Language and Linguistics Compass 6/5 (2012): 310–329, 10.1002/lnc3.338

Typological Universals as Reflections of Biased Learning: Evidence from Artificial Language Learning

Jennifer Culbertson*
University of Rochester

Abstract

The study of language typology has played a critical role in revealing potential constraints on possible linguistic systems. Such constraints, often called ''typological universals'' have long been used to support a foundational premise of generative linguistics—that languages share a set of underlying commonalities. However, recent research has challenged the idea that typological universals—which (at least on the surface) are often statistical tendencies rather than absolute laws—reflect meaningful biases in the linguistic or cognitive system. In part as a response to these critiques, new behavior methods have been developed to probe the link between recurrent typological patterns and the linguistic (and broader cognitive) system. This article focuses on the novel findings which have resulted from this trend, in particular those which use Artificial Language Learning (ALL) paradigms. This exciting strand of research suggests the viability of experimental methods for investigating constraints on human language, and points to new ways of gaining traction on critical questions in cognitive science.

1. Introduction

1.1. TYPOLOGICAL UNIVERSALS AS LEARNING BIASES?

The study of language typology has played a critical role in revealing potential constraints on possible linguistic systems. Such constraints, often called ''typological universals'' have long been used to support a foundational premise of generative linguistics—that languages share a set of underlying commonalities. The place of typological generalizations or universals in linguistic theory is therefore prominent, although the details are undeniably contentious. For many formal linguists, typological universals are the manifestation of deep linguistic principles which constrain human language. Conceived of as learning biases, these principles or constraints delimit the space of hypotheses entertained by the learner, and therefore facilitate language acquisition (Chomsky 1965; Lightfoot 1997; Tesar and Smolensky 1998). Corroborating evidence for the existence of learning biases has traditionally been sought in derivations of typological predictions from abstract linguistic principles (Baker 2001; Cinque 2005, Prince and Smolensky 1993, 2004), studies of first and second language acquisition (Hawkins 2007; Slobin 2004; White 2003) and work on pidgin and creole formation (Bickerton 1981; McWhorter 1998; Mufwene 1990).[1]

However, recent research has underscored several issues which are potentially problematic for this view. First, some universals may in fact reflect *domain-general* cognitive constraints rather than principles specific to the linguistic system. Second, alternative theories have been formulated which place the causal burden on cognition-external forces, explaining typological asymmetries in purely diachronic or cultural terms (Bybee 2009; Blevins 2004; Evans and Levinson 2009; Culicover and Nowak 2002. That many

so-called universals are actually statistical tendencies—admitting (at least apparent) exceptions—has also provided fodder for those skeptical that universals provide a convincing source of evidence for underlying cognitive principles that constraint the linguistic system.

These issues have prompted researchers to turn to new experimental tools designed to probe the link between recurrent typological patterns and the linguistic (and broader cognitive) system. This article focuses on the novel findings which have resulted from this trend of research, in particular those which use *Artificial Language Learning* (ALL) paradigms. These findings help to clarify how typological patterns emerge and what they are evidence for. In some cases, whether the biases revealed are cognition-general or specific to language is specifically targeted. In other cases, the main goal is to take the crucial first step of showing that biases parallel to some typological pattern *exist*. Taken together this exciting strand of research suggests the viability of experimental methods for investigating constraints on the space of human languages, and points to new ways of gaining traction on critical questions in cognitive science.

This article is organized as follows: in the remainder of the introduction, I outline the traditional view of typological universals as evidence for linguistic constraints or learning biases, and discuss in more detail a number of empirical problems with this view, and alternative explanations which have been proposed to account for the existence of typological universals. In §2 I provide an overview of experimental evidence using ALL in syntax, morphology, and phonology, which suggests that typological universals do indeed reflect constraints on learning. I conclude in §3 by discussing the implications of this research and outlining a number of promising directions for future work in ALL, as well as some challenges faced.

1.2. ABSOLUTE VS. STATISTICAL UNIVERSALS AND ALTERNATIVE THEORIES OF TYPOLOGICAL PATTERNS

The Chomskian hypothesis of Universal Grammar (UG), traditionally conceived of as a set of innate and inviolable linguistic principles, has prevailed in generative linguistics as a solution to the so-called *poverty of the stimulus* problem—the idea that given *only* experience with primary linguistic input, it would be impossible to reliably acquire a language(Chomsky 1965, 1975). The motivation behind the UG hypothesis is that a set of constraints on possible languages delimits the space of hypotheses learners must entertain in order to converge on the correct grammar, and therefore makes the learning problem tractable (Gold 1967). The existence of a set of principles constraining possible linguistic systems not only facilitates learning, but predicts that systems which *conflict* with these principles will be impossible to learn, and therefore not expected to arise in human language. Typological universals thus provide a potentially strong confirmation of this prediction, and accordingly are often the starting point from which to formulate specific UG principles.

Mainstream generative linguistic theories typically formulate linguistic principles as inviolable, therefore typological universals which reflect those principles are expected to be absolute[2], or exceptionless, in some sense.[3] Recent critiques of the UG hypothesis have pointed out that, as we gain access to larger language databases from which to take samples, universals once thought to be absolute have increasingly been shown to admit exceptipns. Take Greenberg's (1963) Universal 18 (discussed further below), which states that a language which has pre-nominal adjectives must also have pre-nominal numerals. Whereas in Greenberg's initial sample of 30 languages Universal 18 had no

exceptions, according to the World Atlas of Language Structures Online, 32 languages (4% of the sample) in fact have pre-nominal adjectives but *post*-nominal numerals (Dryer 2008a, b). Evidence along these lines led, for example, Evans and Levinson (2009) to argue that there are in fact no meaningful universal properties specific to the linguistic system.

While evidence that typological universals are statistical rather than absolute is potentially problematic for generative linguistics, the field has as one of its main goals the development of theories which can derive typological tendencies. In many cases, this includes showing that the universal principles which underly them are absolute at an abstract level of representation, even if there are surface exceptions (e.g. Chomsky 1965; Baker 2001; Huang 1982; Rizzi 1990; Cinque 2005; Biberauer et al. 2008; among many others).[4] Some theories of grammar have also responded to the apparent statistical nature of some universals by treating them as the result of probabilistic biases—under these accounts, languages which violate universals are argued to be *possible*, but are predicted to be more difficult to learn and therefore less likely to arise or be acquired veridically by new generations of learners (Culbertson et al. 2012; Wilson 2006).

Nevertheless, two main classes of alternatives to generative theories of typology exist. Under the first, either cognition-general constraints, or constraints on systems that interface with grammar, but not principles specific to the linguistic system, dictate whether a particular language pattern is more or less likely. For example, typologically recurrent patterns might be explained by positing general cognitive constraints which have particular consequences when operating within the domain of language (Christiansen 2000; Christiansen and Chater 2008; Evans and Levinson 2009; Hupp et al. 2009). Alternatively, the properties of the comprehension or production systems that interface with grammar might influence which language structures are preferred or dispreferred (Haspelmat 2008; Hawkins 1994, 2004).

Under the second, the explanation—either for *all* typological universals, or only for those which turn out to be statistical rather than absolute—is placed outside the cognitive system, either as an accident of history, due to cultural factors (Bybee 2009; Christiansen and Chater 2008; Evans and Levinson 2009; de Lacy 2006), or the result of transmission errors from one generation to the next—i.e. the channel *between* speakers rather than within the cognitive system of a single speaker (Blevins 2004; Ohala 1992).

The experiments discussed below are each designed to provide direct behavioral evidence linking biases in the cognitive systems of individuals—either in the grammatical system or in broader cognition—to typological patterns. Where possible I will also point the reader to evidence from natural language acquisition. The results support the view that cognition-external forces alone cannot explain typological asymmetries, and suggest that even statistical generalizations may reflect cognitive biases. Importantly, this work also points toward a way to investigate the locus of biases as either specific to the linguistic system or cognition-general. Cross-linguistic evidence from projects designed to create large databases like the World Atlas of Language Structures (Haspelmath et al. 2008), from in-depth linguistic analysis of individual languages and from the study of natural language acquisition is undoubtedly important for evaluating the alternative explanations reviewed above. However, the goal of this article is to suggest that behavioral methods like ALL can provide both corroborating evidence, and the ability to explore hypotheses that would be difficult or impossible to test without the controlled conditions offered by experimental methods.

2. Experimental Evidence Linking Typological Universals to Learning Biases

2.1. GENERAL INTRODUCTION TO ARTIFICIAL LANGUAGE LEARNING PARADIGMS

Perhaps the most obvious place to look to uncover learning biases is in the acquisition of natural languages. However, ALL paradigms have been used as an alternative since the 1960s to investigate the process of language learning under a controlled laboratory setting. The pioneering studies in Reber (1967, 1989) showed that experimental participants trained on strings generated by artificial finite-state grammars extend the patterns they learn to novel stimuli, suggesting what he called a process of *implicit learning*.[5]

More recently, various ALL paradigms have been created or modified to test the relative ease with which adult and child learners acquire particular patterns. This has been especially useful in cases where languages which exhibit exceptions to a typological universal of interest are difficult to find, or intractable to study. Using these paradigms generally involves constructing miniature artificial languages designed to allow controlled comparison of structures or patterns of interest—a crucial factor which allows researchers to study learning while minimizing confounding factors in a way which is not possible in studies of natural language acquisition.

Depending on the domain targeted, the learning task may expose participants to a lexicon with or without a corresponding semantic component; most ALL experiments targeting phonological patterns expose learners to words but do not require them to learn any semantic mapping (although see e.g. Gupta et al. 2005), by contrast when syntactic patterns are targeted the lexicon and sentences in the miniature language typically correspond to objects, actions, or events the learner is required to attend to ALL paradigms have also be adapted to use *non-linguistic* stimuli, such as patterned sequences of shapes or tones. Training and testing generally take place during a single, relatively short experimental session, but some ALL paradigms involve training and testing over multiple days (e.g.Hudson Kam and Newport 2009).

Tests of learning in ALL paradigms vary. Commonly, learning is gauged in whole or in part using grammaticality or forced-choice judgments testing participants' ability to distinguish patterns which correspond to the training language from those that do not (e.g. Christiansen 2000). In the *Poverty of the Stimulus Paradigm* and similar approaches (Finley and Badecker 2010; Wilson 2003, 2006), learners are trained on a subset of relevant data and tested on their success at learning that data, *and* their ability to generalize to held-out data. In the *Mixture-Shift Paradigm* (Culbertson et al. 2012; Hudson Kam and Newport 2009), learners are exposed to input which contains a mix of patterns, and the extent to which their *productions* shift that mixture towards or away from the options in the input is taken as a measure of their preferences.

The main advantage of ALL paradigms, aside from the ability to observe learning of language patterns that are not attested (or very rare), resides in the ability to control the learning environment in order to rule out potentially confounding factors. Newer studies have generally been more careful in achieving this by following standard practices in experimental design (i.e. those used in psychology). Perhaps most importantly, many recent studies go to some lengths to account for and partial out previous language experience. For example, Finley (2011) makes use of a group of participants who are tested *without* exposure training in order to control for inherent preferences that are not a result of the learning condition. Similarly, (Culbertson et al. 2012) employ a control condition exposed to a random mix of patterns (in their case the particular word orders being compared) in order to gauge prior preferences (e.g. for orders corresponding to English).

Without such controls, the effect of prior language experience remains an issue for the interpretation of ALL results, particularly because many typological universals of interest are present to some degree in all languages (in the case of absolute universals this is necessarily the case). Critics of ALL have also raised the question of whether (adult) ALL participants employ conscious strategies for learning that are not active during natural language acquisition (by children). However, this has remained a speculative criticism rather than one which overs concrete alternative explanations for the data. Generally, then, if an ALL experiment uncovers clear evidence of learning biases parallel to typology (not due solely to prior language experience), the most parsimonious explanation is that the same bias underlies behavior in the experiment, and the typological asymmetry.

2.2. SYNTACTIC UNIVERSALS

Artificial Language Learning paradigms have been used to investigate typological universals in the domain of syntax, in particular word order, for several decades at least, alongside complementary investigations of first- and second-language learning. For example, in their well-known work on Christopher, a so-called polyglot savant, Tsimpli and Smith (1991) and Smith et al. (1993) explored how Christopher learned natural languages, and conducted ALL experiments testing whether he could learn an ''impossible language'' which incorporated several unattested word order patterns.[6] Cook (1988) tested whether child learners would extrapolate word order in novel phrase types as expected based on universals formulated by Greenberg (1963) and others concerning word order correlations across phrases. Lujan et al. (1984) explored a related set of universals using evidence from the acquisition of word order in child second-language learners of Spanish. More recent ALL studies have targeted word order correlations including Greenberg's Universal 18—a constraint on ordering in the nominal domain—and the preference for consistent ordering of heads across phrases—also know as harmonic ordering (Christiansen 2000; Culbertson et al. 2012). The hypothesis that such word order universals, which are generally statistical rather than absolute, reflect learning biases has been challenged most recently by Dunn et al. (2011). The authors used computational models of language change to argue that phrasal word order is *not* in fact strongly correlated within or across languages as would be predicted if, for example, Greenberg's word order universals were the result of underlying cognitive or linguistic biases.

2.2.1. Word Order Harmony Across Phrase Types

In one of the earliest tests of the connection between learners' biases and typological patterns in syntax using ALL, Cook (1988)s asked whether children's implicit expectations of word order in novel phrase types would be predicted by order in phrases they were trained on. Typological studies of word order correlations led Greenberg (1963) and subsequent linguists to note that across phrase types the ordering of the head and its complements is often strikingly consistent (Chomsky 1988; Dryer 1992; Hawkins 1983). For example, work by Hawkins (1983) and Greenberg (1963) suggests that languages tend to follow these four implication patterns:

(1) *Implicational statements tested by Cook(1988)*
   - i. Object-Verb order ⇒ Noun-Postposition order
   - ii. Verb-Object order ⇒ Preposition-Noun order
   - iii. Object-Verb order ⇒ Adjective-Noun order
   - iv. Verb-Object order ⇒ Noun-Adjective order

In order to determine the extent to which these implicational relationships are reflected in children's expectations of word order during acquisition, Cook (1988) exposed English-speaking children to a subset of phrase types, and tested whether they would extrapolate according to (1) above. This was partially confirmed: (i) learners were significantly more likely to extrapolate from OV → Adj-N, and from VO → N-Adj, as predicted, and (ii) although learners were (perhaps suprisingly) biased to extrapolate to postpositions regardless of the input, they were *more* likely to extrapolate from OV → N-Post. However when learners were trained on multiple phrase types they did not behave as predicted (e.g. they tended to extrapolate from OV and N-Post → N-Adj, and from VO and Prep-N → Adj-N.)

Although the task has some methodological shortcomings[7], the results are of interest since they provide partial confirmation that learners' expectations are in line with typological tendencies—in this case word order correlations of precisely the type which were argued to be lacking by (Dunn et al. 2011).

Christiansen (2000) also used an ALL paradigm, although with adult learners, to investigate whether the typological preference for consistent ordering across phrases reflects a learning bias. Participants (40 adults) were taught either an artificial harmonic language with consistent head-final ordering, or a non-harmonic language in which ordering was inconsistent (some phrases were head-initial, some head-final). At test, participants had to identify whether novel strings followed the rules of the language they were exposed to. Learners in the harmonic language condition correctly classified novel strings significantly more often than learners in the non-harmonic condition. This difference was driven by the fact that learners in the non-harmonic condition were significantly less likely to reject ungrammatical strings (learners in both conditions classified grammatical strings equally well), reflecting their difficulty with a typologically dispreferred pattern of inconsistent or non-harmonic head ordering. Christiansen follows Hawkins (1983, 1990) in proposing to explain this tendency as a processing bias—consistent head order is preferred because it is easier to process.

2.2.2. Greenberg's Universal 18

Culbertson et al. (2012) targeted Greenberg's Universal 18, which states that if a language has pre-nominal adjectives it will also have pre-nominal numerals. Data from the World Atlas of Language Structures (Dryer 2008a, b), shown in Table 1 confirm the typological rarity of the Universal 18-violating pattern: only 4% of languages in the sample have Adj-N *and* N-Num order. The data also reveal that harmonic languages, which preserve the order of the noun with respect to both adjectives and numerals are by far the most common type (more than 80% of the languages). Culbertson et al. (2012) tested 65 English-speaking adults in an ALL task in order to determine whether learners have a bias in favor of harmonic languages (as Christiansen (2000) found), and whether they have a bias *against* the Universal 18-violating pattern of Adj-N, N-Num.

Table 1. Evidence for Universal 18 in the WALS sample.

	Noun–Adj	**Adj–Noun**
Noun-Num	443 (52%)	32 (4%)
Num-Noun	149 (17%)	227 (27%)

Using what they call the ''Mixture-Shift Paradigm'', (adapted from Hudson Kam and Newport 2005, 2009), learners were trained on miniature languages with variable patterns of noun, adjective and noun, numeral order—each language had a dominant pattern following one of the four in Table 1, along with some variation. Hudson Kam and Newport (2009) showed that under certain circumstances, learners tend to regularize variation—choosing one pattern and using it (more) consistently despite multiple options in the input. This ''regularization bias'' has also been found in studies of natural language learning (e.g. Singleton and Newport 2004). Hypothesizing that learners will regularize a pattern only when they do not have a substantive bias against it, Culbertson et al. (2012) predicted that participants in their experiment would be most likely to regularize harmonic patterns, and least likely to regularize the Universal 18-violating pattern. This is precisely what they found; learners exposed to a language with the dominant pattern Adj-N, N-Num did not regularize, while learners in all three other conditions did—using the dominant input pattern significantly more regularly than it was found in the input. As expected, learners exposed to a language with a dominant harmonic pattern regularized the most.[8]

Culbertson et al. (2012) also showed that individual participants in the Universal 18-violating condition actually shifted their language toward one of the two harmonic patterns, bringing it more in-line with the hypothesized biases. They proposed a Bayesian model of these results, arguing that the typological asymmetries in Table 1 are the result of *probabilistic* learning biases. The origin of the bias and its locus within the cognitive system remains to be explored further. The bias in favor of harmonic languages has been claimed to be a reflex of a domain-general preference for consistency by Christiansen (2000) (although see Hawkins 1994; Gibson 2000).for an explanation that relies on a language-specific concept of grammatical domain minimization).

Explanations which specifically target the bias against the particular non-harmonic Universal 18-violating pattern have favored of a language-internal solution, for example relating to the language processing system (Hawkins 1994), to constraints of syntactic movement (Biberauer et al. 2008), or to syntactic repercussions of the semantic difference between adjectives and numerals (Culbertson et al. 2012). In the latter case, work by Kamp and Partee (1995) argues that adjectives can be interpreted only after the noun they modify, regardless of word order, since the precise meaning of the adjective in fact depends on the noun (e.g. gradable adjectives like ''small'' are interpreted differently in e.g. ''small mouse'' vs. ''small building''). Placing the adjective after the noun structurally therefore facilitates this since the adjective can be immediately interpreted, while if it is first it cannot be. Since the same requirement does not hold of numerals, Culbertson et al. (2012) suggest that combining a preference for post-nominal adjectives with a preference for harmonic patterns together predicts the pattern Adj-N, N-Num to be disfavored.

In addition to work on word order universals, other well-known syntactic principles have been argued to follow from learning biases found in laboratory and natural language learning. Ellefson and Christiansen (2000) report ALL experiment and computational simulation results which they argue suggest that the subjacency principle—a set of restrictions on dependencies between syntactic elements which are separated by intervening phrasal boundaries or nodes (Rizzi 1990)—may follow from cognition-general biases in learning of sequential information. More generally, research on first- and second-language learning, has shown that proposed markedness hierarchies in syntax like the Noun Phrase Accessibility Hierarchy (Hawkins 2007; Izumi 2003; Keenan and Comrie 1977; Keenan and Hawkins 1974), implicational hierarchies of predication (van Lier 2005), preposition

stranding and pied-piping (French 1985), among others, can predict ease of acquisition and cross-linguistic frequency. These studies present clear opportunities for future research using ALL to target possible syntax-internal constraints, in particular since this approach has the advantage of controlling properties of the input.

2.3. MORPHOSYNTACTIC UNIVERSALS

2.3.1. The Suffixing Preference

Turning to the domain of morphology, one of the most well-known typological universals concerns the attachment site of inflectional and derivational affixes—the so-called *suffixing preference* (Greenberg 1963). Typological evidence for this comes from the rarity of exclusively prefixing languages (compared to exclusively suffixing languages). Further, in languages which feature a mix of both suffixes and prefixes, there are typically more suffixes than prefixes (Hawkins and Gilligan 1988)—e.g. English, which has 181 suffixes but only 56 prefixes (Fudge 1984).Table 2 provides support for the suffixing preference from the World Atlas of Language Structures sample (Dryer 2008c).

Hawkins and Gilligan (1988) suggested that the suffixing preference follows from a preference to have the portion of the word that determines its category at the end, and Hawkins and Cutler (1988) formulated a model of language processing which encodes this as an assumption that identifying each lexical item in continuous speech as early as possible enables faster processing.[9] Since affixes provide little information about lexical identity, processing of prefixed words will be delayed compared to suffixed words. The idea that the suffixing preference results from a learning bias is further corroborated by findings in natural language acquisition; for example, it has been reported that children acquire inflectional morphology at a slower rate in prefixing-only compared to suffixing-only languages, and that children have an easier time learning suffixes compared to prefixes (Clark 2007; Mithun 1989; Slobin 1973).

Several recent studies have taken this hypothesis as a starting point and have sought to provide experimental evidence of a suffixing-preference in a laboratory language learning setting. For example, St. Clair et al. (2009) used an ALL paradigm to investigate the hypothesis that suffixes allow learners to more successfully determine the grammatical category of a word than prefixes. In their experiment, participants (24 adults) were taught novel words, each belonging to one of two categories (A and B), which shared some phonological properties[10] and were marked by two different novel affixes—for one group of participants these were suffixes, for the other group they were prefixes. Participants were trained on utterances comprised of two category word + affix pairs, and were then asked to judge the similarity of novel pairs to the utterances they heard during training (novel pairs were compatible (e.g. category A word + A affix) or incompatible (e.g. category A word + B affix) with training). Participants in the suffixing condition performed significantly better than participants in the prefixing condition (although performance in

Table 2. Suffixing vs. Prefixing in Dryer (2008c).

	Suffixing	Prefixing	Both
Weakly	114	92	–
Strongly	382	54	–
Total	496	146	130

Table 3. Hupp et al. (2009) label extension task.

Trial 1	Trial 2
♥	♥ ★
''This is a Ta Te''	''This is a Ta Te Be''

both conditions was above chance), in line with the prediction that suffixing cues serve as better predictors than prefixes for relationships between functional items, like affixes, and grammatical category words.[11]

Hupp et al. (2009) report results from a number of ALL tasks in which learners (20 adults) were taught novel labels for objects and were tested on their willingness to extend prefixed or suffixed versions of these labels to those same objects. Table 3 below shows an example set of (suffix) trials. Participants first saw a picture and heard the corresponding label. A second object was then added, and participants heard the affixed form of the label, and were asked to choose the picture they thought that label referred to although participants were reluctant to extend modified labels to the same object, they were nevertheless significantly more likely to do so if the modification involved adding a suffix as opposed to a prefix. They also used the same stimuli to conduct a similarity judgment task, where participants heard a target stem followed by both the suffixed or prefixed versions, and were instructed to choose which modification was most similar to the target stem. Participants were more likely to choose the suffixed form compared to the prefixed form.

The ''suffixes'' and ''prefixes'' in these experiments were not productive—each label was paired with a unique novel affix—and the affixes were not associated with any systematic change in meaning. However, Bruening (2010) replicates the suffixing-preference results from the label extension task with productive novel affixes. Bruening (2010) taught English-speaking children and adults novel productive affixes that could attach to a set of novel and familiar animals as either suffixes or prefixes (e.g. *ko-dog* and *dog-ko* in a label extension task. Although contrary to Hupp et al. (2009) both children and adults usually interpreted the inflected words as referring to the same animals as the uninflected words, both groups were again significantly more likely to accept the word-form modification when the inflection was a suffix.

Hupp et al. (2009) hypothesized that the suffixing preference is not specific to language, but due to the increased salience of the beginnings of temporal sequences in any domain.[12] To investigate whether a suffixing-like-preference could also be found in non-linguistic domains as this account would predict, Hupp et al. (2009) conducted several follow up experiments. In one, they constructed stimuli parallel to the word stem + affix stimuli described above, using sequences of musical notes (e.g. a ''stem'' was a sequence of two arpeggiated notes, with a third note either preceding or following—parallel to an affix). In another experiment they used sequentially presented shapes (e.g. a ''stem'' was a sequence of two shapes, with a third ''affix'' shape preceding or following). Using this stimuli, Hupp et al. (2009) again found that participants were significantly more likely to judge the ''suffixed'' form as more similar to the target stem alone. They also used the non-linguistic shape stimuli in a label-extension task, and replicating the results from their novel word task, found that participants were more likely to extend modified labels to an object if the labels included a shape ''suffix''.

The results from these ALL experiments support the suffixing preference as a robust phenomenon which can be explained by appealing to (possibly domain-general) biases in

the cognitive system. If learners are biased in favor of processing lexically contentful parts of morphologically complex words first, then they may be willing to treat a following weak element as suffixed to its host lexical category but prefer to treat preceding weak elements as nevertheless lexically independent. Alternatively, if suffixed forms are more easily processed than prefixed forms, the latter may gradually disappear from the lexicon of a language (e.g. the content expressed by a particular *prefixed* form may come to be expressed using an alternative grammatical strategy with greater probability than suffixed formed). It is worth noting that the suffixing-preference may have originated as the result of a domain-general learning bias that has come to be encoded as a bias in the linguistic system (this kind of argument has been made by Hawkins 1994, 2004).

2.3.2. Case Marking: Reducing Ambiguity

Another morphological domain in which typological research has revealed robust systematic patterns is case marking. It has been noted that case marking systems appear to be functionally designed in order to improve communicative success—i.e. to indicate grammatical roles explicitly via case-marking—while simultaneously avoiding redundancy (e.g. Comrie 1989; Greenberg 1963; Jager 2007). For example, Greenberg (1963) formulated Universal 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.'' Universal 38 makes intuitive sense; there is no need to differentiate the subject in an intransitive sentence, since there are no other participants with which it could be confused. Rather, case marking systems should be used to differentiate subjects of *transitive* verbs from objects of transitive verbs.

Using an ALL paradigm similar to the Mixture-Shift Paradigm (described above Culbertson et al. 2012; Hudson Kam and Newport 2009), Fedzechkina et al. (2011), explored whether learners show evidence of a bias to avoid systematic ambiguity—one side of the functional motivation for case cited by Jager (2007) and others. Adult English-speakers were exposed to two languages which both had variable word order (63% of sentences used SOV order, 37% used OSV order), however one also had case-marking on all object nouns. The two languages therefore differed crucially; in the case language, the grammatical roles of agents and patients were unambiguously marked, but in the no-case language all sentences were potential ambiguous (word order varied, and the set of nouns included only human referents). Participants in the task were trained and tested on one of other these two languages over four consecutive days.

If learners are indeed biased against systematic ambiguity of the type found in the no-case language, they have a simple recourse in the experiment—they could ''regularize'', or fix, the variable word order present in the input. On the other hand, learners in the case language have no such motivation to fix word order. Accordingly, the results of the experiment showed that, when interpreting sentences in the language, learners in the no-case condition were significantly more likely to assume the dominant word order (SOV). That is, if they were shown two pictures with reversed grammatical roles, and asked to indicate which one matched the sentence they heard, they tended to choose the picture corresponding to an SOV interpretation. Similarly, when they produced sentences describing scenes displayed to them, learners in the no-case condition were very likely to use the dominant word order, and were marginally more likely to do so than learners in the case condition. These results suggest that case and word order interact during learning in the experiment; when word order is variable, and there is no case-marking to indicate grammatical roles, learners will regularize word order to establish a reliable cue.

2.3.3. Agreement: The Definiteness Hierarchy

In the domain of agreement, typological studies have revealed an implicational hierarchy which dictates the relationship between the appearance of an overt subject-verb agreement marker and the definiteness of the agreement controlling noun phrase (Corbett 2006; Poletto 2000; Siewierska 2004). Specifically, the higher a subject noun phrase is on the scale of definiteness in (2) below, the more likely it is to trigger agreement, and further if a given subject noun triggers agreement in a language, all subject types which are higher on the definiteness scale will also trigger agreement. This predicts that a system in which agreement is triggered by definite subjects but not indefinite subjects is possible, but the *opposite* system where agreement is triggered by indefinites but not definites is not.

(2) *Definiteness hierarchy:*
Pronoun > proper name > definite NP > specific indefinite NP > nonspecific indefinite NP

To investigate learners' sensitivity to this implicational hierarchy, Culbertson and Legendre (forthcoming) exposed adult English-speaking learners, in a single one hour session, to languages which instantiate these two patterns—the well-attested ''natural'' pattern, and the ''unnatural'' pattern predicted impossible by the hierarchy. However, similar to the input learners might be exposed to during a period of language change, each language used agreement only variably (see Culbertson 2010, In the natural language, agreement was triggered by definite subject noun phrases 75% of the time, but never by indefinite subjects, and in the unnatural language, agreement was triggered by indefinites 75% of the time and never by definites. The results of the experiment revealed that learners exposed to the natural language regularized the pattern of variable agreement, using the agreement morpheme with definite subject DPs more than 75% of the time when asked to produce sentences in the language. Learners in the unnatural condition, on the other hand, did not regularize the pattern of agreement they were exposed to rather, they used agreement with indefinite subjects significantly less often than it was present in the input, and further, they were more likely to over-generalize case to the non-agreement-triggering subject type—in this case definites. The bias against the unnatural language therefore appeared to lead participants to *shift* the input language toward a language predicted possible by the definiteness hierarchy.

2.4. PHONOLOGICAL UNIVERSALS

Typological universals in the phonological domain have been investigated by a number of researchers using ALL paradigms. Universals or tendencies concerning epenthesis (Morley forthcoming), vowel harmony Pycha et al. 2003; Koo and Cole 2006; Finley 2008; Finley and Badecker 2010; Moreton 2008), consonant harmony (Koo and Cole 2006; Wilson 2003), dependency length (Koo and Cole 2006; Newport and Aslin 2004; Pacton and Perruchet 2008), and velar palatalization (Wilson 2006) have been studied using ALL, here I discuss experiments targeting several of these. Although there is robust typological evidence and in some cases evidence from natural language acquisition (e.g. Pater and Werle 2001; on consonant harmony in child language), of particular interest is the explanation for asymmetries in these systems. In particular, in the domain of phonology, at issue is whether cognitive or grammatical learning biases underlie typological asymmetries which could alternatively be explained by phonetic factors. Such factors may

influence how a given phonological pattern is perceived and transmitted between speakers, affecting the likelihood with which it is eventually phonologized—this has been called *channel bias* (Moreton 2008). Typological universals resulting from channel bias have been argued *not* to reflect biases in the cognitive or linguistic system.

Here I follow proponents of these theories (e.g. Ohala 1992), and researchers who argue against them (e.g. Moreton 2008), in characterizing biases localized in the channel between speakers as under the influence of cognition-external forces, but this is not entirely straightforward. In particular, if the perceptual system which characterizes this channel reacts differently to certain properties of the speech-stream compared to others, this requires an explanation. The perceptual system is certainly part of the cognitive system of an individual speaker, therefore it is not completely clear that these channel biases should be treated as cognition-external, although they may be external to the *linguistic* system.

2.4.1. Velar Palatalization

Take for example a well-known typological universal in the domain of phonology, first proposed by Bhat (1978): if a language has palatalization conditioned on back vowels (e.g. /k/ → /tʃ/ before /e/) then it will also have palatalization before front vowels (e.g. /i/). Ohala (1992), Guion (1998) and others have argued that this universal is the result of phonetic factors; velar stops and palatoalveolar affricates are acoustically and perceptually more similar before front vowels compared to back vowels, therefore velar-palatalization will be more likely in the context of the former. If the typology reflects the effect of this phonetic asymmetry on diachronic change, then the force behind the universal need not be localized in the cognitive system of an individual speaker (Blevins 2004; Ohala 1992).

To investigate whether this universal *is* in fact the result of a substantive cognitive bias, Wilson (2006) developed the Poverty of the Stimulus Paradigm, in which learners are exposed to a subset of data instantiating a typologically relevant pattern of interest, and are then tested on their willingness to generalize that pattern to a new set of data. Subjects were taught a language game in which they heard novel word pairs involving velar palatalization either only before the mid-vowel /e/(e.g./kenə/... /tʃenə/) or only before the high-vowel /i/(e.g./kinə/... /tʃinə/). They were then tested on their willingness to generalize this process to a different vowel; participants in the mid-vowel condition were tested with the high-vowel and participants in the high-vowel condition where tested with the mid-vowel.

Wilson (2006) hypothesized that if learners have an implicit cognitive bias favoring alternations between perceptually similar sounds—that is, a bias which is shaped by phonetic factors, but is nevertheless encoded in the mental grammar—then they should show different patterns of generalization across conditions. The results of the study confirm this prediction; parallel to the typological pattern, participants were more likely to generalize velar palatalization from the mid-vowel /e/ to the high-vowel /i/, than in the opposite direction. Under the hypothesis that channel bias alone is responsible for the typological universal, this asymmetrical pattern of generalization is unexpected.

2.4.2. Height-based Vowel Dependency Systems

Moreton (2008), also targeting the explanatory adequacy of channel bias accounts of typology, compared learning of two phonological patterns which differ *only* in terms of typological frequency, not in terms of phonetic precursors for phonologization. In particular, dependencies between the height of vowels in different syllables (height-height

patterns like height harmony or disharmony) are typologically much more common than dependencies between the height of a vowel and the voicing status of following consonants (height-voice patterns). Crucially however, Moreton (2008) shows that the phonetic factors (here coarticulation between segments of interest) which would support a channel bias explanation for the phonologization of these dependencies are equally strong—predicting no typological asymmetry.

(3) *Example training items in Moreton(2008):*
Height-Height: /tikæ/, /tukɔ/
Height-Voice: /tugi/, /tukɔ/

To test whether substantive learning biases can explain the typology, when channel bias cannot straightforwardly do so, learners (24 adults) were taught miniature languages with either a height-height or height-voice dependency (instantiated in a set of CVCV nonce lexical items, as in (3) above). At test, participants heard a pair of novel words, only one of which followed the training pattern, and were instructed to choose which word was in the language they studied. Each participant was taught both languages (ordered randomly), with a break in between during which they listened to music. A cognitive bias favoring height-height over height-voice dependency predicts a difference in performance across these condition, while a cognition-external channel bias would predict no difference. The results of the study in fact parallel the typological asymmetry, showing that choice of the correct item at test was only marginally above chance performance for the height-vowel language, but was significantly better for the height-height language.

2.4.3. Vowel Harmony

The connection between learning biases and typological tendencies among vowel harmony systems have been studied extensively using the Poverty of the Stimulus paradigm by Finley (2008). Finley and Badecker (2010) investigate whether general cognitive constraints operating in the domain of language might explain why a particular logically possible harmony system is unattested. Learners (36 adults) were taught miniature artificial languages with ambiguous patterns of vowel harmony. The training input featured pairs of words, here illustrating a harmony process which was consistent with both a well-attested type of system—directional vowel harmony in which features, e.g. backness or rounding, of a vowel on one edge of a word spread to the other vowels in the word—and an unattested system—so-called ''majority rules'', in which the particular harmony-triggering feature is the one which would ensure harmony while making the fewest number of changes to the underlying form (see (4) below for example stimuli).

(4) *Example training pair in Finley and Badecker (2010)*:
/bo du ti/ → /bodutu/

Learners in the experiment were more likely to generalize to—that is, to judge as belonging to the language they were trained on—new words which unambiguously used a directional rather than a majority rules system, suggesting that they are biased against such systems. Finley and Badecker (2010) also investigated whether the bias against ''majority rules'' patterns can be found with non-linguistic stimuli. They created a set of stimuli which was parallel to their linguistic stimuli, but which used colors and shapes; ''harmony'' in this case involved spreading the shape feature. They found that if participants were trained on a visual analogue of harmony which was consistent with both right-to-left

directional spreading and ''majority rules'' they consistently inferred the directional system, similar to what was found using linguistic stimuli. However, if the training stimuli was consistent with both *left-to-right* directional spreading and ''majority rules'', learners showed no preference for either system. The experiment therefore partially confirms the idea that the bias against ''majority rules'' patterns may be domain-general rather than specific to the language faculty. However, the differences which were found suggest that the way speech is processed may affect how the bias is revealed in vowel harmony systems.

3. Conclusion: Summary, Issues, and Future Directions

The results of the ALL experiments discussed here provide evidence linking typological universals in phonology, morphology and syntax to biases operating during language learning by both adults and children. This evidence strengthens claims about the underlying explanation for typological asymmetries made on the basis of linguistic analysis and natural language acquisition, and has important implications for the debate over how universals should be explained and what factors constrain human language. Importantly, in many cases the universals targeted were *statistical tendencies* rather than being absolute. Nevertheless, learners still exhibited biases parallel to these tendencies, suggesting that the mere existence of counter-examples cannot be used to argue that a universal does not reflect anything about the cognitive system. Culbertson et al. (2012) found that learners are biased against a particular pattern of order in the nominal domain—in line with Greenberg's Universal 18—despite documented languages which use that order. Hupp et al. (2009) and others found evidence for the suffixing preference, which is just that—a preference. Wilson (2006) found that learners were more likely to generalize from a process of palatalization before mid-vowels to palatalization before high-vowels than vice versa. This bias is in line with a typological asymmetry, albeit one which has known exceptions.

Generative linguists have long argued that surface exceptions do not necessarily constitute evidence against principles operating at abstract levels of representation, in part because many mainstream frameworks operate under the assumption that constraints on representations are strict (not violable) rules. Experimental work of this type has opened the door for a new source of evidence and potentially new theories of the nature of learning biases. For example, Wilson (2006) used the results of an ALL study to argue for a theory of phonology in which grammatical principles are encoded as *cognitive biases* rather than absolute restrictions—a bias against a particular pattern does not make it impossible, rather learners' predisposition make certain patterns *more difficult* to acquire. Along similar lines, Culbertson et al. (2012) showed that a statistical typological preference among word order patterns in the nominal domain can be predicted by learning biases revealed in an ALL experiment, and argue for a model of learning that is constrained by probabilistic biases (see also Culbertson and Smolensky forthcoming).

Under this view, if a bias puts pressure on learners to acquire grammars with particular properties, grammars will tend to change over time to satisfy that bias. In other words, a grammar which satisfies some learning bias may act as a kind of magnet, pulling learners exposed to grammars which *do not* satisfy that bias to acquire a grammar which does (or at least which does to a greater extent). Depending on the strength of the bias, the resulting typological universal could be statistical *or* absolute. It remains to be seen, especially as researchers begin to test learners within the critical period for language acquisition (Johnson and Newport 1989), how strong these biases are, and whether in some cases

they will function as absolute constraints. For example, one could imagine that adults in the lab are able to overcome, to some extent at least, biases that are effectively absolute during natural language learning.

In the domain of syntax, sophisticated ALL studies are still in their infancy. A number of influential absolute constraints claimed to hold at abstract levels of representation[13], e.g. binding principles, island constraints, constraints on question formation, remain to be investigated using ALL. One difficulty lies of course in the fact that these constraints are claimed to hold of all languages, therefore an experiment showing that they also hold of learners in the lab may be dismissed on the grounds that the result is due simply to prior language knowledge rather than any innate bias. If therefore seems that the most promising contribution of ALL studies to these issues would use the methodology to probe more deeply the origin and scope of these biases (for example, see on subjacency Ellefson and Christiansen 2000, on subjacency).

A number of studies discussed here already touch upon the origin and scope of underlying biases—issues of particular interest to the broader cognitive science community. For example, Hupp et al. (2009) undertook to show that the so-called ''suffixing preference'' may have broad scope, reflecting a general attentional bias which applies to sequential information in language, music and vision. Finley and Badecker (2010) explored whether a universal preference for certain vowel harmony systems also applied to learning of non-linguistic ''harmony'' patterns involving sequences of shapes. Moreton (2008) worked to show that phonetic precursors which might have encouraged phonologization of some patterns rather than others did not predict the behavioral findings in an ALL task. In these cases, the experimental results suggest that typological tendencies reflect underlying learning biases. This does not, however, exclude the idea that these biases were shaped by, or originated as the result of phonetic or perceptual factors. In the domain of syntax, both Christiansen (2000) and Culbertson et al. (2012) Culbertson, Smolensky, and Legendre suggest that a bias in favor of harmonic word orders (orders that preserve the position of the head as first or last across difference phrase types) found in their ALL experiments may reflect general cognitive processes applied to language. However Culbertson et al. (2012) Culbertson, Smolensky, and Legendre argue that the asymmetry among two *non-harmonic* patterns may in fact reflect a bias specific to the linguistic system. Future work will undoubtedly further investigate how general any uncovered biases are, and where they are localized in the cognitive or perceptual system, e.g. in perception, comprehension, production.

A related issue concerns whether the biases uncovered in ALL studies truly reflect something about *learning* in particular. For example, the idea that language *learners* play a special role in shaping typology (e.g. by shifting input grammars in certain directions rather than others) is not uncontroversial. Some socio- and historical linguists have claimed that adult language users are the dominant force in language change (e.g. Labov 1994), and therefore biases they have might be of equal or more importance. Most ALL studies providing evidence for biases during laboratory learning do not explicitly address this, however many of them are consistent with the idea that biases operating during learning persist into later language use.

The studies discussed here underscore the utility of ALL experiments in expanding the current dialog surrounding typological universals, and understanding the place of the latter in cognitive science. They also serve to highlight areas where future research can focus. In particular, research using ALL will continue to provide perhaps the clearest sources of evidence corroborating (or not) the existence of constraints on the space of human language, and probing the scope of these constraints as specific to language or

cognition-general. Relatedly, this work has the potential to influence our conception of how biases at various levels can work together during learning, and how models of probabilistic biases can be integrated into theories of cognition. ALL also offers an opportunity to further the already rich dialog surrounding the critical period for language acquisition, and to investigate differences between adults and children as they impact the shape of language.

Short Biography

Jennifer Culbertson's research investigates the extent to which constraints on language change can explain synchronic properties of grammars manifest, for example, in typological patterns and linguistic universals. Her work develops new approaches to addressing this question by deploying tools that specifically target the role of the learner in constraining how languages are acquired and thus how they change over time. The work addresses issues in syntax (e.g. word order) and morphosyntax (e.g. clitics, agreement), for which further development of these new tools has been necessary. Currently a postdoctoral fellow [] in the Center for Language Sciences at the University of Rochester, Culbertson received her Ph.D. in Cognitive Science from Johns Hopkins University.

Notes

* Correspondence address: Jennifer Culbertson, 246 Meliora Hall, Rochester, New York, 14627, USA. E-mail: jculbertson@bcs.rochester.edu

[1] The citations given here are of course in no way meant to be exhaustive; there is extensive work targeting each of these sources of evidence for deep principles of the linguistic systems.

[2] The term ''absolute'' is sometimes used to contrast implicational (if a language has property *x*, it must also have property *y*) from non-implicational universals (all languages must have property *x*). Here, I use ''absolute'' to refer to inviolable or exceptionless universals (whether implicational or not), which contrast with ''statistical'' universals—robust tendencies which nevertheless have exceptions.

[3] Even in Optimality Theory PrinceSmolensky04, where constraints themselves are violable, a hard-and-fast line is nevertheless drawn between possible and impossible languages.

[4] One of the most well-known early examples of this concerns wh-question formation. While many languages (e.g. English) overtly move wh-phrases to form questions, prompting Chomsky (1981) and others to formulate principles governing such dependencies, some languages (e.g. Chinese) do not have surface movement. Nevertheless, as Huang (1982) argued, languages like Chinese show the same types of syntactic restrictions on the interpretation of wh-questions as languages with overt movement, leading him to proposed a unified analysis positing *covert* movement in languages like Chinese.

[5] This work prompted an ongoing debate about whether learners in laboratory settings in fact implicitly induce abstract structure, or instead (explicitly) learn more superficial statistical properties of the input (e.g. Perruchet and Pacteau 1990, 1991). The debate continues today, and now includes research using brain imaging techniques (e.g. Turk-Browne et al. 2009).

[6] The results showed for example that Christopher could not learn negation and past tense systems in which contrasts were marked by word order changes (e.g. SVO for positive, but VSO for negative), both unattested types of systems. However, Christopher showed a general tendency to simply transfer English grammatical patterns to the new languages he learned, for example when tested prior to exposed to Berber he categorically used SVO rather than the appropriate VSO.

[7] The results of the experiment cannot be fully explained by a bias favoring harmony across phrase types, nor by the influence of English (or anti-English for that matter). Cook (1988) suggests that some participants may have been treating the task less like language learning than general problem solving. The methodology used in the experiment makes this a plausible suggestion; children were given explicit translations from the artificial language to English, testing involved written multiple choice questions, and children exposed to two phrase types received only 13 examples of each phrase type. These are not typical features of more recent ALL paradigms.

[8] As in all the studies reported here, Culbertson et al. (2012) and Christiansen (2000) tested English-speakers, an important consideration in evaluating the results. Given the robust status—typologically, in natural language learning, and in laboratory learning experiments—of the preference for consistent ordering, it seems reasonable to

conclude that the bias found is not the result of experience with English. For example, Culbertson et al. (2012) find no preference for the harmonic English-like pattern (Adj-N, Num-N) over the harmonic non-English-like pattern (N-Adj, N-Num). They also argue that the behavioral difference found between the two non-harmonic conditions in their experiment—one of which violates Universal 18, while the other is fairly well-attested typologically—cannot be attributed to English knowledge; most importantly, both share one phrasal order with English.

[9] Greenberg (1957) originally suggested that the suffixing-bias might be related to the fact that affixes make up a small (closed) set whereas the words they are attached to make up a larger (potentially open class) set. Prefixing therefore involves a small number of affixes preceding a large class of stems, while suffixing provides the opposite situation, wherein a small set of suffixes follows a large set of stems. Although he actually suggests using ALL to test this, Greenberg cites Osgood (1949) as showing that learning the latter type of system is faster and less error prone than the former

[10] An influential early ALL study suggests that phonological similarities among words in an otherwise semantically arbitrary class may be important for learning. Brooks et al. (1993) showed that learning of gender systems was facilitated when words in a given class had phonologically similar endings (although not when similarity was spread across the word)

[11] There are some potentially confounding factors in the study. For example, participants are English speakers, but the authors provide corpus evidence that suffixes are generally more reliable at predicting category information *in English*—an alternative explanation for why participants more successfully use suffixes for categorization in the context of the experiment.

[12] Additional evidence for this general salience comes from work on music perception (Repp 1992). In the case of words, extrapolating from Hawkins and Cutler (1988), perhaps if lexical identity is encoded in this salient portion of the word it will be processed faster, but if an affix is placed there instead, the added salience will be wasted, and processing will be slowed.

[13] One should note that many if not most of these constraints are heavily debated among theoretical linguistics, e.g. see Hofmeister and Sag (2010) and references therein on whether island-constraints should be treated as purely grammatical.

Works Cited

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

Bhat, D. N. S. 1978. A general study of palatalization. Universals of human language, ed. by Joseph Greenberg, volume 3, 47–92. Stanford, CA: Stanford University Press.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2008. Structure and linearization in disharmonic word orders. Proceedings of the 26th West Coast Conference on Formal Linguistics, ed. by Charles B. Chang and Hannah J. Haynie, 96–104. Somerville, MA: Cascadilla Press.

Bickerton, Derek. 1981. Roots of language. Ann Arbor, MI: Karoma Publishers.

Blevins, Juliette. 2004. Evolutionary phonology: the emergence of sound patterns. New York: Cambridge University Press.

Brooks, Patricia J., Martin D. S. Braine, Lisa Catalano, Ruth E. Brody, and Vicki Sudhalter. 1993. Acquisition of gender-like noun subclasses in an artificial language: the contribution of phonological markers to learning. Journal of Memory and Language 32. 76–95.

Bruening, Paul Reeves. 2010. Children's tolerance of word-form variation. Ph.d., New York: The City University of New York.

Bybee, Joan. 2009. Language universals and usage-based theory. Language universals, ed. by Morten H. Christiansen, Chris Collins and Shimon Edelman, 17–39. Oxford: Oxford University Press.

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

——. 1975. Reflections on language. New York: Pantheon.

——. 1981. Lectures on government and binding. Dordrecht: Foris.

——. 1988. Language and problems of knowledge: the Managua lectures. Cambridge, MA: MIT Press.

Christiansen, Morten H. 2000. Using artificial language learning to study language evolution: exploring the emergence of word order universals. The evolution of language: 3rd international conference, ed. by J. L. Dessalles and L. Ghadakpour, 45–8. Paris: Ecole Nationale Superieure des Telecommunications.

——, and Nick Chater. 2008. Language as shaped by the brain. Behavioral and Brain Sciences 31. 489–509.

Cinque, Guglielmo. 2005. Deriving greenberg's universal 20 and its exceptions. Linguistic Inquiry 36. 315–32.

Clark, Eve. 2007. Morphology in language acquisition. The handbook of morphology, ed. by A. Spencer and Arnold M. Zwicky, 374–89. Oxford: Blackwell.

Comrie, Bernard. 1989. Language universals and linguistic typology. Chicago: University of Chicago Press.

Cook, Vivian. 1988. Language learners extrapolation of word order in micro-artificial languages. Language Learning 38. 497–529.

Corbett, Greville. 2006. Agreement. New York: Cambridge University Press.

Culbertson, Jennifer. 2010. Convergent evidence for categorial change in French: from subject clitic to agreement marker. Language 86. 85–132.
——, and Géraldine Legendre. forthcoming. Investigating the evolution of agreement systems using an artificial language learning paradigm. Proceedings of the 2010 Western Conference on Linguistics. Fresno: Department of Linguistics, California State University.
——, and Paul Smolensky. forthcoming. A Bayesian model of biases in artificial language learning: the case of a word-order universal. Cognitive Science.
——, ——, and Géraldine Legendre. 2012. Learning biases predict a word order universal. Cognition 122. 306–29.
Culicover, P. W., and A. Nowak. 2002. Markedness, antisymmetry and the complexity of constructions. Language Variation Yearbook 2. 5–30.
Dryer, Matthew. 1992. The greenbergian word order correlations. Language 68. 81–183.
——. 2008a. Order of adjective and noun. The world atlas of language structures online, chapter 87, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie. Munich: Max Planck Digital Library.
——. 2008b. Order of numeral and noun. The world atlas of language structures online, chapter 89, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie. Munich: Max Planck Digital Library.
——. 2008c. Order of numeral and noun. The world atlas of language structures online, chapter 26, ed. by Martin. Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie. Munich: Max Planck Digital Library.
Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson, and Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. Nature 473. 79–82.
Ellefson, Michelle R., and Morten H. Christiansen. 2000. Subjacency constraints without universal grammar: evidence from artificial language learning and connectionist modeling. The Proceedings of the 22nd Annual Conference of the Cognitive Science Society, ed. by Lila R. Gleitman and Aravid K. Joshi, 645–50. Mahwah, NJ: Lawrence Erlbaum.
Evans, Nicholas, and Stephen C. Levinson. 2009. The myth of language universals: language diversity and its importance for cognitive science. Behavioral and Brain Sciences 32. 429–48.
Fedzechkina, Maryia, T. Florian Jaeger, and Elissa Newport. 2011. Functional biases in language learning: evidence from word order and case-marking interaction. Proceedings of the 33rd Annual Conference of the Cognitive Science Society, ed. by Laura Carlson, Christoph Hoeschler and Thomas F. Shipley, 318–23. Austin, TX: Cognitive Science Society.
Finley, Sara. 2008. Formal and cognitive restrictions on vowel harmony. Doctoral Dissertation, Baltimore, MD: Johns Hopkins University.
——. 2011. Generalization to novel consonants in artificial grammar learning. Proceedings of the 33rd Annual Conference of the Cognitive Science Society, ed. by Laura Carlson, Christoph Hoeschler and Thomas F. Shipley, 318–23. Austin, TX: Cognitive Science Society.
——, and William Badecker. 2010. Linguistic and non-linguistic influences on learning biases for vowel harmony. Proceedings of the 32nd Annual Conference of the Cognitive Science Society, ed. by S. Ohlsson and R. Catrambone, 706–11. Austin, TX: Cognitive Science Society.
French, M. 1985. Markedness and the acquisition of pied-piping and preposition stranding. McGill Working Papers in Linguistics 2. 131–44.
Fudge, E. C. 1984. English word-stress. London: George Allen and Unwin.
Gibson, Edward. 2000. The dependency locality theory: a distance-based theory of linguistic complexity. Image, language, brain, ed. by A. Marantz, Y. Miyashita and W. O'Neil, 95–126. Cambridge, MA: MIT Press.
Gold, E. Mark. 1967. Language identification in the limit. Information Control 10. 447–74.
Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. Universals of language, ed. by Joseph Greenberg, 73–113. Cambridge, MA: MIT Press.
Greenberg, Joseph H. 1957. Order of affixing: a study in general linguistics. Essays in linguistics, ed. by. Joseph H. Greenberg, 86–94. Chicago: University of Chicago Press.
Guion, Susan Guignard. 1998. The role of perception in the sound change of velar palatalization. Phonetica 55. 18–52.
Gupta, P., J. Lipinski, and E. Aktunc. 2005. Reexamining the phonological similarity effect in immediate serial recall: the roles of type of similarity, category cuing, and item recall. Memory & Cognition 33. 1001.
Haspelmat, Martin. 2008. Parametric versus functional explanations of syntactic universals. The limits of syntactic variation, ed. by. Theresa Biberauer, 75–107. Philadelphia: John Benjamins.
Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie. 2008. The world atlas of language structures online. Munich: Max Planck Digital Library.
Hawkins, John A. 1983. Word order universals. New York: Academic Press.
——. 1990. A parsing theory of word order universals. Linguistic Inquiry 21. 223–61.
——. 1994. A performance theory of order and constituency. Cambridge: Cambridge University Press.
——. 2004. Complexity and efficiency in grammars. Oxford: Oxford University Press.
——. 2007. Acquisition of relative clauses in relation to language universals. Studies in Second Language Acquisition 29. 337–44.

——, and Anne Cutler. 1988. Psycholinguistic factors in morphological asymmtery. Explaining language universals, ed. by. John A. Hawkins, 280–317. Oxford: Blackwell.
——, and Gary Gilligan. 1988. Prefixing and suffixing universals in relation to basic word order. Lingua 74. 219–59.
Hofmeister, P., and I. A. Sag. 2010. Cognitive constraints and island effects. Language 86. 366–415.
Huang, C. T. James. 1982. Move wh in a language without wh-movement. The Linguistic Review 1. 369–416.
Hudson Kam, Carla, and Elissa Newport. 2005. Regularizing unpredictable variation. Language Learning and Development 1. 151–95.
——, and ——. 2009. Getting it right by getting it wrong: when learners change languages. Cognitive Psychology 59. 30–66.
Hupp, Julie M., Vladimir M. Sloutsky, and Peter W. Culicover. 2009. Evidence for a domain-general mechanism underlying the suffixation preference in language. Language and Cognitive Processes 24. 876–909.
Izumi, Shinichi. 2003. Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. Language Learning 53. 285–323.
Jager, Gerhard. 2007. Evolutionary game theory and typology: a case study. Language 83. 74–109.
Johnson, J. S., and Elissa L. Newport. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of english as a second language. Cognitive Psychology 21. 60–99.
Kamp, Hans, and Barbara Partee. 1995. Prototype theory and compositionality. Cognition 57. 129–91.
Keenan, E., and S. Hawkins. 1974. The psychological validity of the accessibility hierarchy. Universal grammar: 15 essays, ed. by. E. Keenan, 60–85. London: Routledge. LSA Summer Meeting.
Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. Linguistic Inquiry 8. 63–99.
Koo, H., and J. Cole. 2006. On learnability and naturalness as constraints on phonological grammar. Proceedings of ISCA Ttutorial and Research Workshop on Experimental Linguistics, 28-30 August 2006 Athens, Greece, ed. by. Antonis Botinis, 165–8. Athens: University of Athens Press.
Labov, William. 1994. Principles of language change, vol. 1: internal factors. New York: Blackwell.
de Lacy, P. V. 2006. Markedness: reduction and preservation in phonology. Cambridge: Cambridge University Press.
van Lier, Eva. 2005. The explanatory power of typological hierarchies: developmental perspectives on non-verbal predication. Morphosyntactic expression in functional grammar, ed. by. Casper de Groot and Kees Hengeveld, 249–80. New York: Mouton de Gruyter.
Lightfoot, David. 1997. Catastrophic change and learning theory. Lingua 100. 171–92.
Lujan, Marta, Liliana Minaya, and David Sankoff. 1984. The universal consistency hypothesis and the prediction of word order acquisition stages in the speech of bilingual children. Language 60. 343–71.
McWhorter, John. 1998. Identifying the creole prototype: vindicating a typological class. Language 74. 788–818.
Mithun, Marianne. 1989. The acquisition of polysynthesis. Journal of Child Language 16. 285–312.
Moreton, Elliott. 2008. Analytic bias as a factor in phonological typology. Proceedings of the 26th West Coast Conference on Formal Linguistics, ed. by. Charles B. Chang and Hannah J. Haynie, 393–401. Somerville, MA: Cascadilla Proceedings Project.
Morley, Rebecca. forthcoming. From phonetics to phonology: learning epenthesis. Proceedings of the 47th Annual Meeting of the Chicago Linguistics Society.
Mufwene, Salikoko S. 1990. Creoles and universal grammar. Linguistics 28. 783–808.
Newport, Elissa L., and Richard N. Aslin. 2004. Learning at a distance I. Statistical learning of non-adjacent dependencies. Cognitive Psychology 48. 127–62.
Ohala, James J. 1992. What's cognitive, what's not, in sound change. Diachrony within synchrony: language history and cognition. Duisberger arbeiten zur sprach- und kulturwissenschaft 14, ed. by. G. Kellermann and M. Morrissey, 309–55. Frankfurt: Peter Lang.
Osgood, Charles E. 1949. The similarity paradox in human learning: a resolution. Psychological Review 56. 132–43.
Pacton, S., and P. Perruchet. 2008. An attention-based associative account of adjacent and nonadjacent dependency learning. Journal of Experimental Psychology: Learning, Memory, and Cognition 34. 80.
Pater, J., and A. Werle. 2001. Typology and variation in child consonant harmony. Proceedings of HILP, University of Potsdam, ed. by. Antony Dubach Green Caroline Féry and Ruben van de Vijver, volume 5, 119–39. Potsdam: University of Potsdam.
Perruchet, P., and C. Pacteau. 1990. Synthetic grammar learning: implicit rule abstraction or explicit fragmentary knowledge?. Journal of experimental psychology: General 119. 264.
——, and ——. 1991. Implicit acquisition of abstract knowledge about artificial grammar: some methodological and conceptual issues. Journal of Experimental Psychology: General 120. 112–6.
Poletto, Cecelia. 2000. The higher functional field evidence from northern Italian Dialects. New York, NY: Oxford University Press.

Prince, Alan, and Paul Smolensky.1993/2004. Optimality theory: constraint interaction in generative grammar. Technical Report, New York, NY: Rutgers University and University of Colorado at Boulder, 1993. Rutgers Optimality Archive 537, 2002. Revised version published by Blackwell 2004.

Pycha, Anne, Pawel Nowak, Eurie Shin, and Ryan Shosted. 2003. Phonological rule-learning and its implications for a theory of vowel harmony. Proceedings of the 22nd West Coast Conference on Formal Linguistics, ed. by. M. Tsujimura and G. Garding, 101–14. Somerville, MA: Cascadilla Press.

Reber, A. S.. 1989. Implicit learning and tacit knowledge. Journal of Experimental Psychology: General 118. 219.

Reber, Arthur S. 1967. Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior 6. 855–63.

Repp, Bruno H. 1992. Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. Cognition 44. 241–81.

Rizzi, Luigi. 1990. Relativized minimality. Cambridge: MIT Press.

Siewierska, Anna. 2004. Person. New York, NY: Cambridge University Press.

Singleton, Jenny L., and Elissa L. Newport. 2004. When learners surpass their models: the acquisition of American sign language from inconsistent input. Cognitive Psychology 49. 370–407.

Slobin, Dan. 2004. From ontogenesis to phylogenesis: what can child language tell us about language evolution? Biology and knowledge revisited: from neurogenesis to psychogenesis, ed. by. J. Langer, S. T. Parker and C. Milbrath, 255–86. Mahwah, NJ: Lawrence Erlbaum Associates.

Slobin, Dan I. 1973. Cognitive prerequisites for the acquisition of grammar. Studies of child language development, ed. by. C. A. Ferguson and Dan I. Slobin, 175–208. New York: Holt, Rinehard and Winston.

Smith, Neil V., Ianthi-Maria Tsimpli, and Jamal Ouhalla. 1993. Learning the impossible: the acquisition of possible and impossible languages by a polyglot savant. Lingua 91. 279–347.

St. Clair, Michelle C., Padraic Monaghan, and Michael Ramscar. 2009. Relationships between language structure and language learning: the suffixing preference and grammatical categorization. Cognitive Science 33. 1317–29.

Tesar, Bruce, and Paul Smolensky. 1998. Learnability in optimality theory. Linguistic Inquiry 29. 229–68.

Tsimpli, Ianthi-Maria, and Neil V. Smith. 1991. Second-language learning: evidence from a polyglot savant. UCL Working Papers in Linguistics 3. 149–69.

Turk-Browne, Nicholas B., Brian J. Scholl, Marvin M. Chun, and Marcia K. Johnson. 2009. Neural evidence of statistical learning: efficient detection of visual regularities without awareness. Journal of Cognitive Neuroscience 21. 1934–45.

White, Lydia. 2003. Second language acquisition and universal grammar. Cambridge: Cambridge University Press.

Wilson, Colin. 2003. Experimental investigation of phonological naturalness. Proceedings of the 22nd West Coast Conference on Formal Linguistics, ed. by. G. Garding and M. Tsujimura, 101–14. Somerville, MA: Cascadilla Press.

——. 2006. An experimental and computational study of velar palatalization. Cognitive Science 30. 945–82.