> (e.g. "I present to know" = 'I know at present')

conclusion: the unidirectionality of grammaticalization (a kind of mutational constraint) cannot explain systematically asymmetric coding

6. Functional-adaptive constraints do not lead us to expect uniform mechanisms of change

Functional-adaptive constraints are **result-oriented** – the path by which the result has been achieved is irrelevant for the result (Haspelmath 2019). What matters is that the system works efficiently.

If it is plausible that synchronic generalizations are due to functional constraints (due to functional adaptation), **we do not expect uniform ways** in which the results have come about.

Cf. evolutionary biology:
wings are adaptive, and we do not expect that wings arise in uniform ways
(wings of birds, bats and insects have diverse origins and arose by diverse paths of change)

Shortness of frequent words is functionally adaptive – and there are diverse paths to shortness:

Zipf (1935): shorter words are shorter because of **clipping** (e.g. *laboratory* > *lab*)

Bybee (2007: 12):

> "My own view of Zipf's finding ... is that high-frequency words undergo reductive changes at a faster rate than low-frequency words... the major mechanism is **gradual phonetic reduction**."

But in most cases, rarer words are longer because they are complex elements, consisting of multiple morphs, e.g.

horse	*vs.*	*hippopotamus*
car	*vs.*	*cabriolet*
church	*vs.*	*cathedral*

The idea of allowing multiple paths to the "desired result" may sound too **teleological** – is grammatical change really goal-oriented?

Ronneberger-Sibold (2014) on "shortening techniques" in German:

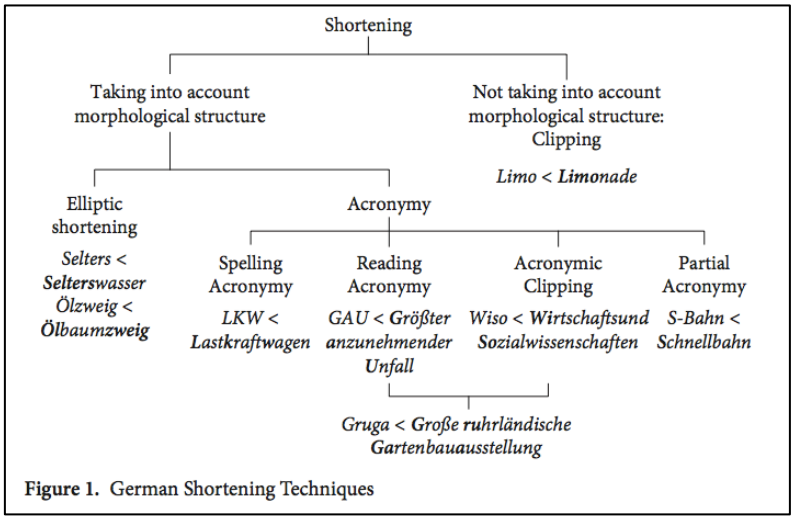

Figure 1. German Shortening Techniques

Thus, occasionally there may even be goal-directed processes, but most of the time, the relevant processes are unconscious (cf. Keller 1994).

> Jespersen (1941: 44)
> "In such cases there can have been no actual wish to improve language (his mother-tongue) on the part of the speaker of the moment, but **his general wish to be understood** as fully and unmistakably as possible... **may gradually lead to** giving up altogether the infelicitous expression."
>
> (1941: 22)
> "most changes are produced inadvertently, and yet they may aid to bring about something that may be called beneficial... Even a **long cumulation through centuries** of small changes... may constitute a considerable gain to the language in question..."

Of course, an asymmetric pattern may arise through **grammaticalization**, e.g.

FREQUENT MEANING:	RARER MEANING:	
present	future	*go – **will** go*
affirmative	negative	*go – **don't** go*
instrumental	comitative	Welsh *a – **gyda***
present tense	past tense	*play – play-**ed***
active	passive	*plays – **is** play**ed***
disjoint	reflexive	*her – her**self***
inanimate object	animate object	Spanish *veo la casa – veo **a** la mujer*
person-downstream	person-upstream	French *il me le présente – *me lui (**à lui**)*
inalienable possessor	alienable possessor	Maltese *id-i – il-ktieb **tiegħi***
noncausal	causal	French *bouillir – **faire** bouillir*

There must be **other pathways**, however, e.g.

FREQUENT MEANING:	RARER MEANING:	
singular	plural	*book – book-**s***
nominative (A/S)	accusative (P)	*he – hi-**m***
allative	ablative	*to – **from***
positive	comparative	*small – small-**er***
male	female	German *König – König-**in***
cardinal	ordinal	*seven – seven-**th***
active	passive	*plays – **is** play**ed***
3rd person indicative	2nd person	Spanish *canta-Ø – canta-**s***

A completely parallel contrast may arise in different ways in different languages, e.g.

coreferential vs. disjoint adpossessive constructions

Danish	*Hun elsker **sin** mand.* 'She$_1$ loves her$_1$ husband.'	*Hun elsker **hendes** mand.* 'She$_1$ loves her$_2$ husband.'
Somersetshire English (Jespersen 1941: 39)	*Bill cut **'s** vinger.* 'Bill$_1$ cut his$_1$ finger.'	*Bill cut **ees** vinger.* 'Bill$_1$ cut his$_2$ finger.'

inalienable vs. alienable adpossessive constructions

Maltese	*id-**i*** 'my hand'	*il-ktieb **tiegħi*** 'my book' (< 'the book my-possession')
Old Tuscan	*moglie-**ma*** 'my wife' (< *mulier mea*)	*terra **mia*** 'my land' (< *terra mea*)

7. Summary

• Many universals of language structure appear to be motivated by **efficiency** of language use, providing an optimal tradeoff between speaker and hearer interests.

• Grammatical marking is often **systematically asymmetric**, throughout all areas of grammar, and these patterns seem to be explainable by efficiency
("the efficiency theory of asymmetric coding in grammar")

• Efficiency can be seen as a **functional-adaptive constraint** that becomes effective in language use, as speakers tend to gradually shift their conventions toward an efficient result.

• In principle, Bybee-style **mutational constraints** may also exist, but they cannot be shown to explain asymmetric coding (they may explain the greater number of oral vowels compared to nasal vowels).

• Jespersen (1941) invoked efficiency in much the same way, but he tried to explain specific instances of language change – this is a speculative enterprise, because specific changes are often accidental and random; it is only **language universals** that we can hope to explain.

• Functional-adaptive explanations need not specify a particular pathway of change – **the "desired" results can come about in various ways** (deliberate change, phonological reduction, expansion via grammaticalization).

References

Anderson, Stephen R. 2005. Morphological universals and diachrony. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 2004*, 1–17. (Yearbook of Morphology). Springer.

Anderson, Stephen R. 2016. Synchronic versus diachronic explanation and the nature of the Language Faculty. *Annual Review of Linguistics* 2(1).

Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11(1). 239–251.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: Benjamins.

Bybee, Joan L. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan L. 2003. Los mecanismos de cambio como universales lingüísticos. In Ricardo Mairal & Juana Gil (eds.), *Entorno a los universales lingüísticos*, 245–263. Cambridge: Cambridge University Press.

Bybee, Joan L. 2006. Language change and universals. In Ricardo Mairal & Juana Gil (eds.), *Linguistic universals*, 179–194. Cambridge: Cambridge University Press.

Bybee, Joan L. 2012. Domain-general processes as the basis for grammar. *The Oxford handbook of language evolution*.

Bybee, Joan L. & Dahl, Östen. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13(1). 51–103.

Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Schmidtke-Bode, Karsten & Levshina, Natalia & Michaelis, Susanne Maria & Seržant, Ilja A. (eds.), *Explanation in typology*, 25–46. Berlin: Language Science Press. (http://langsci-press.org/catalog/book/220)

Croft, William. 1990. *Typology and universals*. Cambridge: Cambridge University Press.

Croft, William. 2003. *Typology and universals*. 2nd edition. Cambridge: Cambridge University Press.

Dahl, Östen. 2001. Inflationary effects in language and elsewhere. In Bybee, Joan B. & Hopper, Paul (eds.), *Frequency and the emergence of linguistic structure*, 471–480. Amsterdam: Benjamins.

Diessel, Holger. 2019. *The grammar network*. Cambridge: Cambridge University Press.

Gordon, Matthew Kelly. 2016. *Phonological typology*. Oxford: Oxford University Press.

Greenberg, Joseph H. 1966. *Language universals: With special reference to feature hierarchies*. The Hague: Mouton.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819. (doi:10.2307/413373)

Haspelmath, Martin. 1998a. Does grammaticalization need reanalysis? *Studies in Language* 22(2). 315–351.

Haspelmath, Martin. 1998b. The semantic development of old presents: New futures and subjunctives without grammaticalization. *Diachronica* 15(1). 29–62.

Haspelmath, Martin. 1999a. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18(2). 180–205.

Haspelmath, Martin. 1999b. Why is grammaticalization irreversible? *Linguistics* 37(6). 1043–1068.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 25–70. (doi:10.1017/S0022226705003683)

Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive linguistics* 19(1). 1–33.

Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by constraints on change? In Schmidtke-Bode, Karsten & Levshina, Natalia & Michaelis, Susanne Maria & Seržant, Ilja A. (eds.), *Competing explanations in linguistic typology*, 1–23. Berlin: Language Science Press. (http://langsci-press.org/catalog/book/220)

Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. New York: Oxford University Press.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jespersen, Otto. 1941. *Efficiency in linguistic change*. Copenhagen: Munksgaard.

Keller, Rudi. 1994. *On language change: The invisible hand in language*. London: Routledge.

Langacker, Ronald W. 1977. Syntactic reanalysis. In Li, Charles (ed.), *Mechanisms of syntactic change*, 57–139. Austin: University of Texas Press.

Lehmann, Christian. 2015. *Thoughts on grammaticalization*. Berlin: Language Science Press. (http://langsci-press.org/catalog/book/88)

Lüdtke, Helmut. 1980. Auf dem Wege zu einer Theorie des Sprachwandels. *Kommunikationstheoretische Grundlagen des Sprachwandels*, 182–252. de Gruyter Berlin.

Lüdtke, Helmut. 1986. Esquisse d'une théorie du changement langagier. *La Linguistique*. Presses Universitaires de France 22(1). 3–46.

Nettle, Daniel. 1999. *Linguistic diversity*. Oxford: Oxford University Press.

Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.

Plank, Frans. 2007. Extent and limits of linguistic diversity as the remit of typology – but through constraints on what is diversity limited? *Linguistic Typology* 11(1). 43–68. doi:10.1515/LINGTY.2007.005.

Regier, Terry & Carstensen, Alexandra & Kemp, Charles. 2016. Languages support efficient communication about the environment: Words for snow revisited. *PLOS ONE* 11(4). e0151138. (doi:10.1371/journal.pone.0151138)

Ronneberger-Sibold, Elke. 2014. Tuning morphosemantic transparency by shortening. In Rainer, Franz & Gardani, Francesco & Luschützky, Hans Christian & Dressler, Wolfgang U. (eds.), *Morphology and meaning: Selected papers from the 15th International Morphology Meeting, Vienna, February 2012*, 275–287. Amsterdam: Benjamins.

Zipf, George K. 1949. *Human behavior and the principle of least effort: An introduction to human ecology*. New York.

Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philogy*. Cambridge: MIT Press.