ICLC 16, Düsseldorf, 2023 August 7-11
(Theme session: "Words and Meanings: Cross-linguistic variability and regularity in the lexicon")

Coexpression and synexpression patterns in lexical and grammatical typology

MARTIN HASPELMATH
Max Planck Institute for Evolutionary Anthropology
and Moscow State University[1]

1. Take-home message 1: The unexplained Root Size Constraint

In all languages, **roots are preferably monosyllabic or bisyllabic**, and longer roots are less preferred the longer they are.

e.g. *dog, cat, horse, bee*
pigeon, beetle, zebra, sparrow

uncommon: *chimpanzee, flamingo, caribou*
caterpillar, alligator, cassowary

But why is this so? There is no good explanation, it seems, but we need one.

2. Take-home message 2: Cross-linguistic comparison conflicts with traditional linguistics

Traditional linguistics (de Saussure, Jakobson, Chomsky…) is about discovering **language structures**.

But cross-linguistic comparison is based on **substance**: phonetic substance and conceptual/functional substance.

(In other words: traditional linguistics is **emic**, comparison is **etic**.)

Thus, cross-linguistic comparison is not as relevant to traditional linguistics as one might hope, and as many people think. It allows us to find **universals** and to make **general statements** about Human Languuage (Haspelmath 2021a), but these have no immediate consequence for the analysis of particular languages.

3. Variable "packaging" of meanings into lexical forms: Colexification and syllexification (Haspelmath 2023a)

– Italian *sentire* packages 'hear' and 'feel' together **colexification**

– English *look* packages 'sight' and 'activity' together **syllexification**

[1] I was a guest scientist at Moscow State University in 1989-1990 and have kept my affiliation ever since.

Cross-linguistic regularities in perception verb colexification

Elisabeth Norcliffe and Asifa Majid
University of Oxford,
elisabeth.norcliffe@psy.ox.ac.uk; asifa.majid@psy.ox.ac.uk

Keywords: Perception verbs, Colexification, Lexical typology, Sensory language

Previous typological studies of perception verbs have suggested there are regularities in how sensory meanings are packaged into words, pointing to our shared biology—in particular, a biological hierarchy of the five senses—as a universal constraint on the lexical expression of sensory meanings (Evans &

"Packaging together" is too vague:

colexification (of two meanings A and B) = expression of either A or B in a root (alternatively)

syllexification (of two meanings A and B) = expression of both A or B in a root (simultaneously)

Crucially, the two meanings A and B are **comparison meanings** – they are etic (conceptual/functional substance), not emic (language-particular structure).

(See Haspelmath 2010; 2018 on **comparative concepts**)

Colexification and syllexification are about **roots** ("minimal lexical forms"), but the phenomenon is more general, extending to grammatical markers:

coexpression (of two meanings A and B) = expression of either A or B in a form (alternatively)

synexpression (of two meanings A and B) = expression of both A or B in a form (simultaneously)

It seems that Vanhove (on Beja, at this conference) means *synexpression* here:

In addition to the two superordinate WALK verbs, *hiːreːr* 'to walk' and *libas* 'to walk at night', three salient manner features are co-expressed in self-propelled WALK verbs: (a) movements of the body, arms, and legs, (b) speed and rhythm, and (c) length of steps, which can combine within one lexical item. In addition, a few verbs also include (d) the aim of locomotion or (e) its direction.

Note that coexpression and synexpression are in some sense opposites:

– ***coexpression*** means that a form **does not differentiate** where it might be expected to differentiate

– ***synexpression*** means that a form **differentiates** where it might be expected not to differentiate

Consider kinship terms (Evans 2011: 509):

(elder)	brother	sister	English
(younger)			
(elder)	kakak		Indonesian
(younger)	adik		
(elder)	ani	otōto	Japanese
(younger)	ane	imōto	

– From the perspective of English, Japanese *ani* **syllexifies** 'brother' and 'elder' (it overdifferentiates).

– From the perspective of Japanese, Indonesian *kakak* **colexifies** 'ani (elder brother)' and 'otōto (elder sister)'

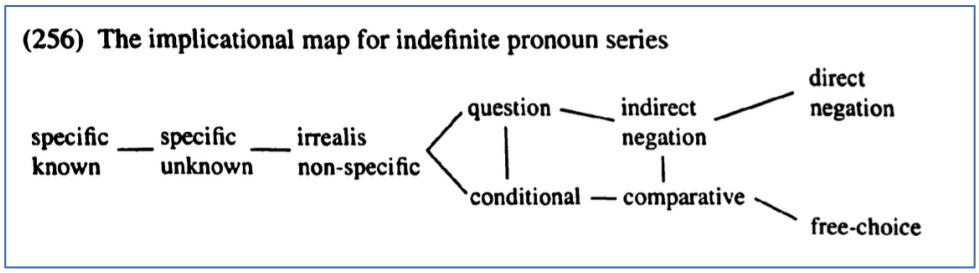
– For a comparison of English and Indonesian, we need to adopt a fine-grained set of comparison meanings (Japanese-style).

4. Only cross-linguistic comparison allows us to state and test universals

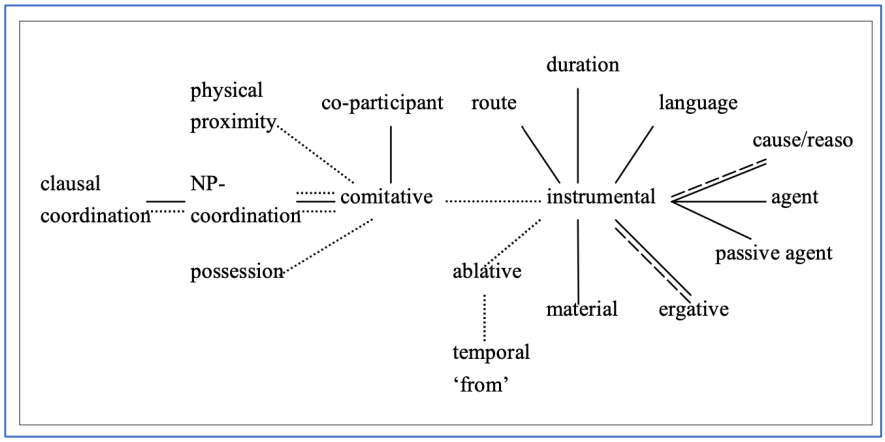
4.1. Universals of colexification (more generally, coexpression)

Expressed in ***coexpression diagrams*** ("semantic maps")

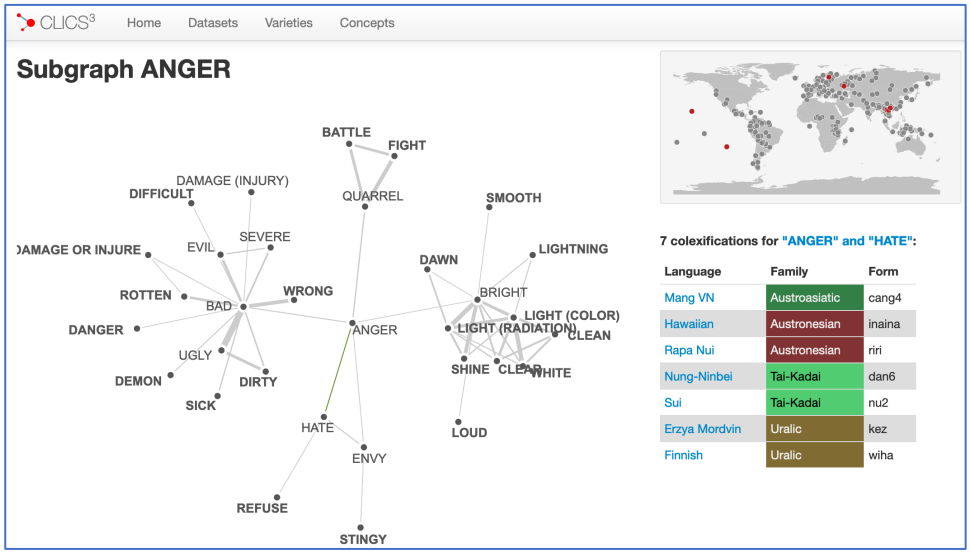
e.g. Haspelmath (1997: 119):



e.g. Narrog & Ito (2007: 282)



CLICS subgraphs: (https://clics.clld.org/graphs/subgraph_82)



4.2. Universals of syllexification (more generally, synexpression)

Syllexification patterns have often been discussed under the heading of "lexicalization patterns" (Talmy 1985; Levin & Rappaport Hovav 2019), but primarily for verbal event representation.

> **Mańczak's Law of Differentiation** (1966: 84)
> More frequently used linguistic elements are generally more differentiated than less frequently used elements.

Table 1. Syllexification in higher-frequency words (Mańczak 1966; 1970)

	highly frequent			***less frequent***		
English	*drink*	*drank*		*consume*	*consum-ed*	
French	*aller*	*va*	'go (INF/3SG)'	*marcher*	*marche*	'walk (INF/3SG)'
French	*père*	*mère*	'father/mother'	*directeur*	*directr-ice*	'director'
Polish	*dwa*	*drugi*	'two/second'	*dziesięć*	*dziesiąt-y*	'ten(th)'
Italian	*buono*	*migliore*	'good/better'	*nuovo*	*più nuovo*	'newe(er)'
Russian	*idët*	*šel*	'goes/went'	*igraet*	*igra-l*	'play(ed)'
German	*Hengst*	*Stute*	'stallion/mare'	*Löwe*	*Löw-in*	'lion(ess)'

4.3. These cross-linguistic generalizations are based on comparison meanings

Thus, the term "polysemy" is not really appropriate, cf. Youn et al. (2016: 1766) ("On the universal structure of human lexical semantics"):

translations reveal cases where a particular language uses a single "polysemous" word to express multiple concepts that another language represents using distinct words. We use the frequency of such polysemies linking two concepts as a measure of their semantic proximity and represent the pattern of these linkages by a weighted network. This network is highly structured: Certain concepts are far more prone to polysemy than others, and naturally interpretable clusters of closely related concepts emerge. Statistical analysis of the polysemies observed in a subset of the basic vocabulary shows that these structural properties are consistent across different language groups, and largely independent of geography, environment, and the presence or absence of a literary tradition. The methods

Polysemy is standardly determined via language-particular tests such as zeugma, e.g.

German *Tasche* 'pocket, bag':

Sie hat das Geld in ihrer Tasche, und ich auch (genauer: in meiner Hosentasche).
'She has the money in her bag/pocket, and me too (more exactly: in my pocket).'

Coexpression is neutral with respect to polysemy or indeterminacy (or even homonymy).

But this also means that colexification studies do not allow us to draw to language-particular conclusions:

cf. Jackson et al. (2019: 1518)

Past research has used colexification patterns across languages to examine the semantic structure of non-emotion concepts. Youn and colleagues coded dictionaries from 81 languages to show that concepts such as "sun," "river," "mountain," and "hill" had universal patterns of colexification that reflected concepts' material and functional properties (*21*).

For instance, languages were more likely to colexify concepts such as "water" and "sea," than concepts such as "sun" and "water," implying that speakers of these languages viewed "water" and "sea" as semantically similar concepts and "sun" and "water" as distinct. We use a similar approach to estimate the variation and structure of emotion semantics across language families.

Traditional linguists often aim for **within-language** generalizations, and these are incompatible with **cross-linguistic** generalizations, cf.

Koptjevskaja-Tamm (at this conference):

A central issue in systematic cross-linguistic or typological comparison, which is dependent on comparable data from different languages, is cross-linguistic identification of phenomena. This, in turn, presupposes a rigorous procedure that ensures we compare like with like. This is, however, difficult to achieve with Conceptual Metaphor Theory's insistence on metaphors as more general conceptual associations that do not boil down to individual metaphorical uses or linguistic convention, but can manifest themselves in many different ways. To quote Gibbs (2015:183), "cognitive linguists, and others, should articulate criteria for identifying metaphoric patterns

Cross-linguistic generalizations must be "**atomistic"**, sacrificing the goal of "doing justice to each language".

5. Lexification typology and grammification typology

Identifying "words" is problematic (but see now Haspelmath 2023b)

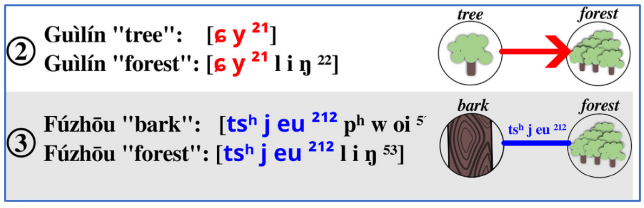
Bentz (at this conference):
orthographic words are (somewhat) arbitrary units of writing (Haspelmath, 2011; Wray, 2015),

Thus, it seems best to use the terms ***root*** and ***(grammatical) marker***, rather than "word":

root =a minimal form that denotes an object, an action, or a property
(and that can occur as part of a free form without another such form)

grammatical **marker** = a minimal bound form that is not a root
(i.e. an affix or a clitic; Haspelmath 2021b; 2023c)

Note also that **partial colexification** (List 2023) concerns the overlap of roots or affixes, i.e. the coexpression always concerns minimal forms:



(from List 2023)

There are thus two subtypes of coexpression and synexpression:

coexpression:	(examples:)
colexification	German *Tasche* colexifies 'bag' and 'pocket'
cogrammification	Latin *-ae* cogrammifies 'genitive' and 'dative'
synexpression:	(examples:)
syllexification	English *bequest* syllexifies 'give' and 'as inheritance'
syngrammification	Latin *-ibus* syngrammifies 'dative' and 'plural'

In addition, François (2022) coined the term **dislexification** (as the opposite of *colexification*), for which we can create the counterpart **circumlexification** (as the opposite of *syllexification*):

E.g. English ***dislexifies*** 'Tasche': *bag, pocket*

German ***circumlexifies*** 'syntax': *Satz-lehre* [sentence-study]

Alternative terms for ***cogrammification***:

morphological syncretism, grammatical polysemy, multifunctionality

Alternative terms for ***syngrammification***:

portmanteau expression, cumulative exponence (sometimes: *conflation*)

6. Explaining coexpression: the role of similarity (?)

For colexification, general explanations have recently become prominent, e.g.

> Xu et al. (2020), Brochhagen & Boleda (2022)
> – in general, the idea is that colexification is due to **semantic similarity**
>
> This echoes Haiman (1974) on cogrammification being due to semantic similarity

Or perhaps a better explanation: **diachronic extendability**

> cf. Cristofaro (2010): coexpressed meanings are often related by **metonymy**,
> not by similarity (e.g. English *while* 'during; whereas')
>
> This also explains why *left*/*right, north*/*south, six*/*seven* are never colexified
> (not because they are "too similar"; pace Brochhagen & Boleda 2022)

The explanation of coexpression would then perhaps be a **mutational explanation**, based on constraints on possible changes (cf. Haspelmath 2019).

7. Toward an explanation of synexpression: the role of frequency

A very puzzling phenomenon: synexpression of "unrelated" meanings

e.g. French	*au garçon*	(*au* [o], < *à le*)	'[to [the boy]]'
e.g. English	*they're coming*	[ðɛə kʌmɪŋ]	'[they [are coming]]'

When meanings are highly frequent, they can be synexpressed even when they are not directly related.

In inflection, this is called *cumulative exponence*, e.g.

Latin	SG	PL	
NOM	*can-is*	*can-es*	'dog(s)'
GEN	*can-is*	*can-um*	
DAT	*can-i*	*can-ibus*	
ACC	*can-em*	*can-es*	
ABL	*can-e*	*can-ibus*	

Note: the number and case meanings are not directly related either – but they occur with high frequency.

Like syngrammification, syllexification occurs primarily in words with high token frequency, e.g. (see above)

Table 1'. Syllexification in higher-frequency words (Mańczak 1966; 1970)

	highly frequent			***less frequent***		
English	*drink*	*drank*		*consume*	*consum-ed*	
French	*aller*	*va*	'go (INF/3SG)'	*marcher*	*marche*	'walk (INF/3SG)'
Polish	*dwa*	*drugi*	'two/second'	*dziesięć*	*dziesiąt-y*	'ten(th)'
Italian	*buono*	*migliore*	'good/better'	*nuovo*	*più nuovo*	'newe(er)'

But how exactly does high frequency explain synexpression?

I would like to suggest:

> Lower-frequency meanings must be circumlexified because of the Root Size Constraint: They are **too rare to be expressed as minimal forms**, because they would be too long.

8. Efficient lexification patterns: Simplicity vs. informativeness?

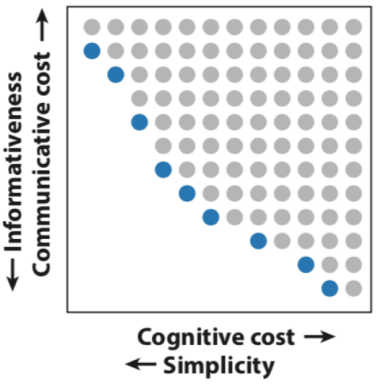
Kemp et al. (2018) (and much recent related work):

Cross-linguistic lexification patterns in multiple domains are **efficient** (they "support efficient communication", "optimize the simplicity/informativeness trade-off"):

– colour terms
– words for snow (Regier et al 2016)
– numerals (Xu et al. 2020b)
– indefinite pronouns (Denić et al. 2022, based on Haspelmath 1997)
– etc.

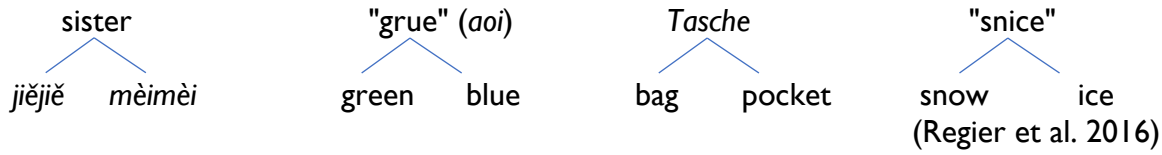
> "At the core of this explanation is the idea that attested numeral systems near-optimally trade off the competing demands of **informativeness** and **simplicity**, given a set of motivated semantic primitives, and a need distribution grounded in linguistic usage." (Xu et al. 2020b: 68)

Kemp et al. (2018):



> "The achievable region of possible semantic systems may be determined by creating a wide range of hypothetical systems out of semantic primitives for a given domain. We then predict that attested semantic systems will tend to lie along or near **the optimal frontier of this region**: the curve defined by systems that are **as informative as possible** for their level of simplicity, and **as simple as possible** for their level of informativeness." (Kemp et al. 2018: 114)

coexpression patterns:



(Regier et al. 2016)

synexpression patterns (more specifically, ***syllexification***):

elephant	young elephant	
dog	*puppy* (= young + dog)	(*puppy* synexpresses both meanings)
fifty	fifty-two	
ten	*twelve* (= ten + two)	(*twelve* syllexifies both meanings)
cousin	female cousin	
sibling	*sister* (= female + sibling)	(*sister* syllexifies both meanings)

What is going on here? I suggest:

> Languages prefer colexifications when the meanings in question are **not** expressed very frequently (= when the communicative need is not high).
>
> Thus, "snice" words are used primarily by societies that do not talk frequently about snow and ice. And *jiějiě*/*mèimèi*-neutralizing kin terms are found especially in languages whose speakers do not use kinship terms frequently.
>
> On the other hand, languages prefer syllexifications when the meanings in question **are** expressed frequently.
>
> For example, 'twelve' is more frequent than 'fifty-two', and 'sister' is more frequent than '(female) cousin'.

What is the role of **frequency of use** here?

9. System complexity ("cognitive cost") vs. coding length ("articulatory cost")

I would like to suggest that it is not cognitive cost, but articulatory cost that trades off with informativeness.

In the simplicity/informativeness literature, the authors seem to have system complexity in mind:

> Regier et al. (2016: 2)
> "From the perspective of efficient communication, a system of fine-grained categories is both more **informative** than a single broad category, and **more complex, requiring more effort to store and process**. The added complexity of a fine-grained system may be worth the investment if the gain in informativeness is compounded by frequent use of the fine-grained categories. This reasoning predicts that semantically fine- grained categories will tend to appear in frequently referenced parts of semantic space."

However, how big is **the cost of storing large numbers of elements**? Since we can store the words of quite a few different languages in our mind, storage space does not seem to be a limitation.

The bottleneck seems to be articulatory speed:

> Levinson & Torreira (2015: 19)
> "Speech production is a bottleneck on the whole language system: at about an average of seven syllables per second, speech can be estimated to have a bit-rate of under 100 bps (Levinson 2000: 28). Studies of language production show that **pre-articulation processes run three or four times faster** than actual articulation (Wheeldon & Levelt 1995). Studies of language comprehension under compression show that people can parse and **comprehend speech at three or four times the speed** of speech production (Calvert 1986: 178; Mehler et al. 1993)."

If frequency of use determines **coding length** (articulatory cost) rather than **system complexity** ("cognitive cost"), we might expect very long opaque expressions for rarely expressed meanings, e.g.

12 =	*twelve*	dog =	*dog*
50 =	*fifty*	elephant =	*elephant*
52 =	**slampingoon*	young elephant =	**ummigrondap*

sister =	*sister*
elder sister =	**cassiluppa*
younger sister =	**tirendombung*

And indeed, we occasionally find very long morphs denoting rarely used concepts, e.g.

cassowary
caterpillar
asparagus

But in general, **languages tend to have monosyllabic or bisyllabic morphs** (occasionally trisyllabic, rarely quadrisyllabic)

– the **Root Size Constraint**

If morphs cannot be too long, this means that less frequently occurring meanings **must be expressed by composite forms**, e.g.

fifty-two
young elephant
elder sister

The overall result is that:

– frequently occurring **grammatical meanings** are expressed by **zero or a short form**
(cf. asymmetric coding universals, *book* vs. *book-**s***; Haspelmath 2021c)

– frequently occurring **lexical meanings** are expressed by **short morphs**
(Zipf's Law of Abbreviation, e.g. *dog* vs. *elephant*)

– frequently cooccurring meanings (lexical or grammatical) are **synexpressed** in unanalyzable morphs (e.g. *snow* 'soft + snice', *puppy* 'dog + young', twelve '10 + 2')

– meanings that do not occur frequently, and hence do not cooccur frequently with other meanings, are lexified by **more coexpressant forms** than frequently occurring meanings

		more coexpressant than	
e.g.	*cousin*	>	*sister* (vs. *brother*)
	elephant	>	*dog* (vs. *puppy*)
	they	>	*he* (vs. *she*)
	sister	>	*jiějiě* (vs. *mèimèi*)
	"snice"	>	*snow* (vs. *ice*)

10. Concluding remarks

– There is no reason to think that with respect to the mapping of meanings onto forms, lexical and grammatical morphs show different cross-linguistic behaviour.

– Thus, *colexification* and *cogrammification* should be considered together (as **coexpression**).

– The inverse of coexpression is **synexpression**: the simultaneous expression of two meanings in a single minimal form.

– Coexpression patterns (whether lexical or grammatical) can be explained by semantic relatedness, maybe more specifically by diachronic extendability.

– Synexpression patterns can be explained by frequency of use: When two meanings occur very frequently together, they tend to be synexpressed.

– In a sense, synexpression can be taken as the default (e.g. in language acquisition), but given the Root Size Constraint, circumexpression must kicked in oncve a certain frequency threshold is no longer reached.

– We do not have an explanation for the Root Size Constraint, but if this is an independent constraint, it explains the frequency condition on synexpression.

References

Brochhagen, Thomas & Boleda, Gemma. 2022. When do languages use the same word for different meanings? The Goldilocks principle in colexification. *Cognition* 226. 105179. (doi:10.1016/j.cognition.2022.105179)

Bybee, Joan L. & Perkins, Revere & Pagliuca, William. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: The University of Chicago Press.

Cristofaro, Sonia. 2010. Semantic maps and mental representation. *Linguistic Discovery* 8(1). 35–52. (doi:10.1349/PS1.1537-0852.A.345)

Denić, Milica & Steinert-Threlkeld, Shane & Szymanik, Jakub. 2022. Indefinite pronouns optimize the simplicity/informativeness trade-off. *Cognitive Science* 46(5). e13142. (doi:10.1111/cogs.13142)

Evans, Nicholas. 2011. Semantic typology. In Song, Jae Jung (ed.), *The Oxford handbook of linguistic typology*, 504–533. Oxford: Oxford University Press.

François, Alexandre. 2022. Lexical tectonics: Mapping structural change in patterns of lexification. *Zeitschrift für Sprachwissenschaft* 41(1). 89–123. (doi:10.1515/zfs-2021-2041)

Haiman, John. 1974. Concessives, conditionals, and verbs of volition. *Foundations of Language* 11(3). 341–359.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press. (https://library.oapen.org/handle/20.500.12657/30819)

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. (doi:10.1353/lan.2010.0021)

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-2017-1005)

Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Van Olmen, Daniël & Mortelmans, Tanja & Brisard, Frank (eds.), *Aspects of linguistic variation: Studies in honor of Johan van der Auwera*, 83–113. Berlin: De Gruyter Mouton. (https://zenodo.org/record/3519206)

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)

Haspelmath, Martin. 2021a. General linguistics must be based on universals (or nonconventional aspects of language). *Theoretical Linguistics* 47(1–2). 1–31. (doi:10.1515/tl-2021-2002)

Haspelmath, Martin. 2021b. Bound forms, welded forms, and affixes: Basic concepts for morphological comparison. *Voprosy Jazykoznanija* 2021(1). 7–28. (doi:10.31857/0373-658X.2021.1.7-28)

Haspelmath, Martin. 2021c. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. (doi:10.1017/S0022226720000535)

Haspelmath, Martin. 2023a. Coexpression and synexpression patterns across languages: Comparative concepts and possible explanations. *Frontiers in Psychology (to appear)*. (https://ling.auf.net/lingbuzz/007402)

Haspelmath, Martin. 2023b. Defining the word. *WORD* (to appear).

Haspelmath, Martin. 2023c. Types of clitics in the world's languages. *Linguistic Typology at the Crossroads (to appear)*. (https://ling.auf.net/lingbuzz/007071)

Kemp, Charles & Xu, Yang & Regier, Terry. 2018. Semantic typology and efficient communication. *Annual Review of Linguistics* 4(1). 109–128. (doi:10.1146/annurev-linguistics-011817-045406)

Levin, Beth & Rappaport Hovav, Malka. 2019. Lexicalization patterns. In Truswell, Robert (ed.), *The Oxford handbook of event structure*, 395–425. Oxford: Oxford University Press.

Levinson, Stephen C. & Torreira, Francisco. 2015. Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*. Frontiers Media SA 6. 731.

List, Johann-Mattis. 2023. Inference of partial colexifications from multilingual wordlists. *Frontiers in Psychology, to appear* (doi:10.3389/fpsyg.2023.1156540)

Mańczak, Witold. 1966. La nature du supplétivisme. *Linguistics* 4(28). 82–89. (doi:10.1515/ling.1966.4.28.82)

Mańczak, Witold. 1970. Sur la théorie die catégories "marquées" et "non-marquées" de Greenberg. *Linguistics* 8(59). 29–36. (doi:10.1515/ling.1970.8.59.29)

Narrog, Heiko & Ito, Shinya. 2007. Re-constructing semantic maps: The comitative-instrumental area. *STUF–Sprachtypologie und Universalienforschung* 60(4/2007). 273–292. (doi:https://doi.org/10.1524/stuf.2007.60.4.273)

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)

Talmy, Leonard. 1985. Lexicalization patterns. In Shopen, Timothy (ed.), *Language typology and syntactic description (Volume III)*, 57–149. Cambridge: Cambridge University Press.

Wray, Alison. 2015. Why are we so sure we know what a word is? In Taylor, John R. (ed.), *The Oxford handbook of the word*. Oxford: Oxford University Press. (doi: 10.1093/oxfordhb/9780199641604.013.032)

Xu, Yang & Duong, Khang & Malt, Barbara C. & Jiang, Serena & Srinivasan, Mahesh. 2020. Conceptual relations predict colexification across languages. *Cognition* 201. 104280. (doi:10.1016/j.cognition.2020.104280)

Xu, Yang & Liu, Emmy & Regier, Terry. 2020. Numeral systems across languages support efficient communication: From approximate numerosity to recursion. *Open Mind* 4. 57–70. (doi:10.1162/opmi_a_00034)

Xu, Yang & Regier, Terry & Malt, Barbara C. 2016. Historical semantic chaining and efficient communication: The case of container names. *Cognitive Science* 40(8). 2081–2094. (doi:10.1111/cogs.12312)

Youn, Hyejin & Sutton, Logan & Smith, Eric & Moore, Cristopher & Wilkins, Jon F. & Maddieson, Ian & Croft, William & Bhattacharya, Tanmoy. 2016. On the universal structure of human lexical semantics. *Proceedings of the National Academy of Sciences* 113(7). 1766–1771. (doi:10.1073/pnas.1520752113)