

Frequency shapes syntactic structure*

HOLGER DIESEL

There is abundant evidence from a wide range of studies that frequency of occurrence influences language use and language development (cf. Diessel, 2007). In light of this evidence, the hypotheses proposed by Ambridge, Kidd, Rowland, and Theakston (this issue) should not be controversial. As they point out, “all but the most hardened classist[s]” agree that frequency of occurrence has some effect on language acquisition. However, there is no consensus among linguists and psychologists as to how the various frequency effects can be explained. In particular, it is unclear how frequency affects the language user’s grammatical (or syntactic) knowledge. As I see it, there are two general strategies to explain the influence of frequency on the development of syntactic structure.

In the generative (and structuralist) approach, syntactic categories are independent of the language users’ linguistic experience. They are defined prior to syntactic analysis as primitive concepts that are universally applicable to all languages (Pinker & Jackendoff, 2005). Building on this view of syntax, some researchers have argued that the generative model of grammar should be augmented by a probabilistic component in order to account for frequency effects in sentence processing and grammar learning. Specifically, these researchers suggest that syntactic categories (and syntactic rules) should be assigned ‘probability values’ that reflect the language users’ experience with particular categories (and rules) and that determine their ease of activation in comprehension and production and their development in L1 acquisition. There is a large body of research demonstrating that statistically enriched grammars of this sort, often referred to as ‘stochastic grammars’, can predict the likelihood of garden-path effects in sentence processing and the occurrence of certain types of mistake in children’s speech (cf. Jurafsky, 1996; Chater & Manning, 2006; Bod, 2009).

Challenging the generative view of syntax, usage-based linguists have argued that syntactic structure is emergent from the language users’ experience with language. In the usage-based approach, grammar is a dynamic system of fluid categories and flexible constraints that are shaped

[*] Address for correspondence: Professor Holger Diessel (PhD), University of Jena, Department of English, Ernst-Abbe-Platz 8, 07743 Jena. tel.: +49-3641-944530; e-mail: holger.diessel@uni-jena.de

by general cognitive mechanisms such as categorization, priming, and automatization (or entrenchment), which are crucially driven by frequency of occurrence (cf. Bybee, 2010; see also Diessel, 2011a).

The emergence of syntactic structures can be studied in two different time frames: in language acquisition and diachronic language change (see Diessel 2011b, 2012a, for discussion). Research on grammaticalization indicates that bound morphemes, grammatical categories, and constructions are derived from frequent strings of linguistic elements. For instance, a large number of studies have shown that inflectional affixes such as tense, aspect, and mood are derived from free function morphemes that have fused with adjacent content words (cf. Bybee, Perkins, & Pagliuca, 1994), and that (free) grammatical markers such as prepositions, auxiliaries, and determiners are based on frequent lexical expressions, notably on nouns and verbs (cf. Hopper & Traugott, 2003), or spatial deictics (cf. Diessel, 2012b), that are routinely used in particular structural positions. Moreover, it has been argued that syntactic categories for grammatical relations such as subject and object are emergent from the language users' (unconscious) analysis of verb–argument constructions with overlapping structural and semantic features (cf. Croft, 2001), and that phrase structure categories such as NP and PP are based on recurrent sequences of linguistic expressions that have developed into 'processing units' through 'repetition' or 'chunking' (Bybee, 2010). Generalizing across these findings, some researchers have argued that all aspects of syntactic structure are derivative. In the usage-based approach, syntax is an 'emergent phenomenon' (Hopper, 1987) grounded in the language user's experience with frequent strings of linguistic expressions (cf. Bybee & Hopper, 2001).

These strings provide the 'input' for L1 acquisition. A fundamental problem young children face is to segment the linguistic sequences they encounter in the ambient language and to identify the units of adult speech (cf. Jusczyk, 1997). Frequency plays an important role in this process. There is good evidence from a number of studies that young children are very sensitive to distributional regularities. They recognize transitional probabilities in recurrent strings of speech sounds, which helps them to 'break into linguistic structure' and to 'unpack' the processing units they experience in adult language (e.g., Saffran, Newport, & Aslin, 1996). Moreover, a number of studies have argued that distributional regularities play an important role in the acquisition of syntactic categories (and schemas). Evidence for this hypothesis comes from corpus-based research on children's errors in spontaneous speech (e.g., Rowland, 2007) and from experimental research on statistical grammar learning (Aslin and Newport, 2012, for a review). For instance, Gómez and Gerken (1999) conducted a series of learning experiments in which they taught

twelve-month-old infants an artificial language consisting of monosyllabic nonce words (e.g., *VOT*, *PEL*, *YIC*) that appeared in a set of particular sentence types (defined by word order and the number of words they include). After training, the nonce words were replaced by a novel set of nonce words and the infants were tested under two conditions. In condition 1, they were exposed to the same sentence types as during training, and in condition 2 they listened to a different set of sentence types. Although the infants had not heard the words of the test sentences before, they recognized the different structural patterns in the two conditions, indicating that children as young as one year of age are able to abstract beyond strings of concrete linguistic expressions and to form abstract categories (and constructional schemas). Similar results have been obtained in several other studies with both infants and adult speakers who were tested under various conditions and with a wide range of different stimuli (cf. Aslin & Newport, 2012).

Taken together, this research suggests that the acquisition of syntax does not need the support of pre-established (or innate) categories and rules. On this account, frequency does not only determine the activation value of syntactic categories (and rules), but also affects the way syntactic concepts are formed and organized in mental grammar.

REFERENCES

- Aslin, R. N., & Newport, E. L. (2012). Statistical learning: from acquiring specific items to forming general rules. *Current Directions in Psychological Science*, **21**, 170–177.
- Bod, R. (2009). From exemplar to grammar: a probabilistic analogy-based model of language learning. *Cognitive Science*, **33**, 752–793.
- Bybee, J. (2010). *Language, usage, and cognition*. Cambridge: Cambridge University Press.
- Bybee, J., & Hopper, P. (eds) (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
- Bybee, J., Perkins, R. D., & Pagliuca, W. (1994). *The evolution of grammar: tense, aspect and modality in the languages of the world*. Chicago: Chicago University Press.
- Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Science*, **10**, 335–344.
- Croft, W. (2001). *Radical Construction Grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, **25**, 108–127.
- Diessel, H. (2011a). Review article of ‘Language, usage and cognition’ by Joan Bybee. *Language*, **87**, 830–844.
- Diessel, H. (2011b). Grammaticalization and language acquisition. In B. Heine & H. Narrog (eds), *Handbook of grammaticalization* (pp. 130–141). Oxford: Oxford University Press.
- Diessel, H. (2012a). Language change and language acquisition. In A. Bergs & L. Brinton (eds), *Historical linguistics of English: an international handbook*, Vol. 2 (pp. 1599–1613). Berlin: Mouton de Gruyter.
- Diessel, H. (2012b). Buehler’s two-field theory of pointing and naming and the deictic origins of grammatical morphemes. In T. Breban, L. Brems, K. Davidse, & T. Mortelmans (eds), *New perspectives on grammaticalization: theoretical understanding and empirical description* (pp. 35–48). Amsterdam: John Benjamins.

- Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific abstract knowledge. *Cognition*, **70**, 109–135.
- Hopper, P. (1987). Emergent grammar. In *Proceedings of the thirteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 139–157). Berkeley: Berkeley Linguistic Society.
- Hopper, P., & Closs Traugott, E. (2003). *Grammaticalization* [2nd ed.]. Cambridge: Cambridge University Press.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, **20**, 137–194.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, **95**, 201–236.
- Rowland, C. F. (2007). Explaining errors in children's questions: auxiliary DO and modal auxiliaries. *Cognition*, **104**, 106–134.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language*, **35**, 606–621.