UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Functional Biases in Language Learning: Evidence from Word Order and Case-Marking Interaction

Permalink
https://escholarship.org/uc/item/0194j2z3

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Fedzechkina, Maryia
Jaeger, T. Florian
Newport, Elissa L.

Publication Date
2011

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Functional Biases in Language Learning: Evidence from Word Order and Case-Marking Interaction

Maryia Fedzechkina (mashaf@bcs.rochester.edu), T. Florian Jaeger (fjaeger@bcs.rochester.edu), Elissa L. Newport (newport@bcs.rochester.edu)
Department of Brain and Cognitive Sciences, University of Rochester
Rochester, NY 14627 USA

Abstract

Why do languages share structural properties? The functionalist tradition has argued that languages have evolved to suit the needs of their users. By what means functional pressures may come to shape grammar over time, however, remains unknown. Functional pressures could affect adults' production; or they could operate during language learning. To date, these possibilities have remained largely untested. We explore the latter possibility, that functional pressures operate during language acquisition. In an artificial language learning experiment we investigate the trade-off between word order and case. Flexible word order languages are potentially ambiguous if no case-marking (or other cues) are employed to identify the doer of the action. We explore whether language learners are biased against uncertainty in the mapping of form and meaning, showing a tendency to make word order a stronger cue to the intended meaning in no-case languages.

Keywords: Language acquisition; language universals; acquisition biases; word order; case-marking

Introduction

Despite a variety of obvious differences, languages show striking underlying commonalities at all levels of linguistic organization (Dryer, 1992; Greenberg, 1963). Such regularities have been the subject of a long-standing debate (e.g., Christiansen & Chater, 2008; Chomsky, 1965; Evans & Levinson, 2009): Are the observed phenomena shaped by linguistic-specific constraints on structure and acquisition, or do they result from cognitively and communicatively motivated constraints? The latter possibility is intriguing as it would reduce the number of linguistic-specific and hence arbitrary properties that need to be accounted for and would result in a more parsimonious explanation of cross-linguistic regularities.

It has long been hypothesized that grammatical structures that reduce the complexity associated with acquisition or processing of language tend to persist diachronically and hence cross-linguistically (e.g., Bates & MacWhinney, 1982; Bever, 1970; Hawkins, 2004; Newport, 1981; Slobin, 1973). The hypothesis that pressures on language use over time affect what grammatical properties of a language survive is supported by evidence that existing languages have properties that have relatively low average processing cost and high efficiency of information transfer. For example, dependency length is known to correlate with processing difficulty. Evidence from English and German suggests that the average dependency length in these languages is close to the theoretical minimum and far below what would be expected by chance (Gildea & Temperley, 2010). Cross-linguistic evidence also suggests that languages have properties that are beneficial for efficient information transmission (Cancho, 2006; Piantadosi, Tily, & Gibson, 2011) and that these properties reflect preferences of incremental production (Aylett & Turk, 2004; Jaeger, 2010).

Further evidence for the idea that functional pressures may constrain the space of possible language structures comes from studies showing a strong correlation between the phenomena categorically required by the grammar in some languages and speakers' gradient preferences in other languages where the grammar allows choices. One example comes from the effects of animacy on word order. The animacy of the grammatical object is an obligatory determinant of word order in the ditransitive construction in Sesotho (Morolong & Hyman, 1977) and Mayali (Evans, 1997). The same factor influences speakers' gradient preference between the two permissible orders in ditransitive alternation in English (Bresnan, Cueni, Nikitina, & Baayen, 2007).

In short, there is evidence (a) that grammatical constraints reflect gradient processing preferences and (b) that languages have properties that facilitate language use (processing and communication) compared to what would be expected if functional pressures did not affect the shape of languages over time. What remains unknown, however, is *by what means* functional pressures come to shape grammar over time and affect the transmission of language from generation to generation. The two broad logical possibilities are that functional pressures throughout life affect language production in adults, causing them to subtly change the input provided to the next generation, or that functional pressures operate during language acquisition itself, biasing learners to deviate slightly from the input they receive. Despite the long history of these claims, direct tests of these two hypotheses have been rare.

We address the latter possibility, that functional pressures affect language acquisition, using an artificial language learning paradigm in which we expose learners to experimentally designed miniature languages. Artificial language learning studies have several properties that make them ideally suited for the current purpose. Generally, adult learners acquire the statistical patterns in the input and reproduce them with roughly the same frequency as the input – 'probability-matching' (Hudson Kam & Newport, 2009). However, some studies have shown that learners preferentially acquire typologically attested patterns (Finley & Badeker, 2008; Newport & Aslin, 2004; Tily, Frank, & Jaeger, 2011). When learners deviate from the artificial language input they receive, their productions reflect typologically more frequent patterns

(Culbertson & Smolensky, 2009).

These results suggest that the artificial language learning paradigm can be used to study biases in language acquisition. However, previous work has not directly tested the prediction that these biases are functional in nature. Here we test whether a functional pressure to avoid systemic ambiguity (i.e., systemic high uncertainty about the intended meaning) causes learners to deviate from the input in such a way that it reduces the uncertainty about form to meaning mappings in the language. Such a deviation would introduce the functional bias into the learned language, providing a vehicle for language change over generations. While some degree of temporary ambiguity is expected if languages are shaped by the pressure of efficient communication (Cancho, 2006), systemic global ambiguity would hinder communication and is indeed rarely (if at all) observed cross-linguistically.

As a simple test of our prediction, we explore the trade-off between flexible word order and case-marking. We expose different groups of participants to two artificial languages that have flexible word order but differ in the presence of case-marking on the object. If language learners indeed have a bias against systemic ambiguity and case-marking is not available, word order should become a stronger cue to the intended meaning.

Method

Participants

Nineteen adult monolingual native speakers of English participated in the experiment. All were undergraduate students or recent graduates from universities in Rochester. Participants were paid daily for their participation and received a bonus upon completion of the entire study.

Description of the two Languages

Lexicon Nouns The two languages used in the experiment contained 6 human nouns: *flugit* (MOUNTIE), *glim* (CHEF), *melnawg* (REFEREE), *norg* (CONDUCTOR), *bliffen* (HUNTER), and *zub* (BANDIT). There were no lexical restrictions on the nouns: Each of them occurred once in the Subject and Object position with each of the verbs.

Verbs There were 10 verbs that denoted various transitive actions: *shen* (CHOP), *daf* (HUG), *kleidum* (HEADBUTT), *slergin* (KICK), *jentif* (KNOCK OVER), *blerfee* (PICK UP), *zamper* (POKE), *prog* (PUNCH), *geed* (ROCK), *mawg* (TAP).

Case-marker The case language had an accusative case-marker *kah* that marked the object of the action.

Grammar Both languages had flexible word order. SOV word order was dominant and occurred in 63% of the input sentences; OSV was the minority word order and occurred in 37% of the input sentences. Within these overall percentages, a system of verb-specific biases was introduced into the language to increase the complexity of the system. Based on previous work (Hudson Kam & Newport, 2009), we hypothesized that a more complex and irregular system would limit the learners' probability matching behavior and would reveal their biases in acquisition. The 10 verbs in the language were divided into four classes approximating verb subcategorization constraints found in natural languages (see Table 1).

Table 1: Verb classes used in the languages.

Number of verbs associated with dominant word order			
100%	83%	50%	0%
3	**3**	**2**	**2**

The assignment of the verbs to the verb classes was counterbalanced between the participants to prevent any accidental associations between a verb and a particular bias class.

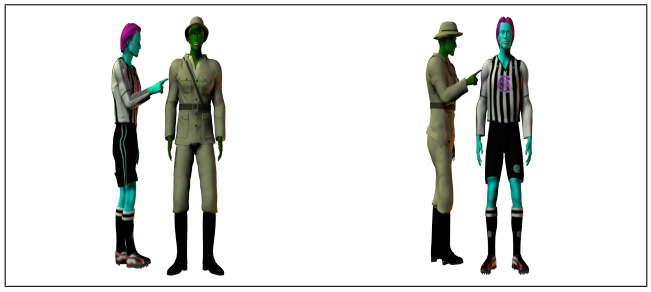

Figure 1: Still images of two sample animations from the experiment.

The two languages differed in the presence of case-marking on the object. In Language 1 no objects were case-marked; in Language 2 all objects were case-marked. Subjects were never case-marked. Thus, all sentences in the no-case input language were globally ambiguous. For example, in the no-case language, the two scenes shown in Figure 1 could be described as either *Flugit glim daf* (SOV) or *Glim flugit daf* (OSV). In the case language, however, the patient was always unambiguously identified by the accusative case-marker *kah*, as in *Flugit glim kah daf* (SOV) or *Glim kah flugit daf* (OSV).

Prediction

Given the design of our language, learners have two ways to reduce ambiguity: They can take advantage of verb-specific information and regularize word order (i.e., make variable word order more consistent and regular) within a verb class, or they can regularize word order across verb classes. As a first step, we predict that in the absence of case-marking, participants will be more likely to regularize word order in the language overall. In contrast, if case-marking is available and language users use it as a cue to the meaning of the sentence, then this should limit word order regularization and learners should acquire greater word order variation.

Procedure

Participants were shown short computer-generated videos depicting transitive actions (cf. Figure 1). Each video was

accompanied by a short aural description in the novel language. The sentences were created by concatenating prerecorded single words normalized for duration, pitch and intensity. This procedure ensured that the scene descriptions in the two languages were identical except for the case-marker *kah* and that no other cues (e.g., intonation) were available to identify the doer of the action.

Participants were trained and tested on one of the languages over four consecutive days. The same procedure was followed on days 1-3, and a slightly different procedure was adopted for the final day of training.

Days 1-3

Noun Training Participants viewed static pictures of the characters and heard their names in the novel language. The initial exposure was followed by a series of short tests where participants were asked to choose the character whose name they heard from a group of characters and to name the character shown on the screen. Feedback on performance was provided after each trial.

Sentence Presentation Participants viewed 60 short videos depicting transitive actions, one at a time, and heard an accompanying sentence describing the event in the novel language. They were instructed to repeat each sentence aloud to facilitate learning. On day 1, participants could replay the first 12 scenes as many times as they wished, to familiarize themselves with the language; no repetitions were allowed at any other time.

Noun Test Participants viewed static pictures of the characters (once for each character) and named them. They received feedback from the experimenter on their performance.

Comprehension Test In each trial, participants heard a novel sentence in the language and were shown two scenes in which the actor and patient were reversed. They were asked to choose the scene that matched the sentence. Reaction times were recorded. All scenes (60 total) contained novel combinations of familiar nouns and verbs.

Production Test Participants were shown a novel short transitive scene and were instructed to describe it in the language learned during the experiment, using the provided verb prompt. All scenes (60 total) contained novel combinations of familiar nouns and verbs, different from the ones used in the comprehension test.

Day 4 On the final day of training the comprehension and production tests included 12 additional scenes depicting previously unseen actions. These scenes were introduced to assess participants' word order choice for novel actions. To avoid effects of fast implicit learning of verb-specific biases for these novel verbs through exposure during the comprehension test, the order of the tests was reversed: Participants performed the production test first and then were presented with the comprehension test. Apart from these changes the procedure remained the same.

Results

Several behavioral measures were collected. To assess the degree of word order regularization, we scored participants' word order choices in production and comprehension. Reaction times were recorded for the forced choice comprehension test to investigate the differences in processing the two languages. Mistakes made in production were scored as well.

Accuracy of Acquisition

Both languages were acquired with a high degree of accuracy, providing evidence that the task was feasible. The total number of grammatical mistakes (i.e., incorrect use of case-marker or word order) was below 5%; lexical mistakes (such as using an incorrect noun or verb) were also rare (around 2% across all days of training).

Word Order Regularization

We assessed the effect of the presence of the case-marker on the extent of word order regularization in comprehension and production. Since both word orders are acceptable in comprehension in the no-case language, we scored the participants' deviation from verb-specific biases. All analyses presented below are conducted on the previously unseen scenes containing familiar verbs; scenes containing novel verbs presented on Day 4 were analyzed separately. The effects reported below also hold when both categories of items are collapsed.

Word Order Regularization in Comprehension We used a mixed logit model (Jaeger, 2008) to analyze the extent of word order regularization on the final day of training, given the presence of case-marking in the language and verb-specific information. We also included the maximum random effect structure for subjects and items justified by the data. There was a significant negative effect of case on word order regularization (β =-.7, z=-2.07, $p<.05$). In the case condition, participants' word order choices closely mirrored the input they received, suggesting that they relied on case-marking as a cue to the meaning of the sentence. In the absence of case-marking, learners tended to generalize the dominant word order. This tendency appeared most strongly when the majority word order was most variable in the input language (i.e., when the verb-specific constraints allowed essentially free choice of word order), as shown in Figure 2.

The pattern of differential word order regularization is also evident in participants' treatment of the two novel verbs presented on Day 4, one of which was heard exclusively with the SOV word order and the other with the OSV word order. The learners of the case language rely on case-marking in interpreting the meaning of the novel scenes and treat the novel verbs as familiar SOV-only and OSV-only verbs (97% and 11% SOV responses respectively). However, the learners of the no-case language use the overall dominant word order for both verbs around 90% of the time.

We also investigated word order preferences of individual participants (see Figure 3). Learners of the case language consistently chose the dominant word order in around 60%

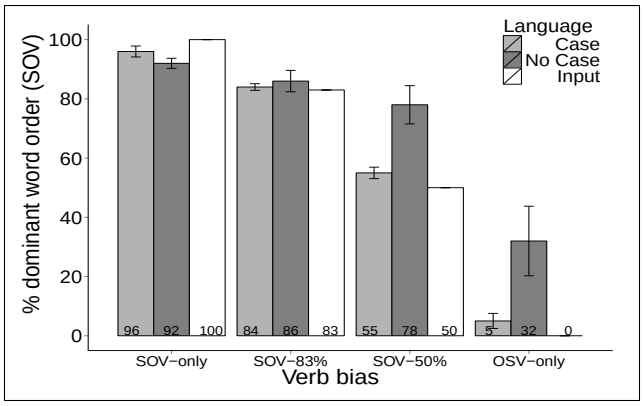

Figure 2: Word order regularization in comprehension. Numbers at the bottom of the bars represent condition means.

of their responses, which closely mirrors its input frequency (63%). This pattern of behavior suggests that the learners acquired case-marking and were using it as a cue to the meaning of the sentence. Learners of the no-case language behaved differently. First, there were more individual differences in this condition: Only 2 participants probability-matched, while the majority of participants generalized the dominant word order. We also observed 3 'super-regularizers' in this condition, who had a strong preference for the dominant word order and used it in about 90% of their responses.

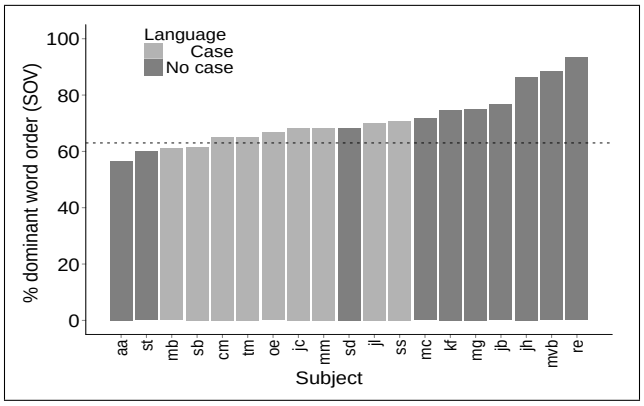

Figure 3: Individual word order preferences in comprehension.

Word Order Regularization in Production The between-language mixed model analyses of Day 4 performance revealed a marginal effect of case on word order regularization (β = -1.1, z=-1.74, p<.09). No learners exactly reproduced the input frequencies of the two word orders[1]. The learners of the no-case language consistently generalized the majority word order. There was no clear pattern for the case language: Learners sometimes overgeneralized (e.g., for equi-biased and OSV-only verbs, see Figure 4), but they sometimes undergeneralized as well (see SOV-only and SOV-biased verbs in Figure 4).

The learners of both types of languages behaved similarly on newly encountered verbs: They all used the overall dominant word order more frequently with these actions.

The word order preferences of individual subjects differed depending on the condition as well. Most learners of the no-case language consistently used the dominant word order more frequently than it was present in the input. We also observed two 'super-regularizers' (same as the 'super-regularizers' discussed above, cf. Figure 3), who essentially fixed word order in their productions. No such word order regularizers were present among the learners of the case language. There was more spread in word order consistency in the case condition: While one half of the participants tended to generalize the dominant word order, the other half tended to under-generalize. These behaviors suggest that the presence of case-marking limits word order regularization and allows word order flexibility in the language.

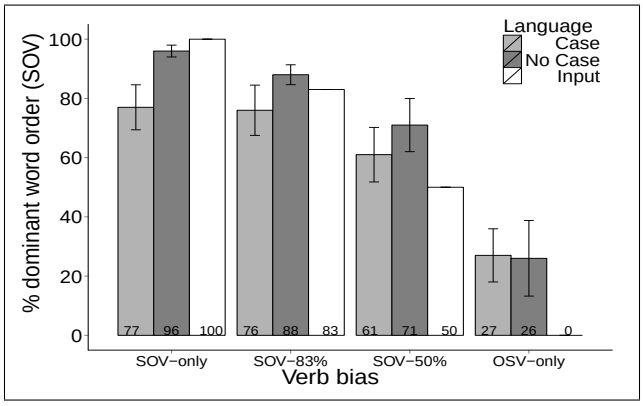

Figure 4: Word order regularization in production. Numbers at the bottom of the bars represent condition means.

Processing Advantage of Case In principle, grammatical functions can be signaled using a single cue, but most languages employ several redundant cues, including word order and case-marking. Can several cues to meaning be advantageous for language processing? To address this question we used a linear mixed-effects model to predict the adjusted[2] response time given the presence of case, chosen word order, and bias-consistency.

[1] All participants achieved 100% accuracy on the vocabulary test on Day 4. Occasionally, they mispronounced the name of one of the two referents (in less than 0.3% of the trials). Given their low frequency and participants' perfect performance on the vocabulary test, the trials containing pronunciation errors were not excluded from the analysis.

[2] Since in our design, all sentences in the case language had the same duration as the sentences in the no-case language + the duration of the case-maker *kah*, we subtracted the duration of the case-marker from the response times for all learners of the case language to make the performance across conditions comparable. There was no effect of case on raw reaction times since our participants generally responded after hearing the stimulus till the end.

There was no effect of case on reaction times. This null effect is unlikely to be due to insufficient power or lack of sensitivity of the timed forced choice task we employed in the comprehension tests. We found that learners exhibited processing effects for the artificial languages that resemble those found in natural languages (e.g. Trueswell, Tanenhaus, & Kello, 1993): Participants were significantly faster to respond if the word order they chose was congruent with verb-specific biases.

Further analysis revealed that the observed fast reaction times for the no-case language were primarily driven by the same three 'super-regularizers' who essentially fixed word order (see Figure 3). These learners seem to have a different hypothesis about the language: They responded to stimuli significantly faster, presumably because they added fixed word order to the language and thus had little or no uncertainty about the form-meaning mapping. Once these participants were removed from the analysis, the effect of case-marking reached significance (β = -.1, t=-1.53, p=.055).

Discussion

We find that case-marking and word order interact during acquisition. If case-marking is present in the language, it serves as a strong cue to the meaning of the sentence and limits word order regularization. In the case language, the case-marker was obligatory in our setup and always unambiguously identified the doer and patient of the action, rendering word order a less important cue to acquire and allowing more freedom in word order variation.

However, if grammatical functions are signaled exclusively by word order, as in our no-case language, flexible word order does not serve as a sufficiently informative cue to meaning. Put differently, the uncertainty about form-meaning mappings (or, more formally, the entropy of the distribution of possible meanings given a sentence) is much higher in the no-case input language than in the case input language. If learners are biased to reduce excessive uncertainty of this type, they can do so by regularizing word order - either in verb-specific ways or, as tested here, over the entire language. Indeed, in the absence of case-marking, learners tend to gradually give up word order flexibility and generalize the dominant word order in both comprehension and (to a lesser extent) in production. This effect is partially driven by a group of learners of the no-case language who almost completely fix word order. One possibility to be addressed in future work is whether such 'super-regularizers' are more likely in the no-case language.

The observed tendency to regularize word order more in the no-case language is not due to arbitrary mistakes since the participants in the experiment *did* learn verb biases: Word order preferences in the production of both languages are clearly conditioned on verb class (cf. Figure 2 and 4). Our results also cannot be attributed to the fact that the participants' native language is a language with a relatively fixed word order (English), since this would not explain why differential word order regularization is observed depending on the condition.

Our participants systematically generalize towards SOV word order, but we do not observe any learners who consistently use OSV word order in the majority of their productions. Several factors can explain this pattern of behavior. First, SOV word order is the dominant word order in our language, and learners are regularizing the most frequent patterns in the input. Second, since our participants are native English speakers and a subject before object (SO) word order is typical in English, this might be a partial reflection of the native language bias (cf. Tily et al., 2011). However, SO is a cross-linguistically preferred word order as well, so the observed tendency to generalize the majority word order could indicate a cognitive bias favoring SO ordering. Ongoing follow-up experiments are teasing apart these factors by investigating the learning of languages with the inverse word order patterns (e.g. OSV 63%, SOV 37%).

Our findings suggest that at least some learning biases are functional in nature. The participants in our experiments showed a bias against systemic ambiguity in the language and more readily learned the structures that reduced uncertainty in the form-meaning mapping. These acquisition biases are reflected in typologically recurrent grammatical patterns and might account for structural similarities found in natural languages. Our participants do not induce cross-linguistically unattested structures, but they acquire and generalize towards typologically frequent patterns in the language they learn: They tend to regularize word order in the absence of case-marking, but maintain a more flexible word order when case is available. This behavior mirrors natural phenomena, such as the loss of free word order from Old to Modern English. Old English was a highly inflected free word order language that gradually lost case-marking and acquired relatively fixed word order during its transition to Modern English.

Having several redundant cues to signal grammatical relationships is potentially advantageous for processing. We find, in particular, that case-marking provides a processing advantage for languages with flexible word order, presumably because it offers an earlier disambiguation point: Stimuli in the case language are processed significantly faster. However, this might not be true for languages with fixed word order: The participants who fixed the word order in the no-case language were just as quick to respond as the learners of the case language. This points to a hypothesis articulated most clearly by Hawkins (2004) that ease of processing is a factor in typological variation.

Our experiment has explored one specific prediction of the functionalist account, but there are several outstanding issues which we will investigate in follow-up work. First, we have only investigated the most simple strategy of ambiguity reduction in the absence of case-marking – overall word oder regularization in the language. It is, however, possible that language learners employ more sophisticated ambiguity reduction techniques and take advantage of the system of verb-

specific biases. We address this possibility in ongoing work. Second, in the no-case language used in this study, all sentences were potentially globally ambiguous. However, in natural languages global ambiguities are rare – at least once context is taken into consideration. In an ongoing experiment we investigate how learners reduce uncertainty when it is limited to certain types of sentences, as is more common cross-linguistically. We have so far addressed the question of whether language learners reduce systemic ambiguity during language acquisition, but this is only one of many functional biases that have been proposed in the literature.

Conclusions

We have investigated the role of functional biases in language acquisition, focusing on the trade-off between flexible word order and case. We find that word order and case-marking interact during acquisition, with learners relying on one of these cues to reduce ambiguity. In the absence of case-marking, learners tend to generalize the dominant word order in both comprehension and in production, suggesting a bias to reduce systemic ambiguity in the language. When case is present, it limits word order regularization since learners use case to resolve ambiguity and word order becomes a less important cue to meaning.

Acknowledgements

This work was supported in part by NSF grant BCS-0845059, an Alfred P. Sloan Fellowship and a Wilmot Award to TFJ and by NIH grant DC00167 to ELN. We thank Colleen Dolan and Keryn Bernstein for their help collecting subject data.

References

Aylett, M. P., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Lang. Speech*, *47*(1), 31-56.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art.* New York: Cambridge University Press.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and development of language* (p. 279-362). New York: Wiley.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Boume, I. Kraemer, & J. Zwarts (Eds.), *Cognitive foundations of interpretation.* Amsterdam: Royal Netherlands Academy of Sciences.

Cancho, R. Ferrer-i. (2006). When language breaks into pieces. A conflict between communication through isolated signals and language conflict between communication through isolated signals and language. *Biosystems*, 242-253.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behav Brain Sci*, *31*, 489–509.

Culbertson, J., & Smolensky, P. (2009). Testing Greenberg's Universal 18 using an artificial language learning paradigm. In *Proceedings of NELS 40.*

Dryer. (1992). The Greenbergian word order correlations. *Language*, *68*, 81-183.

Evans, N. (1997). Role or cast? Noun incorporation and complex predicates in Mayali. In J. Bresnan & P. Sells (Eds.), *Complex predicates* (p. 397-430). Stanford.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Beh. and Brain Sciences*, *32*, 429-492.

Finley, S., & Badeker, W. (2008). Substantive biases for vowel harmony. In J. Bishop (Ed.), *WCCFL 27* (p. 168-176).

Gildea, D., & Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Science*, *32*(2), 286-310.

Greenberg, J. (1963). *Universals of language*. Cambridge, MA: MIT Press.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hudson Kam, C. L., & Newport, E. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*, 30-66.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology Psych.*, *61*(1), 23-62.

Morolong, M., & Hyman, L. M. (1977). Animacy, objects and clitics in Sesotho. *Studies in African Ling.*, *8*, 199-218.

Newport, E. (1981). Constraints on structure: Evidence from American Sign Language and language learning. In W. A. Collins (Ed.), *Aspects of the development of competence. Minnesota symposia on child psychology* (Vol. 14). Hillsdale, NJ: Erlbaum.

Newport, E., & Aslin, R. N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127-162.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proc. Natl Acad. Sci. USA*, *108*(9), 3526-3529.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development.* New York: Holt, Rinehart and Winston.

Tily, H., Frank, M. C., & Jaeger, T. F. (2011). The learnability of constructed languages reflects typological patterns. In *Proceedings of CogSci 2011.*

Trueswell, J., Tanenhaus, M., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *J. of Exp. Psych.: Learning, Memory, and Cognition*, *19*, 528-553.