Tylex course "Explaining language universals", Voronovo (Moscow), 2017 September 3

1. Explanation in research on grammar

MARTIN HASPELMATH
MPI-SHH Jena & Leipzig University

1. Some language universals

(1) Greenberg 1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object. (Greenberg 1963)

(2) Greenberg 14. In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages.

(3) Greenberg 35. There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero.

(4) Greenberg 38. Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb.

(5) All languages have roots denoting things, roots denoting actions, and roots denoting properties (such as dimension, age, or value).

(6) All languages have morphemes denoting negation.

(7) No language has a rule that involves counting elements or features.

(8) If a language has noun-possessor order, it tends to have preposition-NP order, and if it has possessor-noun order, it tends to have NP-postposition order (Dryer 2005).

(9) If a language has OV order, then it tends to have no question-word fronting (Bach's generalization, Bach 1971, Roberts 2007: §1.5.1).

(10) In almost all cases, the ergative case is overtly marked while the absolutive case is not overtly marked (Dixon 1979).

(11) If a language with basic SV order has non-overt independent subject pronouns, it allows postverbal position of the overt subject (pro-drop parameter, Rizzi 1986, Holmberg 2010).

(12) Inflectional morphology always occurs outside derivational morphology (Greenberg 1963, Universal 28)

(13) If a language allows question-word fronting from an adverbial clause, it also allows fronting from a complement clause.

(14) If a marker in a language expresses locative and dative roles, then it also expresses the allative role (Blansitt 1988).

(15) If the reflexive pronoun is distinct for the anaphoric pronoun for disjoint reference, it is longer than the anaphoric pronoun (often derived from it by an additional marker), or equally long (e.g. English *him-self vs. him-Ø*) (Haspelmath 2008c).

(16) Lexicalist Hypothesis: The syntax neither manipulates nor has access to the internal structure of words. (Anderson 1992: 84)

(17) Principle A of the Binding Theory: An anaphor must be bound in its governing category (Chomsky 1981).

(18) If a language has nominal suppletion, it is found in the most frequent nouns (Vafaeian 2013)

On universals

• Some regularities are so strong that we call them UNIVERSALS, because they occur with much greater than chance frequency.

• Universals are sometimes ABSOLUTE, but more often STATISTICAL if there are exceptions (Bickel 2011). A more general term for absolute and statistical universals is UNIVERSAL TENDENCIES.

• RECURRENT PATTERNS (e.g. the use of the same word for 'moon' and 'month') are non-accidental similarities in the sense that there must be something in the human condition that makes it possible for very similar linguistic categories to appear independently in languages that have no historical connection. However, the discovery of a recurrent pattern does not lead us to make predictions about further languages.

• By contrast, the discovery of a universal implies A CLAIM ABOUT ALL OTHER LANGUAGES: If a universal holds (i.e. is found with much greater than chance frequency in a reasonably representative sample), it is claimed that it also holds in any other representative sample. Thus, universal tendencies are claims that can be tested by examining data from the world's languages.

• Universal tendencies need to be distinguished, in particular, from family-specific or region-specific trends, so they need to be based on world-wide sample (e.g. in all major world regions, languages with OV order tend to have postpositions, and

languages with VO order tend to have prepositions; Greenberg 1963: Universal 2; Dryer 1992: 83), even though many languages are exceptions.

• Universal tendencies can also be identified within patterns that are quite rare, e.g. universals of infixation (Yu 2007), because universal tendencies can be implicational ("If a language has infixation, then it has adfixation").

2. What do we mean by *EXPLANATION*?

Explanation is the same as answering "why" questions.

Some "why" questions, with increasing generality:

– why do Russian speakers say *Ja pomogla otcu* (and not **Ja pomogla otco*)?
because the Dative of *otec* 'father' is *otcu*
– why does Russian require the Dative in *Ja pomogla otcu*?
because the verb *pomoč'* 'help' governs a Dative 'helper' argument
– why does *pomoč'* 'help' govern a Dative 'helper' argument?
? because the Dative is preferred with animate, non-patient non-agents
– why is the Dative is preferred with animate, non-patient non-agents?
???

"descriptive explanation" vs. **causal explanation**

descriptive explanation (= analysis) =
explanation of language-particular observations of various levels of generality by means of language-particular rules

> (Descriptive explanation is not controversial – every language-learning textbook formulates descriptive rules, and in fact, every dictionary is a kind of descriptive explanation. The only thing that is controversial is the generality of the rules.)

causal explanation =
explanation of properties of Human Language by causal factors outside of Human Language

Causal explanation is a problem – there is no standard approach to causal explanation in linguistics.

An example:

(19) Inflectional morphology always occurs outside derivational morphology (Greenberg 1963, Universal 28)

e.g. *king-dom-s* *real-ize-d*
ROOT-DERIV-INFL ROOT-DERIV-INFL

Coptic *p-ref-rnobe*
DEF-AGT-sin 'sinner'
INFL-DERIV-ROOT

Causal explanation 1 (functional-adaptive):
Derivational affixes are more relevant to the meaning of the root, so for reasons of iconicity, they occur more closely to the root (Bybee 1985).

Causal explanation 2 (mutational, i.e. change-based):
Derivational affixes always arise by root compounding, whereas inflectional affixes arise from all kinds of function words that occur on the margins of the compounds in the original structures.

Causal explanation 3 (representational):
Derivational morphology is part of the Lexical Component of the mental grammar, while inflectional morphology is part of the Syntactic Component. When a lexical item is inserted into a syntactic tree, it is fully formed and the inflectional markers can only be added outside (Anderson 1992).

How can we tell which of these explanations (if any) is correct?

3. Some fairly clear cases of causal explanation

Functional-adaptive explanation

(6) All languages have morphemes denoting negation.

A language without negation marking would be seriously deficient in communication.

Mutational explanation

(8) If a language has noun-possessor order, it tends to have preposition-NP order, and if it has possessor-noun order, it tends to have NP-postposition order (Dryer 2005).

Adpositions almost always come from possessed nouns in adpossessive constructions, so a noun-possessor construction must give rise to a preposition-NP construction (assuming persistence of ordering). There is no need to invoke any more far-reaching explanation.

e.g. *by (the) cause (of) the weather > because of the weather*

Representational explanation:

(7) No language has a rule that involves counting elements or features.

Nobody has ever suggested a functional-adaptive or change-based explanation for this universal. Locality restrictions are widely found in languages (e.g. **a proud of her son mother*), but the rules always seem to be work in terms of categories, not in terms of numbers.

A hierarchy of causal explanations: weaker explanatory factors take precedence over stronger (less likely) explanatory factors

– If a mutational explanation is available, do not appeal to a functional-adaptive or representational explanation.

– If a functional-adaptive explanation is available, do not appeal to a representational explanation.

– If no other explanation is available, a representational explanation is the only possibility.

4. Other cultural universals

Human groups differ in their languages in striking ways, just as they differ in their cultural practices. Languages are very often thought of as manifestations of different cultures, so are there parallels in culture as well?

(cf. Brown 1991, on human universals)

– all cultures have tools, all cultures have houses
– all cultures have exchange practices
– all cultures have music and myths
– all cultures have marriage
– all complex societies (tribes and upwards) have a ruler
– all states have an upper class
– all states have a moralizing high god (cf. Watts et al. 2015)

Which of these can be explained by functional-adaptive constraints?
For which of these do we need to appeal to representational constraints?

5. What about theories?

Linguists are used to using the word "theory", much more so that the word "explanation". But while it is clear what an explanation is, it is not so clear what a theory is.

Perhaps one could say that an important explanatory factor (such as iconicity) that explains many universals is an "explanatory theory".

But de facto, linguists tend to use the word "theory" in a different sense: To describe a set of claims about the best descriptive framework (e.g. S. Müller 2016, who describes a number of "formal grammatical theories").

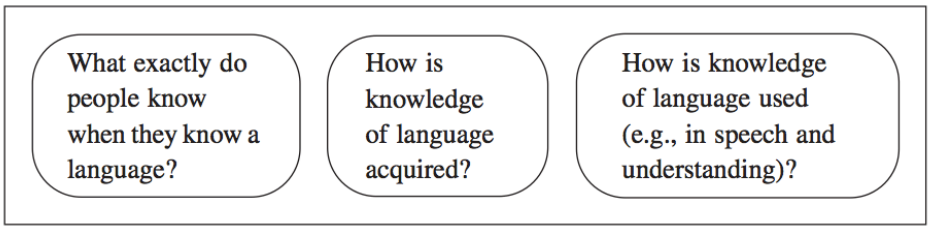
But what are the adequacy criteria for descriptive frameworks? (unclear)

In addition, linguists often use "linguistic theory" as a way of referring to the activity of creating descriptive frameworks. These are sometimes said to be explanatory, but it is more typical for generative linguists to claim that the main task is to **characterize knowledge of language** and the **acquisition of knowledge of language**:

Haegeman (1997: 1):

"The comparative approach in the generative tradition addresses the following questions: (i) what is knowledge of language? (ii) how is this knowledge acquired?... In order to answer these questions we have to identify which linguistic properties can vary across languages and which are constant."

Larson (2010: 11):



6. Observational, descriptive and explanatory adequacy

Chomsky (1965)

– a theory is **observationally adequate** if each observation is reflected in the description

– a theory is **descriptively adequate** if it reflects all the generalization that speakers know

– a theory is **explanatorily adequate** if it provides a principled choice between competing descriptions

What does the latter have to do with "why" questions?

Perhaps: Explanatory adequacy answers **Plato's Problem**:
"Why can we acquire language even in the absence of evidence for the rules that we acquire?"

In other words, explanatory adequacy does not aim to explain language universals, but Plato's Problem.

For the study of language universals, it is not necessary to choose between different descriptions
– observational adequacy is sufficient (Haspelmath 2004).

7. A concrete case: A functional-adaptive explanation of plurative and singulative patterns (cf. Haspelmath & Karjus 2018)

7.1. Introduction

Cross-linguistic trends in number-marking asymmetries can be explained by general usage preferences: Plurals are marked overtly because they are rarer.

explanatory principle: Zipfian efficiency (or "economy"), a kind of functional-adaptive explanation

(3) **Singular** nouns tend to be **zero**-coded, while **plural** nouns tend to be **overtly**-coded (Greenberg 1963, 1966), because the singular tends to be more frequent than the plural

(Greenberg 35. "There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero.")

– explained by significantly higher overall usage frequencies of singular nouns

e.g.	*day*	*day-s*	
	59,298	31,542	(BNC)

Explanatory principle:

(4) **Minimize Form** (Hawkins 2014)
Grammatical systems should have short forms or zero for meanings that are easier to predict, and overt or longer forms for meanings that are less easy to predict.

Singulars are more frequent than plurals, so singular meaning is easier to predict.

7.2. Markedness explanation vs. frequency explanation

markedness explanation:
– the singular tends to be formally basic/unmarked because it is semantically basic/unmarked (Mayerthaler 1981: Ch. 1)

frequency explanation:
– the singular tends to be formally basic because it is more frequent than the plural

Greenberg (1966: 32):

TABLE XVII

Language	*Size of Sample*	*Singular*	*Plural*	*Dual*
Sanskrit	93,277	70.3	25.1	04.6
Latin (Terence)	8,342	85.2	14.8	
Russian	8,194	77.7	22.3	
French	1,000	74.3	25.7	

Objection: Couldn't it be that the singular is more frequent because it is semantically basic/unmarked? (cf. Mayerthaler 1981: 136-140)

Answer: No – this is shown by singulative lexemes, where the "singular" is NOT more frequent. **Singulative lexemes** (with overt singular forms and zero plural forms) are correctly predicted to exist by the frequency explanation, but incorrectly ruled out by the markedness explanation (Mayerthaler 1981: 51-53).

– explained by significantly higher frequencies of plurals for those (kinds of) nouns that tend to have overtly-coded singulars

e.g.	*naħal*	*naħl-a*	(Maltese)
	'bees'	'bee'	
	729	507	(BNC frequency for *bee/bees*)

7.3. Basic comparative concepts: notional and formal

(cf. Haspelmath 2010)

notional categories: **uniplex** vs. **multiplex** nominals (Talmy 1988)

Multiplex nominals are nominals that can be used to refer to phenomena which can readily be conceived of as (internally homogeneous) **groups** of things (and which therefore are expressed by overt plural forms in some languages).

UNIPLEX	MULTIPLEX
day-Ø	*day-s*
bee-Ø	*bee-s*
(a) fish	*(many) fish*

(a) hair	*(she has black) hair*
Maltese *naħl-a* 'bee'	*naħal* 'bees'
Welsh *moron-en* 'carrot'	*moron* 'carrots'

Nominal meanings which frequently occur in multiplex use (e.g. 'bees') can be called **multiplex-prominent meanings**.

formal comparative concepts:

basic/plurative nominal pairs vs. **singulative/basic** nominal pairs

Basic/plurative pairs are pairs where one member is an ***unmarked uniplex*** nominal, and the other member is a ***marked multiplex*** nominal.

Singulative/basic pairs are pairs where one member is an ***unmarked multiplex*** nominal, and the other member is a ***marked uniplex*** nominal.

BASIC/PLURATIVE PAIRS (=*plurative lexemes*)	SINGULATIVE/BASIC PAIRS (= *singulative lexemes*)	
German	Maltese	
Schuh / Schuh-e	*zarbun-a / zarbun*	'shoe/shoes'
Fisch / Fisch-e	*ħut-a / ħut*	'fish (sg.)/fish (pl.)'
Apfelsine / Apfelsine-n	*laringġ-a / laringġ*	'orange/oranges'
Estonian	Welsh	
tigu / teo-d	*malwod-en / malwod*	'snail/snails'
karv / karva-d	*blew-yn / blew*	'hair/hair(s)'
hernes / herne-d	*pys-en / pys*	'pea/peas'

Singulative lexemes are found especially with the following kinds of meanings:

– paired body-parts
– small animals
– fruits/vegetables
– groups of people

These can be called **singulative-prominent meanings**.

7.4. Restating the main claim

The coding of uniplex/multiplex pairs depends on frequency of use:
• Uniplex-prominent meanings tend to be expressed as plurative lexemes
– plurative-prominent meanings tend to be frequently uniplex
• Multiplex-prominent meanings tend to be expressed as singulative lexemes
– singulative-prominent meanings tend to be frequently multiplex

e.g. 'day(s)' tends to be expressed as in English (*day/day-s*)
e.g. 'bee(s)' tends to be expressed as in Maltese (*naħal /naħl-a*)

7.5. Expression tendencies: singulative-prominent meanings

Welsh (King 1993: 67-69)

fruits/vegetables	*madarch*	*maderch-en*	mushrooms
	mwyar	*mwyar-en*	blackberries
	ffa	*ffä-en*	beans
	bresych	*bresych-en*	cauliflower
small animals	*cacwn*	*cacyn-en*	wasps
	clêr	*cler-en*	flies
	hwyaid	*hwyad-en*	ducks
	llygod	*llygod-en*	mice
groups of people	*plant*	*plent-yn*	children
other	*sêr*	*ser-en*	stars
	dillad	*dilled-yn*	clothes
	plu	*plu-en*	feathers

Maltese (Mifsud 1996)

paired body-parts	*zarbun*	*zarbun-a*	shoes
	buz	*buz-a*	boots
fruits/vegetables	*amħ*	*amħ-a*	corn
	lewz	*lewz-a*	almonds
	tuffieħ	*tuffieħ-a*	apples
small animals	*dubbien*	*dubbien-a*	flies
	gawwi	*gawwi-a*	swallows
	wizz	*wizz-a*	geese
other	*taraġ*	*taraġ-a*	stairs
	ravyul	*ravyul-a*	ravioli

7.6. The corpus data

– four languages (**Estonian, Latvian, Norwegian, Russian**), large corpora
– 28 lexemes in each language: three lexemes from semantic groups with six singulative-prominent meanings:

paired body-parts	**ear, foot, lung**
paired items:	**glove, shoe, ski**
fruits:	**apple, potato, strawberry**
small animals:	**bee, pigeon, sheep**
people:	**child, boy, girl**
ethnic groups:	**European, English, speaker of (resp. language)**

plus 10 random lexemes for each language

7.7. Results

– The random lexemes tend to be more frequent in the singular.
– The singulative-prominent lexemes tend to be more frequent in the plural.

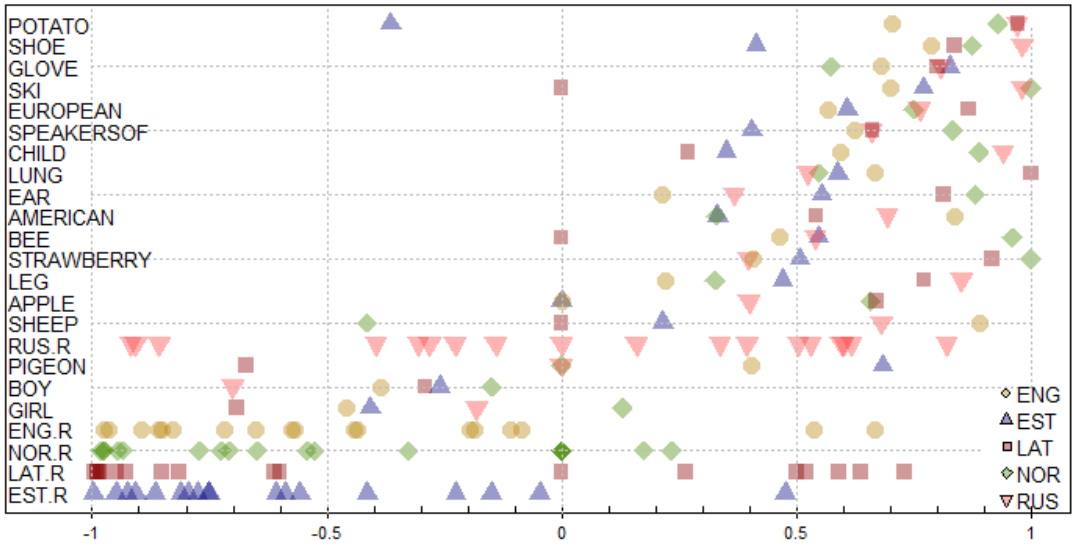


Figure 1: The sample of 180 nouns – 18 preselected nouns and 18 randomly sampled nouns from 5 languages – arranged along the vertical axis by the median asymmetry index value of the concepts. 'R.' marks the random groups. The horizontal axis represents the number asymmetry index, discussed above, so the uniplex-prominent nouns lean to the left, and the multiplex-prominent nouns to the right side of the plot.

The deviation from the average singular/plural ratio of the random lexemes (intercept of the linear regression model below) is statistically significant for all concepts, except for 'pigeon', 'boy' and 'girl'.

7.8. Explanation

The tendency for singulative lexemes to be multiplex-prominent (and for plurative lexemes to be uniplex-prominent) is due to a highly general principle of grammatical coding:

(1) **The grammatical form-frequency correspondence principle:**
When two grammatical forms that differ minimally in meaning (i.e. forms that form a semantic opposition) occur with significantly different frequencies, the less frequent pattern tends to be overtly coded (or coded with more coding material), while the more frequent pattern tends to be zero-coded (or coded with less coding material).

This principle has a well-known explanation in terms of coding efficiency (Zipf 1935, Fenk-Oczlon 1991, Hawkins 2004, Haspelmath 2008a)

The correspondence between form and frequency is implemented by **diachronic mechanisms** which tend to make frequent forms short, because frequent forms are predictable (Zipf 1935):

(2) **Frequency causes predictability, which causes short form:**
In human language, there are recurrent (**usage-based**) diachronic mechanisms which create patterns in which frequently used meanings are expressed by short forms because of their predictability (e.g. Bybee 2007).

Thus, the causal effect is very indirect:

We cannot say that the *day/day-s* pattern in Modern English is due to the fact that *day* is more frequent than *days* in Modern German.

The causal effect is relatively weak, so it cannot be seen in all languages (many languages lack form distinctions between uniplex and multiplex nouns), and especially the tendency for multiplex-prominent nouns to occur as singulatives is manifested only very rarely. (In most languages, all lexemes join the majority pattern, due to **system pressure**, cf. Haspelmath 2014.)

Explanatory mode:

Universal corpus asymmetries explain universal form asymmetries, via diachronic mechanisms.

In this way, corpus data from Norwegian and Russian can be used to explain morphological asymmetries in Maltese and Arbore.

Critical question:
Couldn't it be that singulative lexemes are conceptualized differently in languages with singulative marking, as "less individualized", or "collective", or "masses"?
(This would salvage the meaning-based explanation, cf. Grimm 2012.)

Reply:
Who knows? The frequency-based explanation does not rely on vague concepts such as "conceptualization as less individualized". It relies on corpus frequencies, which can be easily falsified.

8. Three types of grammatical universals

– coexpression universals

"If form A expresseses meaning 1 and meaning 3, then it also expresses meaning 2" (semantic map: 1-2-3)

e.g. if a case form expresses locative and dative, then it also expresses allative (semantic map dative – allative – locative; Blansitt 1988)

– domain universals

(i) "If a dependent form can occur in a larger domain, then it can also occur in a smaller domain"

e.g. if a question-word can be extracted from a finite clause, then it can also be extracted from an infinitival clause (Hawkins 1999; 2004)

(ii) "If a language has left-branching phrases of type A, then it also has left-branching phrases of type B, to minimize recognition domains"

e.g. if a language has left-branching verb phrases (O-V), then it also has left-branching nominal phrases (Poss-N)" (Hawkins 2004; 2014)

– coding universals

"If a language has grammatical coding of meaning α, then it codes meaning β with at least as long a form" (Greenberg 1966; Croft 2003: Chapter 4; Haspelmath 2018)

e.g. if a language has overt singular marking, then it also has overt plural marking

coexpression universals	**domain universals**	**coding universals**
(14) If a marker in a language expresses locative and dative roles, then it also expresses the allative role	(9) If a language has OV order, then it tends to have no question-word fronting (13) If a language allows question-word fronting from an adverbial clause, it also allows fronting from a complement clause.	(3) Greenberg 35. There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero. (4) Greenberg 38. Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb. (10) In almost all cases, the ergative case is overtly marked while the absolutive case is not overtly marked (15) If the reflexive pronoun is distinct for the anaphoric pronoun for disjoint reference, it is longer than the anaphoric pronoun (often derived from it by an additional marker), or equally long (18) If a language has nominal suppletion, it is found in the most frequent nouns

Topic-first universals:

(1) Greenberg 1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object. (Greenberg 1963)

(2) Greenberg 14. In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages.

Effability universals:

(5) All languages have roots denoting things, roots denoting actions, and roots denoting properties (such as dimension, age, or value).

Not valid or not well-defined universals:

(11) If a language with basic SV order has non-overt independent subject pronouns, it allows postverbal position of the overt subject.

(16) Lexicalist Hypothesis: The syntax neither manipulates nor has access to the internal structure of words.

(17) Principle A of the Binding Theory: An anaphor must be bound in its governing category.

9. Regularities and causal factors: Concepts and technical terms

– General terms such as *restriction, constraint, preference, tendency, bias,* and *motivation* have been used in diverse and sometimes confusing ways by linguists.

– One needs to distinguish between terms for **regularities** and terms for **causal factors**, and within the terms for regularities, I distinguish between **language-particular regularities** and **cross-linguistic regularities**.

9.1. Language-particular regularities (covered by "descriptive explanation")

– regularities of language use (e.g. corpus patterns)

– regularities of particular language systems ("descriptive explanatory devices"):

forms, constructions (schemas),
clause, *noun phrase*, *suffix*, *dative case*, or terms for relations between constructions *alternation*, *derivation*

The term *constraint* is sometimes found for language-particular regularities, e.g. in CONSTRAINT-BASED formalisms such as HPSG, and optimality theory also uses constraints for language-particular regularities.

9.2. Universal tendencies

The discovery of a universal (tendency) implies a claim about all other languages: If a universal holds (i.e. is found with much greater than chance frequency in a reasonably representative sample), it is claimed that it also holds in any other representative sample. Thus, **universal tendencies are claims that can be tested** by examining data from the world's languages.

Universal tendencies are characteristics of Human Language. When one appeals to a causal factors that are relevant to any language, then the explanations must be explanations of universal tendencies.

9.3. Causal factors: Preferences, constraints, restrictions

Our explanatory devices are called
causal factors, or
(system-external) motivations, or
constraints

(Two other terms that are used commonly as well, especially outside core linguistics, are *force* and *pressure*. It seems that all these terms are basically synonymous.)

If a constraint is very strong, it can also be called ***restriction***,
if it is weaker, it can be called ***preference***.

References

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Bach, Emmon. 1971. Questions. *Linguistic Inquiry* 2. 153–166.

Bickel, Balthasar. 2011. Absolute and statistical universals. In Patrick Colm Hogan (ed.), *The Cambridge encyclopedia of the language sciences*. Cambridge: Cambridge University Press.

Bickel, Balthasar. 2013. Linguistic diversity and universals. In Nick J. Enfield, Paul Kockelman, & Jack Sidnell (eds.), *Cambridge Handbook of Linguistic Anthropology*. Cambridge: Cambridge University Press, to appear.

Blansitt, Edward L., Jr. 1988. Datives and allatives. In: Michael Hammond, Edith A. Moravcsik, & Jessica R. Wirth (eds.), *Studies in syntactic typology*, 173–191. (Typological Studies in Language, 17). Amsterdam: Benjamins.

Brown, Donald E. 1991. *Human universals*. New York: McGraw Hill.

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

Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Chomsky, Noam A. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam A. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Dryer, Matthew S. 2005. Order of genitive and noun. In *WALS*, 350–353.
Dryer, Matthew S. 2011b. Order of adposition and noun phrase. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 85. Available online at http://wals.info/chapter/85
Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.
Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.
Grimm, Scott. 2012. Individuation and inverse number marking in Dagaare. In Diane Massam (ed.), *Count and mass across languages*, 75–98. Oxford: Oxford University Press.
Haegeman, Liliane. 1997. Introduction: on the interaction of theory and description in syntax. In: Haegeman, Liliane (ed.) *The new comparative syntax*. London: Longman, 1-32.
Haspelmath, Martin & Andres Karjus. 2018. Explaining asymmetries in number marking: Singulatives, pluratives and usage frequency. *Linguistics (to appear)*.
Haspelmath, Martin, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.
Haspelmath, Martin. 2004. Does linguistic explanation presuppose linguistic description? *Studies in Language* 28. 554–579.
Haspelmath, Martin. 2008c. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1). 40–63.
Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663–687.
Haspelmath, Martin. 2014. On system pressure competing with economic motivation. In Brian MacWhinney, Andrej L. Malchukov, & Edith A. Moravcsik (eds.), *Competing motivations in grammar and usage*, 197–208. Oxford: Oxford University Press.
Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75. 244–285.
Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. New York: Oxford University Press.
Holmberg, Anders. 2010. Null subject parameters. In Theresa Biberauer, Ian G. Roberts, Anders Holmberg, & Michelle Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.
Larson, Richard K. 2010. *Grammar as science*. Cambridge, Mass.: MIT Press.
Mayerthaler, Willi. 1981. *Morphologische Natürlichkeit*. Wiesbaden: Akademische Verlagsgesellschaft Athenaion.
Müller, Stefan. 2016. *Grammatical theory: From transformational grammar to constraint-based approaches*. (Textbooks in Language Sciences 1). Berlin: Language Science Press. http://langsci-press.org//catalog/book/25.
Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17(3). 501–557.
Roberts, Ian G. 2007. *Diachronic syntax*. Oxford: Oxford University Press.
Talmy, Leonard. 1988. The relation of grammar to cognition: A synopsis. In Brygida Rudzka-Ostyn (ed.), *Topics in cognitive linguistics*, 166–205. Amsterdam: Benjamins.
Vafaeian, Ghazaleh. 2013. Typology of nominal and adjectival suppletion. *Sprachtypologie und Universalienforschung* 66(2). 112–140.
Watts, Joseph, Simon J. Greenhill, Quentin D. Atkinson, Thomas E. Currie, Joseph Bulbulia & Russell D. Gray. 2015. Broad supernatural punishment but not moralizing high gods

precede the evolution of political complexity in Austronesia. *Proceedings Royal Society B* 282(1804). 20142556. doi:10.1098/rspb.2014.2556.

Yu, Alan C. L. 2007. *A natural history of infixation*. Oxford: Oxford University Press.

Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, Mass.: M. I. T. Press.

Tylex course "Explaining language universals", Voronovo (Moscow), 2017 September 4

2. Universals of grammatical coding

MARTIN HASPELMATH
MPI-SHH Jena & Leipzig University

1. Three types of grammatical universals

– coexpression universals

"If form A expresseses meaning 1 and meaning 3, then it also expresses meaning 2" (semantic map: 1-2-3)

> e.g. if a case form expresses locative and dative, then it also expresses allative
> (semantic map dative – allative – locative; Blansitt 1988)

– domain universals

(i) "If a dependent form can occur in a larger domain, then it can also occur in a smaller domain"

> e.g. if a question-word can be extracted from a finite clause, then it can also be extracted from an infinitival clause (Hawkins 1999; 2004)

(ii) "If a language has left-branching phrases of type A, then it also has left-branching phrases of type B, to minimize recognition domains"

> e.g. if a language has left-branching verb phrases (O-V), then it also has left-branching nominal phrases (Poss-N)" (Hawkins 2004; 2014)

– coding universals

"If a language has grammatical coding of meaning α, then it codes meaning β with at least as long a form" (Greenberg 1966; Croft 2003: Chapter 4; Haspelmath 2018)

> e.g. if a language has overt singular marking, then it also has overt plural marking

coexpression universals	domain universals	coding universals
(14) If a marker in a language expresses locative and dative roles, then it also expresses the allative role	(9) If a language has OV order, then it tends to have no question-word fronting (13) If a language allows question-word fronting from an adverbial clause, it also allows fronting from a complement clause.	(3) Greenberg 35. There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero. (4) Greenberg 38. Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb. (10) In almost all cases, the ergative case is overtly marked while the absolutive case is not overtly marked (15) If the reflexive pronoun is distinct for the anaphoric pronoun for disjoint reference, it is longer than the anaphoric pronoun (often derived from it by an additional marker), or equally long (18) If a language has nominal suppletion, it is found in the most frequent nouns

Topic-first universals:

(1) Greenberg 1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object. (Greenberg 1963)

(2) Greenberg 14. In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages.

Effability universals:

(5) All languages have roots denoting things, roots denoting actions, and roots denoting properties (such as dimension, age, or value).

Not valid or not well-defined universals:

(11) If a language with basic SV order has non-overt independent subject pronouns, it allows postverbal position of the overt subject.

(16) Lexicalist Hypothesis: The syntax neither manipulates nor has access to the internal structure of words.

(17) Principle A of the Binding Theory: An anaphor must be bound in its governing category.

2. Coding asymmetries in grammar

• universal coding asymmetries in grammar ("markedness asymmetries")

singular	plural	(*book* – *book-s*)
present	future	(*go* – *will go*)
3rd person	2nd person	(Spanish *canta* – *canta-s*)

nominative	accusative	(Hungarian *ember – ember-t*)
active	passive	(Latin *cantat – cantat-ur*)
affirmative	negative	(*go – don't go*)
allative	ablative	(*to – from*)
positive	comparative	(*small – small-er*)
predicative verb	nominalized verb	(*go – go-ing*)
action word	agent noun	(*bake – bak-er*)
property	change of state	(*red – redd-en*)

• observation: the zero-coded member of an opposition is generally more frequent

(1) **The form-frequency correspondence principle**
Languages tend to use less coding material for more frequent expressions.

This is uncontroversial for word length (e.g. Zipf 1935: 23), but it is **also generally valid for grammatical patterns**. The insight of this principle is originally due to Greenberg (1966) (see also Croft 2003: Ch. 4; Haspelmath 2008a; 2008b).

(2) **The grammatical form-frequency correspondence principle**
When two grammatical patterns that differ minimally in meaning (i.e. patterns that form a semantic opposition) occur with significantly different frequencies, the less frequent pattern tends to be overtly coded (or coded with more coding material), while the more frequent pattern tends to be zero-coded (or coded with less coding material).

• since we observe strong **universal tendencies** (sometimes even exceptionless), there must be a **highly general** explanatory factor

• key proposal: predictability > shortness of coding
lack of predictability > length of coding

• This is thus a kind of **economy explanation**: Speakers, and hence language systems, favour economical patterns.

• Language systems can adapt to the users' needs because they are **malleable** (= somewhat flexible), and we can study the ways in which economical patterns arise in language systems, again and again.

• More generally, many universal properties of language systems can be seen as **efficiency-based** (cf. Hawkins 2014). (Many linguists think that functional explanation should focus on meaning and function; I think that functional explanations of universals are mostly based on efficiency.)

• Alternative explanations are less comprehensive or do not provide the necessary causal links:

– shortness cannot be the cause of frequency, because the same frequency asymmetries are found when the two opposed forms are equally long

– **iconicity** does not work, because of cases where there are no meaning differences, but only frequency differences (Haspelmath 2008a)

– **markedness** explains neither coding asymmetries nor frequency of use because there are no causal links, and reversals are direct counterevidence

cf. also imperatives:			2nd person	3rd person
	Spanish	ind	*canta-**s***	*canta-Ø*
		impv	*canta-Ø*	***que** cante*

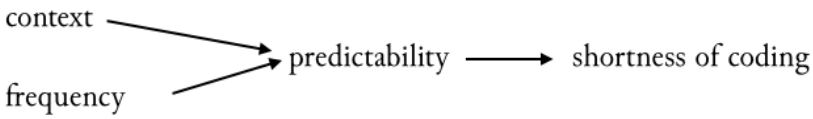
3. The efficiency-based explanation of form-frequency-correspondences

• speakers can afford to use **short or zero forms** for predictable meanings, but they have to make a greater effort for unprdictable meanings

• **higher-frequency forms** are more **predictable** than lower-frequency items because of their frequency

• but **context** can also lead to higher predictability, e.g. when a sentence contains a referent that has been used just before (e.g. *the girl went to the river and (she) looked for fish*) – the result is the same (shortness of coding or zero-coding)

• causal chains:



• since the explanation is based on efficiency, there is no expectation that the **categories** should match across languages – what needs to match is the meanings (so that they are comparable); this theory is thus fully compatible with **categorial particularism** (Haspelmath 2010)

• the preference favouring zero coding of frequent meanings is counterbalanced by the **general tendency favouring overt coding** of meanings – as a result, there are many situations where there is **no coding asymmetry**; what is still predicted is that there should be no "anti-efficient" coding, with the longer form

Non-efficient (symmetric):

	Greek	Mandarin
SG	*vivlí-o*	*shū*
PL	*vivlí-a*	*shū*

Anti-efficient (a hypothetical language, "Martian"):

	Martian
SG	*vivlí-o*
PL	*vivlí*

4. More cases of coding asymmetries

4.1. Nominative vs. accusative (Greenberg 1963)

	English	German	Quechua
NOM	*he*	*Herr Kim*	*wasi* 'house'
ACC	*hi-**m***	*Herr**n** Kim*	*wasi-**ta***

4.2. Male vs. female occupational terms

	Latin	German	Hungarian
MALE	*rex*	*König*	*király*
FEMALE	*reg-**ina***	*König-**in***	*király-**nő***

4.3. Allative vs. ablative marking (Michaelis 2017)

	English	Sri Lanka P.	Principense
ALLATIVE	*to*	*maaket* 'to the market'	*fya* 'to the m.'
ABLATIVE	***from***	*kaaza **impa*** 'from home'	***fo** fya* 'from the m.'

4.4. Spatial marking on place names vs. inanimate nouns vs. animate nouns (Aristar 1997; Creissels & Mounole 2011)

	Basque	Tswana	Tamil
PLACE NAME	*Bilbo-n* 'in Bilbao'	*Gaborone* 'at G.'	
INANIMATE	*mendi-**tan*** 'at the mountain'	*toporo-**ng*** 'in town'	*N-il*
ANIMATE	*neska-**rengan*** 'at the girl's'		*N-**iṭam***

4.5. Directionality vs. configuration (Lestrade et al. 2011)

	Finnish	Lezgian	English	German
	päälle 'onto'	*winel* 'on'	*from below*	*unter (de-n Tisch)*
DIRECTIONALITY	*-lle*	*-el*	*from*	*-n*
CONFIGURATION	***pää-***	***win-***	***below***	***unter***

4.6. Instrumental vs. comitative

	Russian	Hungarian	Welsh	Chinuk Wawa
INS	*myš'-ju* 'with a mouse'	*tol-lal* 'with pen'	*a*	*káka*
COM	***s*** *myš'-ju*	*gyerek-**estül*** 'with child'	***gyda***	***kánamakwst***

4.7. Definiteness with vs. without possessor (Haspelmath 1999)

	German	Welsh	Hebrew
POSSESSED	*mein Ø Buch*	*Ø car y meddyg*	*Ø-sifr-i*
UNPOSSESSED	***das*** *Buch*	***y*** *car*	***ha****-sefer*

4.8. Allophoric (3rd person) vs. locuphoric (1st/2nd) person indexes

(Siewierska 2009)

	Spanish	Polish	Imb. Quechua
3rd	*canta-Ø*	*śpiewa-Ø*	*shamu-rka-Ø*
2nd	*canta-**s***	*śpiewa-**sz***	*shamu-rka-**ngui***
1st	*cant-**o***	*śpiewa-**m***	*shamu-rka-**ni***

4.9. Addressee person index vs. allophoric (= 3rd sg) indexes in imperatives

	Latin	English	Turkish
2nd	*lauda-Ø*	*praise!*	*bak-Ø* 'look'
3rd	*lauda-**to***	***let*** *him praise!*	*bak-**sın*** 'let him look'

4.10. Positive vs. comparative vs. superlative (Bobaljik 2012)

	English	Hungarian	French
positive	*small*	*kis*	*petit*
comparative	*small-**er***	*kis-**ebb***	***plus*** *petit*
superlative	*small-**est***	***leg****-kis-**ebb***	***le*** *plus* ***petit***

4.11. Present tense vs. future tense (cf. Greenberg 1966)

(17)

	English	Latin	Kiribati
PRS	*they praise*	*lauda-nt*	*e taetae* 'he speaks'
FUT	*they **will** praise*	*lauda-**b**-unt*	*e **na** taetae* 'he will speak'

4.12. Present tense vs. past tense (cf. Greenberg 1966)

(18)

	Greek	German	Lezgian
PRS	*ksér-is*	*weiß-t*	*či-zwa*
PST	***í****-kser-es*	*wuss-**te**-st*	*či-zwa-**j***

4.13. Affirmative vs. negative (Miestamo 2005)

(20)		Hebrew	English	Egyptian Arabic
	AFF	*katavti*	*I wrote*	*šuf-t* 'I saw'
	NEG	***lo** katavti*	*I did**n't** write*	***ma** šuft-ti-š* 'I didn't see'

4.14. Disjoint anaphoric vs. reflexive (Haspelmath 2008c)

(21)		English	Hebrew	M. Chinese	Japanese
	DISJOINT	*her*	*oto*	*tā*	*Ø*
	REFLEXIVE	*her**self***	*et ʕ**acmo***	*(tā) **zìjǐ***	***zibun***

4.15. Grooming reflexive vs. extroverted reflexive (Haspelmath 2008c)

(22)		Russian	Dutch	Greek	English
	GROOMING	*moet-sja*	*wast zich*	*plen-ete*	*he washes Ø*
	EXTROVERTED	*vidit **sebja***	*ziet **zichzelf***	*vlép-i **ton eavtó tu***	*he sees **himself***

5. Role-reference association universals

5.1. Introduction

• In many languages, the coding of core arguments depends on their REFERENTIAL PROMINENCE in one way or another

• referential prominence has become widely known by names such as "**animacy hierarchy**" or "**empathy hierarchy**"

• the (a) example below shows overt coding of an argument, and the (b) example shows zero coding of the same argument

(1) split subject (A) marking: Kham (Watters 2002)

a. *no-ra-**e*** *zihm* *jə-ke-rə*
he-PL-ERG house.ABS make-PFV-3PL
'They made a house.'

b. *ŋa:-Ø* *zihm* *ŋa-jəi-ke*
I-NOM house.ABS 1SG-make-PFV
'I made a house.'

(2) split object (P) marking: Sakha (Baker 2015: 4-5)

a. *Masha* *salamaat-**y*** *türgennik* *sie-te.*
Masha porridge-ACC quickly eat-PST.3SG.SBJ
'Masha ate the porridge quickly.'

b. *Masha türgennik salamaat-Ø sie-te.*
Masha quickly porridge-Ø eat-PST.3SG.SBJ
'Masha ate porridge quickly.'

• prominence-conditioned splits are cross-linguistically regular in a way that is surprising but apparently robust (Silverstein (1976), Comrie (1978), Moravcsik (1978a, 1978b), Dixon (1979), Bossong (1985; 1998) Lazard (2001))

In addition, argument coding splits may depend on the referential prominence properties of the SCENARIO:
– in monotransitive constructions, A and P
– in ditransitive constructions, R and T

(3) split marking of P only if A is third person: Teop (Mosel 2007: 10)
a. (3 > 3)
A beiko tenaa paa asun=u ***ben*****-*e* *guu.***
art child my TAM kill=imm OBJ-ART pig
'My child has killed the pig.'

b. (1 > 3)
Enaa paa dee ma=u e guu.
1SG TAM carry DIR=IMM ART pig
'I have brought a pig.'

(4) split marking of R obligatory if T is a personal pronoun, and R is a full nominal: English
a. (N > N)
She gave Kim the money. (≈ She gave the money to Kim.)
b. (pers > pers)
She gave him it. (≈ She gave it to him.)
c. (N > pers)
**She gave Kim it.*
d. (N > pers)
She gave it to Kim.

All these are special cases of the high-level generalization in (5).

(5) **Universal 1: The role-reference association universal**
Deviations from canonical associations of role rank and referential prominence tend to be coded by longer grammatical forms.

• proposed explanation of this universal:
a special case of the form-frequency correspondence universal

• The basic idea is that additional coding such as the accusative marker in (2a) or the dative marker in (4d) is **required when it is least predictable and hence needed the most**, i.e. that argument coding splits reflect a functional motivation.

Across a range of different situations,

– arguments with a **high-ranked role** (the transitive A-argument and the ditransitive R-argument) are **referentially prominent** in the canonical case (or in other words, the most frequent case), and
– arguments with a **low-ranked role** (the transitive P-argument and the ditransitive T-argument) canonically exhibit **lower referential prominence**.

Special coding by longer forms is used when a construction deviates from these canonical associations.

5.2. Referential prominence and association with role rank

Referential prominence is defined by the scales of INHERENT PROMINENCE and DISCOURSE PROMINENCE in (6).

(6) scales of referential prominence
a. inherent prominence

person scale:	locuphoric (1st/2nd) > allophoric (3rd person)
full nominality scale:	person form (independent or index) > full nominal
animacy scale:	human (> animal) > inanimate

b. discourse prominence

specificity scale:	definite (> specific indefinite) > indefinite nonspecific
givenness scale:	discourse-given > discourse-new
focus scale:	background > focus

• one could ask what these different kinds of semantic and discourse-pragmatic notions have in common that justifies subsuming them under referential prominence (cf. Shibatani (2006): "discourse relevance")

• The use of the the term *prominence* in connection with split argument coding seems to have been introduced by Aissen (1999; 2003), inspired by phonological terminology, but it is by now well-established (e.g. Malchukov 2008; Bornkessel-Schlesewsky & Schlesewsky 2009; Lockwood & Macaulay 2012).

(6) canonical role-reference association
Arguments with higher-ranked roles (A, R) tend to be more referentially prominent than arguments with lower-ranked roles (P, T).

This tendency is not meant as an abstract aspect of language structures (like the related notion of "harmonic alignment" of Aissen 2003), but as **a concrete claim about discourse frequencies**.

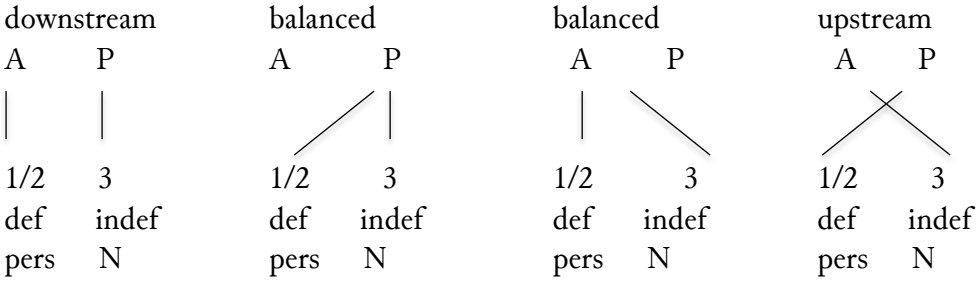
5.3. Canonical associations and universals of coding splits

5.3.1. Two types of canonical associations

(9) ***single-argument*** association tendencies
- a. the A and the R tends to be referentially prominent
- b. the P and T tends to be referentially non-prominent

For example, clauses such as 'The dog found a bone' and 'She gave the boy a key', with definite A and R, and indefinite P and T, are more usual than 'A rock hit the hiker' or 'She gave a boy the key'.

(10)



We can distinguish three kinds of scenarios:

(11) a. **downsteam scenario** (most canonical):
when A/R is referentially more prominent than P/T (e.g. 'I caught a rabbit')

b. **upstream scenario** (least canonical):
when A/R is referentially less prominent than P/T (e.g. 'A dog bit you')

c. **balanced scenario** (intermediate):
when A/R and P/T are equally prominent (e.g. 'I love you', 'The wind opened the window')

• The upstream/downstream metaphor is introduced here because the most common scenarios are expressed in the "easiest" way, while the least common scenarios are more "difficult" (like swimming upstream) and thus need more "coding energy".

5.3.2. Two types of coding splits

(I)

(12) **single-argument** coding split
A single-argument treferential split is an argument coding split for which only the referential prominence of the coded argument is relevant

(14) **Universal 2: differential P flagging** (≈ DOM)
If a language has an asymmetric P flagging split, then the flagging is longer for prominent P-arguments.

(13) **Universal 3: The single-argument coding universal**
If a language has an asymmetric single-argument coding split, then the coding is longer for prominent P/T-arguments and for non-prominent A/R arguments.

• an asymmetric split is a split in which one of the coding types involves **longer** coding. In the majority of cases, this means overt coding contrasting with zero coding, but we sometimes also see examples of shorter and longer overt coding.

• simple examples illustrating the universal are **overt accusatives** in more prominent P-arguments (e.g. animates vs. inanimates, cf. English *he/hi-m* vs. *it/it*), or **overt datives** for less prominent R-arguments (e.g. *give the money to a beggar* vs. *give the beggar money*).

The more general formulation also includes split A flagging ("differential subject marking"), as seen in (1), as well as split R and T flagging (illustrated below).

(II)

(15) scenario split
A scenario split is an argument coding split for which the properties of both arguments in a scenario (A-P, or R-T) are relevant.

(16) **Universal 4: The scenario universal**
If a language has an asymmetric scenario split, then the coding is longest for upstream scenarios, shortest for downstream scenarios, and intermediate for balanced scenarios.

For example, in example (3) from Teop, the non-downstream scenario (3 > 3) requires special coding with the Object marker:

(3) split marking of P only if A is third person: Teop (Mosel 2007: 10)

a. (3 > 3, balanced)
A beiko tenaa paa asun=u ***ben*****-*e guu.***
art child my TAM kill=imm OBJ-ART pig
'My child has killed the pig.'

b. (1 > 3, downstream)
Enaa paa dee ma=u e guu.
1SG TAM carry DIR=IMM ART pig
'I have brought a pig.'

In English, the upstream scenario (pers > N) requires the Dative preposition *to*.

(4) split marking of R obligatory if T is a personal pronoun, and R is a full nominal: English
   a. (pers > N, downstream)
      She gave him the money. (≈ ?She gave the money to him.)
   b. (N > N, balanced)
      She gave Kim the money. (≈ She gave the money to Kim.)
   c. (pers > pers, balanced)
      She gave him it. (≈ She gave it to him.)
   d. (N > pers, upstream)
      She gave Kim it. (OK: She gave it **to Kim.)*

6. Single-argument splits in monotransitive constructions

6.1. Split P-flagging

As mentioned earlier, differential P-flagging, traditionally known as "differential object marking", has been widely discussed in the earlier literature.

6.1.1. Animacy-conditioned split P-flagging

(17) Spanish (García García 2007)
   a. *Conozco *(a) este actor.*
      know:PRS.1SG this:M.SG
      'I know this actor.'

   b. *Conozco (*a) esta película.*
      know:PRS.1SG this:F.SG
      'I know this film.'

6.1.2. Specificity-conditioned split P-flagging
illustrated above in (2) from Sakha; also Punjabi:

(18) Punjabi (Bhatia 1993: 172–174)
   a. *Kataab vekho.*
      book look.IMP.2PL
      'Look at a book.'

   b. *Kataab nũũ vekho.*
      book DAT look.IMP.2PL
      'Look at the book.'

6.1.3. Nominality-conditioned split P-flagging. Some languages have P-flagging only on personal pronouns, but not on full nominals. A well-known example of such a language is English (*he* vs. *him*, *she* vs. *her*, etc.).

6.1.4. Givenness-conditioned split P-flagging. It has been known since Thomson (1912) that split P-flagging is sometimes conditioned by givenness (or "topicality"). Dalrymple & Nikolaeva (2011) discuss a number of relevant cases in some detail.

(19) Persian (Dalrymple & Nikolaeva 2011: 108-112)

a. *man ketâb-râ xarid-am.*
I book-ACC buy.PST-1SG
'I bought the book.'

b. *man sib-i(?*-râ) xord-am.*
I apple-INDEF(-ACC) eat.PST-1SG
'I saw an apple.'

c. *ki mašin-i-?*(râ) did?*
who car-INDEF(ACC) see.PST[3SG]
'Who saw a car?'

6.1.5. Person-conditioned split P-flagging. In Abruzzese (an Italo-Romance variety), special P-flagging by the preposition a occurs only with locuphoric (1st and 2nd) personal pronouns (D'Alessandro 2017: 8).

(20) Abruzzese (dialect of Arielli)

a. *So vistə a mme/ a tte.*
be.1SG seen to me/ to you
'I have seen me/you.'

b. *Semə vistə a nnu/ a vvu.*
be.1PL seen to us/ to you
'We have seen us/you.'

c. **So vistə a Marije/ a jissə/ a quillə.*
be.1SG seen to Mary/ to them/ to them

6.2. Split A-flagging

6.2.1. Person-conditioned split A-flagging. Systems on which ergative marking is restricted in that it does not occur on locuphoric person forms are found in Australia, South Asia, and in two families of the Caucasus (Kartvelian, Nakh-Daghestanian).

(21) Warrgamay

a. *ngana-Ø gaga-ma*
we-NOM go-FUT
'We will go.'

b. *ngana-Ø ngulmburu-Ø ngunda-lma*

we-ERG woman-ACC see-FUT
'We will see the woman.'

c. *maal-du ngulmburu-Ø ngunda-lma*
man-ERG woman-ACC see-FUT
'The man will see the woman.'

6.2.2. Focus-conditioned split A-flagging: when the ergative marker occurs only when the A-argument is focused (cf. McGregor 2010)

(22) Central Tibetan (Tournadre 1995: 264)
a. *khōng khāla' so-kiyo:re'*
he food make-IPFV.GNOM
'He prepares the meals.'

b. *khōng-ki' khāla' so-kiyo:re'*
he-ERG food make-IPFV.GNOM
'HE prepares the meals.'

7. Single-argument splits in ditransitive constructions (cf. Haspelmath 2007)

7.1. Split T coding

7.1.1. Person-conditioned single-argument split of T

(23) Georgian (Harris 1981: 48-49)
a. *Vano-m Anzor-i še-Ø-adara Givi-s.*
Vano-ERG Anzor-NOM PVB-3.OBJ-compared Givi-DAT
'Vano compared Anzori to Givi.'

b. **Vano-m (šen) še-g-adara Givi-s.*
Vano-ERG (you) PVB-2.OBJ-compared Givi-DAT
('Vano compared you to Givi.')

c. *Vano-m šen-i tav-i še-Ø-adara Givi-s.*
Vano-ERG your-NOM self-NOM PVB-3.OBJ-compared Givi-DAT
'Vano compared you to Givi.'

7.1.2. Nominality-conditioned split T-flagging

(24) Ewe (Essegbey 2010)
a. nominal T
Kosí ná [*ga lá*] [*nyɔ́nuví-á*].
Kosi give money DEF girl-DEF
'Kosi gave the money to the girl.' (=11c)

b. person-form T (double-object construction ungrammatical)
*Kosí ná-e Amí.
Kosi give-3SG.OBJ Ami
('Kosi gave it to Ami.')

c. person-form T (with auxiliary *tsɔ́* 'take')
Kosí tsɔ́-e ná Amí.
Kosi take-3SG.OBJ give Ami
'Kosi gave it to Ami.' (lit. 'Kofi took it, gave-to Ali')

7.2. Split R-flagging

7.2.1. Person-conditioned split R-marking

(26) French person clitics

	T (ACC)	R (DAT)		T (ACC)	R (DAT)
1SG	*me*	*me*	1PL	*nous*	*nous*
2SG	*te*	*te*	2PL	*vous*	*vous*
3SG	*le, la*	***lui***	3PL	*les*	***leur***

Table 2: Object person forms (1st person singular and 3rd person masculine singular)

language	1st person R/T '(to) me'	allophoric R 'to him'	allophoric T 'him'
French	*me*	*lui*	*le*
Tangale[a]	*-no/-nọ*	*-ni/-nị*	*mbḛ́ẹ̀ndâm*
Yimas[b]	*ŋa-*	*-(n)akn*	*na-*
Krongo[c]	*àʔàŋ*	*àníŋ*	*ìʔìŋ*

[a] Jungraithmayr (1991: 36) [b] Foley (1991:...) [c] Reh (1985: ...)

7.2.2. Nominality-conditioned split R-flagging

(28) Northeastern Neo-Aramaic of Telkepe (Coghill 2010)
(full nominal R)
a. *wəl-lə pārə* ***ta*** *xa-məskenɒ*
gave-he money to a.certain-poor.person
'He gave money to a certain poor person.' (= Coghill's 11b)

(person-form R)
b. *kəm-yāwəl-lə hadiynɒ*
PST-he.give-3SG.M.OBJ present
'He gave him a present.' (= 14c)

A very similar situation is found in **Bulgarian**, where the Dative preposition *na* is obligatory with full nominals, while clitic pronouns have different forms for Dative and Accusative.

7.2.3. Animacy-conditioned split R-flagging. In Yakkha, a Kiranti language of Nepal, R-arguments are in the Locative case when inanimate, but otherwise in the (zero-coded) Absolutive case.

(29) Yakkha (Schackow 2012: 161-162)

a. *ka* *nniŋda* *photo-ci* *ham-bi?-me?-nenin=ha*
1SG[NOM] 2PL[NOM] photo-PL[NOM] distribute-BEN-NPST-1>2PL-PL
'I distribute the photos to you all.'

b. *sarkar=ŋa* *yaŋ* *tenten=be* *ŋ-hapsu-bi-ci=ha*
government=ERG money[NOM] villages=LOC 3PL.A-distribute-BEN-3PL.P=PL
'The government distributed the money to the villages.'

7.2.4. Specificity-conditioned split R-flagging

In Wolof, an Atlantic language of Senegal, a dative flag is required on R when it is indefinite.

(30) Wolof (Becher 2005: 19)

a. *Jox* *naa* *xale bu jigéén ji* *benn* *velo.*
give 1SG girl DEF INDF bicycle
'I gave the girl a bicycle.'

b. **Jox* *naa* *benn* *xale bu jigéén* *velo* *bi.*
give 1SG INDF girl bicycle DEF
'I gave a girl the bicycle.'

c. *Jox* *naa* *velo* *bi* ***ci*** *benn* *xale bu jigéen.*
give 1SG bicycle DEF to INDF girl
'I gave the bicycle to a girl.'

8. Scenario splits in monotransitive constructions

8.1. Person-conditioned. Yukaghir: accusative flag is required on P when the A is allophoric (like Teop, see above).

(31) Kolyma Yukaghir (Maslova 2003)

a. *met* *es'ie* *tet* *pulut-kele* *kudede-m*
my father.NOM your husband-ACC kill-TR.3SG
'My father has killed your husband.'

b. *met* *tolow* *kudede*
I.NOM deer.NOM kill.TR.1SG
'I killed a deer.'

Sahaptin: ergative flag is required on A when P is locuphoric

(32) Sahaptin (Sahaptian; Pacific Northwest) (Rude 2009: 13-14)

a. *ku =š i-q'ínun-a tílaaki-nɨm*
and 1SG 3.NOM-see-PST woman-ERG
'And the woman saw me.'

b. *ku i-q'ínun-a áswan-Ø tílaaki-na*
and 3.NOM-see-PST boy-ABS woman-ACC
'And the boy saw the woman.'

8.2. Definiteness-conditioned. Eastern Khanty has Ergative case on the A-argument when the P-argument is specific (cf. Baker 2015: 128)

(33) Eastern Khanty

a. *Mä t'əkäjəɣlämnä ula mənɣäləm.*
we.DU.NOM younger.sister.COM berry pick.PST.1PL.SBJ
'I went to pick berries with my younger sister.'

b. *Mə-ŋən ləɣə əllə juɣ kanŋa aməɣaloɣ.*
we-ERG them large tree beside put.PST.3PL.OBJ/1PL.SBJ
'We put them (pots of berries) beside a big tree.'

9. Scenario splits in ditransitive constructions

9.1. Special R coding conditioned by person of T

This is what is known in the literature as "person-case constraint" (PCC).

Bulgarian: dative preposition is required on R when T is locuphoric

(34) Bulgarian (Hauge 1976 [1999])

a. (3>3) *Az im ja preporâčvam.*
I 3PL.DAT 3SG.F.ACC recommend.PRES.1SG
'I recommend her to them.'

b. (3>2) **Az im te preporâčvam.*
I 3PL.DAT 2SG.ACC recommend.PRES.1SG
'I recommend you to them.'

c. *Az te preporâčvam **na** tjah.*
I 2SG.ACC recommend.PRES.1SG to them
'I recommend you to them.'

(35) Shambala (Bantu-G, Tanzania; Duranti 1979: 36)

a. (1>3) *A-za-m-ni-et-ea.*
3SG.SBJ-PST-3SG.THM-1SG.REC-bring-APPL
'S/he has brought him/her to me.'

b. (3>1) **A-za-ni-mw-et-ea.*
3SG.SBJ-PAT-1SG.THM-3SG.REC-bring-APPL
'S/he has brought me to him/her.'

c. *A-za-ni-eta* *kwa* *yeye.*
3SG.SBJ-PST-1SG.THM-bring to him/her
'S/he has brought me to him/her.'

9.2. Special R coding conditioned by nominality of T

In many varieties of English (especially American, it seems), the R cannot be coded in the simplest way when the T is a person form rather than a full nominal. In these varieties, (37a) is unacceptable *(*Pat showed him it)*.

(36) a. (pers > nom, downstream)
Kim showed me his house.

b. (nom > nom, balanced)
Lee showed her brother her new house.

(37) a. (pers > pers, balanced)
**Pat showed him it.*

b. OK: *Pat showed it to him.*

c. (nom > pers, upstream)
**Pat showed his wife it.*

d. OK: *Pat showed it to his wife.*

9.3. Special T coding conditioned by nominality of R

While the use of a special R marker to code the upstream scenarios 3 > 1 and 3 > 2 is perhaps the most widespread pattern, some languages use special forms of T when the R is a person form.

For example, Modern Greek has a set of Genitive (i.e. dative) and Accusative proclitics used in downstream and allophoric balanced scenarios, as seen in (38a).

(38) a. *Tu* *to* *é͡ðose.*
him.GEN it.ACC he.gave
'He gave it to him.'

b. **Tu* *me* *é͡ðose.*
him.GEN me.ACC he.gave
('He gave me to him.')

c. OK: *Tu* *ẽðose* *eména.*
him.GEN he.gave me.FULL.ACC
'He gave me to him.'

9.4. Special R coding conditioned by animacy of T

Icelandic: the preposition *fyrir* is required on R when T is animate (Siewierska & van Lier 2013)

(39) a. *Hann kynnti mér þessa gerð skáldsagna.* (indirective I)
he.NOM introduced me.DAT this type fiction
'He introduced this type of fiction to me.'

b. *Ég mun kynna þig fyrir henni.* (indirective II)
I.NOM will introduce you.ACC to her
'I will introduce you to her.'

9.5. Nominality-conditioned scenario splits. A split of this kind was presented in §1 above (example (4)): In English, the Dative preposition to is required on R if the T is a full nominal, so that instead of *She gave Kim it, one must say She gave it to Kim.

10. Argument-coding vs. verb-coding

special coding can also be verb-coding:

monotranitives: A number of languages use the basic verb form in person-downstream scenarios, but a specially marked verb form in upstream scenarios. These markers are generally called INVERSE markers (Jacques & Antonov 2014).

(46) Itonama (Crevels 2010: 680, 682)
a. (2 > 3)
ke'-sewane
2SG.F-see
'you (F) see him/her'

b. (3 > 2)
ka'-k'i-kamo
2SG.F-INV-hit.face
'he hit you in the face'

Verb-coding is very rare in ditransitive constructions, but there is at least one case in Makassarese:

(47) Makassarese (Jukes 2006: 341)
a. *La-ku-sare-ko doe'.*
FUT-1-give-2.F money
'I'll give you some money.'

b. *La-ku-saré-ang-ko doe-kku*
FUT-1-give-APPL-2.F money-1.POSS
'I'll give you my money.'

c. *La-ku-saré-ang-ko.*
FUT-1-give-APPL-2.F
'I'll give it to you.'

These constructions are not coding splits in the strict sense, of course (because the relevant arguments are always coded in the same way), but the pattern is clearly closely related.

11. Alternations

11.1. Classical passive and dative alternations

Another widespread phenomenon in languages, closely related to coding splits, is ARGUMENT-CODING ALTERNATIONS. An alternation is a situation where two different coding patterns can be used alongside each other, with roughly the same meaning.

(48) passive alternation in English
a. *The men cut down the tree.*
b. *The tree was cut down by the men.*

(49) dative alternation in English
a. *The girl gave the boy the bag.*
b. *The girl gave the bag to the boy.*

By **asymmetric coding**, I mean a situation where one of the alternates either has special verb coding, as is typically the case in passives (Haspelmath 1990), or the argument flagging in one of the alternates is clearly shorter. The latter is the case in the English Dative alternation, where the Double Object construction (in 72a) shows no preposition, while the Prepositional Dative construction has an extra dative preposition. Some alternations are asymmetric in both ways at the same time: Thus, the English passive alternation has special verb coding (the passive auxiliary *be* plus the Past Participle form of the verb), and in addition the argument flagging is longer (the preposition *by* on the agent argument).

(50) **Universal 5**
In an asymmetric argument-coding alternation, the longer alternant tends to be used in situations that deviate from canonical associations of roles and referential prominence.

Here the most relevant subtype of referential prominence is **topicality** or **givenness**. For passives, which are by definition asymmetric, this means that they tend to be used when the A is not given/topical, and/or when the P is not new information. We can even formulate this as a universal:

(51) **Universal 6**
If a passive alternation is sensitive to givenness, then the passive alternant tends to be used when the A is not given information and/or the P is not new information.

That this is indeed the case has been known for quite some time (e.g. Siewierska 1984; Shibatani 1985), although I am not aware of any formulations that are as general as Universal 6.

For dative alternations, which are also by definition asymmetrical, we can likewise say that the longer alternant occurs when unexpectedly the R is not given/topical, and/or when the T is not new information.

(52) **Universal 7**
If a dative alternation is sensitive to givenness, then the dative alternant tends to be used when the R is not given information and/or the T is not new information.

For English, this is clearly the case (e.g. Thompson 1990; Collins 1995), and the situation in related languages is not very different (see van der Beek 2004 for Dutch, for example). However, dative alternations are not very common in the world's languages (Siewierska 1998).

11.2. Split alternations

In addition to coding splits and coding alternations, we also find an intermediate phenomenon that provides further conformation for the present approach: Some construction pairs alternate under some conditions, but are in complementary distribution in other conditions. I call these situations SPLIT ALTERNATIONS.

• in Lummi (Jelinek & Demers 1983), the ordinary Active construction is used only when the scenario is nominality-downstream (as in 53a) or nominality-balanced (as in 53b). When the scenario is upstream, the Passive construction (with the verb suffix *-ŋ,* and the Oblique preposition on the A) is obligatory (see 53c).

(53)a. (pers > N)
x̣čit-s *cə* *swəyʔqəʔ*
know-3SG DET man
'He knows the man.' (NOT: 'The man knows him.')

b. (pers > pers)
x̣čit-s
know-3SG

'He knows it.'

c. (N > pers)
x̣čit-ŋ *ə* *cə* *swəyʔqəʔ*
know-PASS OBL DET man
'He is known by the man.' (= The man knows him)

(54)a. (N > N)
x̣čit-s *cə* *swəyʔqəʔ* *cə* *swiʔqoʔəɬ*
know-3SG DET man DET boy
'The man knows the boy.'

b. (N > N)
x̣čit-ŋ *cə* *swiʔqóʔəɬ* *ə* *cə* *swəyʔqəʔ*
know-PASS DET boy OBL DET man
'The boy is known by the man.'

ditransitives: Koyra Chiini:
• postposition *+se* is the only possibility in not fully canonical situations
• the shorter Double-Object construction (50d) is possible only in nominality-downstream patterns

(55) Koyra Chiini Songhay (Heath 1999: §9.1.2)
a. (nom > pers, nominality-upstream)
Ay *noo* *ga* [*woy* *di* *se*].
1SG.SBJ give 3SG.OBJ woman DEF DAT
'I gave it to the woman.' (= Heath's 445b)

b. (nom > nom, balanced)
Ay *noo* [*woy* *di* *se*] *hari.*
1SG.SBJ give woman DEF DAT water
'I gave the woman some water.' (= 445d)

c. (pers > pers, balanced)
No-o *noo* *ga* [*i* *se*].
2SG.SBJ-IMPF give 3SG.OBJ 3PL DAT
'You give it to them.' (= 449b)

d. (pers > nom, downstream)
No-o *noo* *gi* *njerfu.*
2SG.SBJ-IMPF give 3PL.OBJ money
'You give them some money.' (= 447b)

However, in the nominality-downstream scenario, the Double-Object construction is not obligatory, but either construction is possible:

(56) a. *Yer o har i se i ma noo yer se idumbo.*
1PL.SBJ IMPF say 3PL DAT 3PL.SBJ SBJV give 1PL DAT piece
'We tell them to give us a piece.' (= 448)

b. *boro kul kaa hin ka noo yer a wane fahaamey di*
person all REL can INF give 1PL 3SG POSS understanding DEF
'anyone who can give us information about it' (= 447a)

12. Explanation

(57) **Universal 1: The role-reference association universal**
Deviations from canonical associations of roles and referential prominence tend to be coded by longer grammatical forms.

Frequency-based coding efficiency:

• Languages tend to use efficient coding, i.e. zero or short coding for frequently occurring meanings and functions, and overt and long coding for rarely occurring functions.

• Through adaptability in language use, languages come to have or restore efficient patterns.

For role-reference associations, this kind of explanation has long been advocated, for example by Caldwell (1857: 276), who deserves to be quoted again here (cf. Filimonova 2005: 78):

> "[. . .] the principle that it is more **natural** for rational beings to act than to be acted upon; and hence when they do **happen to be** acted upon – when the nouns by which they are denoted are to be taken objectively – it becomes necessary, in order to **avoid misapprehension**, to suffix to them the objective case-sign"

References

Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language & Linguistic Theory* 17. 673–711.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483.

Aristar, Anthony Rodrigues. 1997. Marking and hierarchy types and the grammaticalization of case-markers. *Studies in Language* 21(2). 313–368. doi:10.1075/sl.21.2.04ari.

Baker, Mark C. 2015. *Case*. Cambridge: Cambridge University Press.

Becher, Jutta. 2005. Ditransitive Verben und ihre Objekte im Wolof: Positionsregeln und Kombinierbarkeit. *Hamburger afrikanistische Arbeitspapiere (HAAP)* 3. 13–27.

Bickel, Balthasar. 1995. In the vestibule of meaning: Transitivity inversion as a morphological phenomenon. *Studies in Language* 19(1). 73–127. doi:10.1075/sl.19.1.04bic.

Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge, MA: MIT Press.

Bossong, Georg. 1985. *Differenzielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bossong, Georg. 1998. Le marquage différentiel de l'objet dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, 193–258. Berlin: Mouton de Gruyter.

Bybee, Joan. 2007. Introduction. *Frequency of use and the organization of language*, 5–22. Oxford: Oxford University Press.

Creissels, Denis & Mounole. 2011. Animacy and spatial cases: Typological tendencies, and the case of Basque. In Seppo KittilÃ¤, Katja VÃ¤sti & Jussi Ylikoski (eds.), *Case, animacy and semantic roles*, 155–182. Amsterdam; Philadelphia: John Benjamins Pub. Co.

Cristofaro, Sonia. 2013. The referential hierarchy: Reviewing the evidence in diachronic perspective. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 69–93. Berlin: De Gruyter Mouton.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

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

Essegbey, James. 2010. Inherent complement verbs and the basic double object construction in Gbe. In Enoch Oladé Aboh & James Essegbey (eds.), *Topics in Kwa syntax*, 177–193. Dordrecht: Springer.

Georgi, Doreen. 2012. A local derivation of global case splits. In Artemis Alexiadou, Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax*, 305–336. (Linguistische Arbeiten). Tübingen: Niemeyer.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.

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

Haiman, John. 1983. On some origins of switch-reference marking. In John Haiman (ed.), *Switch-reference and universal grammar*, 105–128. Amsterdam: Benjamins.

Haspelmath, Martin & Susanne Maria Michaelis. 2017. Why independent possessive person-forms are longer: Diverse sources conspiring towards a uniform result. *forthcoming*.

Haspelmath, Martin, Andreea Calude, Michael Spagnol, Heiko Narrog & Elif Bamyacı. 2014. Coding causal-noncausal verb alternations: A form-frequency correspondence explanation. *Journal of Linguistics* 50(3). 587–625. doi:10.1017/S0022226714000255.

Haspelmath, Martin. 1990. The grammaticization of passive morphology. *Studies in Language* 14(1). 25–72.

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity*, 87–120. Amsterdam: Benjamins.

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

Haspelmath, Martin. 1999. Explaining article-possessor complementarity: Economic motivation in noun phrase syntax. *Language* 75(2). 227–243.

Haspelmath, Martin. 2004. Explaining the Ditransitive Person-Role Constraint: A usage-based approach. *Constructions* 2.

Haspelmath, Martin. 2008a. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1). 40–63.

Haspelmath, Martin. 2008b. Creating economical morphosyntactic patterns in language change. In Jeff Good (ed.), *Linguistic universals and language change*, 185–214. Oxford: Oxford University Press.

Haspelmath, Martin. 2008c. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1). 1–33. doi:10.1515/COG.2008.001.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.

Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Lingustic Typology* 15(3). 535–567.
Haspelmath, Martin. 2013. On the cross-linguistic distribution of same-subject and different-subject "want" complements: Economic vs. iconic motivation. *SKY Journal of Linguistics* 26. 41–69.
Haspelmath, Martin. 2014. On system pressure competing with economic motivation. In Brian MacWhinney, Andrej L. Malchukov & Edith A. Moravcsik (eds.), *Competing motivations in grammar and usage*, 197–208. Oxford: Oxford University Press.
Haspelmath, Martin. 2015. Ditransitive constructions. *Annual Review of Linguistics* 1. 19–41.
Haspelmath, Martin. 2017a. Can cross-linguistic regularities be explained by change constraints? *forthcoming*.
Haspelmath, Martin. 2017b. Explaining alienability contrasts in adpossessive constructions: Predictability vs. iconicity. *forthcoming*.
Haspelmath, Martin & Andres Karjus. 2018. Explaining asymmetries in number marking: Singulatives, pluratives and usage frequency. *Linguistics*.
Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. New York: Oxford University Press.
Jäger, Gerhard. 2007. Evolutionary game theory and typology: A case study. *Language* 83(1). 74–109.
Keine, Stefan. 2010. *Case and agreement from fringe to core: A minimalist approach*. Berlin: De Gruyter.
Lestrade, Sander, Kees de Schepper & Joost Zwarts. 2011. The distribution of labor between adpositions and case within complex spatial PPs. *STUF - Language Typology and Universals* 64(3). 256–274. doi:10.1524/stuf.2011.0018.
Lockwood, Hunter T. & Monica Macaulay. 2012. Prominence hierarchies. *Language and Linguistics Compass* 6(7). 431–446. doi:10.1002/lnc3.345.
Malchukov, Andrej. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118. 203–221.
Maslova, Elena. 2003. *A grammar of Kolyma Yukaghir*. . Vol. 27. (Mouton Grammar Library). Berlin / New York: Mouton de Gruyter.
Mayerthaler, Willi. 1981. *Morphologische Natürlichkeit*. Wiesbaden: Athenaion.
Miestamo, Matti. 2005. *Standard negation: The negation of declarative verbal main clauses in a typological perspective*. Berlin: Mouton de Gruyter.
Moravcsik, Edith A. 1978a. On the case marking of objects. In Joseph H Greenberg (ed.), *Universals of human language*, vol. 4: Syntax, 249–289. Stanford: Stanford University Press.
Moravcsik, Edith A. 1978b. On the distribution of ergative and accusative patterns. *Lingua* 45. 233–279.
Piantadosi, S. T., H. Tily & E. Gibson. 2011. Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences* 108(9). 3526–3529.
Rude, Noel. 2009. Transitivity in Sahaptin. *Northwest Journal of Linguistics* 3(3). 1–37.
Schmid, Hans-Jörg. 2015. A blueprint of the Entrenchment-and-Conventionalization Model. *Yearbook of the German Cognitive Linguistics Association* 3(1). 3.
Schmidtke-Bode, Karsten. 2009. *A typology of purpose clauses*. Amsterdam: Benjamins.
Siewierska, Anna. 2009. Person asymmetries in zero expression and grammatical function. In Franck Floricic (ed.), *Essais de Linguistique Generale et de Typologie Linguistique offerts au Professeur Denis Creissels à l'occasion de ses 65 ans*, 425–438. Paris: Presses de L'Ecole Normale Superieure.
Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian National University.
Stroński, Krzysztof. 2009. Variation of ergativity patterns in Indo-Aryan. *Poznań Studies in Contemporary Linguistics* 46(2). 237–253. doi:10.2478/v10010-010-0012-8.
Watters, David E. 2002. *A grammar of Kham*. Cambridge: Cambridge University Press.
Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, Mass.: M. I. T. Press.
Zúñiga, Fernando. 2006. *Deixis and alignment: Inverse systems in indigenous languages of the Americas*. . Vol. 70. (Typological Studies in Language). Amsterdam: Benjamins.

Tylex course "Explaining language universals", Voronovo (Moscow), 2017 September 5

3. Diachronic adaptation and mutational constraints

MARTIN HASPELMATH
MPI-SHH Jena & Leipzig University

1. The debate with Stephen Anderson and Sonia Cristofaro

Anderson, Stephen R. 2016. Synchronic versus diachronic explanation and the nature of the Language Faculty. *Annual Review of Linguistics* 2(1). doi:10.1146/annurev-linguistics-011415-040735. http://dx.doi.org/10.1146/annurev-linguistics-011415-040735

Haspelmath, Martin. 2016. Stephen Anderson on "diachronic explanation" (of what?). *Diversity Linguistics Comment*. https://dlc.hypotheses.org/888.

Cristofaro, Sonia. 2017. Implicational universals and dependencies. In N. J. Enfield (ed.), *Dependencies in language: On the causal ontology of linguistic systems*, 9–22. Berlin: Language Science Press. http://langsci-press.org/catalog/book/96.

Haspelmath, Martin. 2018. Can cross-linguistic regularities be explained by change constraints? *forthcoming*.

Anderson (2016):
"there are at present no convincingly demonstrated substantive universals governing the set of possible regularities in phonology",
citing Juliette Blievins's book *Evolutionary Phonology*.

an example of diachronic explanation in phonology:

final voicing in Lezgian

čep-edi	*čeb*	'day'
gat-u	*gad*	'summer'

(discussed by Yu 2004; Blevins 2006; Kiparsky 2008)

an example of diachronic explanation in syntax:

Icelandic lacks nominative reflexive pronouns

Now my question to Anderson is: **Can diachronic explanations also account for universal tendencies** (not only for idiosyncrasies, as seen earlier)? Anderson is not as clear about this as he should be. One of his examples in the 2016 paper, the lack of a nominative reflexive in Icelandic, clearly concerns an idiosyncratic phenomenon of one language, Icelandic. The other examples do seem to concern universal tendencies (and the 2005 paper is explicitly about morphological universals), but Anderson does not say clearly that **the universal tendencies are due to universal factors**. He repeatedly talks about "common paths of diachronic development", but "common" is not the same as "universal". If tone "commonly" develops from the loss of a syllable-final consonant, then this may lead us to expect that tone languages have open syllables, but if there is another, equally (or more) common source for tone (maybe a syllable-initial consonant quality), then the expectation disappears. In order to explain a universal tendency by diachrony, one needs to claim that there is a **diachronic asymmetry**: Not only is the diachronic path A > B common, but the reverse diachronic path B > A is impossible (or uncommon). In other words, we need a notion of **universal directionality of language change** if we want to explain universals by diachrony.

Thus, in order to convince me that we should pursue **diachronic explanations of universals**, I need to see evidence that **the diachronic mechanism is universal**, too. Accidental diachronic developments cannot give rise to universal patterns, by definition.

2. Four types of constraints (Haspelmath 2018)

functional-adaptive constraints:	what facilitates communication for speakers and hearers
representational constraints:	what is cognitively preferred or necessary (within the innate language faculty)
mutational constraints:	what is preferred or necessary in language change (= change constraints)
acquisitional constraints:	what is preferred or necessary in acquisition by children

• **Functional-adaptive constraints** are the kinds of factors that have been invoked by functionalists to explain cross-linguistic distributions (e.g. Tomlin 1986; Malchukov 2008; Hawkins 2014; among many others), e.g.

• phonological inventories favour five-vowel systems because these make the best use of the acoustic space (De Boer 2001)

• case systems favour overt ergatives for low-prominence nominals and overt accusatives for high-prominence nominals because of the association between roles and prominence status (Dixon 1994).

• These constraints are called *functional-adaptive* rather than merely *functional* to emphasize their role in explaining **systems**, not **usage** (the "functioning" of language). Functional linguists often focus on understanding the functioning of language in usage, but here my interest is in explaining how systems come to have properties that facilitate communication. (Moreover, I am not talking about the functions of individual parts of the system.)

• Good (2008) uses the term "external explanation" in roughly this sense (cf. also Newmeyer 1998), but all four types of constraints are external in that they are not part of the system. ("System-internal explanation" is just another word for what I called "descriptive explanation" earlier; I do not think that the notion of causality is relevant for such statements, so all causal explanatory factors are external.)

• **Representational constraints** are the kinds of factors that have been invoked by generativists to explain cross-linguistic universals. In the principles and parameters framework (Chomsky 1981), they were called the *principles of universal grammar*.

• For example, the principles of X-bar theory or binding theory have been regarded as representational constraints, as well as universal features and hierarchies of functional categories (e.g. Cinque 1999). The general idea is that "the unattested patterns do not arise as they cannot be generated in a manner consistent with Universal Grammar" (Smith et al. 2016).

• Representational constraints are usually regarded as very strong, i.e. as restrictions (and thus universal grammar is said to be *restrictive*; cf. also Haspelmath 2014). In Good's (2008) survey, these constraints are treated under the label of "structural explanations", but this term (like "system-internal explanations") is better reserved for general statements of regularities of language-particular systems.

• **Mutational constraints** (or change constraints) are constraints on possible diachronic transitions or possible diachronic sources, which can have an effect on synchronic distributions.

e.g. if nasal vowels only ever arise from VN sequences, this explains that all languages with nasal vowels also have nasal consonants, and that nasal vowels are rarer than oral vowels in the lexicon (Greenberg 1978).

e.g. if infixes only ever arise by metathesis from adfixes (= prefixes or suffixes), this explains that they only occur in peripheral position (Plank 2007: 51).

e.g. adpositions only arise from nouns in possessor-noun constructions, this explains that their position correlates with the position of possessed nouns

• The notion of mutational constraints is not completely new (Plank 2007: §2 calls them "diachronic laws"), but I introduced a new term in order to make clear that **the causal factor is located within the process of change**, rather than diachronic change merely realizing a pattern that is driven by functional-adaptive constraints.

• One could also frame the contrast between mutational constraints and functional-adaptive constraints in terms of ***source-oriented*** vs. ***result-oriented*** factors (Cristofaro 2017), or one could say that mutational constraints locate the causal factors within the ***mechanisms*** of change (Bybee 2006). These are just alternative ways of saying that cross-linguistic distributions are due to mutational constraints.

Informally, instead of talking about "result-oriented factors", one could also say that functional-adaptive constraints are "pull forces" that **attract** the variable development into a certain preferred state.

(• ACQUISITIONAL CONSTRAINTS are factors that impact the acquisition of language and that have an effect on cross-linguistic distributions. Such constraints are briefly discussed by Anderson (2016), but they do not seem to play a big role in linguistics. Generative linguists who are concerned with learnability issues generally assume that what can be represented can also be learned, so that there is no distinction between

representational constraints and what can be learned. This type of constraint is mentioned here only in passing, for the sake of completeness.)

3. Functional-adaptive constraints become part of the conventional system through diachronic change

Adaptation presupposes a process of evolution – there is no such thing as functional adaptedness without prior change (Bybee 1988).

(Creationism is a logical option, as in biology, i.e. one might suppose that adaptedness became part of language through the purposeful actions of a benevolent creator. But nobody has proposed this for language, as far as I know, perhaps because rapid change is too obvious.)

This is the same in other areas of cultural adaptation. Institutions like rulers and upper classes take a while to evolve, and likewise linguistic adaptation evolves.

Functionalists often say things like:

> "grammars code best what speakers do most" (DuBois 1985)
>
> "hard constraints mirror soft constraints" (Bresnan et al. 2001)
>
> "grammatical rules are crystallized usage preferences" (Schmidtke-Bode 2018)

• But how can language use have an effect on the language system?

de Saussure 1916	*langue*	= language structure (conventions shared by the community)
	parole	= speech/language use

> "[O]n peut comparer la langue à une symphonie, dont la réalité est indépendante de la manière dont on l'exécute; les fautes que peuvent commettre les musiciens qui la jouent ne compromettent nullement cette réalité." (de Saussure 1916[1972]:36)

Chomsky 1965	*competence*	= language structure (mental state of an individual)
	performance	= speech/language use

> "(*Colorless green ideas sleep furiously/Furiously sleep ideas green colorless*...) Evidently, one's ability to produce and recognize grammatical utterances is not based on notions of statistical approximation and the like."
> (Chomsky 1957: 15-16)

"usage-based linguistics": language structure is considered together with language use

language structure is
– a coherent abstract system (to some extent)
– constrained by the properties of the innate language faculty (to some extent)
– shaped by (the needs of) language use (to a large extent)

• Language structure is malleable

– Speakers are not rigidly limited by the conventions, but they can occasionally violate them

– If a convention violation is picked up by others, the violation becomes a new convention

– Innovation and propagation (which together constitute language change) are not random processes
  – innovations respond to usage preferences (Croft 2000)
  – propagation also responds to usage preferences (Haspelmath 1999)

– Language change is a type of evolutionary process:

  altered replication (mutation, innovation) + selection (diffusion, propagation)
  (cf. Haspelmath 1999, Nettle 1999, Croft 2000, Ritt 2004, Blevins 2004)

• Recent experimental work on optional marking

Kurumada & Jaeger (2015) on optional case-marking in Japanese:

> The results of our recall experiments on Japanese optional case-marking showed that Japanese speakers were more likely to produce case-marking when the grammatical function assignment indicated by the case-marker would otherwise be hard to infer. This preference contributes to robust information transfer by providing additional cues to the intended sentence interpretation when they are most needed. In addition, we observed indepen-

4. The case for "diachronic explanation of language universals" (Bybee, Cristofaro, Anderson)

4.1. Recurrent paths of change

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

Similarly, sound changes occur across languages under similar conditions (lenitions, assimilations, etc.).

A more radical position:

"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 1999),
(b) or from locally motivated preferred **pathways of change** (Bybee 2001, Blevins 2004, grammaticalization literature)"

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

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"

e.g. Indo-Aryan: ergative pattern develops from a passive-like pattern in past/perfective forms:

Hindi

(a) present/imperfective (accusative/neutral)

Ram	*garii*	*caalaa-taa*	*(hai)*
Ram.M.SG.NOM	car.F.SG.NOM	drive-IMPF.M.SG	be.PRES.3SG

'Ram drives a car.'

(b) past/perfective (ergative < **passive-like construction**)

Ram=ne	*garii*	*caalaa-yii*	*(hai)*
Ram.M.SG=ERG	car.F.SG.NOM	drive-PERF.F.SG	be.PRES.3SG

'Ram has driven a/the car.' (< "The car was driven by Ram.")

Georgian:
accusative pattern develops from an antipassive-like pattern
in present/imperfectve forms.

4.2. Some terminology: systems, cross-linguistic distributions, motivations

descriptive system regularities — rules, schemas, constructions

cross-linguistic distributions — (statistical) universals, tendencies

causal motivating factors — **constraints**:
restrictions (very strong), preferences (weaker)

types of constraints:

functional constraints:	what is preferred in language use for communication
representational constraints:	what is cognitively possible (within the language faculty)
mutational constraints:	what can happen in language change
(**acquisitional constraints**:	what can be acquired)

"the true universals of language are not synchronic patterns at all, but the mechanisms of change that create these patterns" (Bybee 2006)	There are no (true) universals or tendencies. **Mutational constraints** lead to apparent universals or tendencies. (MH: There are true universals/tendencies, which can be demonstrated; but possibly some of them are due to **mutational constraints**.)
"universal preferences may result (a) from preference principles that **guide** (or "select") the result of diachrony" (Bickel 2007)	Universals/tendencies may result from **functional constraints**, which guide change into directions preferred for communication.
"universal preferences may result (b) from locally motivated preferred **pathways of change"** (Bickel 2007)	Universals/tendencies may result from **mutational constraints**, i.e. certain kinds of changes are not possible.
"achronic laws vs. diachronic laws" (Plank 2007)	**universals** vs. **mutational constraints** (or: functional/representational constraints vs. mutational constrants) (or: synchronic universals vs. diachronic universals?)
"there are no (or at least very few) **substantive universals** of language" (Anderson 2016)	There are no or few **representational constraints**.
"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" (Anderson 2016)	Universals/tendencies are generally due to **mutational constraints**, not to **representational constraints**. (MH: or due to functional constraints, and maybe acquisitional constraints)

"the [referential hierarchy patterns] do not obviously originate from the **mechanisms** that have been postulated to account for these patterns **on synchronic grounds**, for example animacy and individuation" (Cristofaro 2013: 87)	The referential hierarchy universals are not obviously due to **mutational constraints** involving animacy and individuation. (MH: Right, they are more likely due to functional constraints.)

4.3. Recurrent changes cannot explain universals

We all know that diachrony can beautifully account for idiosyncrasies
(e.g. remnant alternations like *thief/thieves*).

But can diachrony also explain universals?

Anderson (2005; 2016): "common paths of change"

> "[aspect-based split ergativity] It has been observed that in all familiar cases of such a split, nominative/accusative marking is found in clauses where the verbal tense/aspect is imperfective (or continuative, progressive, etc., or some tense form that is a reflex of such an aspect at an earlier stage), whereas ergative/absolutive marking is found in clauses with perfective aspect or some tense form descended from that." (Anderson 2016: §4.1)
>
> "the observed generalization ... is actually the result of the accidental convergence of a number of logically independent paths of historical development ... As it happens, common sources for a new perfective, on the one hand, and for a new imperfective, on the other, converge on similar patterns of split ergativity, although they are quite unrelated to each other." (Anderson 2016: §4.1)

Can universal patterns be due to "accidental convergence"?

> **No**: Anderson would have to claim that the alleged "universal patterns" are only APPARENTLY universal.

Recurrent change (= "paths of change") is not the same as a **mutational constraint**: if a change A > B is very **common**, this cannot explain any observed synchronic universal tendency unless we also know that the opposite change B > A is inexistent (or very rare).

4.4. Mutational constraints

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).

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

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

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

4.5. Evidence for adaptation: Convergence of diverse sources/processes/mechanisms

If several unrelated sources converge on the same result, we have **a remarkable coincidence**:

"As it happens, common sources for a new perfective, on the one hand, and for a new imperfective, on the other, converge on similar patterns of split ergativity, although they are quite unrelated to each other." (Anderson 2016; cf. Anderson 1977: "mechanisms")

"Different instances of the same configuration can also be a result of very different processes. For example, phonological erosion, the transfer of plural meaning from a quantifier to an accompanying element, and the grammaticalization of distributives into plural markers can all give rise to a configuration with zero marking for singular and overt marking for plural, yet these processes do not obviously have anything in common." (Cristofaro 2017)

phonological erosion:	English *day/day-s* < Proto-Germanic *dag-z/dag-az*
from quantifier:	Bengali *chēlē-rā* '(some) of child(ren)' > 'children'
from distributive:	Southern Paiute *qa'nɪ* 'house', *qaŋqa'nɪ* 'house(s) here and there' > 'houses'

The convergence of diverse processes on a uniform result could be accidental, but in that case it would not really be "convergence", and it could not explain a universal tendency.

True convergence can be explained only by some "pull force" – most straightforwardly, it can be explained by a **functional-adaptive constraint**:

– ergative alignment in perfective clauses is functionally adaptive, and
– accusative alignment in imperfective clauses is functionally adaptive (e.g. DeLancey 1981)
– zero singulars vs. overt plurals are functionally adaptive (cf. Haspelmath & Karjus 2017)

5. Functional constraints do not lead us to expect uniform mechanisms or pathways of change

5.1. Do we need evidence for an explanation from the pathways of change?

> "in language universals, causal factors are linguistic changes that create particular synchronic states, and the existence of massive cross-language similarity in synchronic states implies powerful parallels in linguistic change. ... **the validity of a principle as explanatory can only be maintained if it can be shown that the same principle that generalizes over the data also plays a role in the establishment of the conventions** described by the generalization" (Bybee 1988: 352)

> "These [functional] explanations ... have mainly been proposed based on the synchronic distribution of the relevant grammatical phenomena, not **the actual diachronic processes** that give rise to this distribution in individual languages. In what follows, it will be argued that **many such processes do not provide evidence for the postulated dependencies between grammatical phenomena**, and suggest alternative ways to look at implicational universals in general." (Cristofaro 2017)

This is wrong.

If a generalization is 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)

• In general, we do not know much about language change and how and why it happens. The primary evidence for functional-adaptive explanations is the fit between the causal factor and the observed outcome.

• If there is **a good fit**, e.g. if languages overwhelmingly prefer the kinds of word orders that allow easy parsing (Hawkins 2014), or if they tend to show economical coding of grammatical categories (Haspelmath 2008), the best explanation is in functional-adaptive terms, **as long as there is a way for languages to acquire these properties**.

• The latter requirement is always met, as there are no synchronic states which cannot have arisen from other states.

5.2. Mechanisms of change in other fields

Explanations of regularities in the world-wide distribution of cultural traits often appeal to functional-adaptive factors in adjacent fields as well. For example, **anthropologists** sometimes explain religion by prosociality, or monogamy by group-beneficial effects (e.g. Paciotti et al. 2012; Henrich et al. 2012). The issue here is whether better explanations are available, **not whether there is a way for religion or marriage to develop**.

We know little about how religion and marriage first arose or generally arise in societies, and it is very difficult to study the diachronic developments. But we can try to **correlate structural traits of human societies with other traits and draw inferences about possible causal factors**. There is no perceived need in this literature to show that the mechanisms by which religion or monogamy arise must be of a particular type.

Basically, **when the result is preferred, any kind of change can give rise to the result**, and we do not need to understand the nature of the change, let alone show that the change was motivated by the result.

5.3. Diverse paths to shortness of coding

Shortness of frequent words is functionally adaptive –
but what are the pathways that lead to length asymmetries?

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

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

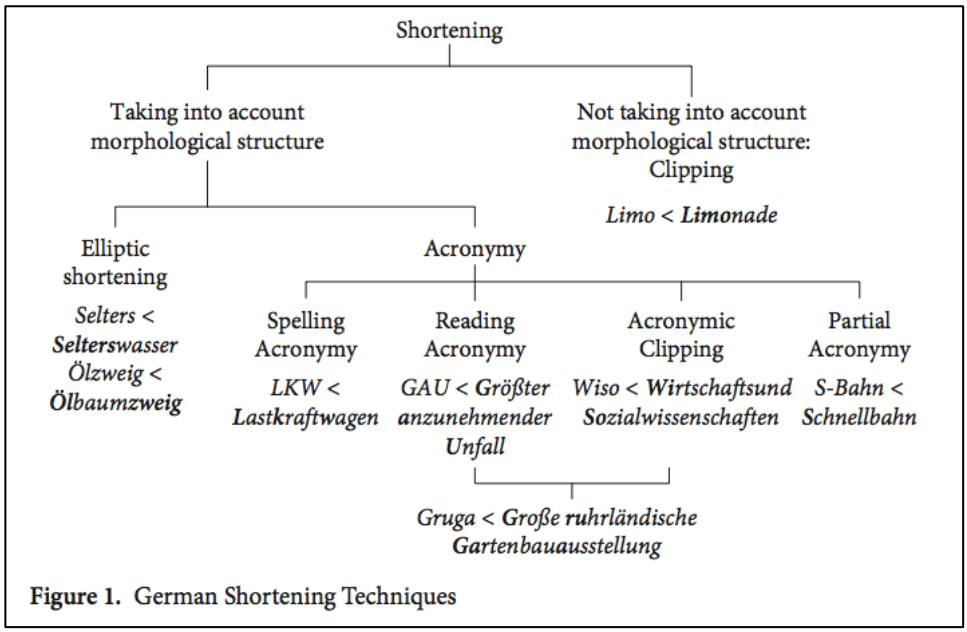


Figure 1. German Shortening Techniques

Some of these may be fully conscious developments, but even in this domain, there are many proposals that are not accepted by the speaker community (e.g. *SSW* for *Sommerschlussverkauf*, 'summer sale').

• similarly, again contra Bybee, in a large number of cases, **grammatical coding asymmetries** arise not by shortening of the more frequent form, but by **"lengthening" of the less frequent form**, e.g.

her	*her-self*
come	*will come*
praise!	*let him praise*

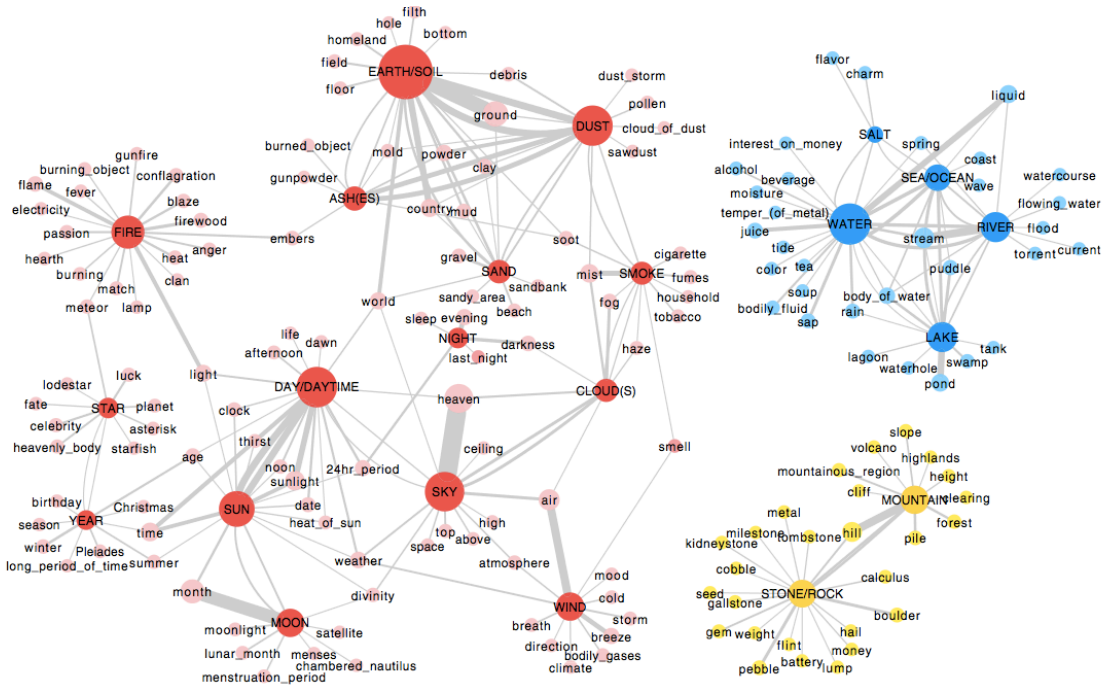
• cases of shorterning are also found, however, e.g.

my	*mine*	
śpiewa-Ø	*śpiewa-sz*	(Polish 'sings', 'you sing')

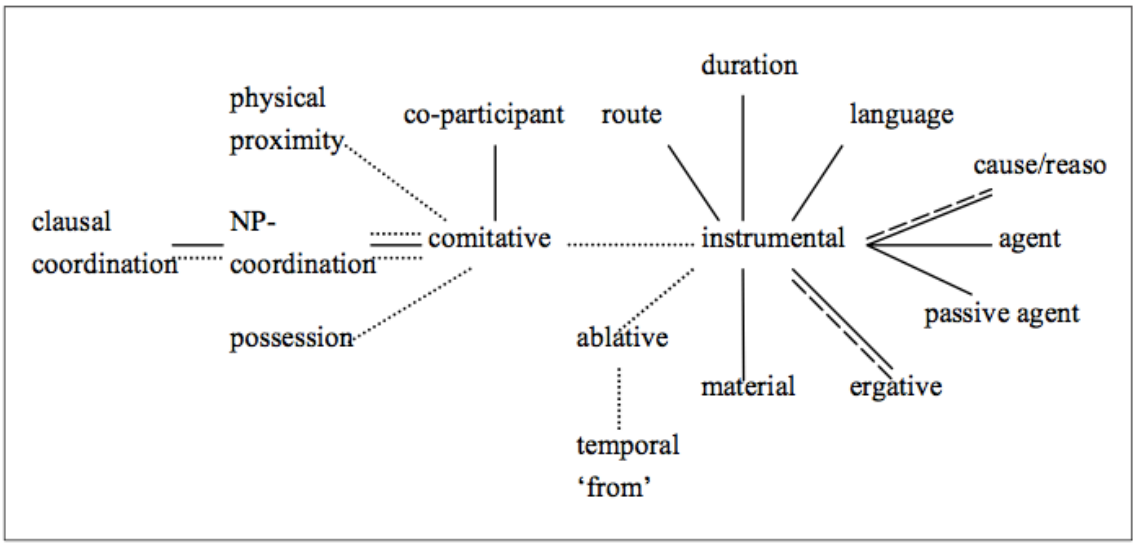
• thus, there is **multi-convergence of different pathways** by which coding asymmetries can come about, which indicates that there is a preferred result (and the change is "goal-oriented".

6. Mutational constraints seem to be the right explanation for coexpression universals

e.g. coexpression universals seem to be due to general tendencies of semantic change (Youn et al. 2016)



Narrog & Ito (2007):



According to Cristofaro (2010), coexpression universals are due to general processes of semantic change, and this may well be so.

Thus:
– Why is instrumental intermediate between comitative and ergative?
– Because comitative can extend to instrumental (but not vice versa), and instrumental can extend to ergative (but not vice versa). Comitative cannot extend directly to ergative, nor can ergative extend to comitative.

Next question: What is the explanation of these semantic extensions?
? Answer: Their "proximity in semantic space" ?

> But what is "semantic space"? Is this a kind of "UG of semantics"?
> Are we dealing in reality with representational constraints?

Cf. Croft (2001: 364): semantic maps provide us with

> "a geography of the human mind, which can be read in the facts of the world's languages in a way that the most advanced brain scanning techniques cannot ever offer us"

But even if this were true, why do many changes only go in one direction, not in the other direction?

– We do not really know, and it may be wise to be modest about our ability to understand language change.

References

Anderson, Stephen R. 1977. On mechanisms by which languages become ergative. In Charles N. Li (ed.), *Mechanisms of syntactic change*. Austin: University of Texas Press.

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. 2008. The logical structure of linguistic theory. *Language* 84. 795–814.

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

Aristar, Anthony R. 1991. On diachronic sources and synchronic pattern: An investigation into the origin of linguistic universals. *Language* 67(1). 1–33.

Bickel, Balthasar, Alena Witzlack-Makarevich, Taras Zakharko & Giorgio Iemmolo. 2015. Exploring diachronic universals of agreement: Alignment patterns and zero marking across person categories. In Jürg Fleischer, Elisabeth Rieken & Paul Widmer (eds.), *Agreement from a diachronic perspective,* 29–51. Berlin: De Gruyter Mouton.

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

Bickel, Balthasar. 2011. Absolute and statistical universals. In Patrick Colm Hogan (ed.), *The Cambridge encyclopedia of the language sciences*. Cambridge: Cambridge University Press.

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

Blevins, Juliette. 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32(2). 117–166.

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

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

Bybee, Joan L. 1988. The diachronic dimension in explanation. In John A. Hawkins (ed.), *Explaining language universals*, 350–379. Oxford: Blackwell.

Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

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

Bybee, Joan. 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. 2006. Language change and universals. In Ricardo Mairal & Juana Gil (eds.), *Linguistic universals*, 179–194. Cambridge: Cambridge University Press. (Originally published as Bybee 2003.)

Bybee, Joan. 2007. Introduction. *Frequency of use and the organization of language*, 5–22. Oxford: Oxford University Press.

Chomsky, Noam A. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic approach*. New York: Oxford University Press.

Collins, Jeremy. 2016. Some language universals are historical accidents. *Humans who read grammars*. http://humans-who-read-grammars.blogspot.com/2016/08/some-language-universals-are-historical.html

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. 2nd edition. Oxford: Blackwell.

Creissels, Denis. 2008. Direct and indirect explanations of typological regularities: The case of alignment variations. *Folia Linguistica* 42(1). 1–38.

Creissels, Denis. 2014. Functive phrases in typological and diachronic perspective. *Studies in Language* 38(3). 605–647. doi:10.1075/sl.38.3.07cre.

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

Cristofaro, Sonia. 2012. Cognitive explanations, distributional evidence, and diachrony. *Studies in Language* 36(3). 645–670. doi:10.1075/sl.36.3.07cri.

Cristofaro, Sonia. 2013. The referential hierarchy: Reviewing the evidence in diachronic perspective. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 69–93. Berlin: De Gruyter Mouton.

Cristofaro, Sonia. 2014. Competing motivation models and diachrony: What evidence for what motivations? In Brian MacWhinney, Andrej L. Malchukov & Edith A. Moravcsik (eds.), *Competing motivations in grammar and usage*, 282–298. Oxford: Oxford University Press.

Cristofaro, Sonia. 2017. Dependencies and explanations of implicational universals. In N. J. Enfield (ed.), *Dependencies in language*. Berlin: Language Science Press.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

Croft, William. 2000. *Explaining language change: An evolutionary approach*. Harlow: Longman.

De Boer, Bart. 2001. *The origins of vowel systems*. Oxford: Oxford University Press.

Dediu, Dan & Bart de Boer. 2016. Language evolution needs its own journal. *Journal of Language Evolution* 1(1). 1–6. doi:10.1093/jole/lzv001.

DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57. 626–657.

Dik, Simon C. 1997. *The theory of functional grammar*. Berlin: Mouton de Gruyter.

Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Dressler, Wolfgang U, Willi Mayerthaler, Oswalt Panagl & Wolfgang U. Wurzel. 1987. *Leitmotifs in natural morphology*. Amsterdam: Benjamins.

Everett, Caleb, Damián E. Blasí & Seán G. Roberts. 2016. Language evolution and climate: The case of desiccation and tone. *Journal of Language Evolution* 1(1). 33–46. doi:10.1093/jole/lzv004.

Foley, William A & Jr. Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Garrett, Andrew. 1990. The origin of NP split ergativity. *Language* xx 261–296.

Gildea, Spike & Fernando Zúñiga. 2016. Referential hierarchies: A new look at some historical and typological patterns. *Linguistics* 54(3). 483–529. doi:10.1515/ling-2016-0007.

Givón, Talmy. 1979. *On understanding grammar*. New York/San Francisco/London: Academic Press.

Good, Jeff. 2008. Introduction. In Jeff Good (ed.), *Linguistic universals and language change*, 1–19. Oxford: Oxford University Press.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.

Greenberg, Joseph H. 1969. Some methods of dynamic comparison in linguistics. In Jan Puhvel (ed.), *Substance and structure of language*, 147–203. Berkeley: University of California Press.

Greenberg, Joseph H. 1978. Diachrony, synchrony and language universals. In Joseph H Greenberg, Charles A Ferguson & Edith A Moravcsik (eds.), *Universals of human language I: Method and theory*, 61 – 92. Stanford: Stanford University Press.

Haspelmath, Martin & Andres Karjus. 2017. Explaining asymmetries in number marking: Singulatives, pluratives and usage frequency. *Linguistics*, to appear

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

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

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. 2004. On directionality in language change with particular reference to grammaticalization. In Olga Fischer, Muriel Norde & Harry Perridon (eds.), *Up and down the cline: The nature of grammaticalization*, 17–44. Amsterdam: Benjamins.

Haspelmath, Martin. 2008. Parametric versus functional explanations of syntactic universals. In Theresa Biberauer (ed.), *The limits of syntactic variation*. Amsterdam: Benjamins.

Haspelmath, Martin. 2014. Comparative syntax. In Andrew Carnie, Yosuke Sato & Dan Siddiqi (eds.), *The Routledge handbook of syntax*, 490–508. London: Routledge.

Haspelmath, Martin. 2016. Stephen Anderson on "diachronic explanation" (of what?). *Diversity Linguistics Comment*. https://dlc.hypotheses.org/888.

Haspelmath, Martin. 2017. Explaining alienability contrasts in adpossessive constructions: Predictability vs. iconicity. *To appear, Zeitschrift für Sprachwissenschaft*

Haspelmath, Martin. 2018. Can cross-linguistic regularities be explained by change constraints? To appear in an edited volume.

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

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Henrich, Joseph, Robert Boyd & Peter J. Richerson. 2012. The puzzle of monogamous marriage. *Phil. Trans. R. Soc. B* 367(1589). 657–669. doi:10.1098/rstb.2011.0290.

Keenan, Edward Louis & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63–99.

Kiparsky, Paul. 2008. Universals constrain change; change results in typological generalizations. In Jeff Good (ed.), *Linguistic universals and language change*, 23–53. Oxford: Oxford University Press.

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.

Labov, William. 1965. On the mechanism of linguistic change. In Charles W. Kreidler (ed.), *Report of the Sixteenth Annual Round Table Meeting on Linguistics and Language Studies*, vol. 18, 91–114. (Georgetown Monographs on Language and Linguistics). Washington, D.C.: Georgetown University Press.

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

Li, Charles N. (ed.). 1977. *Mechanisms of syntactic change*. Austin: University of Texas Press.

Machamer, Peter, Lindley Darden & Carl F. Craver. 2000. Thinking about mechanisms. *Philosophy of Science*. 1–25.

Malchukov, Andrej. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118. 203–221.

McMahon, April & Robert McMahon. 2012. *Evolutionary linguistics*. Cambridge: Cambridge University Press.

Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* 12(26). 6.

Newmeyer, Frederick J. 1998. *Language form and language function*. Cambridge: MIT Press.

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

Paciotti, B., Richerson, P. J., Lubell, M., Waring, T., McElreath, R., Efferson, C., & Edsten, E. (2012). Are Religious Individuals More Generous, Trusting, and Cooperative? An Experimental Test of the Effect of Religion on Prosociality

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and brain sciences* 13(4). 707–727.

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.

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.

Ritt, Nikolaus. 2004. *Selfish sounds and linguistic evolution: A Darwinian approach to language change*. Cambridge: Cambridge University Press.

Ronneberger-Sibold, Elke. 2014. Tuning morphosemantic transparency by shortening. In Franz Rainer, Francesco Gardani, Hans Christian Luschützky & Wolfgang U. Dressler (eds.),

Morphology and meaning: Selected papers from the 15th International Morphology Meeting, Vienna, February 2012, 275–287. Amsterdam: Benjamins.

Sapir, Edward. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace & Co.

Schlegel, Friedrich von. 1808. *Ueber die Sprache und Weisheit der Indier: Ein Beitrag zur Begründung der Alterthumskunde*. Heidelberg: Mohr und Zimmer.

Schleicher, August. 1850. *Die Sprachen Europas in systematischer Uebersicht*. Bonn: H.B. König.

Stassen, Leon. 1985. *Comparison and universal grammar*. Oxford: Blackwell.

Tomlin, Russell S. 1986. *Basic word order: Functional principles*. London: Croom Helm.

Youn, Hyejin, Logan Sutton, Eric Smith, Cristopher Moore, Jon F. Wilkins, Ian Maddieson, William Croft & Tanmoy Bhattacharya. 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.

Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, Mass.: M.I.T. Press.

Tylex course "Explaining language universals", Voronovo (Moscow), 2017 September 6

4. Conflating analysis and explanation

MARTIN HASPELMATH
MPI-SHH Jena & Leipzig University

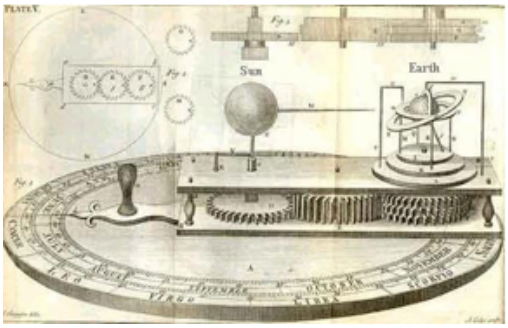
1. Analysis/description as explanation?

Analysis can be seen as part of the same enterprise as causal explanation

– if the analytical/descriptive apparatus is restricted

– and if the restrictions can be attributed to an external cause

cf. Newtonian physics



cf. houses made of Lego:

2. Restricted description as explanation in linguistics

In linguistics, descriptive frameworks are often said to be restrictive, i.e. they allow the description of some phenomena, but not of others.

(I call the approach *restrictivist* in Haspelmath (2014), but it is more or less coextensive with the generative approach.)

(A) First example: possessives and definite articles across languages

three out of four logically possible language types are attested:

in non-possessed NP	*in possessed NP*: no article	*in possessed NP*: article
article	English *Ø my book* *<u>the</u> book*	Italian *<u>il</u> mio libro* *<u>il</u> libro*
no article	Russian *Ø moja kniga* *Ø kniga*	—

(*Anti-English:
<u>the</u> my book
Ø book)

Universal:
If in a language a definite possessed NP has the definite article, then a definite non-possessed NP also has the definite article. (Haspelmath 1999a: 234)

Restrictivist explanation (1980s/1990s):

> UG allows possessives to be determiners or adjectives,
> and allows only one item in the determiner slot. Anti-English would not be acquirable by children (Lyons 1986; Giorgi & Longobardi 1991).

(B) Second example: X-bar Theory

Observation: gaps in attested patterns -- some describable structures don't exist.

(1)	NP —> D [$_{N'}$ N PP]	*the [horse on the meadow]*
	VP —> Adv [$_{V'}$ V NP]	*often [eats a flower]*
	PP —> Adv [$_{P'}$ P NP]	*right [under the tree]*

(but not e.g. *NP —> VP [Adv P])

Redundancy needs to be "expressed" in the descriptive framework:
only phrase structures of the following type are allowed:

(2) **XP —> Y [$_{X'}$ X ZP]** (**X-bar schema**, Jackendoff 1977 etc.)

(More recent incarnation of the same approach: “Merge”)

The non-existence of the unattested structures has been ***explained*** by the new, more ***restrictive*** framework.

• Why don't some languages have rules like "NP --> VP P"?

• Answer 1: Because such structures are not describable by the framework. (???)
• Answer 2: Because the X-bar schema is part of Universal Grammar, i.e. such rules would not be acquirable.

(C) Third example: Inflection outside derivation

Observation: gaps in attested patterns -- some describable structures don't exist.

(3)	ROOT-deriv-infl	*ROOT-infl-deriv	e.g. German	*Handl-ung-en*
	infl-deriv-ROOT	*deriv-infl-ROOT	e.g. Arabic	*ya-ta-kallamu*
				3sg-REFL-speak

Redundancy needs to be "expressed" in the descriptive framework: only morphological structures with inflection outside derivation are allowed, because **derivation is lexical, and inflection is in a post-lexical syntactic component** (Anderson 1992).

Again, this architecture must be innate, because otherwise no explanation has been achieved.

(D) Fourth example: Pro-drop of topical arguments

Observation: gaps in attested patterns -- some describable structures don't exist.

no pro-drop when pronoun = topic:	English	(*She comes./*Ø comes.*)
pro-drop when pronoun = topic:	Italian	(*Ø viene./*Lei viene.*)
no pro-drop when pronoun = focus:	English	(*SHE comes./*Ø comes.*),
	Italian	(*LEI viene./*Ø viene.*)
pro-drop when pronoun = focus:		(unattested)

Redundancy needs to be "expressed" in the descriptive framework: only the constraint DROPTOPIC exists, no constraint DROPFOCUS exists (Grimshaw & Samek-Lodovici 1998). OT constraint tableaux:

(4) English

she$_{TOPIC}$ comes	FAITHFUL	DROPTOPIC
+ *she comes*		*
Ø comes	*!	

(5) Italian

she$_{TOPIC}$ comes	DROPTOPIC	FAITHFUL
lei viene	*!	
+ *Ø viene*		*

3. The ideology

Ultimately, the descriptive framework should be able to describe only the possible languages. The descriptive framework must be innate (= universal grammar), otherwise there is no explanation).

> "The next task [after constructing an explicit mental grammar, M.H.] is to explain why the facts are the way they are, facts of the sort we have reviewed, for example [e.g. binding phenomena, M.H.]. This task of explanation leads to inquiry into the language faculty. A theory of the language faculty is sometimes called universal grammar... Universal grammar provides a genuine explanation of observed phenomena. From its principles we can deduce that the phenomena must be of a certain character, given the initial data that the language faculty used to achieve its current state." (Chomsky 1988: 61-62)

> "The problem that the principles and parameters framework seeks to solve is: How can a grammatical system be flexible enough to account for language variation while at the same time be, to a large extent, restricted in order to account for the relative ease of language acquisition and the impossibility of certain language types?" (Travis 1989: 263)

> "[discussion of formal mechanisms of Distributed Morphology...] On this account, the unattested patterns do not arise as they cannot be generated in a manner consistent with Universal Grammar." (Smith et al. 2016)

Another way of saying this is that the universals should "fall out" from the framework (=the model of UG).

"Unattested languages are cognitively impossible languages" (cf. Newmeyer 2005: §3.3)

The best natural-science analogy: Mark Baker's (2001) *The atoms of language*:

Categories and parameters are analogous to **chemical elements**, and different languages constitute different "mixtures" of parameters, just as different chemical compounds arise from different combinations of elements.
(Chemical elements constitute a restriction that can be attributed to an external cause: physics.)

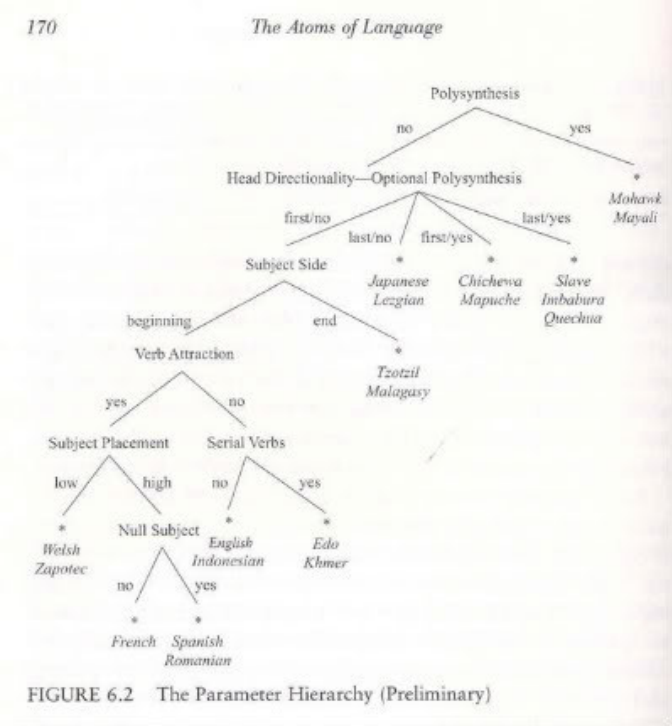


FIGURE 6.2 The Parameter Hierarchy (Preliminary)

From Baker's book, one gets the impression that he would like to be a Mendeleev of linguistics.

But are we on the verge of discovering the "elements of grammar"?

4. Restricted description and language-particular description

In generative linguistics, even when the focus is on **one** phenomenon in **one** language, it is not sufficient to describe the phenomenon completely and correctly.

In addition, one needs to describe the phenomenon **using the restrictive descriptive framework**. This can be seen in many papers, which are divided into a "data" part and an "analysis" part.

From my perspective, this approach is puzzling:

Why would one describe a phenomenon twice?
... once in ordinary language that is readily comprehensible, and once in a complex formal technology?

This approach makes good sense if

(1) the descriptive framework is **innate,** because then the language must indeed be represented mentally in these terms

AND (2) if we are **on the verge** of discovering our "periodic table of categories/parameters"

But

• there is no good independent evidence for innate frameworks (and Hauser, Chomsky, Fitch 2002 say that only recursion is innate)

• and we are apparently not on the verge of knowing all the categories

My approach to analogizing linguistics, biology and chemistry:

	linguistics (unit: language)	**biology** (unit: species)	**chemistry** (unit: compound)
phenomenological description	descriptive grammar	zoological/ botanical description	color, smell etc. of a compound
underlying system	"mental/cognitive grammar"	description of species genome	description of molecular structure
basic building blocks	"cognitive code" (= elements of UG)	genetic code	atomic structure
explanation of phenomenology and system	diachronic adaptation	evolutionary adaptation	?
explanation of basic building blocks	biology (Chomsky: physics)	biochemistry	nuclear physics

Table 1. Basic goals of linguistics, biology and chemistry
(Haspelmath 2004, cf. Baker 2001 for the linguistics/chemistry analogy)

Why do so many linguists adopt the generative approach?

One reason:
It allows linguists who are working on a single language to claim that they are contributing to the discovery of the foundations of Human Language.

But another reason:
It is **intuitively natural** to think that the categories that linguists use are given in advance.

5. The natural-kind presumption

One way of dealing with the generality dilemma is to pretend that differences between languages do not matter much – eacch language presents the child with the same acquisition problem. ("How is knowledge of language acquired?")

Generative grammarians often adopt an explicitly **biological point of view**, a trend that has become stronger in recent decades ("biolinguistics").

It may thus seem natural to regard the **categories of languages as biological objects**, like biological species, or like the biomolecules that are the foundation of living beings.

More abstractly, linguistic categories on this view are **natural kinds**, like chemical elements, which occur on different planets.

And just as it makes sense to ask which molecules occur on Mars (e.g. whether there is water, and if so how much), it may well make sense to ask whether a language on a different continent such as Quechua has "gerunds", or "determiners", or "VP", or "coronal consonants", or any other category that was originally found to exist in English.

If linguistic categories are natural kinds, then it makes perfect sense to hypothesize (or even **assume**, Chomsky 2001: 2) that a category identified for one language (say, "anaphor" in English) will exist in another language.

If linguistic categories are natural kinds, then the following questions make sense (from Haspelmath 2007)

(i) Is English *–like* a stem or a suffix? (c.f. Tuggy 1992, Dalton-Puffer & Plag 2001)
(ii) Is the Romanian definite article a clitic or a suffix? (Ortmann & Popescu 2000)
(iii) Is English *silver ring* a phrase or a compound? (e.g. Bauer 1998, Giegerich 2004)
(iv) Are Mandarin Chinese property words adjectives or verbs (McCawley 1992, Dixon 2004)?
(v) Is the Tagalog *ang*-phrase a subject or a topic? (Schachter 1976)
(vi) Is German *er* a pronoun or a determiner? (Vater 2000)
(vii) Is English *that* in relative clauses a pronoun or a complementizer? (van der Auwera 1985)
(viii) Is the English adverbial *–ly* an inflectional or a derivational suffix?
(ix) Are the two types of intransitive verbs in Jalonke (Mande) unacccusatives and unergatives? (or are they something else?) (Lüpke 2006)
(x) Are French subject clitics (*je, tu, il*...) pronouns or agreement markers? (De Cat 2005)
(xi) Is the German dative a structural case or an inherent case? (Wegener 1991, Woolford 2006)

But again, there is an alternative possibility that needs to be considered:

In biocultural systems, we often see **culture-specific categories** arising.

> **kinship categories** (younger sister, uncle, different-sex sibling, etc.),
> **governance categories** (chief, king, mayor, president, consul, chairman)
> **religious categories** (angel, gospel, surah, stupa, shaman, etc.),
> **food categories** (soup, dumpling, noodle, sandwich, dessert, confectionery)
> etc.

Similarities in social categories across cultures are hardly ever explained by our uniform genetic endowment, and nobody treats such categories are natural kinds.

Likewise, comparative linguistics can be carried out with **comparative concepts** (Haspelmath 2010), a set of concepts that are distinct from the categories used in analyzing particular languages.

(It could of course be that the natural-kinds approach is better, linguistic categories, but **one needs to consider both possibilities.)**

6. Different criteria in different languages

6.1. Nouns vs. verbs (Haspelmath 2012)

If one adopts a natural kinds (or "categorial universalist") position, one must be willing to apply different criteria in different languages. For example, to identify "nouns" in Ancient Greek, English and Mandarin Chinese, quite different criteria are commonly applied. This can be seen in (6a-c).

(6) a. Greek Noun (Dionysius Thrax, Ars minor, 2nd c. BCE)
ὄνομά *ἐστι μέρος λόγου πτωτικόν, σῶμα ἢ πρᾶγμα σημαῖνον*
'a Noun is a case-inflected part of speech that denotes a thing or an action'

b. English Noun (Quirk et al. 1985: 72)
a Noun is a word that can follow determiners like *the, this* and *that*

c. Mandarin Chinese Noun
a Noun is a word that can follow a classifier

As long as SOME criteria can be found for noun status, universalists are content.

But which properties can be taken as evidence for category assignment? There are no constraints on this – each linguist can make their own choices (this is what Croft 2009 calls "**methodological opportunism**").

The method is thus subjective and not rigorous (cf. also Post 2008: 377-378). Rigorous comparison requires that **languages be compared in terms of concepts that apply in the same way to all languages**.

Linguists often ask questions such as those in (7) and (8):

(7) language-particular questions
Does language X have a noun-verb distinction?
Does language X have a verb-adjective distinction?
Does language X have a noun-adjective distinction?

(8) cross-linguistic questions
Do all languages have a noun-verb distinction?
Do all languages have a verb-adjective distinction?
Do all languages have a noun-adjective distinction?

• A lack of a noun-verb distinction seems even more radical than a lack of an adjective class, so this issue seems even more important (cf. Evans & Osada 2005). Austronesian languages, especially Tagalog, have been prominent, most recently in Kaufman (2009).

• In Tagalog, action-roots and thing-roots behave alike in reference and predication constructions, as there is **no copula** in (9b), and the referential use of the action-root does not require more than the nominative marker *ang*.

(9) a. *Nag-íngay* *ang* *áso.* (action-predicate & thing-referent)
AGENTVOICE-noise [NOM dog]
'The dog made noise.'

b. *Áso* *ang* *nag-íngay.* (thing-predicate & action-referent)
dog [NOM AGENTVOICE-noise]
'The one who made noise is a dog.'

property-roots behave in the same way:

(10) a. *Ma-bilis* *ang* *áso.* (property-predicate & thing-referent)
STATIVE-quick [NOM dog]
'The dog is quick.'

b. *Áso* *ang* *ma-bilis.* (thing-predicate & property-referent)
dog [NOM STATIVE-quick]
'The quick one is a dog.'

in attribution, all three root-classes also behave alike, requiring nothing but the **linker** morpheme *-ng/na* between the head and the attribute.

(11) a. *ang* *áso* *-ng* *ma-bilis* (thing-referent & property-attribute)
NOM dog LK STAT-quick
'the quick dog'

b. *ang* *áso* *-ng* *nag-íngay* (thing-referent & action-attribute)
NOM dog LK ACTORVOICE-noise
'the dog who made noise'

c. *ang* *ma-bilis* *na* *nag-ínay* (property-referent & action-attribute)
NOM quick LK ACTORVOICE-noise
'the quick one who/which made noise'

d. *ang* *nag-íngay* *na* *áso* (action-referent & thing-attribute)
[NOM AGENTVOICE-noise] LK dog
'the noise-maker who is a dog' (= 19b)

• Kaufman concludes that Tagalog has a single macroclass of Nouns.

• But if we adopt the principle of **complete identity of behaviour**, then we cannot say that Tagalog has just a single word-class.

• Most strikingly, action-roots take **aspect-modality inflection and voice affixes** (e.g. the prefix *nag-*), while thing-roots do not have these possibilities.

• These morphological differences are very salient, so linguists who have claimed that Tagalog is unlike English with respect to its word-classes have usually said that Tagalog makes no distinction between "syntactic word-classes", only between "morphological word-classes".

• But **aspectual and voice marking is non-uniform across the class of "verbs" in many languages**, and in most languages, verbs have **inflectional subclasses**. So in the absence of clear criteria that determine what constitutes a major class and what constitutes a subclass, one could maintain that the syntactic uniformity seen in (9)-(11) justifies the postulation of a single major world-class Noun, with subclasses based on (less important) morphological criteria.

• However, there is evidence that **syntactically, too, not all roots behave alike**: In some contexts a copula seems to be required with thing-roots (Richards 2009: 141), e.g. when the predicate is a complement of a verb of desire:

(12) a. *Ayo-ko* *na-ng* *l-um-angoy.*
notwant-1SG now-LK swim-ACTORVOICE
'I don't want to swim anymore.'

b. *Ayo-ko* *na-ng* *maging* *doktor*
notwant-1SG now-LK be doctor
'I don't want to be a doctor anymore.' *(*Ayoko nang doktor.)*

• again, this could be described by setting up different subclasses of the broad macro-Noun category, if one decided to give more weight to the criterion of behaviour in ordinary predicative, attributive and referential contexts. **Again, there is no objective, rigorous way of deciding.**

• If one allows different diagnostics (or criteria) for different languages, **one can fish for those diagnostics that one wishes ("methodological opportunism")** – as a result, all languages look like English.

6.2. VP and DP

• For example, in the past, it was not uncommon to argue that **some languages lack a VP**, e.g. German (Haider 1991) or Hungarian (É. Kiss 2002) – these arguments were based on striking differences between German/Hungarian and English.

But not surprisingly, there are also similarities between German/Hungarian and English, and if one focuses on these, **German and Hungarian end up looking like English**.

In the past, it was not uncommon to argue that article-languages like English have a definite article and therefore a DP, while article-less languages like Russian do not have a DP.

But more recently, the similarities between English and Russian have been emphasized, and Russian has been said to have a DP, e.g. by Franks & Pereltsvaig (2004), on the basis of evidence such as the following:

(13) a. *V ètom restorane obedal-i* [$_{DP}$ D [$_{QP}$ *desjat' čelovek*]]. (plural agreement)
b. *V ètom restorane obedal-o* [$_{QP}$ *desjat' čelovek*]. (default agreement)
'In this restaurant had-lunch ten people.'

The possibility of positing zero elements always exists, so one can always take SOME similarity between language X and English as evidence that language X really has the same category as English.

(14) **Uniformity Principle**
"In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances." (Chomsky 2001: 2)

If one allows abstract analyses involving zeroes, movements and so on, then it will never be possible to find "compelling evidence to the contrary" – languages will always be found to be uniform.

7. Optimality Theory: Conflating description and functional explanation?

Optimality Theory appears to appeal to functional factors for descriptive purposes, e.g. the constraints in (15) to describe languages with or without final voicing (Kager 1999: 41)

(15) a. *NoVoicedCoda (codas should be voiceless)
b. Identity(voice) (there should be no alternations)
c. Voiced Obstruent Ban (VOB) (obstruents should be voiceless)

Dutch

input: bɛd	*NoVoicedCoda	Identity(voice)	VOB
→ bɛt		*	*
bɛd	*		**
pɛd	*	*	*
pɛt		**	

English

input: bɛd	Identity(voice)	*NoVoicedCoda	VOB
bɛt	*		*
→ bɛd		*	**
pɛd	*	*	*
pɛt	**		

Bella Coola (or any language lacking voiced obstruents)

input: bɛd	VOB	Identity(voice)	*NoVoicedCoda
bɛt	*	*	
bɛd	**		*
pɛd	*	*	*
→ pɛt		**	

This set of constraints thus predicts that no language can have onset devoicing (the candidate [pɛd] can never emerge as optimal).

This approach has also been applied to syntactic typology (e.g. by Aissen 2003; it was explicitly called "functional optimality theory", e.g. Malchukov 2008).

BUT:

(1) This only works if the functional constraints are assumed to be **innate**, which is even more implausible than innateness of principles-and-parameters categories

(2) as argued by Haspelmath (1999b) and others (e.g. Newmeyer 2005), it does not work for description in general – languages have all kinds of **arbitrary patterns** that are not amenable to this kind of functional analysis.

Phonetically-based phonological change is functionally motivated and often results in "natural" sound patterns (cf. Blevins 2004), but such patterns can be destroyed by subsequent changes. (Cf. Russian *kniga* vs. *knig* [-k], contrasting with the newer pattern *Lida* vs. *Lid* [-d], showing that devoicing is now restricted to older patterns and has become morphologized to some extent.)

The universal that "no language can have onset devoicing" (Kager 1999: 42) is thus probably **due to a mutational constraint**: coda devoicing is a possible (and in fact frequent) change type, whereas onset devoicing does not occur as a change type.

Likewise, in syntax the approach does not work. According to Aissen (2003), referential prominence universals can be explained by innate constraint hierarchies:

(16) a. Vietnamese
*$STRUC_{case}$ >> *Obj/**Hum** & *$Ø_{CASE}$ >> * Obj/**Anim** & *$Ø_{CASE}$ >> *Obj/**Inan** & *$Ø_{CASE}$

b. Spanish
*Obj/**Hum** & *$Ø_{CASE}$ >> *$STRUC_{case}$ >> * Obj/**Anim** & *$Ø_{CASE}$ >> *Obj/**Inan** & *$Ø_{CASE}$

c. Russian
*Obj/**Hum** & *$Ø_{CASE}$ >> * Obj/**Anim** & *$Ø_{CASE}$ >> *$STRUC_{case}$ >> *Obj/**Inan** & *$Ø_{CASE}$

d. Hungarian
*Obj/**Hum** & *$Ø_{CASE}$ >> * Obj/**Anim** & *$Ø_{CASE}$ >> *Obj/**Inan** & *$Ø_{CASE}$ >> *$STRUC_{case}$

A language such as "Anti-Spanish", which only case-marks inanimate objects, cannot be described in this system, because the constraints in the subhierarchy cannot be reranked.

> "OT provides a way ... to reconcile the underlying impulse of generative grammar to model syntax in a precise and rigorous fashion with a conception of DOM which is based on prominence scales. The purpose ... is to develop an approach ... that is formal and at the same time expresses the functional-typological understanding of DOM" (Aissen 2003: 439)

> "It should be acknowledged that constraint conjunction is a powerful operation which, if unrestricted, will generate constraints that are clearly undesirable. For example, if the subhierarchies of [4] were conjoined with *$STRUC_{case}$ rather than with *$Ø_{case}$, all the predictions made by the present analysis would be neutralized. One possibility is to appeal to functional reasoning: although constraints formed by conjunction of the subhierarchies with *$STRUC_{case}$ might exist, grammars in which they were active would be highly dysfunctional since marking would be enforced most strenuously exactly where it is least needed. (Aissen 2003: 447-8, n. 12)"

In other words: Aissen's system is not restrictive enough, but **overgenerates vastly**.
To explain why certain languages predicted by her OT account do not exist, she needs to "appeal to functional reasoning" – thus, to appeal to a functional-adaptive constraint. Her machinery is thus not needed for explanation.

8. Conclusion

While it is quite likely that constraints of the "universal grammar" type very probably exist, in one form or another, **we do not currently know what they are**.

If no mutational or functional-adaptive explanation is available, we need to appeal to innate cognitive constraints, **as a last resort** ("representational constraints", "the cognitive code").

But for language-particular description/analysis, we should work with language-particular categories,

and for comparison, we should work with a separate set of comparative concepts.

References

Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Baker, Mark C. 2001. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

Blevins, Juliette. 2004. *Evolutionary phonology*. Cambridge: Cambridge University Press.

Chomsky, Noam A. 1957. *Syntactic structures*. 's-Gravenhage: Mouton.

Chomsky, Noam A. 1988. *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.

Chung, Sandra. 2012. Are lexical categories universal? The view from Chamorro. *Theoretical Linguistics* 38: 1-56.

Croft, William. 2001. *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William. 2009. Methods for finding language universals in syntax. In Sergio Scalise, Elisabetta Magni, & Antonietta Bisetto (eds.), *Universals of language today*, 145–164. Berlin: Springer.

Dryer, Matthew S. 2006. Descriptive theories, explanatory theories, and basic linguistic theory. In: *Catching Language: Issues in Grammar Writing*, edited by Felix Ameka, Alan Dench, and Nicholas Evans. Berlin: Mouton de Gruyter.

É. Kiss, Katalin. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.

Evans, Nicholas & Toshiki Osada. 2005. Mundari: The myth of a language without word classes. *Linguistic Typology* 9(3). 351–390.

Giorgi, Alessandra, and Giuseppe Longobardi. 1991. The syntax of noun phrases: Configuration, parameters and empty categories. Cambridge: Cambridge University Press.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to

Grimshaw, Jane & Samek-Lodovici, Vieri. 1998). Optimal subjects and subject universals. In: Barbosa, Pilar et al. (eds.) *Is the best good enough?* Cambridge/MA: MITWPL & MIT Press, 193-219.

Haider, Hubert. 1993. *Deutsche Syntax – generativ: Vorstudien zur Theorie einer projektiven Grammatik*. (Tübinger Beiträge Zur Linguistik 325). Tübingen: Gunter Narr Verlag.

Haspelmath, Martin. 1999a. Explaining article-possessor complementarity: economic motivation in noun phrase syntax. *Language* 75.2: 227-43.

Haspelmath, Martin. 1999b. Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18.2: 180-205 (with open peer commentary)

Haspelmath, Martin. 2004. Does linguistic explanation presuppose linguistic description? *Studies in Language* 28.3: 554-579.

Haspelmath, Martin. 2007. Pre-established categories don't exist: consequences for language description and typology. *Linguistic Typology* 11. 119 - 132.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663–687.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(2). 31–80.

Haspelmath, Martin. 2012. Escaping ethnocentrism in the study of word-class universals. *Theoretical Linguistics* 38: 91-102.

Haspelmath, Martin. 2014. Comparative syntax. In Andrew Carnie, Yosuke Sato & Dan Siddiqi (eds.), *The Routledge handbook of syntax*, 490–508. London: Routledge.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 298. 1569–1579. doi:10.1126/science.298.5598.1569.

Jackendoff, Ray. 1977. *X-bar theory.* Cambridge: MIT Press.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35(1). 1–49.

Lazard, Gilbert. 1992. Y at-il des catégories interlangagières? In Susanne Anschütz (ed.), *Texte, Sätze, Wörter und Moneme: Festschrift für Klaus Heger zum 65. Geburtstag*, 427–434. Heidelberg: Heidelberger Orientverlag.

Lyons, Christopher. 1986. The syntax of English genitive constructions. *Journal of Linguistics* 22.123-43.

Malchukov, Andrej. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118. 203–221.

Newmeyer, Frederick J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.

Post, Mark. 2008. Adjectives in Thai: Implications for a functionalist typology of word classes. *Linguistic Typology* 12(3). 339–381.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Richards, Norvin. 2009. Nouns, verbs, and hidden structure in Tagalog. *Theoretical Linguistics* 35(1). 139–152.

Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang & Jonathan David Bobaljik. 2016. Case and number suppletion in pronouns. *Lingbuzz*.

Travis, Lisa. 1989. Parameters of phrase structure. In: *Alternative conceptions of phrase structure*, ed. by Mark R. Baltin and Anthony S. Kroch. Chicago: University of Chicago Press, pp. 263-279.