Chapter 3

New approaches to Greenbergian word order dependencies

Jennifer Culbertson

School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK

1 Cognitive explanations for language typology

1.1 Introduction

Implicational typological universals (e.g., Greenberg 1963) represent a class of dependencies that linguists have been seeking to document, refine and explain for decades. From a functionalist typological viewpoint, the goal of such explorations is to understand how these distributions of patterns arose through a combination of geography, history and cultural evolution. From a generative linguistic viewpoint, the goal is to relate dependencies to features of the human language faculty and thus inform and constrain grammatical theories. While these two perspectives could in principle be mutually informative (Hawkins 2004; Baker & McCloskey 2007), foundational differences have often prevented cross-talk between researchers (Bickel 2007; Haspelmath 2000; Newmeyer 1998). The goal of this chapter is to highlight a strand of behavioral research which can advance the goals of both functionalists and generativists alike. Evidence from controlled laboratory experiments brings to light cognitive biases which might play a causal role in constraining language change, and opens the door to investigating the extent to which they reflect properties of the language faculty narrowly construed, or rather domain-general forces potentially shared across cognitive systems (and even species). This source of evidence therefore adds to our understanding of why language is the way it is–by refining the set of factors likely to have shaped a particular distribution of linguistic patterns–and how we should characterize linguistic competence. I illustrate this with two case studies investigating the

Jennifer Culbertson. 2017. New approaches to Greenbergian word order dependencies. In N. J. Enfield (ed.), *Dependencies in language*, 23–38. Berlin: Language Science Press. DOI:10.5281/zenodo.573778

connection between two Greenbergian word order universals and asymmetrical learning outcomes in the lab.

1.2 Mental universals and typology

Under a traditional nativist view, typological universals are treated as a source of direct evidence from which to make inferences about the content of genetically encoded *mental* universals. The latter are formalized as grammatical constraints ensuring languages change in particular ways and not others, and relatedly, limiting the space of hypotheses entertained by language learners (e.g., Lightfoot 1989; Baker 2001). For example, Greenberg's Universals 3 and 4 state implicational relationships between word order across phrases: if a language is VSO it will have prepositions, by contrast SOV languages tend to have postpositions (Greenberg 1963). If these relations constrain how languages change, then one might expect that if the basic word order changes from VSO to SOV, the order of adpositions will also change (or at least will be more likely to do so).

Perhaps the most problematic aspect of this view is the idea that typology is the observable result of cognitive constraints. Most obviously, this is because distributions of patterns across the world's languages are undoubtedly affected by cognition-external factors–indeed in some cases they may be completely accounted for by appealing to the influence of historical coincidence, areal factors and/or culturally-driven influence. Teasing apart such factors is at best extremely challenging (Cysouw 2005; Ladd, Roberts & Dediu 2015; Piantadosi & Gibson 2014). Further, even if some cognitive constraint *is* part of the explanation for a particular typological universal, a number of questions necessarily remain: Is the underlying mechanism functionally motivated? Is the constraint innately encoded or learned? Is it domain-specific (either evolved specifically for language, or representationally specific to language) or does it operate across cognitive domains? This is particularly important since most typological "universals" are statistical rather than absolute. Universal 4, for example, describes a strong tendency for SOV languages to have postpositions, but this only holds in 472/486 or 97% of cases in a large sample (Dryer 2013c). If this universal is the reflection of an underlying cognitive constraint, it would not immediately be compatible with the notion of inviolable principle employed to formalize constraints in many generative frameworks.

1.3 Probing cognitive explanations experimentally

A growing body of research has begun to investigate the existence and content of mental universals through behavioral experiments, specifically using artificial language learning (ALL) paradigms. Although ALL has been used most extensively to test phonological pattern learning, studies featuring ALL experiments can now be found across all linguistic domains, including syntax (see Moreton & Pater 2012; Culbertson 2012 for literature reviews). This approach treats typology as a source of hypotheses about possible constraints or BIASES in language learning or use rather than as direct evidence for them. While converging evidence supporting a particular hypothesized bias could potentially come from studies of natural language acquisition, ALL paradigms have important advantages. Most obviously, the characteristics of the input language can be precisely controlled and contributions from multiple factors can be independently tested. In addition, it is relatively straightforward to test learning of rare or unattested patterns which might otherwise be very difficult if not impossible to investigate.

These paradigms also make it possible to test the nature and scope of hypothesized biases, for example by instantiating parallel patterns or structures in non-linguistic stimuli. Both domain-general and linguistically specific biases uncovered using these methods could in principle be formalized as inviolable constraints (hard limits on the space of possible languages) of the sort typically posited by mainstream generative linguistic theories. However, just as typological data are often in the form of statistical trends, behavioral data typically reveal probabilistic biases. This suggests they may be better captured by models which allow for probabilistic constraints (e.g., using Maximum Entropy or Probabilistic Harmonic Grammar formalisms; Goldwater & Johnson 2003; Wilson 2006). For example, Culbertson, Smolensky & Wilson (2013) create a probabilistic model of biases in noun phrase word order which also incorporates a bias for regularization – reducing of unconditioned variation – that is outside the grammar itself. Models like this therefore allow biases of different types to combine with one another to predict learning outcomes, and in principle could further take into account non-cognitive factors to more precisely model typological distributions. While many ALL studies focus on learning in individual participants, recent work has involved creating particular social conditions, adding communicative pressures, and transmitting learning outcomes across sets of participants to model language change (e.g., Fay et al. 2010; Kirby et al. 2015; Kirby, Cornish & Smith 2008). These factors can be straightforwardly incorporated into probabilistic models in order to formalize hypotheses and make further predictions about what shapes typology.

To give the reader a clear picture of how ALL works and the kinds of learning biases one can investigate using it, in what follows I discuss in more detail two case studies. These case studies highlight the use of two distinct ALL paradigms in testing the psychological reality of three biases in the learning of nominal word order predicted from Greenbergian Universals 18 and 20.

2 Greenberg's Universal 18

2.1 Introduction

Greenberg's Universal 18 (U18) is stated in (1) below.

(1) If Adj-N then Num-N.

This implicational universal rules out one of the four logically possible patterns in Table 1, namely the one which combines Adj-N with N-Num. The geographic distribution of these four patterns is shown using data from a much larger sample in Figure 1. This map in fact highlights the difficulty with interpreting raw typological frequency data: they may turn out to be misleading once genetic and areal relationship are taken into account. In this case, the larger sample shows that Adj-N & N-Num languages are in fact attested, however they may be over-represented in the raw numbers since the languages are clearly clustered in three small areas. Similarly, many of the languages classified as N-Adj & N-Num (numerically most frequent) are found clustered in Africa. This strongly suggests the need for additional empirical data in understanding this typological tendency.

Beyond Universal 18, Table 1 reveals a second trend in the raw frequency data: ordering patterns which place both Adj and Num on the same side of the noun are by far the most common in the sample. This type of pattern is sometimes called HARMONIC, while the other two are NON-HARMONIC (for discussion of this terminology see Croft 2003: 59-62). A trend toward harmony across phrases is a well

Table 1: Four possible combinations of {N, Adj} and {N, Num} with corresponding frequencies in the languages of the world based on Dryer (2013a,b).

	Adj-N	N-Adj
Num-N	251 (27%)	168 (17%)
N-Num	37 (4%)	509 (52%)

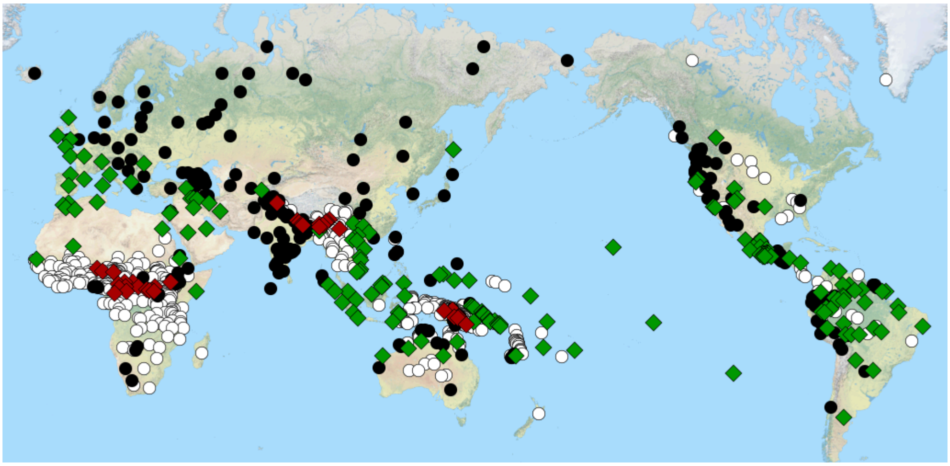


Figure 1: Geographical distribution of ordering patterns based on Dryer (2013a,b). Circles are harmonic (black: Adj-N, Num-N, white: N-Adj, N-Num), diamonds are non-harmonic (green: N-Adj, Num-N, red: Adj-N, N-Num).

known typological universal (many other Greenbergian universals are relevant for this, e.g., 2-6), which has been the subject of much research (e.g., Hawkins 1983; Travis 1984; Chomsky 1988; Dryer 1992; Baker 2001). To summarize then, we can hypothesize two biases based on these typological data: (i) a bias in favor of harmonic patterns, and (ii) a bias against the particular non-harmonic pattern combining pre-nominal adjectives with post-nominal numerals.

2.2 Testing Universal 18

The four patterns in Table 1 are intuitively simple and are all clearly learnable. How, then, might one uncover potentially subtle *differences* in learnability? In Culbertson, Smolensky & Legendre (2012) we did this by introducing variation into the input, essentially allowing us to see which patterns are more easily learnable under noisy conditions. Native-English-speaking adult learners were trained on phrases comprised of a noun and single modifier (adjective or numeral word), the order of which varied between a dominant order–heard in 70% of utterances–and the opposite–heard in 30% of phrases. The dominant order varied randomly across participants in the experiment and instantiated one of the four possible patterns in Table 1. The conditions are represented in Figure 2, with numbers 1–4 indicating the four conditions. For example, in condition 1, learners heard pre-nominal Adj-N and Num-N 70% of the time, and post-nominal N-Adj, N-Num the remaining 30% of the time. Condition 2 has the opposite proportions, and therefore participants heard post-nominal N-Adj and N-Num as the dominant order. Conditions 3 and 4 are non-harmonic; condition 3 participants heard N-Adj and Num-N as the dominant pattern, while condition 4 participants heard the U18-violating Adj-N, N-Num as the dominant pattern.

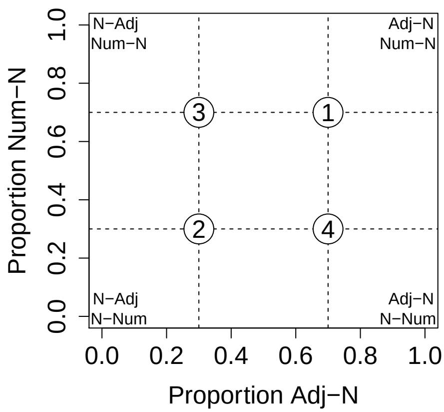
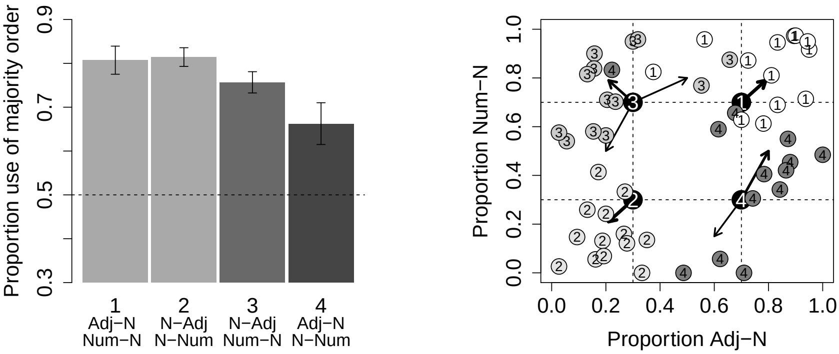


Figure 2: Illustration of experiment conditions. The corners of this space represent deterministic patterns, while inset numbers represent the four variable conditions used in the experiment. Note that condition 1 and 2 are harmonic, while 3, 4 are non-harmonic. Condition 4 is a variable version of the U18-violating pattern.

Independent evidence from natural language and ALL studies (e.g., Singleton & Newport 2004; Hudson Kam & Newport 2009) suggests that learners tend to *regularize* unpredictable (unconditioned) variation of the sort we used in this experiment. We hypothesized that learners would be most likely to regularize variable patterns which conformed to their biases, and would not regularize those they found more difficult to learn. This predicts that participants learning a variable version of one of the two harmonic patterns (1: Adj-N, Num-N, or 2: N-Adj, N-Num) should regularize the majority order, using it more than 70% of the time. By contrast, participants learning the non-harmonic pattern targeted by Universal 18 (4: Adj-N, N-Num) should *not* regularize that pattern.

These predictions were borne out by the results, as shown in Figure 3(a): participants in conditions 1 and 2 regularized the variation in their input–using the majority order substantially more than 70% of the time–while participants in condition 4 did not regularize. Participants in condition 3, who were exposed to



(a) Average use of the majority order in each condition.

(b) Behavioral outcomes for individual participants in each condition relative to their input.

Figure 3: Experiment results.

the non-harmonic pattern not violating U18, show some regularization but not as much as those in the harmonic conditions. Another way to visualize the behavioral outcomes in the experiment is in terms of the space shown in Figure 3(b), which plots individual participants' use of each order relative to their input. This illustrates how learners *shift* or change the language they are exposed to according to their biases. In conditions 1 and 2, learners' tendency to regularize aligns with their bias for harmonic patterns, therefore their output is shifted toward the deterministic corners relative to the input. In non-harmonic condition 3, some learners shift toward a more regular version of their input, but others actively move the language toward one of the two preferred harmonic patterns. In non-harmonic condition 4, this shifting toward a harmonic pattern is much more dramatic and *no* learners have regularized their input pattern. Interestingly, in this experiment native English-speaking participants showed only a small preference for their native-language order: the average regularization was the same across conditions 1 and 2, however more participants in the non-harmonic conditions shifted toward the pre-nominal harmonic pattern (for additional discussion about prior language experience and an alternative explanation of this difference see Culbertson, Smolensky & Legendre 2012; Culbertson & Newport 2015).

To summarize, in Culbertson, Smolensky & Legendre (2012), we started with Universal 18 and generated a set of hypothesized biases. We tested the psychological reality of these biases using an artificial language learning paradigm which exploits learners' tendency to regularize unpredictable variation. We confirmed that regularization of variation is indeed modulated by the particular type of pattern being learned; when the majority pattern in the input conforms to learners' biases, they regularize. When the majority pattern is dispreferred, learners actively change the language to bring it in line with their preferences. With this evidence in hand, researchers interested in constructing explanations for the typological distribution of nominal word order can more confidently add these factors into their models. Moreover, additional research using experimental methods can begin to explore *why* Universal 18 holds in the population tested, and *why* learners might prefer harmonic patterns. This could involve testing structurally similar patterns in non-linguistic domains or investigating the role of language experience in the development of these biases.

3 Greenberg's Universal 20

3.1 Introduction

Greenberg's Universal 20 (U20), as reformulated by Cinque (2005), is stated in (2) below.

(2) In pre-nominal position: Dem-Num-Adj
In post-nominal position: Dem-Num-Adj or Adj-Num-Dem

The explanation for this implicational universal has received significant attention in the literature, particularly after additional typological work by Cinque (2005) and Dryer (2009). Figure 4 plots the frequency of each of the 24 possible combinations of N, Dem, Num, Adj in descending order. The two post-nominal orders picked out by Greenberg are highlighted in black. To account for this distribution, or key aspects it, a number of distinct proposals have been made (e.g., Cinque 2005; Abels & Neeleman 2012; Dryer 2009; Cysouw 2010; Steddy & Samek-Lodovici 2011). All of these proposals include a notion of the semantic or structural distinctions among the modifiers that can be described in terms of scope, as illustrated in Figure 5. In this case, adjectives can be said to take innermost scope since they modify dimension inherent to noun meaning, while numerals serve to count these larger units. Demonstratives take highest scope because they serve to connect the internal material to the surrounding discourse.

These scope relations do not determine linear order, instead a given language can map these structural relations to linear order in various ways. Importantly, of the 24 possible patterns, eight preserve the underlying scope relations in the

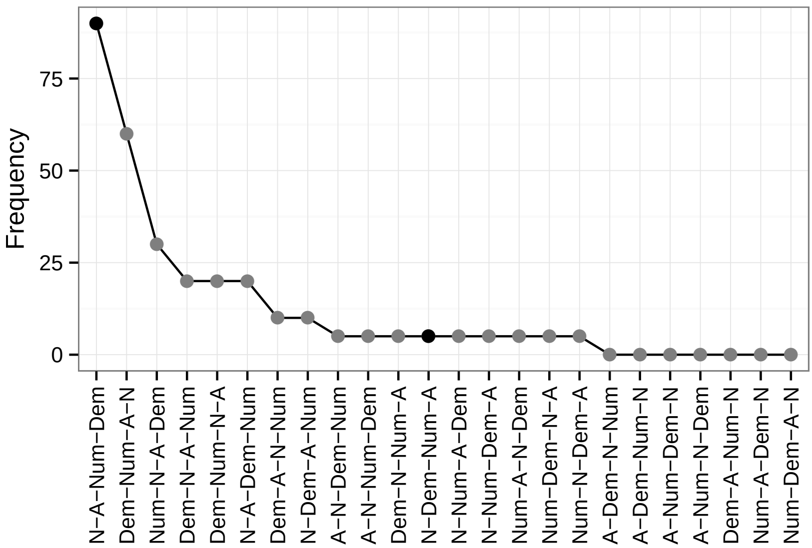
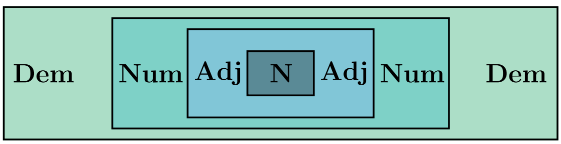
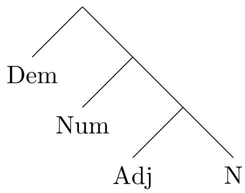


Figure 4: Frequency of 24 possible combinations of N, Dem, Num, Adj as reported in Dryer (2009). Post-nominal orders in Greenberg's Universal 20 are the black points.



(a) Illustration of nested scope relationship among nominal modifiers.



(b) Hierarchical representation of scope. Dem takes widest scope, Adj takes innermost scope.

Figure 5: Scope relationship among nominal modifiers

surface linear order. If in addition to preservation of the scope relations, harmony is also a factor which constrains language change, then we can explain why Dem-Num-Adj-N and the mirror order N-Adj-Num-Dem are the most frequent. Indeed, a principle encoding a harmony preference is present in most analyses of Universal 20, and harmonic patterns were shown to be preferred by learners in Culbertson, Smolensky & Legendre (2012). By the same reasoning, the alternative post-nominal pattern cited by Greenberg, N-Dem-Num-Adj, is expected to be less frequent since it is harmonic but does *not* maintain the isomorphism between scope and the linear order.

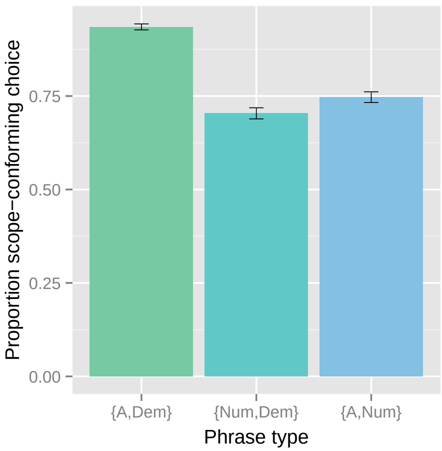
3.2 Testing U20

The two post-nominal orders in Greenberg's Universal, N-Adj-Num-Dem and N-Dem-Num-Adj differ from one another in two important ways. First, as described above, N-Adj-Num-Dem maintains the underlying scope relations in the linear order, while N-Dem-Num-Adj does not (in fact it perturbs them maximally). Second, N-Dem-Num-Adj has the same linear order of the modifiers as English, while N-Adj-Num-Dem does not (in fact it is the opposite). In Culbertson & Adger (2014), we capitalized on this pattern of differences to test whether English speakers learning a new language will transfer their knowledge of linear order, or their knowledge of scope-to-surface isomorphism. We did this by using the poverty-of-the-stimulus paradigm, in which learners are presented with examples from a new language in a way that withholds critical evidence about its structure. At test, learners must generalize to held-out data that will disambiguate the alternative hypotheses. In this experiment, participants heard phrases with a noun and a single post-nominal modifier and then at test were asked about the relative order of modifiers. For example, they might be trained on N-Dem and N-Adj sequences, and then be asked at test whether phrases with N-Adj-Dem or N-Dem-Adj order are most likely in the language.

We trained participants in a number of different input conditions. Here I highlight one set, summarized in Figure 6(a). The results, shown in Figure 6(b), reveal a striking preference at test for orders which are isomorphic to the scope over those with are more surface-similar to English. Interestingly, this preference was most dramatic when the input included Dem and Adj. This suggests that preserving the scope relations among the two most structurally distant modifiers (Dem and Adj) may be more important than the closer ones (either Dem, Num or Num, Adj). This prediction turns out to be typologically accurate; languages which perturb the scope of Adj, Num or Num, Dem are about twice as common as Adj, Dem.

Training order	Testing combo
N-Adj, N-Dem	{Adj, Dem}
N-Num, N-Dem	{Num, Dem}
N-Adj, N-Num	{Adj, Num}

(a) Experimental conditions



(b) Results by condition from Experiment 1

Figure 6: Conditions and results as reported in Culbertson & Adger (2014).

To summarize, this result provides the first experimental evidence for a bias favoring linear orders that maintain an isomorphism with the underlying semantic scope. The evidence is preliminary to the extent that participants' bias may come from knowledge of this abstract property of English. To determine whether the bias can be found in learners without direct experience with it, future work will need to target a population whose language *violates* this preference–for example Kikuyu is one of the few languages with N-Dem-Num-Adj. Nevertheless, combined with a preference for harmony, as shown in Culbertson, Smolensky & Legendre (2012), this provides a promising potential explanation for the typological asymmetry among these 24 ordering patterns. As with Universal 18, the scope of this bias remains an open question which can be investigated further using experimental techniques. It could be the case that the mapping between hierarchical structure and linear order in other domains (i.e. motor/action planning) respects similar kinds of constraints.

4 Conclusion

Research in typology is critical for generative linguistics, where the enterprise is to characterize the human language faculty, including any constraints on the systems it can generate. Although there is disagreement as to whether these constraints must be hard-and-fast limits, or soft biases, and whether they are necessarily special features of language, typology is a source of crucial data. I have suggested here that these data should be used in formulating hypotheses about possible biases rather than treated as their observable result. Accordingly, the goal of much research using ALL paradigms is to provide behavioral evidence for hypothesized connections between typological patterns, like Greenberg's word order universals, and properties of the human cognitive system. The two case studies described above present examples of this kind of research; in both cases, biases are hypothesized on the basis of typological data, and predicted effects on learning are tested using ALL. These experiments corroborate the typological evidence, suggesting that (1) learners are biased in favor of harmonic word order patterns and disfavor one non-harmonic pattern especially (Adj-N, N-Num), and (2) learners tend to infer relative orders of nominal modifiers that preserve the underlying semantic relations among them.

To the extent that connections between typological frequency and ease of learning are borne out, I would argue that the results also bear on major questions addressed by work in functionally-oriented typology; distinctions among patterns in terms of learnability (or use-ability) can be integrated into theories constructed to explain pathways of language change and, ultimately, typological distributions. The methods themselves are also increasingly used to further investigate the content and scope of biases, and whether they might be amplified or altered by social or communicative context. The case studies I have highlighted here illustrate, I hope, the kind of work that is informed by and can make progress in addressing important issues for both typology and generative linguistics.

Acknowledgements

I would like to thank Nick Enfield and attendees of the *Dependencies Among Systems of Language* workshop, held June 2014 in Ardennes, Belgium. I also thank Larry Hyman for comments on a previous version of this paper.

References

Abels, Klaus & Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74.

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

Baker, Mark & James McCloskey. 2007. On the relationship of typology to theoretical syntax. *Linguistic Typology* 11(1). 273–284.

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

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

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36(3). 315–332.

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

Culbertson, Jennifer. 2012. Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass* 6(5). 310–329.

Culbertson, Jennifer & David Adger. 2014. Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences* 111(16). 5842–5847. DOI:10.1073/pnas.1320525111

Culbertson, Jennifer & Elissa L. Newport. 2015. Harmonic biases in child learners: In support of language universals. *Cognition* 139. 71–82.

Culbertson, Jennifer, Paul Smolensky & Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122. 306–329.

Culbertson, Jennifer, Paul Smolensky & Colin Wilson. 2013. Cognitive biases, linguistic universals, and constraint-based grammar learning. *Topics in Cognitive Science* 5(3). 392–424. DOI:10.1111/tops.12027

Cysouw, Michael. 2005. Quantitative methods in typology. In Reinhard Köhler, Gabriel Altmann & Rajmund G. Piotrowski (eds.), *Quantitative linguistics: An international handbook*, 554–578. Berlin: Mouton de Gruyter.

Cysouw, Michael. 2010. Dealing with diversity: Towards an explanation of NP-internal word order frequencies. *Linguistic Typology* 14(2). 253–287.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–183.

Dryer, Matthew S. 2009. On the order of demonstrative, numeral, adjective and noun: An alternative to Cinque. Talk presented at Theoretical approaches to disharmonic word orders, Newcastle University, May-June 2009.

Dryer, Matthew S. 2013a. Order of adjective and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/87.

Dryer, Matthew S. 2013b. Order of numeral and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/89.

Dryer, Matthew S. 2013c. Relationship between the order of object and verb and the order of adposition and noun phrase. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/95.

Fay, Nicolas, Simon Garrod, Leo Roberts & Nik Swoboda. 2010. The interactive evolution of human communication systems. *Cognitive Science* 34(3). 351–386.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the workshop on variation within optimality theory*, 111–120. Stockholm: Stockholm University.

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.

Haspelmath, Martin. 2000. Why can't we talk to each other? *Lingua* 110(4). 235–255.

Hawkins, John A. 1983. *Word order universals*. New York: Academic Press.

Hawkins, John A. 2004. *Complexity and efficiency in grammars*. Oxford: Oxford University Press.

Hudson Kam, Carla & Elissa L. Newport. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59(1). 30–66.

Kirby, Simon, Hannah Cornish & Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105(31). 10681–10686.

Kirby, Simon, Monica Tamariz, Hannah Cornish & Kenny Smith. 2015. Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141. 87–102.

Ladd, D. Robert, Seán G. Roberts & Dan Dediu. 2015. Correlational studies in typological and historical linguistics. *Annual Review of Linguistics* 1. 221–241.

Lightfoot, David. 1989. The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Sciences* 12(02). 321–334.

Moreton, Elliott & Joe Pater. 2012. Structure and substance in artificial-phonology learning, part I: Structure. *Language and Linguistics Compass* 6(11). 686–701.

Newmeyer, Frederick J. 1998. The irrelevance of typology for grammatical theory. *Syntaxis* 1. 161–197.

Piantadosi, Steven T & Edward Gibson. 2014. Quantitative standards for absolute linguistic universals. *Cognitive Science* 38(4). 736–756.

Singleton, Jenny L. & Elissa L. Newport. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* 49(4). 370–407.

Steddy, Sam & Vieri Samek-Lodovici. 2011. On the ungrammaticality of remnant movement in the derivation of Greenberg's Universal 20. *Linguistic Inquiry* 42(3). 445–469.

Travis, Lisa. 1984. *Parameters and effects of word order variation*. MIT Ph.D. dissertation.

Wilson, Colin. 2006. An experimental and computational study of velar palatalization. *Cognitive Science* 30(5). 945–982.