Linear grammar as a possible stepping-stone in the evolution of language

Ray Jackendoff[1] · Eva Wittenberg[2]

Abstract We suggest that one way to approach the evolution of language is through reverse engineering: asking what components of the language faculty could have been useful in the absence of the full complement of components. We explore the possibilities offered by *linear grammar*, a form of language that lacks syntax and morphology altogether, and that structures its utterances through a direct mapping between semantics and phonology. A language with a linear grammar would have no syntactic categories or syntactic phrases, and therefore no syntactic recursion. It would also have no functional categories such as tense, agreement, and case inflection, and no derivational morphology. Such a language would still be capable of conveying certain semantic relations through word order—for instance by stipulating that agents should precede patients. However, many other semantic relations would have to be based on pragmatics and discourse context. We find evidence of linear grammar in a wide range of linguistic phenomena: pidgins, stages of late second language acquisition, home signs, village sign languages, language comprehension (even in fully syntactic languages), aphasia, and specific language impairment. We also find a full-blown language, Riau Indonesian, whose grammar is arguably close to a pure linear grammar. In addition, when subjects are asked to convey information through nonlinguistic gesture, their gestures make use of semantically based principles of linear ordering. Finally, some pockets of English grammar, notably compounds, can be characterized in terms of linear grammar. We conclude that linear grammar is a plausible evolutionary precursor of modern fully syntactic grammar, one that is still active in the human mind.

The question of language evolution is often taken to be: How did language evolve in our species? We prefer to frame the question differently. Languages are learned by children. The course of their learning looks much like other learning, such as learning of motor and social skills. This has led many to think that humans just came to be good at learning complex skills. So the question of language evolution is better stated as: How did humans evolve so as to be able to *learn* language (as well as all those other skills)? That is, what makes the human brain "language-ready," in the sense of Arbib (2012)? And is this at all different from being generally "complex-skill-ready"?

To answer these questions, a prior question has to be addressed: When children learn a language, what do they end up knowing? At the very least, knowing a language is being able to map between patterns of sounds (or gestures) and meanings/concepts/thoughts (whatever *they* are), as in Fig. 1. Speaking involves going upward in Fig. 1, from concepts to the motor system; comprehending language involves going downward, from auditory input to concepts. Figure 1 also includes a bit of the bigger picture of the mind: Conceptual structure is also linked to perception and action, so one can talk about what one sees and about actions that one plans to carry out.

If we remove phonetic patterns and their connections from Fig. 1, leaving only the part outside the ellipse, we get an organism that can perceive and have thoughts, and that can

✉ Ray Jackendoff
  ray.jackendoff@tufts.edu

1   Center for Cognitive Studies, Tufts University, Medford, MA 02155, USA

2   Departments of Linguistics and Psychology and Center for Research in Language, University of California–San Diego, La Jolla, CA, USA
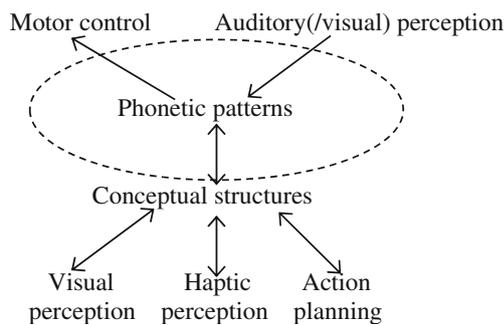
Fig. 1 Within the ellipse is the minimal structure for vocal/gestural communication: a mapping between concepts (or thoughts) and the production and perception of phonetic patterns that signal those concepts. Concepts, in turn, can be related to visual and haptic perception and to action planning

Fig. 2 Within the ellipse are the parts added to Fig. 1 to form the modern language capacity. The mapping between concepts and utterances is mediated by phonological, syntactic, and morphological structure

act on the basis of those thoughts, but cannot communicate those thoughts in any rich fashion. This is a plausible sketch of primate cognition, which is fairly sophisticated—but lacks language (Cheney & Seyfarth, 1990). So the evolution of language had to involve at least a new ability to map concepts to sounds and gestures and to use these communicatively.

However, language actually consists of a good deal more than this: First, there is phonological structure—the systematized organization of sounds (or, in sign languages, gestures). Second is morphology—the internal structure of words, such that the word *procedural* can be seen as built from *proceed* plus *-ure* to form *procedure*, plus *-al* to form *procedural*: [[[*proceed*] [*-ure*]] [*-al*]]. Third is syntax, the organization of words into phrases and sentences. Syntax determines canonical word order, so that, for instance, *The girl kissed the boy* signals who did the kissing and who got kissed. Syntax also allows descriptions of characters and events to be elaborated into phrases:

> [the girl in [[the blue hat] and [red sneakers]] tried to kiss [the boy [that she loved]].

Here, it is still understood that the girl was the potential kisser, even though the word *girl* is nowhere near *kiss*, and it is understood that she loved the boy, even though *girl* is nowhere near *love*.

We conclude that language in modern humans involves the network of cognitive organization shown in Fig. 2. Hence, what has evolved in humans is the ability to learn to turn thoughts into sounds by structuring them into words and phrases, and the ability to learn to pull thoughts out of the sounds that other people make, by making use of these intermediate structures.

There is no direct evidence for when our ancestors started being able to do this. Nor is there evidence for what they had to say: there are no fossil vowels or fossil concepts. The usual way to test evolutionary scenarios, namely comparison to other species, is not too helpful.
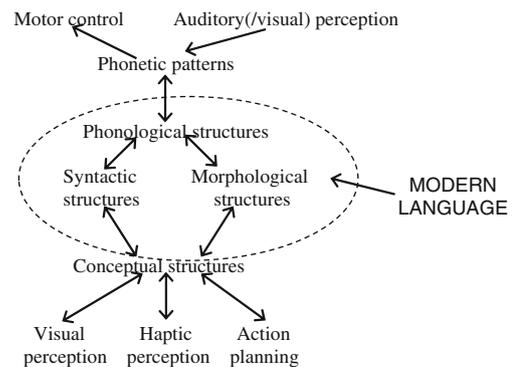
Modern apes do not learn very much in the way of human language (Hedwig, Mundry, Robbins, & Boesch, 2015; Kako, 1999; Scott-Phillips, 2016; Seidenberg & Petitto, 1978; Zuberbühler, 2015). And they certainly do not invent language spontaneously, as deaf children do in the absence of sign language input (Goldin-Meadow, 2003, 2016). So apes' communicative abilities are not very useful in closing the gap between them and humans. Moreover, although being "language-ready" requires an ability for vocal imitation, shared with many other species (though apparently not other primates) through convergent evolution (Feher, 2016; Fitch, 2010; Vernes, 2016), there seems to be little evidence in other species for the ability to acquire human grammars of the sort in Fig. 2.

Another way to form plausible hypotheses about evolution is through reverse engineering: asking what components could have been useful in the absence of others. Consider the eye. A primitive retina would be useful for vision even in the absence of muscles that focus the lens and move the eyeballs—though it would obviously be more limited. On the other hand, without a retina, the muscles couldn't help one see. So it makes sense that something like the retina must have evolved before the lens and the muscles. (Whatever muscles might have existed before, ready to be exapted for focusing, they could not have served that function before there was a lens to deform; see Fernald, 2000, for the course of evolution of the eye.)

We propose something similar for language. A primitive system for communicating thoughts via sound or gestures, along the lines of Fig. 1, is useful without phonology, morphology, or syntax. The latter components can improve an existing communication system, but they are useless on their own. So if the components of language evolved in some order, it makes sense that the connection between phonetics and meaning came first, followed by these further refinements. Our hypothesis, therefore, is that this is the kind of system

that some ancestors of modern humans were able to learn, a stepping-stone on the way from ape minds to human. We can't *prove* that this is the way language evolved, but we will show that simpler systems of this sort exist in the languages of today, and we will offer some ideas about how these systems work.

The basic idea comes from Derek Bickerton (1990). He proposes a form of language that he calls "protolanguage," which surfaces in many different circumstances, and which is a relic of early stages of human or hominid language. We suggest that something like this form of language consists of the subset of the full language system shown in Fig. 1, omitting morphology and syntax (and possibly even phonology, leaving just phonetics). We will call such a system a *linear grammar*.

A system with a linear grammar would have words— that is, stored pairings between a phonological form and a piece of conceptual structure. The linear order of words in an utterance would be specified by phonology, not by syntax. The individual words would map to meanings, but beyond linear order, there would be no further structure—no syntactic phrases that combine words, such as *the girl in the blue hat and red sneakers*, and no structure inside words, such as in the word *procedural*. In such a language, word order could still play a role. One could say *girl kiss boy* and mean the girl did the kissing and the boy got kissed, not vice versa. But the source for this meaning would not be in syntactic structure (i.e., the subject precedes the verb and the verb precedes the object), because, by hypothesis, a linear grammar lacks syntactic categories such as nouns, verbs, subjects, and objects. What it still does have is a set of rules for mapping from *semantic* notions to linear order in phonology. One such rule might specify that the word denoting an agent—here *girl*—precedes the word denoting the action—*kiss*. That is, this kind of language would map directly between linear order in phonology and semantic roles, without the intervention of syntactic categories.

Since a linear grammar is by hypothesis linear, it could not have subordinate clauses such as the relative clause in *the boy that she loved*—that is, syntactic recursion. The thought could still be expressed, but most likely as two successive sentences, for instance *Girl love boy. Girl kiss boy.*

A linear grammar would also lack morphology, so, for instance, it would not have tenses and agreement on verbs or case on nouns, nor could it have passives, which are marked by inflections on the verb. We would also expect a linear grammar not to have functional items like definite and indefinite articles. So it would allow one to say *girl kiss boy*, but not *the girl kissed a boy* or *a girl kisses the boy*. The speaker would have to leave it up to the context to indicate when this kissing took place.

Our hypothesis, then, is that linear grammar could serve on its own as a useful system of communication.

In contrast, syntax and morphology, without linearly ordered words to organize, would be of no use on their own.[1] In other words, linear grammar is like the retina, and syntax and morphology are more like the muscles that move the eyes—powerful refinements of the visual system, but only useful when a more basic system is in place.

What makes this hypothesis of interest is that a variety of modern-day linguistic systems have the flavor of linear grammar. The first is *pidgins*, the early stages of contact languages. Pidgins are often described (e.g., Bickerton, 1981; Givón, 2009) as having no subordination, little or no morphology, no grammatical words like *the*, and unstable word order governed primarily by semantic principles like agent before action. If the context permits, the characters in the action can be left unexpressed. For instance, if the context had already brought the boy to attention, the speaker might just say *girl kiss*, which in English would require a pronoun—*The girl kissed him*. From the perspective of linear grammar, we can ask: Is there any evidence that pidgins have parts of speech like nouns and verbs, independently from the semantic distinction between individuals and actions? Is there any evidence for syntactic phrases, beyond semantic cohesion? The descriptions in the literature suggest that there is not, though a closer examination is called for. If such evidence does not surface, pidgin grammars are a good candidate for real-world examples of our hypothesized linear grammar.

Later on, of course, contact languages add many features of more complex languages, such as conventionalized word order, grammatical categories, and syntactic subordination. Such languages are called *creoles*. We see the transition from a pidgin to a creole not as going from nonlanguage to language, but rather as adding some syntactic and morphological principles that were not present in the pidgin—as it were, some muscles for the lens.

For a second case, involving *late second language acquisition*, Wolfgang Klein and Clive Perdue (1997; see also Dimroth, 2013) did a multilanguage longitudinal study of immigrants learning various second languages all over Europe. They found that all speakers achieved a stage of semiproficiency that they called the "Basic Variety." Many speakers went on to improve on the Basic Variety, but others did not. At this stage, there is no inflectional morphology or sentential subordination, and known characters are freely omitted. Instead, there are simple, semantically based principles of word order including, for instance, agent before action.

---

[1] Thus we take issue with the position (Chomsky, 2016; Hauser, Chomsky, & Fitch, 2002) that recursion was the founding innovation that led to language. Recursion cannot occur unless there is something to recurse on—in the case of language, words. At the same time, the use of recursion in language could well have been preceded by recursion in other mental domains such as conceptual structure, social cognition, and spatial cognition. See Pinker and Jackendoff (2005) and Jackendoff (2011).

From our standpoint, then, the Basic Variety looks like another kind of linear grammar.

A third case is *home signs*, the languages invented by deaf children who have no exposure to a signed language. Susan Goldin-Meadow has shown (2003, 2016) that they have at most rudimentary morphology; they also freely omit known characters. In our analysis, home signs only have a semantic distinction of object versus action, not a syntactic distinction of noun versus verb. Word order is probabilistic and is based, if anything, on semantic roles. Homesigners do produce some sentences with multiple verbs (or action words), which Goldin-Meadow describes as "embedding." We think these are rudimentary serial verb or serial action-word constructions, without embedding, sort of like the compound verb in English expressions such as *He came running*. So this looks like a linear grammar with possibly a bit of morphology. (Full disclosure: Goldin-Meadow does not entirely agree with our assessment.)

Another case is *village sign languages*, which develop in isolated communities with a significant occurrence of hereditary deafness. The best known of these is Al-Sayyid Bedouin Sign Language in the Negev desert of Israel (ABSL; Sandler, Meir, Padden, & Aronoff, 2005); another is Central Taurus Sign Language (CTSL), spoken in two remote villages in the mountains of Turkey (Ergin et al., 2014). CTSL has some minimal morphology, mostly confined to younger speakers. But there is little or no evidence for syntactic structure. In sentences involving one character, the word order is normally agent + action, and two-character sentences are normally (optional) agent + patient + action: *girl ball roll*. But if a sentence involves two animate characters, so that semantics alone cannot resolve the potential ambiguity, word order is not very reliable. For instance, *girl boy hit* is a bit vague about whether the girl hit the boy or vice versa, requiring a huge reliance on pragmatics, common knowledge, and context. In fact, there is a strong tendency to mention only one animate character per predicate, so speakers sometimes clarify by saying things like *Girl hit, boy get-hit*. So CTSL looks like a linear grammar, augmented by a small amount of morphology. Similar results have been obtained in ABSL and the earlier stages of Nicaraguan Sign Language (Kegl, Senghas, & Coppola, 1999).

These less complex systems are not confined to emerging languages; they also play a role in *language processing*. Townsend and Bever (2001) discuss what they call semantically based "interpretive strategies" that influence language comprehension. In particular, hearers tend to rely in part on semantically based principles of word order such as agent precedes action, which is why (in our account) speakers have more difficulty with constructions such as reversible passives and object relatives, in which the agent does not precede the action. Similarly, Ferreira and Patson (2007) discuss "good-enough" parsing, in which listeners apparently rely on linear order and semantic plausibility rather than syntactic structure. It is well known that we see similar though amplified symptoms in language comprehension by agrammatic aphasics (Gibson, Sandberg, Fedorenko, Bergen, & Kiran, 2015). Finally, Van der Lely and Pinker (2014) argue that a particular population of children with specific language impairment behave as though they are processing language through something like a linear grammar. The literature frequently describes these so-called "strategies" or "heuristics" as something separate from language. But they are still mappings between phonology and meaning—just simpler ones.

We conjecture, then, that the language processor makes use of both syntactic grammar and simpler linear grammar (see also Martin, 2016, for how cues on different levels might interact in processing). When the two kinds of rules produce conflicting analyses, interpretation is slower and less stable—even when syntax wins out, as in reversible passives. And when syntactic rules break down under conditions of stress or disability, the linear grammar is still in operation.

We have also encountered a full-blown language whose grammar appears to be close to a linear grammar: *Riau Indonesian*, a vernacular with several million speakers, described by Gil (2005, 2009). Gil argues that this language has no syntactic parts of speech and no inflectional morphology such as tense, plural, or agreement. Known characters in the discourse are freely omitted. Messages that English expresses with syntactic subordination are expressed in Riau paratactically, with utterances like *girl love, kiss boy*. The word order is quite free, but agents tend to precede actions, and actions tend to precede patients. The freedom of this language can be illustrated by Gil's example (in which the last two cases require considerable support from the context):

ayam makan, "chicken eat" can mean

a. "{a/the} chicken(s) {is/are eating/ate/will eat} {something/it}"
b. "{something/I/you/he/she/they} {is/are eating/ate/will eat} {a/the} chicken"
c. "{a/the} chicken that {is/was} eating"
d. "{a/the} chicken that {is/was} being eaten"
e. "someone is eating with/for the chicken"
f. "where/when the chicken is eating"

This collection of symptoms again looks very much like a linear grammar. Hence, this is a language virtually all of whose grammar is syntactically simple in our sense.

Another kind of linear grammar—that is, a system that relies on the linear order of the semantic roles being expressed to form conceptual relations—surfaces when people are asked to express actions or situations in a *nonlinguistic task*, such as in gesture or act-out tasks (Futrell et al., 2015; Goldin-Meadow, So, Özyürek, & Mylander, 2008; Hall, Mayberry,

& Ferreira, 2013). Overall, there is a vast preference to gesture, or act out, the agent first (e.g., *girl*), and then the patient (e.g., *boy*). The action is usually expressed last (*kiss*), but when there is a potential ambiguity, people like to avoid it by expressing the action in the middle, between the agent and patient. Crucially, the ordering preferences in these tasks are remarkably stable, independently of the ordering preferences in test subjects' native language. That seems to indicate that the capacity to map certain semantic notions to certain linear orders is at least partly independent from language itself.

As a final case, traces of something like linear grammar lurk within the grammar of English! Perhaps the most prominent case is *compounding*, in which two words are stuck together to form a composite word (see Jackendoff, 2010a, and references therein). The constituents may be any part of speech: not just pairs of nouns, as in *kitchen table*, but also *longbow*, *undercurrent*, *castoff*, *overkill*, *speakeasy*, and *hearsay*. The meaning of the composite usually includes the meanings of the constituents, but the relation between them is determined pragmatically. Consider examples like these, and how their variety resembles that of the Riau examples above:

collar size = "size of collar"
dog catcher = "person who catches dogs"
nail file = "something with which one files nails"
beef stew = "stew made out of beef"
bike helmet = "helmet that one wears while riding a bike"
bird brain = "person whose brain is similar to that of a bird"
bike girl = "girl who left her bike in the hall" (an off-the-cuff example, understood only in context)

The main constraint is that the second noun usually determines what kind of object the compound denotes; for instance, *beef stew* is a kind of stew, whereas *stew beef* is a kind of beef. But this can be determined solely from the linear order of the nouns and needs no further syntax.

To sum up, remarkably similar grammatical symptoms turn up in a wide range of different scenarios. This suggests to us that linear grammar is a robust phenomenon, entrenched in modern human brains. It provides a scaffolding on top of which fully syntactic languages can develop, either in an individual, as in the case of the Basic Variety, or in a community, as in the case of pidgins and emerging sign languages. Furthermore, it provides a sort of safety net when syntactic grammar is damaged, as we have seen with aphasia and specific language impairment. We have also seen that it is possible to express a great deal even without syntax, for example in Riau Indonesian—though having syntax gives speakers more sophisticated tools for expressing themselves.

Let us now return to the original question about the evolution of the human ability to learn language. We suggested that this question can be approached through reverse engineering, by asking what kind of system could have preceded the modern human language faculty—what kind of system human ancestors could have been capable of learning. We propose that linear grammar is a good candidate, although, as we said at the outset, it is unclear how one could prove it. Nor does this provide any hints about when the hominid line achieved either linear grammar or syntactic grammar. Perhaps someday there will be better evidence from genetics. For now, we are happy to regard it as an intriguing hypothesis.

References

Arbib, M. (2012). *How the brain got language: The mirror neuron hypothesis.* Oxford, UK: Oxford University Press.

Bickerton, D. (1981). *Roots of language.* Ann Arbor, MI: Karoma.

Bickerton, D. (1990). *Language and species.* Chicago, IL: University of Chicago Press.

Cheney, D., & Seyfarth, R. (1990). *How monkeys see the world.* Chicago, IL: University of Chicago Press.

Chomsky, N. (2016). The language capacity: architecture and evolution. doi:10.3758/s13423-016-1078-6

Dimroth, C. (2013). Learner varieties. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics.* Oxford, UK: Blackwell. doi:10.1002/9781405198431.wbeal0673

Ergin, R., Caselli, N., Ilkbasaran, D., Meir, I., Cohen-Goldberg, A., & Jackendoff, R. (2014, July). Argument structure in a newly documented village sign language. Paper presented at the 6th Conference of the International Society for Gestural Studies, San Diego, California.

Feher, O. (2016). Atypical birdsong and artificial languages provide insights into how communication systems are shaped by learning, use and transmission. doi:10.3758/s13423-016-1107-5

Fernald, R. D. (2000). Evolution of eyes. *Current Opinion in Neurobiology, 10,* 444–450.

Ferreira, F., & Patson, N. (2007). The good enough approach to language comprehension. *Language and Linguistics Compass, 1,* 71–83.

Fitch, W. T. (2010). *The evolution of language.* Cambridge, UK: Cambridge University Press.

Futrell, R., Hickey, T., Lee, A., Lim, E., Luchkina, E., & Gibson, E. (2015). Cross-linguistic gestures reflect typological universals: A subject-initial, verb-final bias in speakers of diverse languages. *Cognition, 136,* 215–221.

Gibson, E., Sandberg, C., Fedorenko, E., Bergen, L., & Kiran, S. (2015). A rational inference approach to aphasic language comprehension. *Aphasiology.* Advance online publication. doi:10.1080/02687038.2015.1111994

Gil, D. (2005). Word order without syntactic categories: How Riau Indonesian does it. In A. Carnie, H. Harley, & S. A. Dooley (Eds.), *Verb first: On the syntax of verb-initial languages* (pp. 243–263). Amsterdam: John Benjamins.

Gil, D. (2009). How much grammar does it take to sail a boat? In G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language complexity as an*

evolving variable (pp. 19–33). Oxford, UK: Oxford University Press.

Givón, T. (2009). *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam, The Netherlands: Benjamins.

Goldin-Meadow, S. (2003). *The resilience of language*. Hove, UK: Psychology Press.

Goldin-Meadow, S. (2016). What the hands can tell us about language emergence. doi:10.3758/s13423-016-1074-x

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences, 105,* 9163–9168.

Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition, 129,* 1–17.

Hauser, M., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298,* 1569–1579.

Hedwig, D., Mundry, R., Robbins, M. M., & Boesch, C. (2015). Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Animal Cognition, 18,* 423–435.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences, 3,* 272–279.

Jackendoff, R. (2002). *Foundations of language*. Oxford, UK: Oxford University Press.

Jackendoff, R. (2010a). *Meaning and the lexicon*. Oxford, UK: Oxford University Press.

Jackendoff, R. (2010b). Your theory of language evolution depends on your theory of language. In R. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language: Biolinguistic perspectives* (pp. 63–72). Cambridge, UK: Cambridge University Press.

Jackendoff, R. (2011). What is the human language faculty? Two views. *Language, 87,* 586–624.

Jackendoff, R., & Wittenberg, E. (2014). What you can say without syntax: A hierarchy of grammatical complexity. In F. Newmeyer & L. Preston (Eds.), *Measuring grammatical complexity* (pp. 65–82). Oxford, UK: Oxford University Press.

Kako, E. (1999). Elements of syntax in the systems of three language-trained animals. *Animal Learning & Behavior, 27,* 1–14. doi:10.3758/BF03199424

Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.), *Language creation and language change* (pp. 179–238). Cambridge, MA: MIT Press.

Klein, W., & Perdue, C. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research, 13,* 301–347.

Martin, A. E. (2016). Language processing as cue integration: Grounding the psychology of language in perception and neurophysiology. *Frontiers in Psychology, 7,* 120. doi:10.3389/fpsyg.2016.00120

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition, 95,* 201–236.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar in a new sign language. *Proceedings of the National Academy of Sciences, 102,* 2661–2665.

Scott-Phillips, T. C. (2016). Meaning in great ape communication: Summarising the debate. *Animal Cognition, 19,* 233–238.

Seidenberg, M., & Petitto, L. (1978). Signing behavior in apes: A critical review. *Cognition, 17,* 177–215.

Townsend, D., & Bever, T. (2001). *Sentence comprehension*. Cambridge, MA: MIT Press.

Van der Lely, H. K. J., & Pinker, S. (2014). The biological basis of language: Insight from developmental grammatical impairments. *Trends in Cognitive Sciences, 18,* 586–595.

Vernes, S. (2016). What bats have to say about speech and language. doi:10.3758/s13423-016-1060-3

Zuberbühler, K. (2015). Linguistic capacity of non-human animals. *Wiley Interdisciplinary Reviews: Cognitive Science, 6,* 313–321.