

Empirical approaches to the study of language evolution

W. Tecumseh Fitch¹

Published online: 1 February 2017
© Psychonomic Society, Inc. 2017

Abstract The study of language evolution, and human cognitive evolution more generally, has often been ridiculed as unscientific, but in fact it differs little from many other disciplines that investigate past events, such as geology or cosmology. Well-crafted models of language evolution make numerous testable hypotheses, and if the principles of strong inference (simultaneous testing of multiple plausible hypotheses) are adopted, there is an increasing amount of relevant data allowing empirical evaluation of such models. The articles in this special issue provide a concise overview of current models of language evolution, emphasizing the testable predictions that they make, along with overviews of the many sources of data available to test them (emphasizing comparative, neural, and genetic data). The key challenge facing the study of language evolution is not a lack of data, but rather a weak commitment to hypothesis-testing approaches and strong inference, exacerbated by the broad and highly interdisciplinary nature of the relevant data. This introduction offers an overview of the field, and a summary of what needed to evolve to provide our species with language-ready brains. It then briefly discusses different contemporary models of language evolution, followed by an overview of different sources of data to test these models. I conclude with my own multistage model of how different components of language could have evolved.

Keywords Language evolution · Speech · Syntax · Semantics · Comparative approach · Strong inference · Neurolinguistics · Paleo-DNA · Protolanguage

✉ W. Tecumseh Fitch
tecumseh.fitch@univie.ac.at

¹ Department of Cognitive Biology, University of Vienna, Vienna, Austria

Introduction

Interest in language evolution has surged in the past two decades (Berwick, Friederici, Chomsky, & Bolhuis, 2013; Carstairs-McCarthy, 1999; Christiansen & Kirby, 2003; Dunbar, 1996; Fitch, 2010; Hauser, Chomsky, & Fitch, 2002; Pinker & Bloom, 1990; Pinker & Jackendoff, 2005; Számadó & Szathmáry, 2006; Tallerman & Gibson, 2011; Yang, 2013). This surge has been accompanied by fundamental progress in our understanding of this difficult multidisciplinary problem. While there have also been naysayers who deny this progress, or the very possibility of progress in understanding cognitive evolution (Hauser et al., 2014; Lewontin, 1998), the purpose of this review and the special issue of which it is part is to demonstrate that such pessimism is unjustified. The reasons are simple: We have whole new classes of data that provide new insights into key issues and problems (e.g., paleo-DNA). The field also profits from a productive new inter-disciplinary community that is constructively engaging with these problems (centered around the biennial EvoLang conference series), and a flood of more traditional sorts of data (e.g., regarding animal cognition and communication, genetics, and neuroscience). This combination has led to increasingly sophisticated models of language evolution that make multiple testable predictions, and improved evaluation criteria for assessing such models. The result, I will argue here, is an ongoing transition of scientific research on language evolution from one dominated by speculation and pet hypotheses to “normal” science, marked by attempts to empirically evaluate multiple plausible hypotheses.

Despite the nonexistence of time machines, and the oft-mentioned fact that language does not fossilize, there is no reason in principle that models of language evolution need be any more speculative or untestable than those of other scientific disciplines that deal with singular chains of events

in the distant past. Cosmologists interested in the Big Bang, or geologists studying continental drift and plate tectonics, are quite familiar with such difficulties. They proceed in their study of the past by examining present-day phenomena empirically, and using these data to evaluate explicit contesting models of what might have happened when, and why. Assuming that certain principles of physics, chemistry, or geology remain unchanged, this enables practitioners in these fields to triangulate on adequate models and further refine them by generating further testable predictions. This process has led, for example, to plate tectonics going from a speculative hypothesis, often ridiculed, to something universally accepted in modern geology (Gohau, 1990). There is little preventing the same general scientific process from being effective in the study of language evolution. We have a relatively clear endpoint of the process in the present, and can reconstruct the starting point (our last common ancestor with chimpanzees) in detail using the comparative method with existing species. Making the reasonable assumption that many of the biological principles underlying genetic, neural, cognitive, and behavioral traits have remained constant during the intervening 6 million years, and fed by the fragmentary but important data of the fossil record (and now “fossil” DNA), we enjoy essentially the same preconditions to progress as 20th-century geologists evaluating plate tectonics. This leads to the real possibility of fundamental advances in understanding language evolution in the coming years, building on the progress of the last few decades.

The present article will attempt to concisely summarize this progress and to provide a snapshot of language evolution research as it stood in late 2016. It will also serve as an overview of the current special issue. The article has five main parts. First, I provide a theoretical overview of the conceptual playing field, stressing the importance of a multicomponent approach to language, of strong inference over multiple plausible hypotheses, and of a comparative approach using behavioral, neural, and genetic data from a broad range of living species to inform our understanding of the mechanisms underlying language. These are general points that apply to any problem in cognitive evolution. I then turn in “What evolved?” to language specifically, focusing on three derived components of linguistic cognition that are not shared with chimpanzees, our nearest living cousins (vocal control, hierarchical syntax, and complex semantics/pragmatics). Part 3 “Models of language evolution” gives a brief overview of some of the debates and models currently dominating the conceptual landscape.

Part 4, “Empirical data,” provides a more comprehensive overview of the data that are relevant to testing models of language evolution. I subdivide these into four broad classes: comparative biological data, fossil/archaeological data, neural data, and genetic data. The sheer abundance and diversity of these data is problematic, because few if any scientists are fully competent to evaluate them all. This means that many

“facts” that are accepted and repeated frequently in the secondary literature do not stand up to serious scrutiny by the standards of their specific fields, so that the outsider may be left with the feeling that everything is contested (and therefore nothing can be taken seriously). While skepticism is certainly necessary when evaluating all data, I try to separate the wheat from the chaff and focus on results that seem most solid. Of these four classes of data, the most exciting are genetic data, and particularly paleo-DNA from extinct hominins, which offer the tantalizing hope of explicitly testing and rejecting predictions of current models of language evolution.

Finally the fifth “Synthesis” part of the article attempts to do something I have resisted previously: It offers a comprehensive, testable model of language evolution, from our last common ancestor with chimpanzees to modern humans. I offer this model as an example of a model that is both consistent with existing data and that makes predictions about data yet to be gathered, and compare it in detail with other contemporary models.

Theoretical framework and general overview: Studying cognitive evolution

I will first outline some general principles for studying cognitive evolution, including the need to subdivide any complex trait into component parts, the need to adopt a broad comparative approach to understand the evolution of these components, and the need to adopt a “strong inference,” hypothesis-testing framework to evaluate and test such models.

The multicomponent approach

The first and most obvious theoretical move in understanding the biology and evolution of a complex cognitive ability is to fully acknowledge the multiplicity of mechanisms that underlie it. This is no different in language than in, for example, vision (Hubel, 1988; Marr, 1982), music (Peretz & Coltheart, 2003), aesthetics (Leder, Belke, Oeberst, & Augustin, 2004), or social cognition (Fitch et al., 2010). While this perspective—the *multicomponent approach*—seems obvious in the case of vision, it has been oddly absent from many discussions of language evolution. Perhaps because of the unique nature of human language, language seems to intuitively invite “single cause” thinking, where some particular trait is singled out as “the key” to language and by extension to human uniqueness. Depending on the scholar, this favored trait may be speech (Lieberman, 1984, 2006), syntax (Berwick & Chomsky, 2016; Chomsky, 2010), or shared intentionality (Tomasello, Carpenter, Call, Behne, & Moll, 2005), but in each case one factor is emphasized and other relevant factors are downplayed. I believe that this widespread tendency toward monolithic thinking about language is one of the root causes of dissent in the study of language evolution, since

once a particular factor has been chosen, other factors (and other scholars' thinking) appear to be irrelevant.

The antidote to this persistent problem is to acknowledge that language is made up of multiple separable (but interacting) components, and undertake to analyze them. This set of components—the so-called faculty of language in a broad sense—is large, but divides naturally into two categories: those that are shared (sometimes widely) with other species and those that are recent acquisitions of the human lineage since our evolutionary divergence from chimpanzees. Obviously, since chimpanzees lack language, this subset bears a disproportionately important explanatory burden in understanding language evolution. But a trait may be novel in humans in this sense but *not* be unique to our species. The human capacity for complex vocal learning is a case in point: though basically absent in chimpanzees or other primates (see below), it is shared with a diverse if scattered collection of other bird and mammal species (Brainard & Fitch, 2014; Fitch & Jarvis, 2013; Janik & Slater, 1997). This derived subset of language mechanisms is not synonymous with the “faculty of language in a narrow sense”—those traits that are unique to humans and unique, within humans, to language itself (Fitch et al., 2005; Hauser et al., 2002). I will refer to these core traits *derived components of language* (relative to our last common ancestor with chimpanzees), or “DCLs.” By our current understanding of ape cognition and communication, the set of DCLs contains *at least* three separable components (see [The derived components](#) section): complex vocal learning, hierarchical syntax, and complex semantics/pragmatics (cf. the “three Ss” of speech, syntax and semantics in Fitch, 2010).

The broad comparative approach

Another key component of the framework advocated here is the use of a broad comparative method in studying the evolution of cognitive traits. In this section I will illustrate the comparative approach using the evolution of vision, which clearly demonstrates its value in a domain less controversial than language evolution.

The multiplicity of DCLs implies that each biological mechanism may have different genetic and neural substrates, and often different evolutionary histories. For example, in vision we can clearly separate color vision from the perception of form or movement, both in terms of the retinal and cortical mechanisms involved (Hubel, 1988; Livingstone, 2002) and their evolutionary history and timing (Jacobs & Rowe, 2004). A very rich source of data in understanding this evolutionary history is provided by a broad comparative approach, studying vision in a wide variety of species to develop and test hypotheses about the evolution of particular abilities.

For example, humans and closely related primates (e.g., chimpanzees and macaques) have tri-chromatic color vision

(involving three different cone photo-pigments), in contrast to most other mammals, which have only dichromatic vision. One might thus infer that color vision is an “advanced trait” found only in relatively sophisticated species. But the broad comparative dataset, including many species from insects to fish to birds to New World monkeys, clearly demonstrates that this inference would be spectacularly incorrect. Indeed, it turns out that mammals are the outliers and that, in vertebrates, trichromacy, or even tetrachromacy was the primitive initial state, and still typifies fish, lizards, and birds. During the Mesozoic, due to a primarily nocturnal existence, this rich color vision was lost in the ancestor of modern mammals, only to be regained by some primates in the last 10–20 million years (Jacobs, 1993; Jacobs & Deegan, 1999). What's more, the repeated evolution of complex color vision provides important clues to the function of this trait (Kremers, Silveira, Yamada, & Lee, 2000; Vorobyev, 2004), and sometimes even reveals “deep homology” where the same mutations in the same genes has occurred independently in clades as widely separate as primates and butterflies (Frentiu et al., 2007).

When comparing traits among species, biologists typically recognize two different classes of shared traits: homologies and analogies (technically, analogy is just a subtype of a grab-bag class termed “homoplasy,” which is essentially everything that is not a homology; Lankester, 1870; Sanderson & Hufford, 1996). *Homologies* are derived from a trait present in the common ancestor of the species in question; thus, homologies provide evidence for inferences about the existence of that trait in that common ancestor. Homologies are crucial for interpreting cognitive, neural, and genetic history, since such data typically leave no fossil traces. When some cognitive mechanism, neural circuit, or genetic sequence is observed in multiple close relatives, we can confidently infer its presence in their last common ancestor. *Analogies*, in contrast, are convergently evolved traits—they were not present in the common ancestor but arose independently in multiple lineages. This independence means that only analogies represent statistically independent data points for testing evolutionary hypotheses (Harvey & Pagel, 1991). We can thus legitimately use analogies both to test mechanistic hypotheses and functional hypotheses concerning adaptation and natural selection.

In addition to these two traditional categories, the genetic revolution in the last decades has revealed a third type of similarity: *deep homology* (Shubin et al., 1997, 2009). Deep homology exists when two traits have evolved independently at a phenotypic level (i.e., the trait in question was not present in the common ancestor), but where the genetic and developmental mechanisms underlying the trait are nonetheless shared and homologous. The now classic case concerns complex eyes in insects and vertebrates, which were not present in the common ancestor (and are phenotypically analogies) but nonetheless rely upon many of the same regulatory genes (e.g., Pax-6) for their development (genetic homologies).

Another nice example of deep homology is the importance of FoxP2 in vocal learning in humans and birds (Fitch, 2009b, 2011a; Scharff & Petri, 2011)—the same gene plays an important role in regulating vocal learning in these clades, but the ability itself was not present in the common ancestor of birds and mammals. Deep homology is important because there is increasing evidence that it is common and relevant in the evolution of cognition (Arendt, 2005; Parker et al., 2013; Salas, Broglio, & Rodríguez, 2003).

As these examples make clear, our understanding of cognitive evolution would be seriously incomplete if we focused exclusively on comparisons of humans with other primates (a narrow comparative approach). It is unfortunate that such limited comparisons were the primary source of comparative data concerning language evolution for most of the 20th century, despite a few dissenting voices (e.g., Nottebohm, 1972, 1975). Fortunately, the genomic revolution has led to a widespread recognition of the fundamental conservatism of gene function in very disparate species (e.g., sponges, flies, and humans; Coutinho, Fonseca, Mansurea, & Borojevic, 2003) and there is a rising awareness that distant relatives like birds may have as much, or more, to tell us about the biology and evolution of human traits as comparisons with other primates (Emery & Clayton, 2004).

Strong inference and multiple hypotheses

The final general point concerns the need for simultaneous testing of multiple, plausible hypotheses. There is a long tradition in empirical research, stemming from null-hypothesis testing statistical approaches, that pits a single plausible hypothesis against a null statistical hypothesis that often has little a priori scientific plausibility. Null hypotheses include variants on “the data have no pattern,” “the data are randomly distributed,” “there is no relationship between two variables,” or “two categories do not differ.” Although statisticians have long criticized this approach (Anderson, 2008; Cohen, 1994), and newer approaches, including model-comparison and Bayesian methods, are rising in popularity, old habits die hard, and this approach often leads to a trivial “test” of some favorite “pet” hypothesis against an a priori implausible alternative. This “one hypothesis” approach has, unfortunately, been typical in writings on language evolution, which often simply ignore previous work, or stoop to disparage alternative hypotheses with derogatory nicknames (e.g., the “bow-wow” or “ding-dong” theories, in a tradition initiated by Max Müller’s 19th-century attacks on Darwin (Müller, 1861).

Fortunately, an alternative approach has long been available: the method of empirically testing the predictions of multiple scientifically plausible hypotheses simultaneously. Especially when an attempt is made to have the hypothesis set tested be exhaustive, this method has been dubbed “strong inference” (Chamberlin, 1890), and when thoughtfully

implemented offers a much more efficient path to resolution of scientific debates and apparently discrepant data. It is precisely this approach, and the steady accrual of consistent data, that transformed “continental drift” from a crazy old idea to accepted scientific theory by 1970 (Gohau, 1990).

Language evolution offers an ideal arena for strong inference because decades of speculation have led to many plausible hypotheses about how specific DCLs evolved, and in some cases detailed arguments about the order in which they appeared in human phylogeny. Similarly abundant models exist when we consider the cognitive and neural bases of language and their relationship to traits found in other species. This plethora of existing models (each of which at least one scholar deemed plausible enough to publish) means that we have quite a full roster of explanations and predictions concerning incoming data. Many of these models can be falsified by new data, especially when their predictions contrast with those from other hypotheses. And, as I will document in detail below, there is plenty of relevant data, and more coming in every day. The main problem for this approach is not with data *or* hypotheses, but sociological: There is no well-developed tradition of scholars in language evolution taking each other’s models seriously. Instead the tradition has been one where others’ models are ridiculed or (worse) ignored. Many of the articles in this issue illustrate that scientists are increasingly taking into account each other’s models, and a wide variety of data from many disciplines, when proffering their own hypotheses. And that represents real progress.

The role of simplicity

The role of simplicity and parsimony in considering alternative hypotheses presents a challenge. Obviously, the basic principle of parsimony and Occam’s razor (do not create unnecessarily complex hypotheses when simpler ones suffice) play a role in all scientific discourse. But in biology, and evolution in particular, parsimony certainly never has the final say in adjudicating among hypotheses: Because evolution has no foresight, it often tinkers together solutions that are far from simple or elegant (Jacob, 1977). At best, parsimony provides a default preference for simpler hypotheses, in the absence of further information, but this preference should be temporary (Fitzpatrick, 2008). Simplicity considerations can and should be easily trumped by actual data concerning the biological reality (cf. de Boer, 2016).

A deeper question, considered by several authors in this issue, is “what counts as simple?” (Chomsky, 2016; Johnson, 2016; Perfors, 2017). The core idea of the Minimalist paradigm in linguistics is that linguistic syntax can be reduced to a very simple but powerful core operation, Merge, which serves to combine lexical elements (Chomsky, 1995); this conception opened the door to inquiry into the evolution of such an operator (Chomsky, 2010, 2016). But

simplicity at the computational level of description does not necessarily translate into implementational simplicity at the neural level (or vice versa), and simplicity of neural implementation is arguably the level most relevant to evolutionary discussions of rapid adaptive change (Johnson, 2016; Perfors, 2017). These issues are key open issues in contemporary discussions of language evolution (Berwick, 1998; Berwick & Chomsky, 2016) and not likely to be resolved until we know more about how genes, brains, and language are inter-related (Ramus, 2006; Fisher, 2016). For now, it seems prudent not to rely overly heavily on parsimony in adjudicating between competing hypotheses about language evolution.

What evolved? Shared and derived components of language

The shared foundations

Language is a complex faculty that allows us to encode, elaborate and communicate our thoughts and experiences via words combined into hierarchical structures called sentences. Words are learned, and thus shared by communities, but differ across languages, and their form is mostly arbitrary (Saussure, 1916) despite a nontrivial amount of onomatopoeia and sound symbolism (Blasi, Wichmann, Hammerström, Stadler, & Christiansen, 2016; Sapir, 1929). Humans are born with a capacity to acquire the words and grammars of their local language(s)—an “instinct to learn language” (Fitch, 2011b; Marler, 1991). It is this capacity—sometimes termed “the language faculty”—whose evolutionary history or phylogeny we seek to explain when studying language evolution (rather than historical change or “glossogeny”; cf. below and Hurford, 1990).

The human faculty of language in this broad sense includes all of the various cognitive and physiological mechanisms that support the human capacity to acquire language; most of these mechanisms are shared with other species (the FLB of Hauser et al., 2002). Thus, despite the fact that language *in toto* is unique to our species, most components underlying it are shared, sometimes very broadly and sometimes only with a few other species. Because I have already discussed these many shared capacities in detail in other places (Fitch, 2005, 2010), I will only mention the highlights here.

There are several areas of very significant (nearly complete) overlap, and these form the backdrop for any discussion of the few remaining differences (the “shared foundations” of the language system illustrated in Fig. 1). These include basic physiological mechanisms involved in perception and motor control, along with various cognitive mechanisms involving learning, problem-solving, and memory.

Starting with physiology, *human auditory capacities* are shared with most other vertebrates, including fish. The essential architecture and function of the auditory system, from the

middle ear through cochlea and up to cortex via multiple brainstem waypoints, is shared with other mammals, and there is little evidence of any fundamental differences in human hearing from most other familiar mammals except that adult human hearing occupies a relatively low frequency range (roughly 20 Hz to 15 kHz), and many species go well beyond our upper limit of 20 kHz. Although much has been made recently of differences in the shape of the chimpanzee and human audiogram (Martinez et al., 2013), in an attempt to use the middle ear bones of extinct hominins to reconstruct the evolution of speech perception, I am skeptical of the relevance of these data for two reasons. The first is that the primary determinant of hearing range and acuity is the cochlea, not the outer or middle ears (Ruggero & Temchin, 2002). The second is that the supposed differences between human and chimpanzee audiograms are based on a tiny sample of chimpanzees, which showed considerable differences between individuals (Elder, 1934; Kojima, 1990) and a distinctive so-called W-shaped audiogram (Kojima, 1990) that was not seen in the earlier study. Captive housed animals often suffer noise-induced hearing loss, caused by exposure to loud vocalizations in a reverberant concrete environment, which may explain this “divot” in adult sensitivity. More data will thus be necessary to conclude that this is a real phenomenon in the chimpanzee species as a whole, rather than an isolated problem of one individual.

More important, behavioral tests indicate that both chimpanzees and other species (e.g., dogs) have excellent central abilities to process human speech at the phonetic and lexical levels (Savage-Rumbaugh et al., 1993; Kaminski, Call, & Fischer, 2004), and chimpanzees can even understand bizarre signals such as sinewave speech or noise-vocoded speech (Heimbauer, Beran, & Owren, 2011). A host of other auditory phenomena such as categorical perception or vowel “prototype magnets” have also been documented in other species from birds to chinchillas (Kluender, Diehl, & Killeen, 1987; Kuhl & Miller, 1978). Thus, there is little reason today to accept the old assertion that “speech is special” from a perceptual point of view (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), or to think that new mechanisms of speech perception needed to evolve in the hominid lineage to support language evolution.

Turning to signal output, there are no clear differences between humans and other apes in terms either of vision or manual/ facial control that would prevent an ape from learning signed language if the proper semantic and syntactic “software” were in place—one reason for the popularity of “gestural protolanguage” hypotheses (see below and Corballis, 2002; Hewes, 1973). Thus, discussions of possible differences have focused on *speech output*. The essential functioning of the human lungs, larynx, and tongue is again shared very broadly with other mammals, from bats to elephants, both in terms of anatomy and regarding the physics and physiology of vocal production (Fitch, 2000b; Herbst et al., 2012; Taylor &

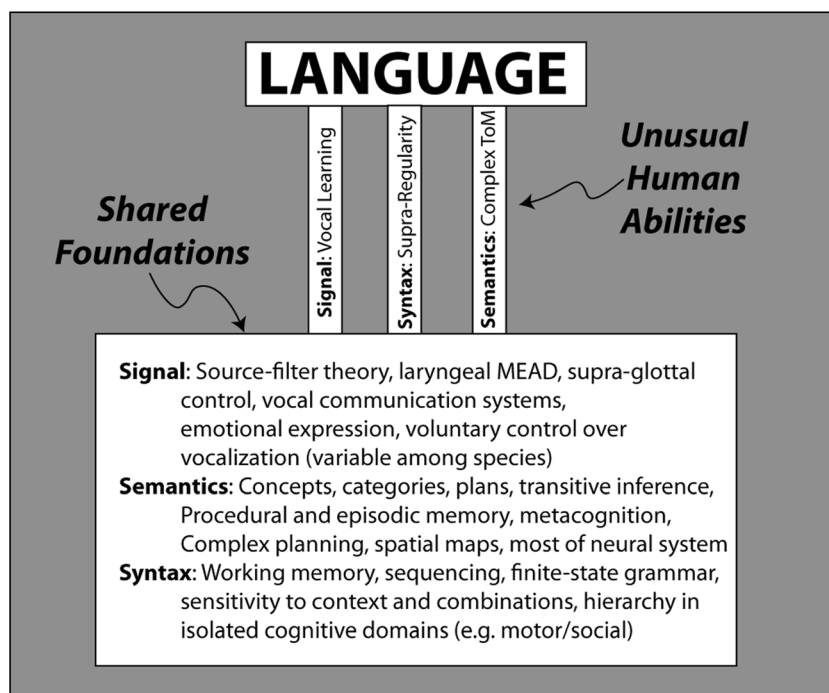


Fig. 1 An architectural metaphor indicating the many different mechanisms underlying the human capacity to acquire language, illustrating that the vast majority are shared with other species (the

“Shared Foundations,” and only a few represent derived characters of our species (relative to our common ancestor with chimpanzees: “Unusual Human Abilities”)

Reby, 2010). The human tongue is similar in anatomy to that in other apes (Takemoto, 2008), and we now know that a mild descent of the larynx and hyoid bone occurs in chimpanzees (Nishimura, Mikami, Suzuki, & Matsuzawa, 2006). Given that all mammalian species examined have a flexible capacity to lower the larynx during vocalization (Fitch, 2000a), and many species possess a descended larynx much more extreme than that in humans (Fitch & Reby, 2001; Weissengruber, Forstenpointner, Peters, Kübber-Heiss, & Fitch, 2002), it is clear that the much-discussed role of the descended larynx in human speech has been greatly overestimated in its importance (Fitch et al., 2016). This means that most attempts to estimate the onset of speech abilities based on fossils are dead ends (reviewed in Fitch, 2009a), with the possible exception of thoracic canal size, indicating an increase in breath control (MacLamon & Hewitt, 1999). Like hearing, the anatomy of the primate vocal tract was essentially “speech ready” whenever the neural control and cognitive capabilities evolved, as concluded long ago by both Darwin and Negus (Darwin, 1871; Negus, 1938). Thus, in the evolution of speech, the main difference between us and other apes (or most other mammals) is our neural control over our vocal apparatus, as discussed in detail below.

Regarding *basic cognition*, the list of cognitive differences between humans and other animals has grown steadily smaller (Vauclair, 1996; Tomasello & Call, 1997; Shettleworth, 2009). Those most relevant to human language are reviewed in (Hurford, 2007; Fitch, 2010). Shared systems include different forms of memory, from working memory to episodic memory

(Emery, 2006; Inoue & Matsuzawa, 2007), a capacity for approximate number (Dehaene, 1992; Dehaene et al., 2015), and a host of more basic capacities including categorization, transitive inference, navigation, and planning. All of these capacities constitute a *basic “cognitive toolkit”* that we share with most other mammalian and avian species. More specialized capacities including *tool use* are also shared, not only with other primates but with a variety of avian and mammalian species (McGrew, 2004; Pruetz & Bertolani, 2007; Tebbich, Taborsky, Fessl, & Blomqvist, 2001; van Lawick-Goodall & van Lawick-Goodall, 1967; Weir, Chappell, & Kacelnik, 2004; Whiten, Horner, & de Waal, 2005). Finally, one of the classic supposed differences between humans and other apes was our possession of a “*theory of mind*”—a capacity to represent the beliefs and desires of other individuals (Povinelli & Eddy, 1996; Premack & Woodruff, 1978). Such a capacity, at a basic level, is now well-documented in chimpanzees and ravens (Bugnyar, Reber, & Buckner, 2016; Hare & Tomasello, 2004). Thus, despite an undoubted increase in intelligence during human evolution, the specific, empirically demonstrated cognitive differences that underlie this quantitative difference have grown increasingly scarce and seem to focus on syntax and semantics/pragmatics, discussed below (cf. Herrmann, Call, Hernández-Lloreda, Hare, & Tomasello, 2007).

Again, the short overview above is by no means an exhaustive list of language-related traits that we share with other species. But I hope it clarifies the essential point that most of the mechanisms involved in language acquisition evolved long before humans

split with chimpanzees. These mechanisms were already in place when the few specific changes required for language evolved. As discussed below, all of these characteristics were preadaptations to language: mechanisms that are *used in* language, but did not specifically evolve *for* language (Fitch, 2012).

Derived components of language

We will now turn, in more detail, to those few changes that can be considered to be key innovations: derived components of language (DCLs) that were both required for language and evolved in a hominin context. These are derived relative to our last common ancestor with chimpanzees and bonobos, the $LCA_{\text{chimpanzee}}$ or LCAc hereafter.

Speech and vocal production learning

As discussed in detail elsewhere (Fitch, 2009a, 2010), there is little evidence that any major changes in the vocal apparatus itself were required for our ancestors to gain the capacity for speech. Nonetheless, the belief remains distressingly widespread that major changes in vocal anatomy, and particularly the descended human larynx, were required prerequisites for speech. For example, many authors (Barney, Martelli, Serrurier, & Steele, 2012; Crystal, 2003; Lieberman & Blumstein, 1988; Raphael, Borden, & Harris, 2007; Yule, 2006) state that differences in vocal anatomy are responsible for the complete failure of chimpanzees raised in a human home to acquire even a few clearly understandable words (Hayes, 1951). As one example, “early experiments to teach chimpanzees to communicate with their voices failed because of the insufficiencies of the animals’ vocal organs” (Crystal, 2003, p. 402). This idea is refuted by virtually all modern data on mammalian vocal production (most recently Fitch et al., 2016), so its persistence in modern discussions is perplexing.

The origin of this idea apparently goes back to early studies by Philip Lieberman and his colleagues suggesting limitations in the range of vowels that could be produced by a macaque or chimpanzee (Lieberman et al., 1972; Lieberman, Klatt, & Wilson, 1969). These studies never claimed that the total lack of vocal control and vocal learning in these species could be explained only by vocal anatomy, and Lieberman’s most recent publications suggest that “a nonhuman SVT [supra-laryngeal vocal tract] can produce all vowels excepting quantal vowels” (Lieberman, 2012, p. 613). Thus, even the strongest adherents suggest that these changes represent adaptive “tweaks” to increase the intelligibility of speech, rather than being prerequisites for any form of speech. Other commentators find even this idea dubious (Boë, Heim, Honda, & Maeda, 2002; de Boer, 2010; Fitch, 2010; Nishimura et al., 2006), in part because nonlinguistic functions for changes such as a descended larynx are now known from nonhuman species (Fitch & Reby, 2001).

Thus, from a modern perspective, it seems clear that changes in complex vocal *control* rather than vocal anatomy were the key innovations required for the evolution of speech. This is fortunately an area where the comparative database is rich, because a level of vocal control enabling the production of novel sounds from the environment has evolved multiple times in birds and mammals (Fitch, 2000b; Janik & Slater, 1997; Jarvis, 2007; Reichmuth & Casey, 2014). The current list of gifted vocal learners includes some species of songbirds, hummingbirds, parrots, cetaceans, seals, bats, and elephants. The conclusion that vocal anatomy is not central to speech is supported by the well-documented cases of speech imitation by nonhuman species such as birds and elephants (Klatt & Stefanski, 1974; Pepperberg, 2005; Stoeger et al., 2012), whose vocal anatomy differs greatly but still can produce easily understandable words and phrases. In the recently documented case of an elephant imitating multiple Korean words, the elephant appears to use movements of its trunk, inserted into the oral cavity, to vary its formant frequencies (Stoeger et al., 2012). In such cases it seems obvious that it is the species’ capacity to control its vocal tract, not vocal anatomy, that is key. In fact, the only case of speech imitation by an animal where the anatomy *is* largely the same as in humans is the harbor seal Hoover, who apparently used his normal mammalian larynx and vocal tract to produce readily intelligible English words and phrases (Ralls, Fiorelli, & Gish, 1985).

Despite this relatively broad set of species or clades where complex vocal production learning is known, humans remain the only known primate capable of vocal production learning. It is important to note that this is *not* to say that other primates have no vocal control, or that their vocalizations are produced robot-like given the proper stimulus—such a caricature does not apply to any bird or mammal. Virtually all species tested are able to bring one or more of their innate vocalizations under control (e.g., to produce a call upon a cue), including many primates (reviewed in Adret, 1992; Fitch & Zuberbühler, 2013). For example, Larson, Sutton, Taylor, and Lindeman (1973) could train four macaques to produce various species-typical calls (barks, grunts, and coos) upon command. When reward was made contingent on producing longer calls, all the monkeys succeeded in doing so by switching to the longer coo calls. This is evidence of “call usage learning” but *not* vocal production learning. Recent data from chimpanzees, where food calling behavior of newcomers to a zoo population gradually became more similar to that of previous residents (Watson et al., 2015), show nothing more than this type of call usage learning (and are not particularly convincing evidence of that; Fischer, Wheeler, & Higham, 2015).

Other recent evidence of intentional (voluntary) vocalization in chimpanzees showed that alarm calls fulfill several criteria for intentional communication developed in the ape gesture literature (Schel, Townsend, Machanda, Zuberbühler, & Slocombe, 2013). Again, however, this is evidence for intentional control over calls within the innate repertoire shared by all chimpanzees, and not evidence of vocal production

learning. Indeed, given the strong and long-standing experimental evidence for similar audience effects in chickens (Evans & Marler, 1994; Gyger & Marler, 1988) and many other species, it is difficult to see why this finding received so much media attention. None of these new studies change the basic fact that chimpanzees and other apes cannot learn to imitate novel vocalizations from their environment; rather, they demonstrate a basic capacity for control over innate vocalizations that is found in most birds and mammals. This form of control is believed to result from “gating connections” from cingulate cortex onto the basic brainstem chassis that controls innate vocalizations (Jürgens, 2002).

The current leading hypothesis for the mechanisms underlying the increased vocal control necessary for vocal production learning is that such control requires *direct synaptic connections* from motor cortical regions (or its equivalent, area RA, in songbirds) onto the motor neurons controlling the larynx (or syrinx in birds; Deacon, 1992; Jarvis, 2004a; Jürgens, 2002; Kuypers, 1958; Ploog, 1988; Simonyan, 2014; Striedter, 2004). Decades of research show that such connections are lacking in most primates (Jürgens, 1998, 2002) but are present in both songbirds and humans (Striedter, 2004; Wild, 1997). These data show that in several groups in which vocal production learning has evolved independently, corresponding direct connections are present, and such connections are *not* present in relatives that are incapable of vocal learning. This exemplifies a mechanistic hypothesis deriving from comparisons of humans and other primates, later tested in clades that have independently evolved an equivalent ability. Although the current data only concern parrots and songbirds, other clades of vocal learners (including both bats and seals) offer similar potential tests of this hypothesis.

The most intriguing new data suggesting sound learning in apes is fully consistent with the direct connections hypothesis (Marshall, Wrangham, & Arcadi, 1999; Reynolds Losin, Russell, Freeman, Meguerditchian, & Hopkins, 2008; Wich et al., 2009). These are data showing that novel nonphonated (nonlaryngeal) sounds can be controlled and learned by great apes including such attention-getting sounds as a lip buzz (or raspberry) in chimpanzees and whistling in an orangutans. Because great apes *do* have strong direct cortical connections with the brainstem motor neurons controlling the jaws and lips (Kuypers, 1973), this supports the hypothesis that vocal production learning requires such direct connections (Jürgens, Kirzinger, & von Cramon, 1982). Thus, it seems that it is specifically control over the larynx (the motor neurons for which lie in the nucleus ambiguus within the medulla) that is lacking in most primates and that therefore needed to evolve *de novo* during recent human evolution.

Hierarchical syntax

As often remarked, an unusual aspect of language is that it can be expressed via multiple modalities: both speech (audio-

vocal) and written communication (visuo-motor) are typical of educated humans, and a smaller community communicates via signed languages (also visuo-motor). Although written language is to some extent “parasitic” on spoken language, scholars now universally accept that signed languages are full linguistic systems, as communicatively adequate as spoken language when acquired from birth (Emmorey, 2002; Klima & Bellugi, 1979; Stokoe, 1960). Clearly, language as a system for expressing thought is not limited to the audio-vocal channel (for further implications of this fact, see de Boer, 2016; Goldin-Meadow, 2016; Kendon, 2016).

What does seem to be both universal in human languages, regardless of modality, and required for their expressive power, is hierarchical syntax (Chomsky, 1957, 1965, 2016). Without this fundamental characteristic, our open-ended ability to map novel thoughts onto understandable signals would be impossible. We might have a somewhat useful language nonetheless (like beginners in a foreign language making their way without grammar), but we would be unable to express complex concepts precisely.

Because “syntax” (or “grammar”) are terms used in many ways, it is useful to characterize what, precisely, is the derived aspect of human syntactic abilities. It is certainly *not* the capacity to produce rule-governed behavior, nor a capacity to produce or interpret signals—virtually any vertebrate does these things. Many organisms can also make “infinite use of finite means” in the restricted sense that they can produce an unlimited variety of signal strings (think of an incessantly barking dog—no day’s barking will be precisely the same as another’s). But of course, to the extent that they “mean” anything, these bark strings all mean the *same* thing. Some birds do more than this, acquiring quite complex and structured song repertoires via vocal learning, with basic elements numbering in the thousands (Kroodma & Parker, 1977) and an unlimited variety of orderings (Hultsch & Todt, 1989; Weiss, Hultsch, Adam, Scharff, & Kipper, 2014), and listeners show clear awareness of different song types (Naguib & Kipper, 2006; Naguib & Todt, 1997). While such complex repertoires are far from trivial, and their strings differ discriminably from one another, they do not communicate equally complex semantic messages. So the capacity to combine learned elements in a rule-governed manner—a basic phonological “syntax” of birdsong—is not sufficient to reach the level of human phrasal syntax and semantics (Marler, 2000).

Recent research using artificial grammar learning to probe the specific types of rules that organisms are capable of learning has offered a clearer picture (see ten Cate, 2016), and the use of formal language theory provides a useful metric and common description language for analyzing these rule types in terms of computational complexity (Jäger & Rogers, 2012). The most important distinction at present appears to be between the levels of regular or “finite state” grammars, some types of which are accessible to multiple animal species, and supra-regular

grammars that go beyond this (including both context-sensitive and context-free grammars, sometimes termed “phrase structure grammars”). At present, there are no convincing studies showing the successful acquisition of a supra-regular grammar in a nonhuman animal: for every claimed success (Abe & Watanabe, 2011; Gentner et al., 2006; Rey, Perruchet, & Fagot, 2012), there is a convincing critique (van Heijningen et al., 2009; Beckers et al., 2012; Poletiek et al., 2016). The results of such research have been summarized by Fitch and Friederici (2012), and further explored in important recent work by Dehaene et al. (2015) and Wang, Uhrig, Jarraya, and Dehaene (2015).

Thus, the shared component of syntax includes capabilities within the regular or finite state domain (also adequate to account for phonological phenomena in language Heinz & Idsardi, 2013). In contrast, the DCL component of syntax involves supra-regular computational capabilities like those underlying hierarchical linguistic structure (Fitch, 2014). From this perspective we could say loosely that animals have phonology but lack hierarchical syntax. Because both syntactic and semantic structures depend on hierarchical structure (rather than linear ordering), this is a crucial derived feature.

Because many of the pieces in this special issue discuss syntax at length, I will not further belabor this: Without hierarchical syntax, we would not have modern language. Explaining how and why we attained this more sophisticated level of cognitive computation must be a central explanandum in understanding the biology of language.

Semantic and pragmatic components of language

The last, and least understood, DCL involves a poorly defined complex of abilities involving the context-dependent semantic interpretation of signals—semantics/pragmatics. This capacity is based on a complex and highly-developed set of cognitive abilities for social cognition (Herrmann et al., 2007).

Again there is a shared component of semantics: an ability to learn to interpret novel signals (words or gestures) is widespread. Chimpanzees, parrots and domestic dogs are all capable of rapidly learning human words and mapping them onto real-world referents in a context-dependent manner (Kaminski et al., 2004; Pepperberg, 1981; Pepperberg, 1991; Pilley & Reid, 2011; Savage-Rumbaugh et al., 1993). In the case of Alex the grey parrot, the words learned include quite abstract qualities such as number, shape and color, not simple objects responded to reflexively (Pepperberg, 1999). In the case of Kanzi the bonobo, differing word orders can be reliably mapped on to different meanings (cf. Lyn, 2017; Pepperberg, 2016). Thus the capacity to *acquire a basic lexicon*, mapping signals to concepts, is an ability we share not just with chimpanzees but also dogs and (some) parrots. Extensive field work with primates (mostly with baboons) also indicates that the capacity to form quite complex and

context-dependent inferences is present in nonhuman primates (see Cheney & Seyfarth, 2007; Seyfarth & Cheney, 2016).

The derived aspects of semantics/pragmatics concern more complex meanings than single word meanings. A precise delineation of what differentiates humans from other animals in this domain remains elusive and debated. One key area where there are both similarities and differences concerns Gricean inference—the use of context, principles of communicative intent, and theory of mind to derive inferences about the implicit meaning of what is said (Grice, 1975). There is broad agreement about the importance of such pragmatic interpretation in language (cf. Moore, 2016; Scott-Phillips, 2016; Sperber & Wilson, 1986), but debate about the degree to which the logical principles laid out by Grice are in fact cognitively represented by ordinary language users (Bar-On, 2013).

A long-standing tradition held that “theory of mind” (ToM) is uniquely human (Penn & Povinelli, 2007; Povinelli et al., 1990; Premack & Woodruff, 1978). However, a series of recent experiments with chimpanzees shows that at least a basic capacity to represent the knowledge of others is present in apes (Bräuer, Call, & Tomasello, 2007; Call & Tomasello, 2008; Hare et al., 2000), and very recent data based on eye-tracking suggests a capacity to represent false beliefs as well (Krupenye et al., 2016). These results indicate that a capacity for first-order ToM was already present in the LCAC. Of course, humans are able to represent much more complex representations of others’ thoughts (e.g., second-order ToM—knowing what one agent thinks about another agent’s thoughts), and such higher order ToM is fundamental to most pragmatic inference. Thus, *higher order ToM* still appears to be a DCL that is key to semantic interpretation.

Another trait that appears to differentiate humans from other apes concerns not our *ability* to communicate, but rather our *proclivity* to do so, a proclivity for which I have borrowed the German word *Mitteilungsbedürfnis* (the drive to share thoughts). Apes given training with gestural or keyboard based communication systems learn to use them to provide answers to questions in return for rewards (food, tickling, etc.; see Pepperberg, 2016; Savage-Rumbaugh, 1986; Savage-Rumbaugh et al., 1993). However, they very rarely use these systems to volunteer information themselves, except for requests (typically for food or tickling!). In contrast, children by age 4 are founts of information, pointing to objects and naming them, observing and commenting on the world around them, and in general using language to share information with others. It is easy to overlook the importance of this trait, but without it the free flow of information that provides a prime benefit of language would slow to a trickle. While not uniquely human (honeybees certainly have a strong desire to share information about food locations with one another), this trait does seem to differentiate us from other apes.

Finally, a fundamental aspect of human communication (not just language) is *ostension*—the signaling of

signalhood—and there is an ongoing debate about the degree to which ostensive signaling is unique to humans or shared with apes (Moore, 2016; Scott-Phillips, 2014, 2016). At present, it seems likely that our human propensity to generate communicative acts, and explicitly mark them as such, is quantitatively more highly developed, and perhaps a true qualitative difference, but more research will be required to demonstrate this conclusively.

Thus, it seems clear that the complex of pragmatic/semantic “tools” available to humans is uniquely well-developed (with ToM, Gricean inference, *Mitteilungsbedürfnis*, and ostension all hypertrophied relative to other apes). But for no one of these subcomponents do we have clear evidence of a definitive qualitative difference: it seems to be the whole package that evolved. This is why I treat this as a single DCL, although future research may cleave this complex into biologically separate components. There can be no doubt that these cognitive abilities are critical for language, but they also play a role in nonlinguistic aspects of human cognition (Scott-Phillips, 2014). They mark the third crucial domain of difference between humans and our nearest primate relatives (Herrmann et al., 2007).

Cultural aspects of language evolution

There is a growing consensus, well reflected in this issue, that cultural change has an important part to play in understanding language (Adger, 2017; Bowling, 2017; Kirby, 2017; Pagel, 2016; Steels, 2016;). This study of cultural change has advanced rapidly in recent years, part of a more general scientific focus on cultural evolution (Boyd & Richerson, 1996; Fitch, 2011c; Laland et al., 2010; Mesoudi et al., 2004). This research has recently taken a strong empirical turn both in humans (Kirby et al., 2008; Morgan et al., 2015; Smith & Kirby, 2008) and animals (Fehér, 2016; Fehér, Wang, Saar, Mitra, & Tchernichovski, 2009; Whiten et al., 1999). Because the term “language evolution” can be interpreted either in terms of biological evolution of the language faculty or the cultural evolution of specific languages, Jim Hurford introduced the useful term “glossogeny” to denote the latter specifically (Hurford, 1990). Although cultural and biological evolution are sometimes considered as mutually exclusive competing explanations (e.g., Christiansen & Chater, 2008), they are increasingly seen as complementary: glossogeny can explain much of the odd, language specific variability we see among different languages, and thus “takes the pressure” off biological explanations from having to explain intricate details of language (Chomsky, 2010; Deacon, 1997; Fitch, 2008, 2011d). The major review article in this issue by Simon Kirby gives an overview of the rapid advance of empirical and modeling approaches for understanding the nature and consequences of glossogeny (Kirby, 2017); Mark Pagel’s

article provides a further overview of how the study of glossogeny has progressed in the last decade (Pagel, 2016).

Models of language evolution: Hypothetical protolanguages

I now turn to a brief overview of proposed models of language evolution, and some of the core debates in its study. Although models like these are often termed “theories” of language evolution, I prefer to use the term “model” because “theory” in science typically connotes a much more well-tested and widely accepted model than any existing models of language evolution.

Previous dichotomies and debates

There are three divisive distinctions that, in my opinion, have been more controversial than deserved (cf. Fitch, 2010). The first concerns the internal (thought) versus external (communication) uses of language. As stated earlier, language is a complex faculty that allows us to encode, elaborate and communicate our thoughts and experiences via learned words combined into hierarchical structures (sentences). One of the persistent debates in the field has been whether the use of language for encoding and elaborating thought is primary (Chomsky, 2010; Newmeyer, 2005)—as an “inner tool”—or whether its use for communication was the consistent core function (e.g., Pinker & Jackendoff, 2005). I think that this is a misleading dichotomy. Contemporary language clearly functions both in our inner mental lives and for communication, and it seems Procrustean to denote one as “primary.” Second, even our inner thoughts occur with a phonology (we can reasonably ask a bilingual whether they are presently thinking in English or German), and so the word forms learned via externalization still play a role internally (Jackendoff, 2011). Finally, we know that both animals and global aphasics can think (engage in complex cognitive processes) without language, so language is not *necessary* for thought to occur (see below). Nonetheless, we clearly use language for thinking, and most people probably spend more time using it internally than externally. I can see no major grounds for enforcing a distinction between these in modern human language; the only genuine issue is whether the *original* evolutionary function of some DCL was for thought or communication.

This leads to a second persistent debate, regarding continuity *versus* change in function. Darwin famously solved the problem raised by the evolution of complex organs with his idea that an organ of some complexity could evolve for one function, developing a certain degree of complexity, and then later change its function (Darwin, 1859; Gould, 1985). This old function was long termed a “preadaptation,” but in recent

years the new function has been dubbed an “exaptation” to avoid implications of evolutionary foresight (Gould & Vrba, 1982). This notion of *change of function* is a central, although often overlooked, plank of evolutionary thought (Pievani & Serrelli, 2011) because it can refute arguments of “irreducible complexity” (Denton, 1985)—that complex organs like wings or eyes could not have evolved from simple beginnings because they need to be complex to have any functionality at all. Applied to language evolution, it seems clear that that syntactic combinatoriality and semantic compositionality might both have their roots in prelinguistic conceptual structures (Berwick & Chomsky, 2016; Bickerton, 1990, 2000a). This idea solves various other long-standing problems, including what I call the “lone mutant” problem (Fitch, 2010; Orr & Cappannari, 1964): Who would some hominin lucky enough to have more advanced syntactic competence talk to? The answer initially is “no one” (at least until offspring were born who shared the mutation), but the ability would nonetheless be useful in private thought. To accept this idea is not to claim that language isn’t (or wasn’t) used in communication—only to say that advances do not always need to serve immediate communicative functions (Fitch, 2011e). The idea that conceptual structure led the way is plausible enough that Newmeyer described it as the “consensus view” of the 1990s (Newmeyer, 2005), but no dichotomous either/or decisions are necessary about this issue.

A final persistent issue worth a brief mention is the saltation versus gradualism distinction (Berwick, 1997), which exemplifies a larger debate in evolutionary theory. Darwin was an extreme gradualist, believing that only the accumulation of small successive variants could lead to adaptation, but he was criticized for this viewpoint by many contemporaries (Eldredge & Gould, 1972; Gould, 1982; Theissen, 2009), and the belief that “Darwinism = gradualism” was one reason that many early geneticists opposed Darwinism (e.g., Bateson, 1894). But by the time of the modern synthesis of evolutionary theory and genetics, it became clear that there is a continuum of both tempo (rate of change) and mode (type of change) in evolution (Gould & Eldredge, 1977; Simpson, 1944), a viewpoint that has become ever clearer as genetic mechanisms have become better understood (Fitch & Ayala 1994). At the level of DNA *all* evolutionary change is discrete, because there are only four discrete bases in the genetic material. The fossil record often exhibits apparently discontinuous bursts of rapid change after long periods of stasis, but discontinuity on a geological time scale does not imply saltation over generations. Furthermore, the fact that a trait like language is discontinuous in terms of extant species carries no implication of phylogenetic saltation (it may have evolved very gradually over millions of years, but no fossils or extant species are left as records of that gradual transition). The only substantive issue in this debate concerns the size of phenotypic change associated with a mutation, and whether large changes can be

(or are likely to be) improvements (cf. Fitch, 2010). It is certainly not implausible that small genetic changes or single mutations could lead to rapid and important rearrangements of neural circuitry (Ramus & Fisher, 2009). There is again no reason to assume that evolution always works in either way: we should consider each trait and its genetic/neural underpinnings individually.

Conceptions of protolanguage

I now turn to issues that I believe *are* central to discussions of language evolution. Any modern model should take for granted the comparative data, and take as its starting point the LCAC, reconstructed via inferences based on comparisons with chimpanzees and bonobos. Although there is much we still don’t know about apes, this research is just “normal” comparative biology, and allows rather confident specification of the starting point of the six million year journey to modern human language. The other fixed point is the end stage: a current understanding of language and cognition, again normal science, but this time linguistics and cognitive science.

The challenging component is the intermediate period: here we are reliant on the sparse data of the fossil record, which is notoriously incomplete and controversial (see [Paleontological data](#) section), and which provides rather few restrictions (Tattersall, 2016). But based on brain size, ecology, and toolkits, it seems clear that the earliest hominins (e.g., australopithecines) did not have modern language, although they might have had some DCL precursors. Although things become less clear within the genus *Homo*, most experts accept that Neanderthals lacked at least some component of modern language (see [Paleontological data](#) section and Mellars, 1989; Tattersall, 2009). It is also clear that by the time modern humans dispersed out of Africa (by 60 thousand years ago), we had the full package of modern language DCLs, since all humans around the world have the same essential capacity to acquire any language. Accepting these assumptions as a reasonable starting point, there is a period of roughly 2 million years during which most of the action must have occurred, with only a few anatomically distinct stages between *Homo habilis* and *Homo sapiens*. A complete model needs to offer explanations for how *all* of the empirically deduced derived components of language evolved during this period. Most existing models attempt to only explain some of the DCLs (e.g., speech or syntax, but not both), and few grapple with the entire package.

The existence of multiple DCLs leads logically to a notion of “protolanguage”—some hypothesized system of thought and/or communication that had some DCL(s) but not the full suite. The only way out of this conclusion is to state that some particular component of language was the key, and that “language evolution” amounts solely to the appearance of that chosen component (e.g., Berwick, 1997; Herrmann et al.,

2007). Many of the previous debates in the field can be dissolved by recognizing that the models being debated attempt to explain different parts of the problem (syntax, speech, social cognition, etc.), and thus are in fact complementary (Fitch, 2010). Remaining debate should concern these differing conceptions of protolanguage, and explanations of when they existed, how they evolved, and why they were adaptive at the time. A key issue in these models becomes the ordering of acquisition of different DCLs.

Lexical protolanguage

This model has many variants; prominent exponents include Derek Bickerton (Bickerton, 1990, 2000b) and Ray Jackendoff (Jackendoff, 1999, 2010). The essential idea is that hominins first developed a word-based protolanguage that was learned, symbolic, and useful, and only later evolved a capacity for hierarchical syntax. Typically the birth of modern hierarchical syntax is seen as the final stage of language evolution, so this is a “syntax last” model. However, such models leave open the origins of the other DCLs (speech and semantic/pragmatic capabilities), except for predicting that vocal learning should have evolved before syntax. Although Bickerton’s version of lexical protolanguage is often considered definitive (or even simply assumed to be “protolanguage”), there are numerous variants on this basic idea (reviewed in Chapter 12 of Fitch, 2010). All of these models imply that syntax was a very late acquisition in language evolution.

Gestural protolanguage

The term “protolanguage” was first used in an evolutionary context by Gordon Hewes in the context of gestural protolanguage (Hewes, 1973)—the idea that during the first stages of language evolution, communication was via gesture, mime, and sign. Current proponents of this view include Arbib (2002), Corballis (2002, 2016), and Tomasello (2008). A key virtue of this model is that it takes the well-attested superiority of apes in gestural versus vocal communication as its starting point. Its prime flaw is that it has difficulties explaining why the transition to speech, which came later in the evolution of language, was so complete (Emmorey, 2005; MacNeilage & Davis, 2005; Seyfarth, 2005), as further discussed by Arbib (2016) and Kendon (2016) in this issue. The clear prediction of this model is that syntax and semantics preceded speech during evolution.

Musical protolanguage

An alternative model for the earliest stages of language evolution, due to Darwin (1871), is that the first DCL to be acquired in phylogeny was the capacity for complex vocal learning. On

the model of birdsong, Darwin suggested that this vocal learning capacity was initially used not for communicating complex propositions, but rather to produce complex vocal performances (cf. Fitch, 2013); a prominent current champion is (Mithen, 2005). Despite its Darwinian origins, this model has a checkered history, apparently having been forgotten and then independently rediscovered multiple times (Brown, 2000; Livingstone, 1973; Richman, 1993). In many ways, Robin Dunbar’s “vocal grooming” hypothesis is consistent with the musical protolanguage hypothesis (Dunbar, 1996, 2016), although it extends beyond song to include more primitive vocalizations such as laughter (see also Dunbar, 2016; Locke, 2016; Provine, 2016). Although Darwin was quite vague about the later stage of semantics (and said nothing about syntax), his ideas were fleshed out by the linguist Otto Jespersen (Jespersen, 1922), who suggested plausible roots for both semantics and syntax (via analysis of previously holistic utterances; cf. Botha, 2009; Tallerman, 2008; Wray, 1998). For critiques of the musical protolanguage hypothesis see Steklis and Raleigh (1973) and Tallerman (2013).

Mimetic protolanguage

A model focused on the intermediate stages of language evolution, and which incorporates aspects of the previous models, is Merlin Donald’s mimetic protolanguage (Donald, 1991, 2016). Similar to musical protolanguage, Donald envisions a crucial role for performative, group-defining rituals, initially devoid of propositional meaning, as an initial stage of language evolution, one he ties to *Homo erectus*. But mimetic protolanguage was an all-inclusive affair, including both gestural and vocal components, and it thus elides the distinction between gestural and vocal protolanguages (cf. Mithen, 2005). This is appealing in that gesture is present in both apes and in modern humans, so we can assume it was present and playing a communicative role throughout hominin evolution. Many commentators side with the idea that pitting vocal and gestural models against each other creates a false dichotomy (de Boer, 2016; Goldin-Meadow, 2016; Kendon, 2016). Donald’s piece in the current issue lays out this model and predictions concisely.

Summary

The key observation about these models is that they make testable predictions; in particular each of these models makes contrasting predictions about the order of acquisition of DCLs. As emphasized in the next section, paleo-DNA can play a central role in testing these predictions. Thus, for example, the musical protolanguage model predicts spoken language as the first DCL to evolve, while gestural protolanguage suggests it was the last. Musical protolanguage models also suggest a close link between speech phonology and song, a

prediction that can be tested using both brain imaging and genetic data. All of these models posit a prolonged period during which the hypothesized protolanguage was the main form of communication, often during the reign of *Homo erectus/ergaster*, suggesting that the DCL(s) involved should be more robust to brain damage, genetic abnormalities and/or developmental delay (cf. Fitch, 2010). Precisely these sorts of testable contrasting predictions offer our best hope of moving this discipline beyond “story telling” and into real science, and provide grounds for believing that major progress in understanding language evolution can be made in the coming decades. I now turn to the sorts of data available for testing between them.

Empirical data relevant to testing models of language evolution

A wide variety of data are directly relevant to understanding language, most obviously those stemming from cognitive science and linguistics, including developmental, comparative, and historical linguistics. Such data have provided a reasonably clear picture of what language is and how it is acquired during infancy and childhood. The core findings, as taught in introductory courses in linguistics or psychology of language, have been reviewed in detail in many places (e.g., Crystal, 2003; Yule, 2006). Although the correct theoretical and philosophical framework for understanding this picture remains a topic of discussion (see, e.g., the articles by Arbib, 2016; Chomsky, 2016; Jackendoff & Wittenberg, 2016; Scott-Phillips, 2016), it all concerns modern humans and thus defines the “end target” of any evolutionary model, rather than the steps required to get to this point. It is only rather recently that detailed theoretical models of modern language have been used to fuel hypotheses about how language evolved (e.g., Berwick & Chomsky, 2016; Givón, 2002; Hurford, 2011; Jackendoff, 2002, 2010; Scott-Phillips, 2014), though earlier efforts include Bickerton (1990) and Pinker and Bloom (1990).

Comparative cognition and cognitive biology

In my opinion, the data that still remain most under-utilized in analyzing the biology and evolution of language are comparative data from nonhuman animals (“animals” hereafter), particularly those from nonprimate species such as birds, bats, or dogs. There is of course a long tradition of using comparisons between humans and other primates, although even these have tended to be superficial or assume that because some trait is seen in some monkey species (e.g., vervet alarm calls) it was present in our lineage before the evolution of language. In fact, a key role of primate comparisons is to reconstruct our LCA with chimpanzees and bonobos in detail. This common ancestor was *not* a chimpanzee, but an extinct species for which we

have no fossil evidence. It is important to note that both humans and chimpanzees have been evolving since this split, and we thus find both behavioral/physiological traits (e.g., sexual swellings) and genetic differences that are due to chimpanzee evolution, and where humans retain the ancestral or “primitive” state (cf. Pääbo, 2014). Thus, reconstructing the LCAC requires broader primate comparisons (e.g., with orangutans and gorillas) to determine the “base state” from which we started. A more detailed characterization of the LCAC is found in Fitch (2010).

Beyond defining this starting point, we can also use comparisons with an ever widening group of related species to determine homologous traits shared by larger groups (see [The broad comparative approach](#) section). This homology-based approach allows us to rebuild our earlier and earlier common ancestors (e.g., with primates, mammals, tetrapods, vertebrates); the comparative method used in this way provides the biologist’s equivalent of a time machine, and (particularly when combined with genetic data) allows us to say with certainty when and how particular cognitive capacities arose during evolution (see section “The Long Time Scale”). This basic approach has been nicely captured in Richard Dawkins’ very accessible *The Ancestor’s Tale* (Dawkins, 2004), and is further discussed or illustrated by several articles in this issue (Byrne & Cochet, 2016; Fischer, 2016; Lyn, 2017; Seyfarth & Cheney, 2016).

A key point sometimes overlooked when considering animal data is that both animal communication *and* cognition are relevant. We cannot assume that, because a species does not show some capability in its communication system, it lacks it (cf. ten Cate, 2016)—obviously a cognitive capability could evolve for and be used in other systems, and the species in question has no need to use these in its communication system. Also, if we accept that some capacities key to language may have evolved in contexts other than communication (and were later exapted during language evolution), then we need to consider data from animal cognition on an equal footing with animal communication.

The other major role of cross-species comparisons, as already discussed in “[The broad comparative approach](#)” section, is to determine analogies—traits that evolved independently in humans and animals. The evolution of vocal learning is perhaps the most obviously relevant to language, already highlighted by Darwin (1871) and further discussed by Hickok (2016), Locke (2016), and Vernes (2016) and detailed in articles in a recent special issue on animal communication and language (Brainard & Fitch, 2014). But equally interesting examples come from syntax (which can be broadly thought of as perception of patterns of various levels of complexity), where great recent progress has been made in defining what birds can do (cf. Fitch, 2014; Fitch & Friederici, 2012; ten Cate, 2016), or social cognition (cf. Bugnyar et al., 2016; Call & Tomasello, 2008; Fitch et al., 2010; Scott-Phillips, 2016; Seyfarth & Cheney, 2016), where the cognitive

parallels of pragmatics and theory of mind can clearly be found in ravens. These examples of analogy provide ways to test adaptive hypotheses (e.g., the social intelligence hypothesis of Byrne & Whiten, 1988 and Humphrey, 1976, or the cooperative breeding hypothesis of Burkart, Hrdy, & van Schaik, 2009 and Lukas & Clutton-Brock, 2012)—both introduced with reference to primates—and to explore the mechanistic basis of these abilities in an independently evolved brain.

This special issue provides a very rich selection of comparative data, including work on nonhuman primates (Arbib, 2016; Byrne & Cochet, 2016; Fischer, 2016; Lyn, 2017; Seyfarth & Cheney, 2016), birds (Fehér, 2016; Okanoya, 2017; ten Cate, 2016), bats (Vernes, 2016), or a mixture of species (Pepperberg, 2016). I will thus delve no further in this introduction into the value and virtues of comparative data.

Neuroscientific data

It is sometimes said that we know almost nothing about how the brain computes language (e.g., Berwick & Chomsky, 2016). While this was true in the 1960s, where aphasia and disorders were the main source of data, it is far from true today. The last decade has witnessed massive advances in our understanding of neural computation in general (e.g., single neuron computation, predictive coding) and in the specializations of the human brain (e.g., the great enlargement of Broca's area and massive increase in its connectivity via the arcuate fasciculus). Particularly welcome developments in human brain imaging have led far beyond the first neophrenological stage of brain imaging (where “the area for x is sought,” x being language, syntax, social intelligence, love, etc.) to the clear recognition of the importance of brain networks (often spread widely throughout cortex) and connectivity in neural computation.

Prominent among these advances is the use of diffusion tensor imaging (DTI) and related techniques to map and measure the white-matter tracts connecting distant brain regions (Anwander, Tittgemeyer, von Cramon, Friederici, & Knösche, 2007; Catani & Mesulam, 2008; Friederici, 2009; Melhem et al., 2002). This allows *in vivo* analysis of the *anatomical connectivity* of different regions, and how they vary between species (e.g., Rilling et al., 2008). Such analyses are complemented by *functional connectivity* analysis (which uses cross-correlation in time across multiple brain areas) to evaluate which regions are coactivated in particular tasks (Sporns et al., 2004; Xiang et al., 2010; Hamilton et al., 2013), an approach also potentially applicable to interspecies comparisons (Rilling et al., 2007). Such approaches have led to important advances in our understanding of the neural computations that underlie language, including the still ongoing debate about whether any are specific to language or not (cf. Fedorenko & Thompson-Schill, 2014).

Another important development is the use of transcranial magnetic stimulation (TMS) and related techniques to noninvasively and temporarily inhibit or enhance brain function in specific, selected brain regions (Pascual-Leone, Walsh, & Rothwell, 2000; Rödel et al., 2009; Udden et al., 2008). These methods allow neuroscientists to actually test the causal hypotheses derived from brain imaging studies, which are typically correlational in nature. Finally, a surge in infant brain imaging has led to unprecedented insights into the surprising degree to which human neural specializations for language are already present and left-hemisphere biased at birth (or before, in premature babies), as ably reviewed by Ghislaine Dehaene-Lambertz in this issue.

Specific major recent advances in neuro-linguistics include the following findings:

1. Data from severe global aphasics clearly demonstrates that language is not necessary for sophisticated thought (Donald, 1991; Fedorenko & Varley, 2016; Varley & Siegal, 2000).
2. The neural basis of syntactic structure-building relies heavily upon an extended network centered on the inferior frontal gyrus, particularly Broca's area (BA 44/45) (Dehaene et al., 2015; Pallier, Devauchelle, & Dehaene, 2011; Friederici, 2016; Wang, Uhrig, et al., 2015).
3. This area is greatly expanded in humans relative to chimpanzees (Schenker et al., 2010).
4. Its dorsal interconnections to parietal and temporal cortices via the arcuate fasciculus have been massively expanded in humans relative to other primates (Rilling et al., 2008).
5. Much of this network is already present, and left lateralized, before birth, as shown by brain imaging of premature and newborn infants (Dehaene-Lambertz & Spelke, 2015), demonstrating that prolonged exposure to speech is not necessary for the development of these human specializations (Dehaene-Lambertz, 2017).
6. The neural basis of semantic composition relies on multiple temporal and frontal areas, with the anterior temporal lobe playing an important role in early semantic composition (Bemis & Pykkänen, 2012; Pykkänen & Marantz, 2003).
7. Higher order semantic combinations are built later, and the ventro-medial prefrontal cortex plays an important role in sustaining such temporary semantic structures for use in other brain regions (Bemis & Pykkänen, 2013).

The brain imaging studies referred to above have been repeatedly replicated, across many languages, and apply to both spoken, written, and signed languages, and in many cases apply to both perception and production (Poeppel, Emmorey, Hickok, & Pykkänen, 2012). While the dataset for comparative comparisons is sparser, we have a rather clear conception

of what precisely changed during the evolution of the human brain: In addition to a general size expansion, there were particular expansions (both in raw size and in terms of connectivity) of brain regions long known to play an important role in linguistic syntax.

Again, I need not belabor this topic in the introduction: it is covered by many experts in this issue (Arbib, 2016; Boeckx, 2016; Friederici, 2016; Hickok, 2016). Together, this body of research clearly refutes the misconception that we know little or nothing about the brain mechanisms underlying language.

Paleontological data

Another important, if sometimes over-estimated, source of data relevant to language evolution is the hominin fossil and paleontological record (the traditional term “hominid” for human ancestors contrasted with “pongid”—a false or paraphyletic clade containing all of the other great apes; “hominid” is now often used to refer to humans and other great apes, and the modern term to refer to humans and our extinct relatives postdating our split with chimpanzees is “hominin”). While the fossil record provides important clues to such things as body size, brain size, and technological abilities, the inferences these allow about language abilities are tenuous, and remain controversial. All commentators agree that the hominin fossil record reveals a quite “bushy” phylogenetic tree, with multiple different species existing at any given time point; there was no simple linear march from *Australopithecus* to us. Nonetheless, I will only mention those extinct hominins believed to be within, or close to, our direct ancestral lineage.

The clearest evidence from the fossil record concerns the timing of bipedalism, large body size, and large brains. It is now clear that one of the first events in the hominin lineage was the habitual assumption of **bipedal posture**, and that this occurred *before* any major increase in brain size (Stringer & Andrews, 2005). Thus, members of the genus *Australopithecus* (australopithecines hereafter) had brains roughly the size of chimpanzees, but habitually walked upright. We can infer from tool use in living chimpanzees that the LCAc used simple tools of stone and plant materials, and it is not until the Oldowan period (starting about 2.6 million years ago, MYA hereafter) that we see more sophisticated stone tools with cutting edges, presumably produced by late australopithecines and certainly by early *Homo*. These tools, while certainly useful, are not far beyond the cognitive capabilities of existing chimpanzees (Toth, Schick, Savage-Rumbaugh, & Sevcik, 1993) and so provide no evidence of language in their makers.

The greatest changes in the hominin lineage occurred with the advent of the genus *Homo*, which clearly exhibited a suite of derived traits marking these as a truly new kind of primate. *Brain size* increased considerably (Holloway, Broadfield,

Yuan, Schwartz, & Tattersall, 2004), sexual dimorphism *decreased*, and the toolkit became much more sophisticated. These were also the first of our ancestors to leave Africa, successfully colonizing much of the Old World, and thus had a much more flexible ability to colonize and exploit new niches than australopithecines. *Homo erectus* is the traditional moniker for the most widespread and successful of these hominins, although authorities often reserve *erectus* for the Asian members of this group and use *Homo ergaster* for those remaining in Africa. The Achulean hand-axes that formed an important component of their durable toolkit were sophisticated tools, far beyond the capabilities of modern apes; indeed, we modern humans cannot make these tools without considerable practice and hard work. Despite this, hand-axe technology remained essentially the same for about one million years, strongly suggesting that *Homo erectus/ergaster* did not have the full cognitive and cultural capabilities of modern humans, and by inference did not have full modern human language. Because of this contrast between features resembling ours (larger brain and body size, sophisticated tools, ecological flexibility) but also unlike us (cultural stasis over a prolonged period), these hominins are often considered a prime candidate for hominins that possessed some sort of protolanguage, with some but not all DCLs of modern language.

Another increase in brain size and cognitive sophistication is represented by a suite of fossils often referred to as *Homo heidelbergensis* in a broad sense (Ruff, Trinkaus, & Holiday, 1997). These hominins produced excellently balanced wooden throwing spears (Thieme, 1997), and are thought to be close to the common ancestor of Neanderthals and modern humans (debate still rages over whether *H. antecessor* may be the more suitable moniker; Bermúdez de Castro et al., 1997; Endicott, Ho, & Stringer, 2010). In Europe and Asia, these hominins were succeeded by *Homo neanderthalensis* and a newly discovered species, Denisovans (see below). Neanderthals had brain sizes identical to or exceeding those of modern humans, and their hunting practices and stone tool kits approached ours in complexity, but many crucial *symbolic* aspects of the artifacts of modern humans are found rarely or not at all in association with Neanderthals (Mellars, 1998b, 2004; Tattersall, 2016). However, despite the excellent fossil and archaeological record left by Neanderthals, and decades of discussion, their cognitive abilities remain highly controversial (e.g., Dediu & Levinson, 2013; Lieberman, 2007; Stringer & Gamble, 1993; Wynn & Coolidge, 2004)—ranging from “Neanderthals were just like us” to “they lacked key cognitive capacities including symbolic language.” Given their utter disappearance in Europe shortly after the arrival of modern humans from Africa, it seems more plausible to assume at least some cognitive differences. Fortunately, high-quality paleogenetic data is now available for both Neanderthals and Denisovans, data which offers clear hope of progress beyond this impasse.

Genetic and paleogenetic data

One of the most exciting empirical developments of the last decade, relevant not only to language evolution but more broadly to our understanding of the human condition, is in molecular genetics and genomics (cf. Fisher, 2016; Pääbo, 2014). Public databases containing detailed, whole-genome sequences for thousands of individuals are freely available, as are the genomes of thousands of nonhuman species, including our nearest ape relatives. We also have high-quality genomes for two extinct archaic hominins, Neanderthals and Denisovans. The fact that all of these resources are freely available means that science has entered a new era where vague talk of “mutations” can transition to hard facts about gene sequences and allele distributions. These new data also offer unprecedented ways to test hypotheses about language evolution and to uncover the timeline through which genes specifically involved in various DCLs changed and spread through our species. If a single empirical development warrants optimism and excitement about the coming decades of language evolution research, it is these advances in genetics and genomics.

Nonetheless, these data pose daunting challenges, even for the professional molecular geneticists and bioinformaticians who produce them and especially for outsiders from linguistics or cognitive science for whom the data themselves and the tools to analyze them are *terra incognita*. It is unfortunate that a simplistic “gene for language” approach still tends to dominate the popular press and many scholarly discussions, when it is now clear that most traits involve multiple alleles, and that the function of any gene needs to be understood within a context of the many other genes with which it interacts. Thus in the same way that today's neuroscientists focus less on individual neurons and more on neuronal circuits (and less on specific brain regions than on global brain functional networks), the new genomic era forces us to think more in terms of genetic circuits, gene regulation, and interactions among multiple alleles than was typical during the pregenomic era. The multicomponent approach also provides a valuable scaffolding for understanding the specificity and/or generality of the effects of particular genetic changes. In addition to these essentially conceptual realignments one must add the vast diversity of genes (the names alone—mostly barely pronounceable acronyms—are daunting), and the rapidly changing analytical approaches for studying them. Thus, integrating linguistic, psychological and neural perspectives with the data rapidly pouring out of sequencing labs and onto the Internet is anything but trivial. But, as I will argue below, it will be worth it, because genes provide the closest thing to “fossils of language” we will ever have, and paleo-DNA in particular provides the analog of a time machine, taking us back nearly half a million years to the split between early Neanderthals and modern humans.

There are essentially three time scales at which genomic data are applicable. The longest is that studied by comparing humans with other living species, from bacteria to chimpanzees, and stretches from more than one billion years ago (10^9 years) to our separation from chimpanzees and bonobos roughly 6 million years ago. The shortest time scale involves, potentially, the genomes of all humans alive today, and is important from a linguistic viewpoint because of the widely accepted observation that humans from all human cultures today have an equivalent capacity for acquiring language, and that the language acquisition system is thus a human universal (barring clinical cases). Examination of the differences among modern humans thus allows the exclusion of a vast amount of genetic diversity as essentially irrelevant to language. More interestingly, we can use signatures of selection derived from comparisons of modern human genomes to gain insights into *when* particular selective events may have occurred, though the accuracy of this method grows poor past 30–50 thousand years ago (Przeworski, 2002). Finally, and most exciting, an intermediate time scale is offered by gene sequences from extinct hominins including Neanderthals, which pushes our time scale back to at least 500 KYA (kilo-years ago) and can be expected to push back further with the acquisition of more fossil DNA. I now review findings from each of these time scales in turn.

The long time scale: Comparative genomics

Starting with the genomes of living species, we can trace back evolution at the level of many millions (for vertebrates) or even several billion years (for all living things). Applying the comparative approach to compare human genomes with other species, we can reconstruct the evolution of most of human biology (because most of human biology is shared with at least some other species). Returning to eye evolution, trichromatic color vision is a trait that is rare in mammals but shared by humans and other apes (as well as various other primates). But it turns out that the ancestral vertebrate had at least trichromatic color vision (as evidenced by color vision in living fish, reptiles, and birds), and that the ancestral mammal lost one or more color-sensitive proteins, presumably as an adaptation to a nocturnal lifestyle (Collin & Trezise, 2004; Jacobs, 1993). Then, via a gene duplication event and subsequent functional divergence of the duplicates, trichromacy was regained in some primates including our own ancestors. This nonintuitive evolutionary trajectory is just one example of how a comparative genomic approach allows the confident reconstruction of our evolutionary past in exquisite detail—and the truth is not simple. This approach puts the vast majority of our evolutionary history within reach, from the origins of life to our separation from chimpanzees about 6 MYA, and it allows us to confidently catalog the many components of

our language faculty that predated the evolution of language in recent humans.

Of course, given the exceptionality of human language relative to ape communication, some of the most pressing questions concern the differences between humans and other apes. The most famous such genetic difference is the FOXP2 gene, a gene involved in oral motor control and speech, discussed in detail in Simon Fisher's contribution to this issue (Fisher, 2016). But unfortunately, finding further clear differences will involve both luck and hard work, because there has been so much genetic change that it is difficult to come to grips with. Roughly 20 million single-nucleotide substitutions (involving roughly 40% of all proteins) differentiate chimpanzees from humans (Pääbo, 2014). Most of these changes are presumed to have no biological effect, but simply represent accumulated genetic "noise." Thus, searching for the key *functional* genetic changes underlying human/chimpanzee differences involves searching for needles in a genetic haystack, requiring probabilistic tools to screen for regions of interest.

One approach is to use bioinformatics tools to seek out signatures of positive selection in the human genome, but due to the relatively rapid time decay of such signatures, this approach is limited to rather recent changes. Still, such comparisons have yielded several genetic changes of interest. A nonlinguistic example involves spines on the penis, which are present in chimpanzees but thankfully absent in humans: This loss appears to involve the loss of function of an enhancer of the gene involved in spine development (McLean, 2011). Loss of function of a particular protein led to inactivation of a muscle protein expressed particularly in the temporalis jaw muscle, MYH16, and may have led to the reduction in this muscle's size and jaw robusticity (Stedman et al., 2004). Finally, and more relevant to language, a similar loss of enhancer function led to inhibition of the GADD45G gene, which appears to limit cell division in the developing brain (McLean et al., 2011). Decreased activity of this inhibitory gene may be one of several changes involved in the expansion of brain size during hominin evolution. As these examples show, the *loss* of gene function may be just as important as gains of functionality to understand human differences. These also illustrate the nonintuitive way genetic circuitry and interactions—"inhibit an enhancer of an inhibitor"—often underlie even apparently simple phenotypic changes.

The central quantitative measure of selection in comparisons between species is the d_N/d_S ratio: the *ratio of nonsynonymous to synonymous* base pair changes. Because the genetic code is redundant, there are multiple three-base pair codons in DNA that yield the same amino acid in the coded protein; these are termed "synonymous" mutations. Only those DNA changes that actually change the resulting protein—the nonsynonymous mutations—are expected to be "visible" to natural selection. Thus, when the d_N/d_S ratio is

high, with an excess of such coding changes, we can infer that the corresponding region has been under selection. This approach yields the strongest signal in cases where a particular region of the genome has been under continued selection, for example, in regions involved in disease resistance.

Another approach to finding the functional genetic needles in the haystack of nonfunctional changes involves searching for so-called *human accelerated regions*, or HARs. These are regions of the genome that are, in general, highly conserved (e.g., among mammals or vertebrates), suggesting that they are functionally important, but have nonetheless changed considerably during human evolution. One such region involves a particular RNA sequence, HAR1F, that is expressed specifically in a subclass of human cortical neurons (Pollard et al., 2006) and may play a role in specifying the six-layered structure of the neocortex. This gene illustrates another surprise of the genomic era: the importance of "noncoding" RNA in biology and development. Classically, RNA was seen as little more than the middleman between DNA and the proteins that do the real work. But it is now clear that even RNA that is not translated into protein can, by itself, play myriad biological roles, and this HARF1 gene is an interesting example, again involved in brain development.

Another theme of the genomic revolution is the importance of *gene duplication* in evolution. This can be either gene duplication with subsequent divergence (already mentioned above for color vision) or simply making more copies of a gene so that more protein product ends up being expressed. A potentially language-relevant example of the first phenomenon is SRGAP2, a gene which has duplicated three times in humans relative to other apes. The protein coded by one human copy has been truncated and binds with the normal proteins in such a way that neurons expressing this human truncated form (Charrier, 2012) have a larger number of long dendritic spines (potentially modifying the receptive field properties and/or operative time scales in neural circuits). As these examples all show, multiple genetic differences between chimpanzees and humans, functionally relevant to human cognition and perhaps language evolution, can already be listed and studied at present, even though the total list of differences involves 20 million base pairs.

The short time scale: Comparing human populations

Comparisons among living human populations and individuals can also play an important role in testing hypotheses about the genetic bases of language and can offer clues about evolution. The most obvious role of contemporary diversity in DNA, as already mentioned, is negative: When some alleles show variation among living human populations, this variation can be assumed to be irrelevant, based on the fact that there is no known significant variation in the language faculty among living humans. Contemporary genetic data reveal,

incontrovertibly, that “racial” differences between existing human populations are literally skin deep and concern appearance but no known genes involved in cognition or neural function (Pääbo, 2014).

Another important role concerns disorders with an inherited genetic basis, where individual variations can offer insights into the role of the genes involved in language development. The discovery of the FOXP2 gene, originally uncovered via a British family with a severe speech and language disorder (Fisher, 2016; Fisher, Vargha-Khadem, Watkins, Monaco, & Pembrey, 1998) is a well-known example. But there are many other disorders with genetic bases where genomes of afflicted individuals shed light on genes involved in language (I will call these “language-related genes,” or LRGs). Relevant disorders can either be rather specific to components of language, like FOXP2, specific language impairment (van der Lely & Pinker, 2014), or dyslexia (Mozzi et al., 2016; St Pourcain et al., 2014; Wang, Chen, et al., 2015), or they can be broader disorders like autism, which have important consequences for, but are not specific to, language (Graham & Fisher, 2015; Raff, 2014; Rodenas-Cuadrado, Ho, & Vernes, 2014). A particularly interesting result involves CNTNAP2, a gene coding for a neuroligin specifically expressed in the human cortex and involved in cortical development. Initially discovered due to its interactions with FOXP2, variants in one portion of CNTNAP2 are associated with language disorders and autism (Arking et al., 2008; Rodenas-Cuadrado et al., 2014; Vernes et al., 2008). Clinical genetic research is now exploding due to the easy availability of individual genomes and the growth of individualized medicine, and the data that it generates will play a key role in understanding the complex genetic basis for language acquisition and processing.

One intriguing aspect of variation among living humans is that, due to the large size of the human population relative to the genome, virtually all mutations compatible with life probably exist, right now, in some human somewhere today (Pääbo, 2014). Thus, mutations in any gene thought to be involved in language could potentially be sought out, and the resulting phenotype studied. Although this may seem far-fetched at present, the rapid decrease in price and increase in usage of whole-genome sequencing means that such targeted searching may well be within reach in a decade.

But the most potentially exciting application of data from living humans is its use to derive inferences about our (relatively recent) evolutionary past. Now classic examples of this include the analysis of the recent evolution of lactose tolerance in some populations (northern Europe and Africa), which results from genetic variants that became common only in dairy farming populations. This is an excellent example of *gene-culture coevolution* (Burger, Kirchner, Bramanti, Haak, & Thomas, 2007; Tishkoff et al., 2007). Another nice example is the AMY1 gene, which codes for the starch-digesting enzyme amylase. In the ancestral state (e.g., in chimpanzees, Neanderthals, or some hunter-gatherers) there is a single copy

of this gene; but most living humans today have multiple copies. The number of copies correlates with reliance on a starch-rich diet (probably initially due to exploitation of roots and tubers, and later due to agriculture), with up to nine copies found in some contemporary populations. Such population-level differences in allele frequencies can be used to test hypotheses about early human movements, such as the origins and early movements of Indo-European language speakers into Europe and India, complementing historical linguistic data (Bouckaert et al., 2012; Cavalli-Sforza & Piazza, 1993; Gray & Atkinson, 2003). But again, none of these differences are likely to provide clues to genes involved specifically in language, because of the homogeneity of the human language faculty. Indeed, given that humans had already occupied many continents by around 50 KYA, with virtually no gene flow between (say) Australia and North America after then, no genes are likely to have achieved species-wide fixation since then. Any genes fixated at the population, but not the species, level can thus be presumed to have little relevance to the biological basis of the language faculty.

Present-day variation can also hold clues to the more distant past; the most prominent example being various “signatures of selection” that result from strong selection on a particular gene. Various measures have been proposed to quantify such within-species positive selection, including Tajima’s *D* or the Macdonald/Kreitman test (Nielsen, 2005). For example, strong positive selection can lead to a “selective sweep” where a particular allele goes to fixation (becomes present throughout the population, replacing any other alleles). Limited recombination around such genes means that the entire region of the chromosome surrounding the selected allele shows less variation than the background level (Hancock & Di Rienzo, 2008; Maynard Smith & Haigh, 1974). A careful analysis of such regions can support inferences about when in the past that gene went to fixation, and thus to (roughly) order the acquisition of different alleles (Nielsen, 2005). Unfortunately, recombination leads this signal to decay quite rapidly in evolutionary terms (Przeworski, 2002), so the time depth of this method is limited and will not allow us to peer back, in humans, more than about 100,000 years. Both selective sweeps and other proposed signatures of selection also tend to confound demographic phenomena with selection (Nielsen, 2005); in the case of humans, we know that there was a population bottleneck (a small founder population of our species, in Africa) followed by a massive, ongoing expansion of the human population over the last 100,000 years, so demographic effects are expected to be strong. All of these factors limit the utility of such population-level analyses for understanding language evolution.

The medium time scale: Paleo-DNA from archaic humans

All of the genetic data discussed above is exciting in that it provides unprecedented insights into evolution over long

(comparative) or short (modern human) time scales, based on genomes of living organisms. These data are central to attempts to understand the complex linkage between genotype and phenotype, which remains a key unsolved challenge in contemporary biology. But the time scales involved leave a gap between around 6 MYA (our divergence with chimpanzees) and about 100 KYA (the rough limit to backward extrapolation from existing humans). This is unfortunate, because this is precisely the time period within which language evolution occurred in our species, and whatever genes make language possible became fixated in the ancestral population(s) of modern humans. Thus, it is extremely auspicious that a new source of data is available for precisely this time period: “paleo-DNA” retrieved from the bones of now-extinct species.

We now have excellent-quality genomes for two extinct archaic hominins, the Neanderthals and Denisovans. These are two related species of humans, whose homelands were Europe and Asia, that evolved independently of modern *Homo sapiens*, in these regions. Neanderthals have a rich fossil and archaeological record, while Denisovans were unknown until their genome was sequenced (and thus represent the first fossil hominin species to be discovered based on DNA alone). I will focus on Neanderthals. Although the bones from which these genomes come are not, themselves, very old (about 50,000 years), they push back the time span covered by genomic approaches to roughly 500,000 years, when modern human populations split away from the Neanderthal/Denisovan lineage.

Because of the rich archaeological record for Neanderthals, we know quite a lot about their culture and its development over time (see this issue’s contribution by Tattersall, 2016). These were very robust and powerful humans, with brain sizes slightly larger in absolute terms than those of modern humans, and with a sophisticated toolkit that enabled a hunter-gatherer lifestyle in the tundra-like conditions of Pleistocene Europe. Despite the excellent record, and their obvious cognitive sophistication, Neanderthals did not show the exponential rate of cultural change characterizing our own species from about 80,000 years ago to the present (Mellars, 1998b, 2004; Tattersall, 1999). As already noted, most archaeologists and anthropologists conclude that, despite their many similarities to us, Neanderthals lacked some cognitive and/or linguistic features of our species (Gunz, Neubauer, Maureille, & Hublin, 2010; Hublin, 2009; Mellars, 1998a; Mithen, 2005; Schepartz, 1993; Shea, 2003; Wynn & Coolidge, 2004; for a dissenting view, see Dediu & Levinson, 2013). While I agree with the premise that Neanderthals lacked *something*, I think it unlikely that they lacked language entirely; rather, a multi-component perspective suggests that they already possessed some aspects of modern language and cognition, but lacked one or more others. In other words, Neanderthals possessed some form of protolanguage, representing a crucial

intermediate step between our common ancestor with chimpanzees and modern humans. The key question then becomes which components of language were shared and which differed, and this is a question that paleo-DNA data are uniquely capable of answering.

Analyses that compare Neanderthal and modern human DNA provide an invaluable window into language evolution. If we take just those genetic differences in which Neanderthals share the ancestral allele with chimpanzees, but all modern humans share a novel, derived allele, we come up with a rather short and manageable list (Prüfer et al., 2014): Although there are about 31,000 base pairs that differ, only about 3,000 are within gene regulatory regions and thus reckoned to influence gene expression, and there are only 96 fixed amino acid changes in 87 different proteins. A disproportionate number of these genes code for proteins involved in brain development, and thus represent plausible candidates for cognition- and language-related genes. Crucially (unlike the chimpanzee/human difference list), this list is short enough that each of these genetic differences can be explored using modern molecular methods (e.g., inserting the modern human allele into genetically engineered cell lines of mice and observing resulting phenotypic differences).

Perhaps the most noteworthy early result of this paleo-DNA work is the finding that Neanderthals (and Denisovans) shared with us the modern derived variant of the FOXP2 gene, indicating (contra the original predictions of Enard et al., 2002, based on selective-sweep methods) that the modern variant evolved and was fixed at least 500 KYA, before the split between us and Neanderthals. Given the importance of FOXP2 in human oro-motor control, particularly of complex sequences (Alcock, Passingham, Watkins, & Vargha-Khadem, 2000; Vargha-Khadem, Gadian, Copp, & Mishkin, 2005), this suggests that Neanderthals already possessed some form of speech. Nonetheless, humans differ from Neanderthals in a regulatory region of the FOXP2 gene, a binding site for the transcription factor POU3F2 (Maricic et al., 2013), which suggests some changes in FOXP2 expression, potentially relevant to spoken language, evolved after our split from Neanderthals. These results support the notion that Neanderthals had some components of spoken language, but not others, and that sophisticated vocal control and speech evolved earlier in human evolution than full modern language; they thus provide support for musical protolanguage hypotheses like that of Darwin and many others. They are evidence against the prediction of the gestural protolanguage hypotheses that speech was a late acquisition, occurring only after other components of language (e.g., syntax and semantics) were already in place (cf. Fitch, 2010).

Despite this exciting progress in testing hypotheses about language evolution, clearly a single gene comparison cannot resolve the myriad debates revolving around such hypotheses,

and many more LRGs will need to be understood before any firm conclusions can be drawn (cf. Mozzi et al., 2016). But the critical point illustrated by the fact that Neanderthals, unlike chimpanzees, shared our version of FOXP2 is that paleo-DNA allows us in principle to test specific evolutionary hypotheses and address long-standing debates in a way that, only a decade ago, would have been unthinkable. The remaining difficulties at this point are not access to the genetic data, which are available free to everyone, but rather understanding the still vexing complexity of the mapping between genes and phenotypic traits of interest in language variants. The good news is that a very large research community, mostly made up of clinicians and molecular biologists with no specific interest in language evolution, is already hard at work resolving those issues. Every increase in our understanding of particular gene variants relevant to language, cognition, or the brain can now be immediately checked against the Neanderthal genome to see whether modern human variants evolved before or after our split from this extinct species (and to support further inferences about the kind of protolanguage Neanderthals spoke).

Synthesis: A staged-protolanguages model of language evolution

To show how the overall hypothesis testing framework outlined above can be concretely put into action, I will now offer my own new model of language evolution, a model that proposes an order of acquisition for each of the key components introduced in the “What evolved?” section and explains in what contexts they were adaptive. I will build freely upon the previous models of protolanguage reviewed in section “Models of language evolution” and offer concrete testable predictions based on the types of evidence discussed in the “Empirical data” section. My aim here is illustrative, to show that a model *can* be constructed that is consistent with all available data, and that makes clear testable predictions. I give attention to and respect for previous research, trying to highlight areas of agreement and disagreement with previous workers.

My model is composed of multiple separate hypotheses about key innovations in language evolution. It proposes four clear stages since our divergence from our LCA with chimpanzees, each one of them associated with the acquisition of one or more specific key novel capabilities: “key innovations” on the way to modern language (cf. Liem, 1973, 1990). I start with a brief overview of these proposed stages and then go into more detail about how my proposal differs (or borrows) from those of others.

Stage 1 Vocal learning: Singing *Australopithecus*

The first stage involved the acquisition of vocal learning capabilities to generate learned vocal sequences without any

propositional meaning: a “prosodic protolanguage” (I tentatively suggest that this stage was associated with *Australopithecus*).

Stage 2 Mimesis: Mimetic *Homo erectus*

In the next, crucial, stage, this vocal learning capacity combined with the preexisting gestural abilities of the LCA to form a richer and more elaborated mimetic protolanguage (Donald, 2016) that combined gesture and expressive learned vocalizations in shared group rituals and information exchange. During this stage, which I associate with *Homo erectus*, pressure to learn and elaborate both vocal and manual sequences left its traces in the more elaborated technology of the Achulean—a one million year period during which this mimetic protolanguage was the main communication system, and these highly successful hominins expanded into the entire Old World. Because of these deep roots, clear traces of this stage remain evident in modern humans in music, dance, mime, and their expression in group rituals. The selective pressures during this stage were essentially about group bonding and mate choice among adults, but children would certainly have participated fully in these activities.

Stage 3 Propositional meaning: Semantic *Homo antecessor*

Although the previous two stages find clear analogs in the animal world, the next proposed stage is unique to the hominid lineage: The key innovation was use of the complex, socially shared mimetic sequences to share *propositional* meanings for the first time. Previous mimetic sequences were meaningful only in the sense that music and dance are: they connote occasions and express moods or aesthetic trajectories, but cannot convey specific abstract thoughts. I suggest that the next (third) stage was associated with the last common ancestor of Neanderthals and humans (*Homo heidelbergensis/antecessor*). The driving force for this key innovation, which added explicit semantics to a preexisting mimetic communication system, was sharing detailed propositional information with close kin, especially between parents and their offspring (this was thus a “mother tongue,” driven by kin selection to raise inclusive fitness; Fitch, 2004), with close similarities to what Kevin Laland terms “teaching” (Laland, 2016, 2017). At this point, the information-transmission capabilities of our ancestors made a huge leap, but protolanguage was still limited in communicative scope by its restriction to unsuspecting use with close kin.

From a structural viewpoint, this protolinguistic system would retain the sequential structures typifying the previous mimetic stage, but these sequences would already reflect a certain amount of hierarchical structure stemming from the preexisting hierarchical structure of thought, without there being any specific syntactic markers of such structure. I posit that this propositional protolanguage continued to be the

system used by Neanderthals and Denisovans. The crucial problem solved at this stage was “honest” communication of accurate information to others, the evolution of which is hard to explain among unrelated adults, but easy to understand in the context of knowledgeable adults sharing information with their offspring and close kin (cf. Fitch, 2004).

Stage 4 Syntactic *Homo sapiens*

The fourth and final stage—modern, fully syntactic language—occurred only in anatomically modern *Homo sapiens*, sometime between 200 and 80 KYA. In stage-three propositional protolanguage, sequences reflected hierarchical structure only accidentally: signalers attempting to express hierarchical thoughts might unconsciously provide cues to the hierarchy (e.g., via word order, or the pauses and prolongations typical of mimetic protolanguage). During this last stage, a rich social communicative “ecosystem” appeared within which children were embedded, creating competition for successful acquisition of the information in these signals, putting selective pressure on children to rapidly and successfully acquire and generalize them (cf. Locke, 2016). I suggest that it was during this last stage that human “dendrophilia” (Fitch, 2014)—our ability and propensity to perceive hierarchical structure in sequences—arose. While this “mental tree reading” ability would initially be selected in children, as adults it would not disappear, and it was in the new context of propositional information transfer among *unrelated adults* that the Machiavellian side of hierarchical mind-reading had its most telling effects, for it began the upward spiral of multiply embedded theory of mind, epistemic vigilance (essentially skepticism and distrust), and tit-for-tat reciprocal sharing of valued information among nonkin that still typifies our species today.

As will be clear to those familiar with the literature, the “SMSS” (song, mimesis, semantics, syntax) model proposed above builds upon and synthesizes many previous ideas about language evolution, and each of the individual stages, key innovations, and selective pressures has been previously discussed. The first “prosodic protolanguage” stage is closely allied to Darwin’s musical protolanguage hypothesis. The key innovation of this first stage was vocal learning, driven by much the same pressures that drove vocal learning in many other vocal learning species, and requiring the mechanistic acquisition of direct cortico-motor connections (Jürgens, 2002). During this “singing ape” stage, sequencing of learned signals was already in place, adding a selective pressure for rapid and accurate sequence learning as typified by many modern songbirds (Jarvis, 2004b; Kroodsma, 2005; Marler & Slabbekoorn, 2004). But the signals produced, though complex and shared among groups, were not “music” in the modern sense (for discussion, see Fitch, 2013). In the second “mimetic protolanguage” stage, vocal learning combined with preexisting gestural capabilities to yield a much richer

communicative system; although this combines aspects of music and gestural protolanguages, it differs from them in not proposing any point at which gesture alone was central: learned vocal displays were there from the beginning. In this I closely follow Donald, Kendon, and others, with acknowledgement of Arbib’s more nuanced “upward spiral” conception of gestural-vocal interaction (Arbib, 2005, 2016).

Once an elaborate, socially shared but nonpropositional communication system is in place, the major problem is how propositional semantics could be added to such a system. In Jespersen’s words (citing Wilhelm von Humboldt): “How did man become, as Humboldt somewhere defined him, ‘a singing creature, only associating thoughts with the tones?’” (Jespersen, 1922; von Humboldt, 1836). Humboldt’s original phrase was “because man, as a type of animal, is a singing creature, but with thoughts bound to the tones” (my translation, p. 76). This is the major problem faced by musical protolanguage theories, and there can be little doubt that both the mimetic potential of gesture and the onomatopoeic nature of sound helped pave the way for this key innovation (Blasi et al., 2016). The key questions then are “why only us?”—of all singing creatures, why are we the only ones to do this—and “what selective pressures?”—what would drive signalers to honestly communicate valuable information (and thus perceivers to attend to them)? I have previously argued that communication among kin was the key selective pressure on signalers, and that the existence of large bodies of learned knowledge was the key selective pressure on perceivers (Fitch, 2004, 2007). Kevin Laland extends this argument in the current issue, and although I find his term “teaching” somewhat misleading as it connotes, to me, a conscious and formal intention to teach, we essentially agree concerning the basic evolutionary story and the problems solved by this approach (see also Nowicki & Searcy, 2014). To my knowledge, my idea that this “mother tongue” was the communicative system typifying *Homo heidelbergensis/antecessor*, and later Neanderthals and Denisovans, is novel.

The basic hypothesis that the final stage of language evolution involved the acquisition of full modern hierarchically embedded syntax is shared by many previous writers (notably Bickerton, 1990; Chomsky, 2010; Jackendoff & Pinker, 2005). While incorporating many of the insights of these previous authors, my hypothesis differs from them in the following ways. First, I do not posit any “one word” stage of language evolution: Protolanguages always involved sequences, and words were “condensed out” of holistic utterances by learners, rather than first appearing as isolated building blocks to be put together later. Thus, in terms of word origins, mine is an “analytic” model, rather than a “synthetic” model (Arbib, 2005; Wray, 2000). Furthermore I posit that much of the syntactic work needed for modern language had already been done via prolonged selection for sequence processing and generation in the previous stages—mimetic and propositional

protolanguages already had essentially all of the neural mechanisms required for modern spoken phonology and prosody in place. Thus, the key innovation at this stage is *dendrophilia* (Fitch, 2014)—a domain-general proclivity to perceive hierarchical structure that came to be applied not only to language but also to music and decorative arts. Here I concur with Chomsky (2010), Berwick and Chomsky (2016) and Chomsky (2016) that it is the flexible and unrestricted capacity for hierarchical embedding that is central to our species' cognitive uniqueness; I differ from them in seeing it as something that was flexible enough to immediately play a role in both structured thought *and* linguistic communication, as well as in other domains (it is only at this stage that music in its modern sense, replete with hierarchical structure, was born; cf. Patel, 2016). Elsewhere, I have explored the neural changes that were necessary to achieve this final stage of dendrophilia (Fitch, 2014)—but in short, dendrophilia requires both extensive connections between temporal and parietal areas and the prefrontal regions surrounding Broca's area (the arcuate fasciculus) as well as a great expansion of Broca's areas (Friederici, 2016; Rilling et al., 2008; Schenker et al., 2010). Finally, I agree with Keller (1995), Deacon (1997), Heine and Kuteva (2002), Steels (2016), Kirby (2017), and many others that much of the complexity evident in the syntax of modern languages has arisen repeatedly by grammaticalization processes of cultural evolution and required no further neural changes beyond those needed for dendrophilia.

In summary, the model proposed above provides a synthesis of many previous ideas about language evolution, incorporating multiple hypotheses from the literature about the order in which the different key innovations of language arose, what neural changes were needed, and why these were selectively advantageous at that time. This model makes a slew of testable predictions, and cashing most of them out will only require a better understanding of neural mechanisms and their genetic basis (normal science, with living organisms) because then we can use the existing genomes of Neanderthals and Denisovans (and hopefully in the future *Homo heidelbergensis* and *erectus*) to validate or invalidate the predictions. Indeed, some of these predictions (e.g., that Neanderthals would share the vocal sequencing capabilities of modern humans) are already supported by the fact that they shared our modified FOXP2 gene (see above and cf. Fisher, 2016). However, most of the predictions, for example that Neanderthals had poor theory of mind (were less Machiavellian) and lacked dendrophilia, will require a better understanding of the genetic basis of these traits (e.g., from the study of autistic spectrum disorder for theory of mind, or the development of the arcuate fasciculus for dendrophilia). But again, these advances are part of modern clinical and developmental genetics: no time machines are required to test these hypotheses.

Conclusions

In this introductory review, I obviously could not list every idea about language evolution or every source of data relevant to its study, but I do hope that this review of hypotheses and data whets the reader's appetite for more, and illustrates a few key points. First, contrary to an oft-stated opinion, there is a refreshingly large volume of data relevant to language origins once the problems and models have been clearly stated, and an open-minded approach based on strong inference is adopted. In particular, different models of language evolution make different predictions in multiple empirical domains, and data capable of discriminating among such well-specified models are often either available or can be gathered with current methods.

Second, the relevant data are almost bewilderingly diverse and voluminous: they span a set of disciplines that no single scholar, however knowledgeable, could hope to individually master. This means that future progress in understanding the evolution of language requires productive collaborations between researchers from different disciplines, something that is easier to talk about than to actually do.

Finally, and most excitingly, this huge amount of data flooding in offers a unique new promise for contemporary researchers. These data are often generated by researchers who have no particular interest in language evolution *per se*, but see themselves rather as broadly studying cognitive neuroscience, paleo-genomics, brain evolution, theoretical linguistics, or human (cultural) evolution. For those willing and able to educate themselves about these data and the issues specific to language evolution, and who adopt a hypothesis-testing framework, such data can allow real progress. A clear example is the abundant genetic data from many living species (including thousands of individual humans) as well as an ever-increasing number of extinct hominids—data freely available online to any interested person. There is also a steady growth in other relevant freely available data (e.g., the WALS database of linguistic structure; Haspelmath, Dryer, Gil, & Comrie, 2005). Thus, despite the technical challenges of accessing and understanding such data, any scholar or team of scholars can in principle use this body of knowledge without undertaking the vast cost and effort of gathering it themselves. The study of language evolution has reached the "big data" stage, and harnessing this will require significant changes in our scientific attitude to developing and testing models of language evolution.

In the final section, I offered an illustration of the kind of theorizing that will be needed to effect these changes: comprehensive models that seek to explain *all* of the key innovations that needed to occur to yield modern language, considered from mechanistic, phylogenetic and adaptive viewpoints. Such models—and mine is only one of many plausible suggestions—consist of multiple specific, testable

hypotheses. The data available for such testing include comparative, linguistic, neural, and paleontological findings, but progress will lean heavily on genomics since the closest thing available today to a time machine is DNA, especially the paleo-DNA from extinct hominins. It is unfortunate that these genomic data do not (currently) extend further back in time (e.g., to *Homo erectus* or *Australopithecus*), but we shouldn't look this gift horse in the mouth. Never before have hypotheses about the linguistic capabilities of Neanderthals been so clearly and empirically testable.

To conclude, the evolution of language has been termed “the most challenging scientific problem of our time”, and from a sociological point of view, the interdisciplinary collaboration that it requires is indeed a challenge. But it is also one of the most interesting problems in modern science, and concerns an issue absolutely central to understanding our own species: how we rose from being an ecologically peripheral African ape, a few million years ago, to the most important (and dangerous) species alive on our planet today. The work reviewed here, and the remaining articles in this special issue, suggest that a deep understanding of this ancient problem may be attainable in the next few decades.

References

- Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, *14*(8), 1067–1074.
- Adger, D. (2017). Restrictiveness matters. *Psychonomic Bulletin & Review*. (In this issue)
- Adret, P. (1992). Vocal learning induced with operant techniques: An overview. *Netherlands Journal of Zoology*, *43*(1/2), 125–142.
- Alcock, K., Passingham, R. E., Watkins, K. E., & Vargha-Khadem, F. (2000). Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain and Language*, *75*, 17–33.
- Anderson, D. R. (2008). *Model based inference in the life sciences: A primer on evidence*. New York, NY: Springer.
- Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., & Knösche, T. R. (2007). Connectivity-based parcellation of Broca's area. *Cerebral Cortex*, *17*(4), 816–825.
- Arbib, M. A. (2002). The mirror system, imitation, and the evolution of language. In C. Nehaniv & K. Dautenhahn (Eds.), *Imitation in animals and artifacts* (pp. 229–280). Cambridge, MA: MIT Press.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, *28*(2), 105–167.
- Arbib, M. A. (2016). Toward the language-ready brain: Biological evolution and primate comparisons. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1098-2
- Arendt, D. (2005). Genes and homology in nervous system evolution: Comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory in Biosciences*, *124*(2), 185–197.
- Arking, D. E., Cutler, D. J., Brune, C. W., Teslovich, T. M., West, K., Ikeda, M., ... Chakravarti A. (2008). A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *American Journal of Human Genetics*, *82*(1), 7–9.
- Barney, A., Martelli, S., Serrurier, A., & Steele, J. (2012). Articulatory capacity of Neanderthals, a very recent and human-like fossil hominin. *Philosophical Transactions of the Royal Society B*, *367*, 88–102.
- Bar-On, D. (2013). Origins of meaning: Must we ‘Go Gricean’? *Mind & Language*, *28*(3), 342–375.
- Bateson, W. (1894). *Materials for the study of variation treated with especial regard to discontinuity in the origin of species*. London, UK: Macmillan.
- Beckers, G. J. L., Bolhuis, J. J., Okanoya, K., & Berwick, R. C. (2012). Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *NeuroReport*, *23*(3), 139–145.
- Bemis, D. K., & Pyllkkänen, L. (2012). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, *23*(8), 1859–1873. doi:10.1093/cercor/bhs170
- Bemis, D. K., & Pyllkkänen, L. (2013). Combination across domains: An MEG investigation into the relationship between mathematical, pictorial, and linguistic processing. *Frontiers in Psychology*, *3*, 583. doi:10.3389/fpsyg.2012.00583
- Bermúdez de Castro, J. M., Arsuaga, J. L., Carbonell, E., Rosas, A., Martínez, I., & Mosquera, M. (1997). A hominid from the lower Pleistocene of Atapuerca, Spain: Possible ancestor to Neandertals and modern humans. *Science*, *276*(5317), 1392–1395.
- Berwick, R. C. (1997). Syntax facit saltum: Computation and the genotype and phenotype of language. *Journal of Neurolinguistics*, *10*(2), 231–249.
- Berwick, R. C. (1998). Language evolution and the minimalist program: The origins of syntax. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language* (pp. 320–340). New York, NY: Cambridge University Press.
- Berwick, R. C., & Chomsky, N. (2016). *Why Only Us: Language and Evolution*. Cambridge, Massachusetts: MIT Press.
- Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*, *17*(2), 89–98. doi:10.1016/j.tics.2012.12.002
- Bickerton, D. (1990). *Language and species*. Chicago, IL: Chicago University Press.
- Bickerton, D. (2000a). Resolving discontinuity: A minimalist distinction between human and non-human minds. *American Zoologist*, *40*, 862–873.
- Bickerton, D. (2000b). How protolanguage became language. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 264–284). Cambridge, UK: Cambridge University Press.
- Blasi, D. E., Wichmann, S., Hammerström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound-meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*. *113*(39), 10818–10823. doi: 10.1073/pnas.1605782113
- Boë, L.-J., Heim, J.-L., Honda, K., & Maeda, S. (2002). The potential Neandertal vowel space was as large as that of modern humans. *Journal of Phonetics*, *30*(3), 465–484.
- Boeckx, C. (2016). The language-ready head: Evolutionary considerations. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1087-5
- Botha, R. (2009). On musilanguage/“Hmmmm” as an evolutionary precursor to language. *Language & Communication*, *29*(1), 61–76.
- Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummond, A. J., ... Atkinson, Q. D. (2012). Mapping the origins and expansion of the first Indo-Europeans. *Science*, *337*(6097), 957–960. doi:10.1126/science.1219669
- Bowling, D. (2017). The continuing legacy of nature versus nurture in biolinguistics. *Psychonomic Bulletin & Review*. (In this issue)
- Boyd, R., & Richerson, P. J. (1996). Why culture is common but cultural evolution is rare. *Proceedings of the British Academy*, *88*, 77–93.

- Brainard, M. S., & Fitch, W. T. (2014). Editorial overview: Communication and language: Animal communication and human language. *Current Opinion in Neurobiology*, 28, v–viii.
- Bräuer, J., Call, J., & Tomasello, M. (2007). Chimpanzees really know what others can see in a competitive situation. *Animal Cognition*, 10, 439–448.
- Brown, S. (2000). The “musilanguage” model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). Cambridge, MA: MIT Press.
- Bugnyar, T., Reber, S. A., & Buckner, C. (2016). Ravens attribute visual access to unseen competitors. *Nature Communications*, 10506. doi:10.1038/ncomms10506
- Burger, J., Kirchner, M., Bramanti, B., Haak, W., & Thomas, M. G. (2007). Absence of the lactase-persistence associated allele in early Neolithic Europeans. *Proceedings of the National Academy of Sciences*, 104, 3736–3741.
- Burkart, J. M., Hrdy, S. B., & van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology*, 18(5), 175–186.
- Byrne, R. W., & Cochet, H. (2016). Where have all the (ape) gestures gone? *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1071-0
- Byrne, R. W., & Whiten, A. (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford, UK: Clarendon Press.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 year later. *Trends in Cognitive Sciences*, 12(5), 187–192.
- Carstairs-McCarthy, A. (1999). *The origins of complex language*. Oxford, UK: Oxford University Press.
- Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state. *Cortex*, 44(8), 953–961.
- Cavalli-Sforza, L. L., & Piazza, A. (1993). Human genomic diversity in Europe: A summary of recent research and prospects for the future. *European Journal of Human Genetics*, 1(1), 3–18.
- Chamberlin, T. C. (1890). The method of multiple working hypotheses. *Science*, 148, 754–759.
- Charrier, C. (2012). Inhibition of SRGAP2 function by its human-specific paralogs induces neoteny during spine maturation. *Cell*, 149, 923–935.
- Cheney, D. L., & Seyfarth, R. M. (2007). *Baboon metaphysics: The evolution of a social mind*. Chicago, IL: University of Chicago Press.
- Chomsky, N. (1957). *Syntactic structures*. The Hague, Netherlands: Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2010). Some simple evo devo theses: How true might they be for language? In R. Larson, V. Deprez, & H. Yamakido (Eds.), *The evolution of human language: Biolinguistic perspectives* (pp. 45–62). Cambridge, UK: Cambridge University Press.
- Chomsky, N. (2016). The language capacity: Architecture and evolution. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1078-6
- Christiansen, M., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31, 489–509.
- Christiansen, M., & Kirby, S. (Eds.). (2003). *Language evolution*. Oxford, UK: Oxford University Press.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49(12), 997–1003.
- Collin, S. P., & Trezise, A. E. O. (2004). The origins of colour vision in vertebrates. *Clinical and Experimental Optometry*, 87(4/5), 217–223.
- Corballis, M. C. (2002). *From hand to mouth: The origins of language*. Princeton, NJ: Princeton University Press.
- Coutinho, C. C., Fonseca, R. N., Mansurea, J. J. C., & Borojevic, R. (2003). Early steps in the evolution of multicellularity: Deep structural and functional homologies among homeobox genes in sponges and higher metazoans. *Mechanisms of Development*, 120, 429–440.
- Crystal, D. (2003). *The Cambridge encyclopedia of language* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Darwin, C. (1859). *On the origin of species*. London, UK: John Murray.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. London, UK: John Murray.
- Dawkins, R. (2004). *The ancestor's tale*. New York, NY: W. W. Norton and Co.
- de Boer, B. (2010). Investigating the acoustic effect of the descended larynx with articulatory models. *Journal of Phonetics*, 38(4), 679–686.
- de Boer, B. (2016). Evolution of speech and evolution of language. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1130-6
- Deacon, T. W. (1992). The neural circuitry underlying primate calls and human language. In J. Wind, B. A. Chiarelli, B. Bichakjian, & A. Nocentini (Eds.), *Language origins: A multidisciplinary approach* (pp. 301–323). Dordrecht, Germany: Kluwer Academic.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York, NY: Norton.
- Dediu, D., & Levinson, S. C. (2013). On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology*, 4(397), 1–17.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44(1/2), 1–42.
- Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The neural representation of sequences: From transition probabilities to algebraic patterns and linguistic trees. *Neuron*, 88, 2–19.
- Dehaene-Lambertz, G. (2017). The human infant brain: A neural architecture able to learn language. *Psychonomic Bulletin & Review*. (In this issue)
- Dehaene-Lambertz, G., & Spelke, E. S. (2015). The infancy of the human brain. *Neuron*, 88, 93–109.
- Denton, M. (1985). *Evolution: A theory in crisis*. Bethesda, MD: Adler and Adler.
- Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard University Press.
- Donald, M. (2016). Key cognitive preconditions for the evolution of language. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1102-x
- Dunbar, R. (1996). *Grooming, gossip and the evolution of language*. Cambridge, MA: Harvard University Press.
- Dunbar, R. I. M. (2016). Group size, vocal grooming and the origins of language. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1122-6
- Elder, J. H. (1934). Auditory acuity of the chimpanzee. *Journal of Comparative and Physiological Psychology*, 17, 157–183.
- Eldredge, N., & Gould, S. J. (1972). Punctuated equilibria: An alternative to phyletic gradualism. In T. J. M. Schopf (Ed.), *Models in paleobiology* (Vol. 3, pp. 115–151). San Francisco, CA: Freeman.
- Emery, N. J. (2006). Cognitive ornithology: The evolution of avian intelligence. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 361, 23–43.
- Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, 306, 1903–1907.
- Emmorey, K. (2002). *Language, cognition and the brain: Insights from sign language research*. London, UK: Erlbaum.
- Emmorey, K. (2005). Sign languages are problematic for a gestural origins theory of language evolution. *Behavioral and Brain Sciences*, 28(2), 130–131.
- Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., ... Paäbo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, 418, 869–872.

- Endicott, P., Ho, S. Y. W., & Stringer, C. (2010). Using genetic evidence to evaluate four palaeoanthropological hypotheses for the timing of Neanderthal and modern human origins. *Journal of Human Evolution*, *59*(1), 87–95.
- Evans, C. S., & Marler, P. (1994). Food-calling and audience effects in male chickens, *Gallus gallus*: Their relationships to food availability, courtship and social facilitation. *Animal Behaviour*, *47*, 1159–1170.
- Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences*, *18*(3), 120–126.
- Fedorenko, E., & Varley, R. (2016). Language and thought are not the same thing: Evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, *1369*(1), 132–153.
- Fehér, O. (2016). Atypical birdsong and artificial languages provide insights into how communication systems are shaped by learning, use, and transmission. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1107-5
- Fehér, O., Wang, H., Saar, S., Mitra, P. P., & Tchernichovski, O. (2009). *De novo* establishment of wild-type song culture in the zebra finch. *Nature*, *459*, 564–568.
- Fischer, J. (2016). Primate vocal production and the riddle of language evolution. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1076-8
- Fischer, J., Wheeler, B. C., & Higham, J. P. (2015). Is there any evidence for vocal learning in chimpanzee food calls? *Current Biology*, *25*(21), 1028–1029.
- Fisher, S. E. (2016). Evolution of language: Lessons from the genome. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1112-8
- Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics*, *18*(2), 168–170.
- Fitch, W. T. (2000a). The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals. *Phonetica*, *57*, 205–218.
- Fitch, W. T. (2000b). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*(7), 258–267.
- Fitch, W. T. (2004). Kin selection and “mother tongues”: A neglected component in language evolution. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 275–296). Cambridge, MA: MIT Press.
- Fitch, W. T. (2005). The evolution of language: A comparative review. *Biology and Philosophy*, *20*, 193–230.
- Fitch, W. T. (2007). Evolving meaning: The roles of kin selection, allomothering and paternal care in language evolution. In C. Lyon, C. Nehaniv, & A. Cangelosi (Eds.), *Emergence of communication and language* (pp. 29–51). New York, NY: Springer.
- Fitch, W. T. (2008). Glossogeny and phylogeny: Cultural evolution meets genetic evolution. *Trends in Genetics*, *24*(8), 373–374.
- Fitch, W. T. (2009a). Fossil cues to the evolution of speech. In R. P. Botha & C. Knight (Eds.), *The cradle of language* (pp. 112–134). Oxford, UK: Oxford University Press.
- Fitch, W. T. (2009b). The biology & evolution of language: “Deep Homology” and the evolution of innovation. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences IV* (pp. 873–883). Cambridge, MA: MIT Press.
- Fitch, W. T. (2010). *The evolution of language*. Cambridge, UK: Cambridge University Press.
- Fitch, W. T. (2011a). “Deep homology” in the biology & evolution of language. In A. M. Di Sciullo & C. Boeckx (Eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty* (pp. 135–166). Oxford, UK: Oxford University Press.
- Fitch, W. T. (2011b). Innateness and human language: A biological perspective. In M. Tallerman & K. R. Gibson (Eds.), *The Oxford handbook of language evolution* (pp. 143–156). Oxford, UK: Oxford University Press.
- Fitch, W. T. (2011c). Genes, language, cognition, and culture: Towards productive inquiry. *Human Biology*, *83*(2), 323–329.
- Fitch, W. T. (2011d). Unity and diversity in human language. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *366*, 376–388.
- Fitch, W. T. (2011e). The evolution of syntax: An exaptationist perspective. *Frontiers in Evolutionary Neuroscience*, *3*(9), 1–12.
- Fitch, W. T. (2012). Evolutionary developmental biology and human language evolution: Constraints on adaptation. *Evolutionary Biology*, *39*(4), 613–637.
- Fitch, W. T. (2013). Musical protolanguage: Darwin’s theory of language evolution revisited. In J. J. Bolhuis & M. B. H. Everaert (Eds.), *Birdsong, speech and language: Exploring the evolution of mind and brain*. Cambridge, MA: MIT Press.
- Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews*, *11*(3), 329–364.
- Fitch, W. M., & Ayala, F. J. (1994). Tempo and mode in evolution. *Proceedings of the National Academy of Sciences*, *91*(15), 6717–6720.
- Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B*, *367*, 1933–1955.
- Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, *97*(2), 179–210.
- Fitch, W. T., Huber, L., & Bugnyar, T. (2010). Social cognition and the evolution of language: Constructing cognitive phylogenies. *Neuron*, *65*, 795–814.
- Fitch, W. T., & Jarvis, E. D. (2013). Birdsong and other animal models for human speech, song, and vocal learning. In M. A. Arbib (Ed.), *Language, music, and the brain: A mysterious relationship* (Vol. 10, pp. 499–539). Cambridge, MA: MIT Press.
- Fitch, W. T., Mathur, N., de Boer, B., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, *2*(12), e1600723. doi:10.1126/sciadv.1600723
- Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London B*, *268*(1477), 1669–1675.
- Fitch, W. T., & Zuberbühler, K. (2013). Primate precursors to human language: Beyond discontinuity. In E. Zimmerman, S. Schmidt, & E. Altenmüller (Eds.), *The evolution of emotional communication* (pp. 26–48). Oxford, UK: Oxford University Press.
- Fitzpatrick, S. (2008). Doing away with Morgan’s canon. *Mind & Language*, *23*(2), 224–246.
- Frentiu, F. D., Bernard, G. D., Cuevas, C. I., Sison-Mangus, M. P., Prudic, K. L., & Briscoe, A. D. (2007). Adaptive evolution of color vision as seen through the eyes of butterflies. *Proceedings of the National Academy of Sciences*, *104*(Suppl. 1), 8634–8640.
- Friederici, A. D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences*, *13*(4), 175–181.
- Friederici, A. D. (2016). Evolution of the neural language network. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1090-x
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*, 1204–1207.
- Givón, T. (2002). *Bio-linguistics: The Santa Barbara lectures*. Amsterdam, Netherlands: John Benjamins.
- Gohau, G. (1990). *A history of geology* (A. V. Carozzi & M. Carozzi, Trans.). New Brunswick, NJ: Rutgers University Press.
- Goldin-Meadow, S. (2016). What the hands can tell us about language emergence. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1074-x
- Gould, S. J. (1982). Punctuated equilibrium—A different way of seeing. In J. Chermas (Ed.), *Darwin up to date* (Vol. 6, pp. 119–130). London, UK: IPC Magazines.

- Gould, S. J. (1985). Not necessarily a wing: Which came first, the function or the form? *Natural History*, 94(10), 12–25.
- Gould, S. J., & Eldredge, N. (1977). Punctuated equilibria: The tempo and mode of evolution reconsidered. 3, 115–151.
- Gould, S. J., & Vrba, E. S. (1982). Exaptation—A missing term in the science of form. *Paleobiology*, 8, 4–15.
- Graham, S. A., & Fisher, S. E. (2015). Understanding language from a genomic perspective. *Annual Review of Genetics*, 49, 131–160.
- Gray, R. D., & Atkinson, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*, 426(6965), 435–439.
- Grice, H. P. (1975). Logic and conversation. In D. Davidson & G. Harman (Eds.), *The logic of grammar* (pp. 64–153). Encino, CA: Dickenson.
- Gunz, P., Neubauer, S., Maureille, B., & Hublin, J.-J. (2010). Brain development after birth differs between Neanderthals and modern humans. *Current Biology*, 20(21), R921–R922. doi:10.1016/j.cub.2010.10.018
- Gyger, M., & Marler, P. (1988). Food calling in the domestic fowl (*Gallus gallus*): The role of external referents and deception. *Animal Behaviour*, 36, 358–365.
- Hamilton, L. S., Sohl-Dickstein, J., Huth, A. G., Carels, V. M., Deisseroth, K., & Bao, S. (2013). Optogenetic activation of an inhibitory network enhances feedforward functional connectivity in auditory cortex. *Neuron*, 80, 1066–1076.
- Hancock, A. M., & Di Rienzo, A. (2008). Detecting the genetic signature of natural selection in human populations: Models, methods, and data. *Annual Review of Anthropology*, 37, 197–217.
- Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59(4), 771–785.
- Hare, B., & Tomasello, M. (2004). Chimpanzees are more skillful in competitive than cooperative cognitive tasks. *Animal Behaviour*, 68, 571–581.
- Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford, UK: Oxford University Press.
- Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (2005). *The world atlas of linguistic structures*. Oxford, UK: Oxford University Press.
- Hauser, M., Chomsky, N., & Fitch, W. T. (2002). The language faculty: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.
- Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. D., Watumull, J., ... Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*. doi:10.3389/fpsyg.2014.00401
- Hayes, C. (1951). *The ape in our house*. New York, NY: Harper.
- Heimbauer, L. A., Beran, M. J., & Owren, M. J. (2011). A chimpanzee recognizes synthetic speech with significantly reduced acoustic cues to phonetic content. *Current Biology*, 21(14), 1210–1214.
- Heine, B., & Kuteva, T. (2002). On the evolution of grammatical forms. In A. Wray (Ed.), *The transition to language* (pp. 376–397). Oxford, UK: Oxford University Press.
- Heinz, J., & Idsardi, W. (2013). What complexity differences reveal about domains in language. *Topics in Cognitive Science*, 5, 111–131.
- Herbst, C. T., Stoeger, A. S., Frey, R., Lohscheller, J., Titze, I. R., Gumpenberger, M., & Fitch, W. T. (2012). How low can you go? Physical production mechanism of elephant infrasonic vocalizations. *Science*, 337(6094), 595–599.
- Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317(5843), 1360–1366.
- Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology*, 14, 5–24.
- Hickok, G. (2016). A cortical circuit for voluntary laryngeal control: Implications for the evolution language. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1100-z
- Holloway, R. L., Broadfield, D. C., Yuan, M. S., Schwartz, J. H., & Tattersall, I. (2004). *The human fossil record, brain endocasts—The paleoneurological evidence*. Hoboken, NJ: John Wiley.
- Hubel, D. H. (1988). *Eye, brain, and vision*. San Francisco, CA: Freeman.
- Hublin, J.-J. (2009). The origin of Neandertals. *Proceedings of the National Academy of Sciences*, 106(38), 16022–16027.
- Hultsch, H., & Todt, D. (1989). Memorization and reproduction of songs in nightingales (*Luscinia megarhynchos*): Evidence for package formation. *Journal of Comparative Physiology A*, 165, 197–203.
- Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge, UK: Cambridge University Press.
- Hurford, J. (1990). Nativist and functional explanations in language acquisition. In I. M. Roca (Ed.), *Logical issues in language acquisition* (pp. 85–136). Dordrecht, Germany: Foris.
- Hurford, J. R. (2007). *The origins of meaning*. Oxford, UK: Oxford University Press.
- Hurford, J. R. (2011). *The origins of grammar: Language in the light of evolution II*. Oxford, UK: Oxford University Press.
- Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17(23), 1004–1005.
- Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, 3(7), 272–279.
- Jackendoff, R. (2002). *Foundations of language*. New York, NY: Oxford University Press.
- Jackendoff, R. (2010). 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(3), 586–624.
- Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, & Chomsky). *Cognition*, 97(2), 211–225.
- Jackendoff, R., & Wittenberg, E. (2016). Linear grammar as a possible stepping-stone in the evolution of language. *Psychonomic Bulletin & Review*.
- Jacob, F. (1977). Evolution and tinkering. *Science*, 196, 1161–1166. doi:10.3758/s13423-016-1073-y
- Jacobs, G. H. (1993). The distribution and nature of colour vision among the mammals. *Biological Reviews*, 68, 413–471.
- Jacobs, G. H., & Deegan, J. (1999). Uniformity of colour vision in Old World monkeys. *Proceedings of the Royal Society of London B*, 266, 2023–2028.
- Jacobs, G. H., & Rowe, M. P. (2004). Evolution of vertebrate colour vision. *Clinical and Experimental Optometry*, 87(4/5), 206–216.
- Jäger, G., & Rogers, J. (2012). Formal language theory: Refining the Chomsky hierarchy. *Philosophical Transactions of the Royal Society B*, 267(1598), 1956–1970.
- Janik, V. M., & Slater, P. B. (1997). Vocal learning in mammals. *Advances in the Study of Behaviour*, 26, 59–99.
- Jarvis, E. D. (2004a). Brains and birdsong. In P. Marler & H. Slabbekorn (Eds.), *Nature's music: The science of birdsong* (pp. 226–271). New York: Academic Press.
- Jarvis, E. D. (2004b). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016, 749–777.
- Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. *Journal of Ornithology*, 148(1), 35–44.
- Jespersen, O. (1922). *Language: Its nature, development and origin*. New York, NY: W. W. Norton & Co.
- Johnson, M. (2016). Marr's levels and the minimalist program. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1062-1
- Jürgens, U. (1998). Neuronal control of mammalian vocalization, with special reference to the squirrel monkey. *Naturwissenschaften*, 85(8), 376–388.

- Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience & Biobehavioral Reviews*, 26(2), 235–258.
- Jürgens, U., Kirzinger, A., & von Cramon, D. Y. (1982). The effects of deep-reaching lesions in the cortical face area on phonation: A combined case report and experimental monkey study. *Cortex*, 18(1), 125–139.
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for ‘fast mapping’. *Science*, 304, 1682–1683.
- Keller, R. (1995). *On language change: The invisible hand in language*. New York, NY: Routledge.
- Kendon, A. (2016). Reflections on the “gesture-first” hypothesis of language origins. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1117-3
- Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin & Review*. (In this issue)
- Klatt, D. H., & Stefanski, R. A. (1974). How does a mynah bird imitate human speech? *Journal of the Acoustical Society of America*, 55(4), 822–832.
- Klima, E. S., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.
- Kluender, K., Diehl, R. L., & Killeen, P. R. (1987). Japanese quail can learn phonetic categories. *Science*, 237, 1195–1197.
- Kojima, S. (1990). Comparison of auditory functions in the chimpanzee and human. *Folia Primatologica*, 55, 62–72.
- Kremers, J., Silveira, L. C. L., Yamada, E. S., & Lee, B. B. (2000). The ecology and evolution of primate color vision. In K. R. Gegenfurtner & L. T. Sharpe (Eds.), *Color vision: From genes to perception* (pp. 123–142). New York, NY: Cambridge University Press.
- Kroodtsma, D. E. (2005). *The singing life of birds*. Boston, MA: Houghton Mifflin.
- Kroodtsma, D. E., & Parker, L. D. (1977). Vocal virtuosity in the brown thrasher. *Auk*, 94, 783–785.
- Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308), 110–114.
- Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America*, 63, 905–917.
- Kuypers, H. G. J. M. (1958). Corticobulbar connections to the pons and lower brainstem in man: An anatomical study. *Brain*, 81(3), 364–388.
- Kuypers, H. G. J. M. (1973). The anatomical organization of the descending pathways and their contributions to motor control especially in primates. In J. E. Desmedt (Ed.), *New developments in EMG and clinical neurophysiology* (Vol. 3, pp. 38–68). Basel, Switzerland: Karger.
- Laland, K. N. (2016). The origins of language in teaching. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1077-7
- Laland, K. N. (2017). *Darwin’s unfinished symphony: How culture made the human mind*. Princeton, NJ: Princeton University Press.
- Laland, K. N., Odling-Smee, J., & Myles, S. (2010). How culture shaped the human genome: Bringing genetics and the human sciences together. *Nature Reviews Genetics*, 11(2), 137–148.
- Lankester, E. (1870). On the use of the term homology in modern zoology, and the distinction between homogenetic and homoplastic agreements. *Annals and Magazine of Natural History*, 6, 34–43.
- Larson, C. R., Sutton, D., Taylor, E. M., & Lindeman, R. (1973). Sound spectral properties of conditioned vocalizations in monkeys. *Phonetica*, 27, 100–112.
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology*, 95, 489–508.
- Lewontin, R. C. (1998). The evolution of cognition: Questions we will never answer. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science: Methods, models, and conceptual issues* (2nd ed., pp. 107–131). Cambridge, MA: MIT Press.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461.
- Lieberman, P. (1984). *The biology and evolution of language*. Cambridge, MA: Harvard University Press.
- Lieberman, P. (2006). *Toward an evolutionary biology of language*. Cambridge, MA: Harvard University Press.
- Lieberman, P. (2007). Current views on Neanderthal speech capabilities: A reply to Boe et al. (2002). *Journal of Phonetics*, 35, 552–563.
- Lieberman, P. (2012). Vocal tract anatomy and the neural bases of talking. *Journal of Phonetics*, 40(4), 608–622.
- Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge, UK: Cambridge University Press.
- Lieberman, P., Crelin, E. S., & Klatt, D. H. (1972). Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man, and the chimpanzee. *American Anthropologist*, 74(3), 287–307.
- Lieberman, P. H., Klatt, D. H., & Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, 164, 1185–1187.
- Liem, K. F. (1973). Evolutionary strategies and morphological innovations: Cichlid pharyngeal jaws. *Systematic Zoology*, 22, 425–441.
- Liem, K. F. (1990). Key evolutionary innovations, differential diversity, and symecomorphosis. In M. H. Nitecki (Ed.), *Evolutionary innovations* (pp. 147–170). Chicago, IL: University of Chicago Press.
- Livingstone, F. B. (1973). Did the australopithecines sing? *Current Anthropology*, 14(1/2), 25–29.
- Livingstone, M. (2002). *Vision and art: The biology of seeing*. New York, NY: Abrams.
- Locke, J. (2016). Emancipation of the voice: Vocal complexity as a fitness indicator. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1105-7
- Lukas, D., & Clutton-Brock, T. H. (2012). Cooperative breeding and monogamy in mammalian societies. *Proceedings of the Royal Society B: Biological Sciences*, 279(1736), 2151–2156.
- Lyn, H. (2017). The question of capacity: Why enculturated and trained animals have much to tell us about the evolution of language. *Psychonomic Bulletin & Review*.
- MacLarnon, A. M., & Hewitt, G. P. (1999). The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology*, 109, 341–363.
- MacNeilage, P. F., & Davis, B. L. (2005). Evolutionary sleight of hand: Then, they saw it; now we don’t. *Behavioral and Brain Sciences*, 28(2), 137–138.
- Maricic, T., Günther, V., Georgiev, O., Gehre, S., Curlin, M., Schreiwies, C., Naumann, R., ... Pääbo, S. (2013). A recent evolutionary change affects a regulatory element in the human FOXP2 gene. *Molecular Biology and Evolution*, 30(4), 844–852.
- Marler, P. (1991). The instinct to learn. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 37–66). Hillsdale, NJ: Erlbaum.
- Marler, P. (2000). Origins of music and speech: Insights from animals. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 31–48). Cambridge, MA: MIT Press.
- Marler, P., & Slabbekoom, H. (2004). *Nature’s music: The science of birdsong*. New York, NY: Academic Press.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: Freeman.
- Marshall, A. J., Wrangham, R. W., & Arcadi, A. C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour*, 58, 825–830.
- Martinez, I., Rosac, M., Quamb, R., Jaraboc, P., Lorenzob, C., Bonmatib, A., ... Arsuagab, J. L. (2013). Communicative capacities in middle

- Pleistocene humans from the Sierra de Atapuerca in Spain. *Quaternary International*, 295, 94–101.
- Maynard Smith, J., & Haigh, J. (1974). The hitchhiking effect of a favourable gene. *Genetics Research*, 23, 23–35.
- McGrew, W. C. (2004). *The cultured chimpanzee*. Cambridge, UK: Cambridge University Press.
- McLean, C. Y., Reno, P. L., Pollen, A. A., Bassan, A. I., Capellini, T. D., Guenther, C., ... Kingsley, D. M. (2011). Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature*, 471, 216–219. doi:10.1038/nature09774.
- Melhem, E. R., Mori, S., Mukundan, G., Kraut, M. A., Pomper, M. G., & van Zijl, P. C. M. (2002). Diffusion tensor MR imaging of the brain and white matter tractography. *American Journal of Radiology*, 178, 3–16.
- Mellars, P. A. (1989). Major issues in the emergence of modern humans. *Current Anthropology*, 30(3), 349–385.
- Mellars, P. (1998a). The fate of the Neanderthals. *Nature*, 395(6702), 539–540.
- Mellars, P. A. (1998b). Neanderthals, modern humans and the archaeological evidence for language. In N. G. Jablonski & L. C. Aiello (Eds.), *The origin and diversification of language* (pp. 89–115). San Francisco: California Academy of Sciences.
- Mellars, P. (2004). Neanderthals and the modern human colonization of Europe. *Nature*, 432(7016), 461–465.
- Mesoudi, A., Whiten, A., & Laland, K. N. (2004). Is human cultural evolution Darwinian? Evidence reviewed from the perspective of ‘The Origin of Species’. *Evolution*, 58, 1–11.
- Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind, and body*. London, UK: Weidenfeld & Nicolson.
- Moore, R. (2016). Meaning and ostension in great ape gestural communication. *Animal Cognition*, 19(1), 223–231.
- Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., ... Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, 6, 6029.
- Mozzi, A., Forni, D., Clerici, M., Pozzoli, U., Mascheretti, S., Riva, S., ... Sironi, M. (2016). The evolutionary history of genes involved in spoken and written language: Beyond FOXP2. *Scientific Reports*, 6, 22157.
- Müller, F. M. (1861). The theoretical stage, and the origin of language. *Lectures on the Science of Language*. London, UK: Longman, Green, Longman, and Roberts.
- Naguib, M., & Kipper, S. (2006). Effects of different levels of song overlapping and singing behavior in male territorial nightingales (*Luscinia megarhynchos*). *Behavioral Ecology and Sociobiology*, 59, 419–426.
- Naguib, M., & Todt, D. (1997). Effects of dyadic vocal interactions on other conspecific receivers in nightingales. *Animal Behaviour*, 54, 1535–1543.
- Negus, V. E. (1938). Evolution of the speech organs of man. *Archives of Otolaryngology*, 28(3), 313–328.
- Newmeyer, F. J. (2005). *Possible and probable languages: A generative perspective on linguistic typology*. Oxford, UK: Oxford University Press.
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, 39, 197–218.
- Nishimura, T., Mikami, A., Suzuki, J., & Matsuzawa, T. (2006). Descent of the hyoid in chimpanzees: Evolution of face flattening and speech. *Journal of Human Evolution*, 51, 244–254.
- Nottebohm, F. (1972). The origins of vocal learning. *American Naturalist*, 106, 116–140.
- Nottebohm, F. (1975). A zoologists’ view of some language phenomena with particular emphasis on vocal learning. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development: A multi-disciplinary approach* (Vol. 1, pp. 61–103). New York, NY: Academic Press.
- Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, 28, 48–53.
- Okanoya, K. (2017). Sexual communication and domestication may give rise to signal complexity necessary for the emergence of language: A hint from songbird studies. *Psychonomic Bulletin & Review*. (In this issue)
- Orr, W. F., & Cappannari, S. C. (1964). The emergence of language. *American Anthropologist*, 66(2), 318–324.
- Pääbo, S. (2014). The human condition—A molecular approach. *Cell*, 157, 216–226.
- Pagel, M. D. (2016). Darwinian perspectives on the evolution of human languages. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1072-z
- Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, 108(6), 2522–2527.
- Parker, J., Tsagkogeorga, G., Cotton, J. A., Liu, Y., Provero, P., Stupka, E., & Rossiter, S. J. (2013). Genome-wide signatures of convergent evolution in echolocating mammals. *Nature*. doi:10.1038/nature12511
- Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience—Virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology*, 10(2), 232–237.
- Patel, A. D. (2016). Using music to study the evolution of cognitive mechanisms relevant to language. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1088-4
- Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a ‘theory of mind’. *Philosophical Transactions of the Royal Society of London B*, 362(1480), 731–744.
- Pepperberg, I. M. (1981). Functional vocalizations by an African grey parrot. *Zeitschrift für Tierpsychologie*, 55, 139–160.
- Pepperberg, I. M. (1991). A communicative approach to animal cognition: A study of conceptual abilities of an African grey parrot. In C. A. Ristau (Ed.), *Cognitive ethology* (pp. 153–186). Hillsdale, NJ: Erlbaum.
- Pepperberg, I. M. (1999). *The Alex studies: Cognitive and communicative abilities of grey parrots*. Cambridge, MA: Harvard University Press.
- Pepperberg, I. M. (2005). Insights into vocal imitation in grey parrots (*Psittacus erithacus*). In S. L. Hurley & N. Chader (Eds.), *Imitation, human development, and culture* (pp. 243–262). Cambridge, MA: MIT Press.
- Pepperberg, I. M. (2016). Animal language studies: What happened? *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1101-y
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6(7), 688–691.
- Perfors, A. (2017). On simplicity and emergence. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1157-8
- Pievani, T., & Serrelli, E. (2011). Exaptation in human evolution: How to test adaptive vs exaptive evolutionary hypotheses. *Journal of Anthropological Sciences*, 89, 9–23.
- Pilley, J. W., & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes*, 86(2), 184–195.
- Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13, 707–784.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What’s special about it? *Cognition*, 95(2), 201–236.
- Ploog, D. W. (1988). Neurobiology and pathology of subhuman vocal communication and human speech. In D. Todt, P. Goedeke, & D. Symmes (Eds.), *Primate vocal communication* (pp. 195–212). Berlin, Germany: Springer-Verlag.
- Poepfel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *Journal of Neuroscience*, 32(41), 14125–14131.

- Poletiek, F. H., Fitz, H., & Bocanegra, B. R. (2016). What baboons can (not) tell us about natural language grammars. *Cognition*, *151*, 108–112.
- Pollard, K. S., Salama, S. R., Lambert, N., Lambot, M.-A., Coppens, S., Pedersen, J. S., ... Haussler, D. (2006). An RNA gene expressed during cortical development evolved rapidly in humans. *Nature*, *443*, 161–172.
- Povinelli, D. J., & Eddy, T. J. (1996). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development*, *61*, 1–151.
- Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *104*, 203–210.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *4*, 515–526.
- Provine, R. R. (2016). Laughter as an approach to vocal evolution: The bipedal theory. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1089-3
- Pruetz, J. D., & Bertolani, P. (2007). Savanna chimpanzees, *Pan troglodytes verus*, hunt with tools. *Current Biology*, *17*, 412–417.
- Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., ... Pääbo, S. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, *505*, 43–49.
- Przeworski, M. (2002). The signature of positive selection at randomly chosen loci. *Genetics*, *160*, 1179–1189.
- Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences*, *7*(5), 187–189.
- Raff, M. (2014). Open questions: What has genetics told us about autism spectrum disorders? *BMC Biology*, *12*, 45.
- Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, *63*, 1050–1056.
- Ramus, F. (2006). Genes, brain, and cognition: A roadmap for the cognitive scientist. *Cognition*, *101*, 247–269.
- Ramus, F., & Fisher, S. E. (2009). Genetics of language. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences IV* (pp. 855–872). Cambridge, MA: MIT Press.
- Raphael, L. J., Borden, G. J., & Harris, K. S. (2007). *Speech science primer: Physiology, acoustics, and perception of speech*. Baltimore, MD: Lippincott Williams & Wilkins.
- Reichmuth, C. J., & Casey, C. (2014). Vocal learning in seals, sea lions, and walrus. *Current Opinion in Neurobiology*, *28*, 66–71.
- Rey, A., Perruchet, P., & Fagot, J. (2012). Centre-embedded structures are a by-product of associative learning and working memory constraints: Evidence from baboons (*Papio papio*). *Cognition*, *123*(1), 180–184.
- Reynolds Losin, E. A., Russell, J. L., Freeman, H., Meguerditchian, A., & Hopkins, W. D. (2008). Left hemisphere specialization for oro-facial movements of learned vocal signals by captive chimpanzees. *PLOS ONE*, *3*(6), e2529.
- Richman, B. (1993). On the evolution of speech: Singing as the middle term. *Current Anthropology*, *34*, 721–722.
- Rilling, J. K., Barks, S. K., Parr, L. A., Preuss, T. M., Faber, T. L., Pagnoni, G., ... Votaw, J. R. (2007). A comparison of resting-state brain activity in humans and chimpanzees. *Proceedings of the National Academy of Sciences*, *104*(43), 17146–17151.
- Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, *11*(4), 426–428.
- Rödel, R. M. W., Olthoff, A., Tergau, F., Simonyan, K., Markus, H., Kraemer, D., & Kruse, E. (2009). Human cortical motor representation of the larynx as assessed by transcranial magnetic stimulation (TMS). *Laryngoscope*, *114*(5), 918–922.
- Rodenas-Cuadrado, P., Ho, J., & Vernes, S. C. (2014). Shining a light on CNTNAP2: Complex functions to complex disorders. *European Journal of Human Genetics*, *22*, 171–178.
- Ruff, C., Trinkaus, E., & Holiday, T. W. (1997). Body mass and encephalization in Pleistocene *Homo*. *Nature*, *387*(6629), 126–127.
- Ruggero, M. A., & Temchin, A. N. (2002). The roles of the external, middle, and inner ears in determining the bandwidth of hearing. *Proceedings of the National Academy of Sciences*, *99*(20), 13206–13210.
- Salas, C., Broglio, C., & Rodríguez, F. (2003). Evolution of forebrain and spatial cognition in vertebrates: Conservation across diversity. *Brain, Behavior and Evolution*, *62*(2), 72–82.
- Sanderson, M. J., & Hufford, L. (Eds.). (1996). *Homoplasy: The recurrence of similarity in evolution*. San Diego: Academic Press.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*, 225–239.
- Saussure, F. D. (1916). *Course in general linguistics* (W. Baskin, Trans.). New York, NY: McGraw-Hill.
- Savage-Rumbaugh, E. S. (1986). *Ape language: From conditioned response to symbol*. New York, NY: Columbia University Press.
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, *58*, 1–221.
- Scharff, C., & Petri, J. (2011). Evo-devo, deep homology and FoxP2: Implications for the evolution of speech and language. *Philosophical Transactions of the Royal Society B*, *366*(1574), 2124–2140.
- Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *PLoS ONE*, *8*(10), e76674.
- Schenker, N. M., Hopkins, W. D., Spocter, M. A., Garrison, A. R., Stimpson, C. D., Erwin, J. M., ... Sherwood, C. C. (2010). Broca's area homologue in chimpanzees (*Pan troglodytes*): Probabilistic mapping, asymmetry and comparison to humans. *Cerebral Cortex*, *20*(3), 730–742.
- Schepartz, L. A. (1993). Language and modern human origins. *Yearbook of Physical Anthropology*, *36*, 91–126.
- Scott-Phillips, T. C. (2014). *Speaking our minds*. New York, NY: Palgrave Macmillan.
- Scott-Phillips, T. C. (2016). Pragmatics and the aims of language evolution. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1061-2
- Seyfarth, R. M. (2005). Continuities in vocal communication argue against a gestural origin of language. *Behavioral and Brain Sciences*, *28*(2), 144–145.
- Seyfarth, R. M., & Cheney, D. L. (2016). Precursors to language: Social cognition and pragmatic inference in primates. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1059-9
- Shea, J. J. (2003). Neandertals, competition, and the origin of modern human behavior in the Levant. *Evolutionary Anthropology*, *12*, 173–187.
- Shettleworth, S. J. (2009). *Cognition, evolution, and behavior*. Oxford, UK: Oxford University Press.
- Shubin, N., Tabin, C., & Carroll, S. (1997). Fossils, genes and the evolution of animal limbs. *Nature*, *388*, 639–648.
- Shubin, N., Tabin, C., & Carroll, S. (2009). Deep homology and the origins of evolutionary novelty. *Nature*, *457*, 818–823.
- Simonyan, K. (2014). The laryngeal motor cortex: Its organization and connectivity. *Current Opinion in Neurobiology*, *28*, 15–21.
- Simpson, G. G. (1944). *Tempo and mode in evolution*. New York, NY: Columbia University Press.
- Smith, K., & Kirby, S. (2008). Cultural evolution: Implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *363*(1509), 3591–3603.

- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford, UK: Blackwell.
- Sporns, O., Chialvo, D. R., Kaiser, M., & Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences*, 8(9), 418–425.
- St Pourcain, B., Cents, R. A., Whitehouse, A. J., Haworth, C. M., Davis, O. S., O'Reilly, ... Smith, G. D. (2014). Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature Communications*, 5, 483.
- Stedman, H. H., Kozyak, B. W., Nelson, A., Thesier, D. M., Su, L. T., Low, D. W., ... Mitchell, M. A. (2004). Myosin gene mutation correlates with anatomical changes in the human lineage. *Nature*, 428, 415–418.
- Stels, L. (2016). Human language is a culturally evolving system. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1086-6
- Steklis, H. D., & Raleigh, M. J. (1973). Comment on Livingstone. *Current Anthropology*, 14(1/2), 27.
- Stoeger, A. S., Mietchen, D., Oh, S., de Silva, S., Herbst, C. T., Kwon, S., & Fitch, W. T. (2012). An Asian elephant imitates human speech. *Current Biology*, 22, 2144–2148.
- Stokoe, W. C. (1960). *Sign language structure: An outline of the communicative systems of the American deaf*. Silver Spring, MD: Linstock Press.
- Striedter, G. F. (2004). *Principles of brain evolution*. Sunderland, MA: Sinauer.
- Stringer, C., & Andrews, P. (2005). *The complete world of human evolution*. London, UK: Thames & Hudson.
- Stringer, C., & Gamble, C. (1993). *In search of the Neanderthals*. London, UK: Thames & Hudson.
- Számádó, S., & Szathmáry, E. (2006). Selective scenarios for the emergence of natural language. *Trends in Ecology & Evolution*, 21(10), 555–561.
- Takemoto, H. (2008). Morphological analyses and 3D modeling of the tongue musculature of the chimpanzee (*Pan troglodytes*). *American Journal of Primatology*, 70(10), 966–975.
- Tallerman, M. (2008). Holophrastic protolanguage: Planning, processing, storage, and retrieval. *Interaction Studies*, 9(1), 84–99.
- Tallerman, M. (2013). Join the dots: A musical interlude in the evolution of language? *Journal of Linguistics*, 49(2), 455–487.
- Tallerman, M., & Gibson, K. (Eds.). (2011). *The Oxford handbook of language evolution*. Oxford, UK: Oxford University Press.
- Tattersall, I. (1999). *Becoming human: Evolution and human uniqueness*. New York, NY: Harcourt-Brace.
- Tattersall, I. (2009). Human origins: Out of Africa. *Proceedings of the National Academy of Sciences*, 106(38), 16018–16021.
- Tattersall, I. (2016). How can we detect when language emerged? *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1075-9
- Taylor, A., & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research. *Journal of Zoology*, 280(3), 221–236.
- Tebbich, S., Taborsky, M., Fessl, B., & Blomqvist, D. (2001). Do woodpecker finches acquire tool-use by social learning? *Proceedings of the Royal Society B*, 268(1482), 2189–2193.
- ten Cate, C. (2016). Assessing the uniqueness of language: Animal grammatical abilities take center stage. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1091-9
- Theissen, G. (2009). Saltational evolution: Hopeful monsters are here to stay. *Theory in Biosciences*, 128, 43–51.
- Thieme, H. (1997). Lower Palaeolithic hunting spears from Germany. *Nature*, 385, 807–810.
- Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., ... Deloukas, P. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics*, 39(1), 31–40.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford, UK: Oxford University Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.
- Toth, N., Schick, K. D., Savage-Rumbaugh, E. S., & Sevcik, R. A. (1993). Pan the tool-maker: Investigations into the stone tool-making and tool using capabilities of a bonobo (*Pan paniscus*). *Journal of Archaeological Science*, 20, 81–91.
- Udden, J., Folia, V., Forkstam, C., Ingvar, M., Fernandez, G., Overeem, S., van Elswijk, G., ... Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: An rTMS study. *Brain Research*, 1224, 68–79.
- van der Lely, H. K. J., & Pinker, S. (2014). The biological basis of language: Insight from developmental grammatical impairments. *Trends in Cognitive Sciences*, 18(11), 586–595.
- van Heijningen, C. A. A., de Vissers, J., Zuidema, W., & ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, 106, 20538–20543.
- van Lawick-Goodall, J., & van Lawick-Goodall, H. (1967). Use of tools by the Egyptian vulture, *Neophron percnopterus*. *Nature*, 212, 1468–1469.
- Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). *FOXP2* and the neuroanatomy of speech and language. *Nature Reviews Neuroscience*, 6(2), 131–138.
- Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. *Current Biology*, 10(12), 723–726.
- Vauclair, J. (1996). *Animal cognition: An introduction to modern comparative psychology*. London, UK: Harvard University Press.
- Vernes, S. C. (2016). What bats have to say about speech and language. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1060-3. Advance online publication.
- Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., ... Fisher, S. E. (2008). A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine*, 359(22), 2337–2345.
- von Humboldt, W. (1836). *Über die Kawi-Sprache auf der Insel Java [On the Kawi Language of Java]*. Berlin, Germany: Druckerei der Königlich-Akademie der Wissenschaften.
- Vorobyev, M. (2004). Ecology and evolution of primate colour vision. *Clinical and Experimental Optometry*, 87(4/5), 230–238.
- Wang, R., Chen, C.-C., Hara, E., Rivas, M. V., Roulhac, P. L., Howard, J. T., ... Jarvis, E. D. (2015). Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *Journal of Comparative Neurology*, 523, 892–906.
- Wang, L., Uhrig, L., Jarraya, B., & Dehaene, S. (2015). Representation of numerical and sequential patterns in macaque and human brains. *Current Biology*, 25(15), 1966–1974.
- Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., ... Slocombe, K. E. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology*, 25, 495–499.
- Weir, A. A. S., Chappell, J., & Kacelnik, A. (2004). Shaping of hooks in New Caledonian crows. *Science*, 297, 981.
- Weiss, M., Huelsch, H., Adam, I., Scharff, C., & Kipper, S. (2014). The use of network analysis to study complex animal communication systems: A study on nightingale song. *Proceedings of the Royal Society B, Biological Sciences*, 281(1785).
- Weissengruber, G. E., Forstenpointner, G., Peters, G., Kübber-Heiss, A., & Fitch, W. T. (2002). Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*), and domestic cat (*Felis silvestris f. catus*). *Journal of Anatomy*, 201, 195–209.

- Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., ... Boesch, C. (1999). Cultures in chimpanzees. *Nature*, *399*, 682–685.
- Whiten, A., Horner, V., & de Waal, F. B. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, *437*, 737–740.
- Wich, S. A., Swartz, K. B., Hardus, M. E., Lameira, A. R., Stromberg, E., & Shumaker, R. W. (2009). A case of spontaneous acquisition of a human sound by an orangutan. *Primates*, *50*(1), 56–64.
- Wild, J. M. (1997). Neural pathways for the control of birdsong production. *Journal of Neurobiology*, *33*, 653–670.
- Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language & Communication*, *18*(1), 47–67.
- Wray, A. (2000). Holistic utterances in protolanguage: The link from primates to humans. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 285–302). Cambridge, UK: Cambridge University Press.
- Wynn, T., & Coolidge, F. L. (2004). The expert Neanderthal mind. *Journal of Human Evolution*, *46*, 467–487.
- Xiang, H.-D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral Cortex*, *20*, 549–560.
- Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), 6323–6327.
- Yule, G. (2006). *The study of language* (3rd ed.). Cambridge, UK: Cambridge University Press.