J. Linguistics **49** (2013), 455–487. © Cambridge University Press 2013
doi:10.1017/S0022226713000017 First published online 13 February 2013

REVIEW ARTICLE

Join the dots: A musical interlude in the evolution of language?[1]

MAGGIE TALLERMAN

Newcastle University, UK

(Received 18 May 2012; revised 4 January 2013)

W. Tecumseh Fitch, *The evolution of language*. Cambridge: Cambridge University Press, 2010. Pp. xii + 611.

This review article examines the major premises concerning the evolution of language offered in Tecumseh Fitch's *The Evolution of Language* (2010). Various aspects of Fitch's argumentation are disputed, specifically including the idea that language was preceded in evolution by a musical protolanguage stage in which there were no words, but which instead exhibited a prosodic, meaningless 'bare phonology'. According to Fitch, this putative stage in language evolution also laid the foundations for syntax, providing musical phrases which were subsequently filled by meaningful words. Counterarguments to these ideas are presented, and an alternative, word-based model of protolanguage is defended.

1. Introduction

Biologist Tecumseh Fitch is probably best known to most linguists as co-author, with Noam Chomsky and a second biologist, Marc Hauser, of two papers in a debate concerning the contents of the language faculty: Hauser, Chomsky & Fitch (2002) and Fitch, Hauser & Chomsky (2005). Though both papers engage extensively with ideas about what might specifically have evolved to give rise to language, in a recent lecture Chomsky (2011) described work on language evolution thus:

> There's a field called 'evolution of language' which has a burgeoning literature, most of which in my view is total nonsense … In fact, it isn't

[1] Various erudite scholars read the first draft and saved me from myself in numerous ways. I am truly grateful to Steve Anderson, Christina Behme, Kathleen Gibson, Jim Hurford, Nigel Vincent and also to three referees for the *Journal of Linguistics*. They may be a teeny bit to blame for the proliferation of footnotes, but all faults and errors are mine.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

even about evolution of language, it's almost entirely speculations about evolution of communication, which is a different topic.

While some of us working in the field may feel a little hurt by these accusations, it is clear that Chomsky's central concern is not the evolution of what most biologists, psychologists or even linguists consider to be 'language'; rather, it is solely 'the faculty of language in the narrow sense' (FLN; Hauser et al. 2002) – namely, whatever is both uniquely human and uniquely linguistic. So what stance on language evolution does Chomsky's sometime co-author take in this book? Fitch concentrates firmly on FLB: the faculty of language in the broad sense, where, unlike in FLN, we WOULD expect to find both homologues and analogues elsewhere in biology (broadly similar features, either with or without a common ancestry; see Section 3 below). In fact, Fitch even states (88) that he's going to use the term 'FLB' in preference to 'UG', which is normally taken to signify our specific language endowment. FLB, on the other hand, encompasses everything that contributes to the language capacity, not merely features 'designed' for language, thus including very general traits such as memory and respiration. This focus on FLB means that Fitch includes many features that aren't even species-specific, such as vocalization, rather than what is uniquely linguistic. In a sense, this choice sets the scene for what follows. Fitch's concerns, then, diverge widely from Chomsky's; for instance, Fitch postulates stages in language evolution before the emergence of full language (i.e. proto-language), and he firmly believes that much can be learned about the evolution of language by studying animal models. Fitch also considers that the most significant fact about language evolution is not syntax or the lexicon or phonological systems, but the emergence of learned vocalization.

Fitch aims in this book to present a wide-ranging account of the evolution of all aspects of language. Ultimately, this aim is not entirely fulfilled, but I'm happy to acknowledge that there is much informative, provocative and entertaining material along the way.

This review is organized as follows. Section 2 presents an overview of the book's contents and some preliminary assessment as to its value. Section 3 discusses what (if anything) can be learned about the language faculty from investigating other animals, a prominent theme in the book under review. I argue that the answer is, substantially less than Fitch believes to be the case. Section 4 provides critique of Fitch's proposal that the earliest form of protolanguage used by our ancestors was a musical protolanguage. In Section 5, I examine Fitch's views of what followed this purported musical stage in language evolution, and ask how syntactic language could have been built from this kind of protolanguage. Section 6 turns to a different topic, that of the evolution of a communication system that is both 'cheap' and 'honest', which Fitch considers to be problematic. This section suggests that

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

the problems are less extensive than assumed. Section 7 briefly considers textbook choices for courses in language evolution.

2. Outline of the book and general assessment

Fitch's goal is to provide 'an unbiased overview of the relevant questions, approaches, and data' in language evolution along with 'a dispassionate survey of the available hypotheses, and an even-handed evaluation of their strengths and weaknesses' (4). In this, I think Fitch succeeds pretty well, to the extent that the book will certainly be used as a text in language evolution courses for many years to come. However, he also promotes some ideas that I consider quite insupportable, as we'll see, including proposals for a musical protolanguage. First, I turn to a broad outline of the contents.

The book is presented as two halves, each with two main sections. A short glossary is provided at the end. The first half of the book is intended to provide introductory 'tutorial' treatment of the biological and linguistic material that the reader needs to grasp in order to understand the basics of language evolution. On the biology side, this includes an overview of what evolution involves, summaries of animal cognition and communication, and of human evolution. On the language side, as we'll see, there is far less detail. Chapter 1, 'Language from a biological perspective', opens with the infant chimpanzee Viki, and her failure to learn language – an excellent starting point. Fitch then goes on to discuss some material that will be familiar to linguists, such as Hockett's design features, the FLN/FLB distinction, E- and I-language; and also other material that may be less familiar, such as gene expression, epigenesis, and canalized behaviours or 'instincts to learn'. Chapter 2, which I found particularly interesting, is 'Evolution: Consensus and controversy'. It summarizes the basic details of evolution by natural selection, and briefly introduces some selection mechanisms that will play an important role in Fitch's argument later on, including sexual selection and kin selection. This chapter also includes a sketch of the comparative method, which is used in biology in exactly the same way as in linguistics, but of course makes comparisons across species – rather than across languages – in order to deduce properties of a common ancestor. There's an outline of what in biology is termed the 'modern synthesis' (between genetics and evolutionary theory), and an introduction to the fascinating topic of DEEP HOMOLOGY, including discoveries that, for instance, 'the principles of cell biology and genetic regulation are so conserved that identical mechanisms regulate wing growth in flies and hand growth in mice and humans' (56). This topic is rather important for Fitch, because one of the main tenets throughout the book is that little in evolution is radically new; thus, he suggests that it's well worth investigating 'which subcomponents of language can be found in the communication systems of other species' (26). Chapter 3, 'Language', provides a very brief overview of modern linguistics, and

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

includes tutorial sections on phonology, syntax, semantics and pragmatics. These topics also recur in other chapters in the first half of the book. Chapter 4, 'Animal communication and cognition', concludes the first section of the book. I'll have more to say about this later.

The next three chapters form a section entitled 'Meet the ancestors'. Chapter 5, also 'Meet the ancestors', takes us from single-celled creatures through early fish, then dinosaurs, the first mammals, and finally primates. If there's one chapter which could have been omitted from a book called *The evolution of language*, this is it. Fascinating though it genuinely is, I don't think that going back to the origins of life on earth is necessary to understand language evolution. Nor does most of the chapter shed any light on language. The following two chapters are more human-centred: Chapter 6, 'The LCA: Our last common ancestor with chimpanzees', looks at the emergence of specifically human traits such as an extended childhood and enhanced sociality, some of which will be relevant to language acquisition because they relate to the trait of traditional transmission. And Chapter 7, 'Hominid paleontology and archaeology', looks at fossils in our lineage, brain size, and also tools and culture. (Going against most modern usage, Fitch deliberately chooses the term 'hominid' for preference, rather than 'hominin', to refer to members of our own lineage.)

The second half of the book 'systematically introduces and evaluates current theories about language evolution' (8). Three chapters on the evolution of SPEECH start this half off: Chapter 8, 'The evolution of the human vocal tract', Chapter 9, 'The evolution of vocal control: The neural basis for spoken language' and Chapter 10, 'Models of the evolution of speech and phonology'. Phonology actually receives little attention – notionally, in the region of ten pages (374–385), but in reality, some of this is taken up with vocal imitation and 'similar processes in nonhuman animals', including dialectal variation (378), and the general properties of pidgins and creoles (379f.). Phonology is clearly not a major interest of Fitch's. In fact, nor is phonetics: there are two dreadful attempts at transcription of English (133, 376).

Finally, there are four chapters in a section entitled 'Evaluating phylogenetic models of language evolution'. More precisely, these chapters concentrate on different models of protolanguage. Chapter 11, 'Historical overview: Western theories of language origin before Darwin', is a somewhat odd choice. It's a very good read and well executed. But it took us too far afield; for instance, I'd freely exchange the page on the Bible for a better sketch of genetics (see below). It's also a shame that space was used in this way when a more recent publication which helped kick-start the field of evolutionary linguistics, Pinker & Bloom (1990), receives so little attention in the book. Along with Bickerton (1990), this paper set the stage for more substantive and less speculative discussion of language evolution, and should be given credit. Chapter 12, 'Lexical protolanguage', is intended as an

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

outline of the kind of word-based protolanguage proposed by Bickerton (1990) and Jackendoff (2002). Gestural origins is the topic of Chapter 13, 'Signs before speech: Gestural protolanguage theories'. And the final main chapter, 14, is 'Musical protolanguage', about which much more later.

There is some well-taken criticism in this section, for example of Bickerton (1990); but basing the discussion of lexical protolanguage on this publication, now well over twenty years old, does not provide the balanced picture that Fitch intends.[2] For instance, Bickerton – who is never slow to acknowledge when he has been wrong – has not believed that 'complex syntax … result[s] from a single genetic mutation' (408) since at least Bickerton (2000). There are also some strange criticisms of the Bickertonian model of protolanguage: '[a] further weakness of lexical models is their assumption that the posited protolanguage has essentially disappeared in modern human society, and protolinguistic "fossils" make their appearance only under extraordinary social circumstances' (433). This is odd because elsewhere, Fitch himself criticizes the view that, in the sphere of language, ontogeny recapitulates phylogeny (405). Actually, though, this is pretty well exactly what Bickerton (1990) proposes: that infants under around two years old produce a modern protolanguage. In fact, Jackendoff (2002: 235) notes approvingly that '[w]hat elevates Bickerton's story above mere speculation is his claim that protolanguage is still present in the modern human brain'. So no, protolanguage in Bickerton's model has not 'disappeared in modern human society'.[3]

Overall, three fundamental problems arise with the content of the book. The first problem is that there isn't enough emphasis on the exceptional nature of language as a human faculty. In particular, the putative parallels with animal communication and cognition are at times exaggerated. Take statements like this: '[e]ven syntax, at least at a simple level, finds analogs in other species (e.g. bird and whale "song") which can help us to understand both the brain basis for syntactic rules and the evolutionary pressures that can drive them to become more complex' (18). While there's SOME truth in the first half, given the existence of both hierarchical structure and simple dependencies in animal 'syntax' (see Hurford 2012 for an excellent survey), I fear that a non-linguist reading the claim that analogues of syntax are found in other animals would get entirely the wrong idea.[4] Grammatical

[2] Fitch would not have had access to Bickerton (2009), the most recent exposition of his ideas.

[3] For the record, I agree wholeheartedly with Hurford (2012), Bickerton, Jackendoff, and others who suggest that evidence of earlier stages of the language faculty can be seen in various contexts in language today. If Fitch supports this view too, then we have a significant point of accord.

[4] A referee for *JL* suggests that it might be useful here to distinguish between 'phonological syntax', where the units that are put together are meaningless, and 'lexical syntax' (Marler 1998), which produces semantic compositionality. Birdsong and whalesong exhibit the

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

systems in language are NOT merely a more complex version of animal communication systems, which are entirely non-compositional, with no duality of patterning, and which do not contain word classes or headed phrases. Moreover, the content of the relative clause on understanding 'the brain basis for syntactic rules' is overstated (overviews of what is currently known and hypothesized about syntax and the brain can be found in Bickerton & Szathmáry 2009). Another example: 'the human capacity for complex vocal learning, allowing children to learn the words of their community's language, is shared with many birds, whales and seals' (5). Again, I see the point here: we're certainly not the only species displaying vocal learning. But since no animals learn anything like a human lexicon, Fitch's claim could easily give the wrong impression; I return to this topic below.

The second problem is that if readers don't already have at least a basic grasp of what Fitch is talking about, I suspect that in many instances, they won't understand all that much more after reading the book. A subsection entitled 'What is syntax?' (103) takes up less than one and a half pages, but since it also includes sketches of the terms 'word' and 'morpheme', there's really very little space left for syntax itself. Recursion is mentioned (104, 116) and defined – I feel inadequately – as concerning 'rules that … refer to themselves in their own definition' (104), but is not illustrated. This is important because I assume that many non-linguist readers will approach Fitch looking for a way to understand the terms of the recent recursion debate, starting with Hauser et al. (2002). As Fitch himself notes (76), laypeople often think they know what language involves, when really they have a fairly vague idea that includes many incorrect notions; for this reason, the sections on the components of language needed to be fuller, more informative and more explanatory. For instance, a very brief sketch of *wanna* contraction concludes without explanation that 'this process is blocked if there is a phrase boundary between *want* and *to*' (96), illustrating with the following data:

(1) a. What do you want to do?
    b. What do you wanna do?
    c. Who do you want to go to the store?
    d. *Who do you wanna go to the store?

These data are presented with no further discussion. What will the reader understand of this? Is it obvious that there's a phrase boundary between *want* and *to* in (1c, d) that isn't there in (1a, b)? Can you see this boundary? I can't; and this is in any case a misleading characterization of the phenomenon. Readers of this journal will recognize the point being

former, but not the latter. The book under review does not clearly distinguish between the two, which is what I think may be misleading for non-linguists.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

made, but (and I checked on this intuition) not even our own students will get it.

Similar comments apply to the biology sections: even though a lot more introductory material is provided from this discipline, which is of course Fitch's own, it is often not that detailed or clear, or is just not basic enough to give the non-specialist reader any kind of handle on the topics under discussion. For instance, students taking my language evolution module often ask me what exactly DNA is, or what genes are. Annually I wheel out my doubtless pathetically inadequate answers, but it would be really good if such basic and vital questions were answered in this book, which is, after all, intended for use as a textbook. Fitch does indeed provide a definition of the genetic code (e.g. 49–50, 210), but I'm probably not the only linguist for whom these pages fail to shed much light. Note, I'm not implying that Fitch gets the fundamental biology wrong, just that he often doesn't help the reader to understand it.[5] Here's a couple of examples among many. In a discussion of vocal anatomy: '[i]f the sternohyoid muscles are tensed, the basihyoid moves downward …, while if the digastric and stylohyoid are tensed, it moves upward' (333) – no definitions are provided. In a discussion of the brain: '[a]t the next level we find a midbrain control center, made up of the periaqueductal gray region and surrounding tegmentum, which serves to elicit vocalization and in some cases to control parameters' (347) (we aren't told what parameters). And it's not the case that there was no space left for a fuller treatment of such topics, if we really need to know about them. For instance, most of Chapter 5, 'Meet the ancestors', could be cut to make room, since, as noted above, it is extremely interesting but scarcely relevant to language.

The third problem, which I expand on in what follows, is that the major focus of the book is vocal imitation and the evolution of speech, rather than on linguistic (i.e. cognitive) features such as syntax, semantics and phonology.

Broadly, then, the contents have a good range and scope, but the book suffers from a lack of detail about what language involves and the crucial ways in which language differs from animal communication systems. I turn next to the question of whether aspects of language have analogues or

[5] In a section on the FOXP2 gene (358–363), Fitch links its role in human vocal learning with the role of the homologous FoxP2 gene in songbird vocal learning (see also Berwick 2011a). Fitch clearly acknowledges that FOXP2 is not a 'magic bullet' for our understanding of speech or language (362), but the discussion gives the impression that we know more about the role of FOXP2 than is actually the case. FOXP2 is what is known as a transcription factor, meaning that it controls the expression of many other genes, turning them on and off. While it is implicated in human vocalization (as well as that of mice, echolocating bats, and birds), its general effects are not yet understood. Fitch unfortunately downplays the morphosyntactic problems seen in KE family members with the relevant FOXP2 mutation, problems which are clearly distinct from their difficulties with vocalization; see also Diller & Cann (2009) and Piattelli-Palmarini & Uriagereka (2011).

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

homologues in other animals, and what, if anything, can be learned from a study of animal cognition and communication.

3. (WHAT) CAN WE LEARN ABOUT THE LANGUAGE FACULTY FROM OTHER ANIMALS?

Various major themes recur throughout the book. One prominent thesis is that many of the abilities and traits subsequently taken over ('exapted') for use in language – meaning FLB in the broadest possible sense – are phylogenetically ancient, certainly common to primates, possibly even to vertebrates generally, including birds. For instance, categorical perception, vital for phonology, is 'widely shared with other species' (99), and even pigeons can learn to categorize pictures not only of trees, but, unexpectedly, of fish and letter shapes. Summarizing a discussion of physical and social intelligence in animals, Fitch states that 'the building blocks of pragmatics were in place long ago, with multiple features ready to be inherited or "exapted" by language as it evolved, including discreteness, hierarchy, open-endedness, and a basic form of propositionality' (160). While there can be no dispute that birds and mammals only distantly related to us do very smart things, claiming that these four critical properties of language are already IN PLACE in animal cognition long before hominins evolved vastly underrates the magnitude of the language evolution problem.

Some of the traits that are part of the language faculty certainly have simpler antecedents in animal cognition. And to be clear, I acknowledge that we largely share a conceptual system with other primates; see Hurford (2007) for discussion. There is good evidence, for instance, of hierarchical social knowledge in various monkey species; see Cheney & Seyfarth (2005) for overview. But it's too large a leap to suggest that such systems are 'precursors to syntax' (118). In fact, as Kathleen Gibson points out to me, there is evidence that humans have significantly greater hierarchical capacity across a wide range of behavioural domains than apes, and in turn, great apes have greater hiearchical capacity than monkeys (e.g. Greenfield 1991; Gibson 1993, 1996). For instance, while humans readily learn to join two blocks of wood by hammering a nail into them, novel CONSTRUCTIONS of this kind are apparently beyond great ape capabilities.

What is problematic in the book is that while Fitch focuses on the ways in which animal cognition is similar to human cognition, and on putative evidence in animal cognition of 'building blocks' for language, the uniqueness of language as a system is downplayed by the lack of detailed consideration. Syntax gets remarkably short shrift, with nothing on the evolution of such critical properties as functional versus lexical vocabulary, structure dependence (a mere mention, 104), grammatical relations, or local and non-local dependencies in phrases and clauses, including anaphoric

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

reference, displacement phenomena, and so on. In the short section entitled 'What is syntax?' (103–105), the property of headedness is not even mentioned. Despite not having discussed these fundamental universals of language, Fitch concludes that '[i]f many of the apparent complexities in linguistic syntax result not from the underlying complexity of an innate syntax module, but rather derive from much older mechanisms …, much of this complexity was already present before human language evolution began' (118; see also 430f.). But underestimating in this way what is needed to build a language faculty simply leaves much of language evolution unaccounted for.

Regarding syntax, Fitch notes that if its evolution involved merely 'one or a few pervasive and powerful operators',[6] this 'would be excellent news from a biologist's perspective' (118). Such a model is explicitly contrasted with a 'complex UG' model requiring 'a large suite of complex, interlocking, and evolutionarily novel mechanisms' (118), which is implied to be less biologically plausible. I take it that the latter model is well illustrated by the Government and Binding framework (Chomsky 1981), which has now been abandoned within the Minimalist paradigm (Chomsky 1995) in favour of a model that is biologically more plausible. Problematically, however, we have no evidence for this view of syntactic evolution either – not in our genes, not in our brain structure, certainly not in animal models. Surely, then, rather than positing ANY major steps that cannot be empirically tested, it is better to demonstrate how natural language syntax might have evolved in gradual adaptive stages from simpler precursors. This is the agenda pursued convincingly by Ray Jackendoff (e.g. Jackendoff 2002, 2010, 2011), which sees the language faculty as 'a layered composite of grammars from many different levels' (2011: 617). Of course the resulting sketches of possible stages in language evolution are still speculative, as Jackendoff well knows, but they are reasonable and promising, and seem both biologically and linguistically credible.

In fact, syntax as a major part of the language faculty is generally downplayed in this book. For Fitch, 'the sharpest distinction between humans and chimpanzees' is not syntax but 'vocal imitation' (503). Now, vocal imitation is certainly a novelty in our lineage, and clearly crucial for vocal (E-)language, so it is desirable to have an account of this discontinuity between *Homo sapiens* and the other great apes. But imitation cannot be the main problem to account for, exactly because it is NOT specifically linguistic. On the contrary, imitation pervades everything humans do, including the learning of social norms, survival skills and cultural practices. Since it's so general in human cognition, and is critical for a learned (proto)language, imitation itself must be ancient in the hominin lineage, predating any kind of

[6] Here I thought, read 'Merge'? Yet on pages 409–410, Fitch seems to disown the idea that 'there is only one … structure-building combinatorial system'.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

language; in biological terms, imitation must already be in place before it can be exapted for language. We still need an account of the evolution of the specifically human kind of imitation (Byrne & Russon 1998), but because it's not uniquely linguistic, this is easier to unravel. Syntax, conversely, is problematic precisely because it apparently contains so much novelty of the kind indicated in earlier paragraphs. An indication of how this might have arisen is what is missing from this book.

Fitch does return to syntax at various points throughout the book, for instance in a three-page section entitled 'Whence syntax?' (429–432). However, this section is oddly critical of the extensive use made in research on language evolution of the concept of GRAMMATICALIZATION: 'Once language is off the ground, it is easy to see how grammaticalization provides grist for the syntactic mill, but it fails to address the basic questions of how the more general properties of syntax arose' (431). First, it's hard to know what 'general properties of syntax' Fitch has in mind, since, as noted above, there is virtually no detail on this topic in the book. Second, this statement displays a fundamental misunderstanding of grammaticalization processes. Certainly, the most critical property of syntax, the ability to combine meaningful signals, must be the first step in syntactic evolution, and this has nothing to do with grammaticalization. But from there, the processes of grammaticalization in fact PRODUCE many of the 'general properties of syntax' and morphosyntax, including distinct word classes (both lexical and functional), subordination, grammatical function-changing constructions such as the passive, tense, agreement and case, and so on (e.g. Hurford 2003, Heine & Kuteva 2007). In a sense it's unsurprising that grammaticalization doesn't figure more prominently in this book, because this is just the kind of mechanism (uniquely human and uniquely linguistic) that we CAN'T learn anything about from animal models, and which thus doesn't really fit Fitch's agenda. But the message here is surely that the agenda is misguided to the extent that it focuses so restrictively on what can be learned from other animals.

Moreover, the case for 'pre-linguistic' abilities in other animals is frequently overstated. For instance, 'language-trained apes demonstrate an ability to acquire a sizeable lexicon' (168). Well, no, they don't. They demonstrate an impressive ability to learn numerous correspondences between arbitrary sounds, shapes or gestures and specific real-world entities and actions (rarely, if ever, including abstract concepts); thus, they display an elementary symbolic capacity. We might call this vocabulary-learning, provided it's understood that this implies something very different from human abilities. Even setting aside the non-trivial issue of size of 'vocabularies' learned in humans vs. apes (which Fitch does mention, 148), this correspondence-matching ability alone is not what constitutes a lexicon. Apes don't learn the full lexical properties of vocabulary items, which critically include both selectional restrictions and subcategorization

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

requirements, as well as semantic collocations.[7] Thus, they never learn word classes and how they are used, or the relationships between heads and their various dependents (such as the transitivity of verbs), which are also lexical properties. Nor do they learn grammatical morphemes or functional vocabulary items; for instance, even Sue Savage-Rumbaugh's star pupil, the bonobo Kanzi, stumbled repeatedly over such items as auxiliary *can* and the conjunction *and* (Savage-Rumbaugh, Shanker & Taylor 1998). These crucial distinctions are not discussed by Fitch. While it is one thing to claim that the ape ability to learn 'words' (an ability shared with parrots; Pepperberg 2000) may be phylogenetically ancient, it is quite another to casually equate it to the human acquisition of a massive lexicon, which also includes many idioms and construction frames such as those seen in *John sang his heart out* or *John yelled his head off* (Culicover & Jackendoff 2005). Here and elsewhere, the non-linguist lay reader will come away with the deceptive impression that very little had to evolve to 'buy' language.

A second leitmotif in the book is the idea that we can learn a great deal about language by studying the CONVERGENT evolution of some of its features in other animals. Convergent evolution is the appearance of functionally similar features or traits in different lineages, without descent from a common ancestor, such as wings in birds, bats and insects: such features are known as analogues. (In contrast, homologues are inherited features shared within a taxon, the result of descent from a common ancestor; an example is the taillessness of the great apes, including us.) Fitch's motivation for extensive use of convergence is that the selection pressures driving the evolution of a feature or capacity should also play a role in the hominin lineage; thus, if we can see why certain traits evolve in other species, we may get an indication of how and why they appear in humans.

However, as Stephen Anderson points out to me (p.c.), the equivalence of the conditions that produced some analogous capacity in different species is not self-evident. Further, not all the analogues proposed here bear scrutiny. Here is an example. Fitch suggests that mockingbird song 'provide[s] a clear parallel to the hierarchical structure of phonological and syntactic phrases in language' and that '[t]he most complex animal "songs" rival human speech in their complexity, and appear to require context-free

[7] Savage-Rumbaugh & Rumbaugh (1978) trained chimpanzees to distinguish between lexigrams (symbols on a keypad) for two distinct actions: 'pour', which applied only to drinks, and 'give', for solid food only. The animals used these to request various foods and drinks, the lexigrams for which were known to them. Then new food and drink lexigrams were introduced, to see if the chimps generalized the 'pour'/'drink'distinction to novel items. Some animals did generalize the rule for the two action lexigrams correctly, while others required extensive retraining. This demonstrates good categorization skills (and the results were later replicated with tools and food), but is not equivalent to the lexical knowledge found in language.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

grammars' (183). These are huge claims, and ones that are not substantiated in the text. There is indeed hierarchical structure in much birdsong, but it doesn't go beyond two layers of hierarchical depth, with humpback whalesong displaying slightly more complexity in terms of dependencies than birdsong – and that is the limit.[8] Crucially, of course, there is no semantic compositionality in animal song. For readers of this journal, that last statement is Linguistics 101. For the uninitiated biologist or palaeontologist reading Fitch's book, perfunctory statements about the complexity of animal song such as those cited above would be seriously misleading. Casually overstated claims such as the ape lexicon example and the implication here that birdsong is parallel in complexity to language spoil much of the book for me; they obscure the wealth of genuinely informative detail. And a focus on animal 'parallels' to features of language also reveals little about the evolution of the highly specific properties of language, of which phonology, the lexicon, semantic compositionality and syntactic recursion are prime examples.

Turning next to vocalization, Chapter 8 on 'The evolution of the human vocal tract' presents much of Fitch's own work on mammalian vocal tracts, again with the aim of studying the convergent evolution of a trait in species that are not closely related to us, and comparing the selective pressures that may have led to its emergence. In this case, one of the main explananda is human hyolaryngeal descent, i.e. the permanent and major descent of the larynx and the hyoid bone, to which the tongue is attached. Along with these changes, humans have evolved a highly mobile curved tongue, rather than the flat tongue of other great apes, and a supralaryngeal vocal tract (SVT) in which the oral and pharyngeal tubes are of equal length, as opposed to the normal mammalian pattern where the oral portion of the tract is much longer (MacLarnon 2012). Together, these modifications result in a vocal tract uniquely able to produce the wide array of human speech sounds. Most researchers have linked the initial changes in the SVT to changes in hominin lifestyle (such as dietary shifts and the advent of cooking, leading to changes in dentition and musculature for chewing) and to obligate bipedalism, which triggers various physical reconfigurations in the head and neck, and which appears to have occurred during the long *Homo erectus* era, beginning some 1.8 million years ago.

[8] Hurford (2012: Chapter 1) has recently provided a detailed and carefully articulated exploration of the formal grammatical properties of animal song. Though context-free or phrase structure grammar is technically not required to capture its properties (Hurford 2012: 97), bird and whale specialists nonetheless recognize discrete PHRASES in the more complex instances of song. A context-free grammar thus provides a better description for these cases, enabling the hierarchical structure to be acknowledged. However, such grammars are always very simple, as indicated in the text. Thanks to Jim Hurford for clarifying these facts for me.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

Fitch has a very different position on the initial selection pressures that led to reconfiguration of the SVT, and indeed, a very surprising take on the earliest uses made of the resulting reshaped hominin vocal tract (Section 4 below). I end this section, then, with an examination of Fitch's view of the selection pressures surrounding speech adaptations in early humans. As anyone who's ever been in the audience at a Fitch talk now knows, various species only distantly related to humans also have a descended larynx, including red deer and koalas; clearly, neither cooking nor bipedalism were factors here. Fitch argues (321–323) that the selection pressure for laryngeal descent in other animal species was SIZE EXAGGERATION. Elongating the larynx produces lower formant frequencies, giving the impression of an animal with greater body size (in simple terms, deep voice = big animal). This can have a territorial function; better defence of territory should lead to more matings and more offspring, and is thus adaptive. All this is perfectly reasonable. But is it true, as Fitch claims, that 'the principles underlying the size exaggeration hypothesis apply equally to humans' (322)? The main evidence offered for this idea is the fact that, in addition to an initial developmental descent of the larynx that occurs in human infants of both sexes in the first months of life, boys undergo a secondary lowering at puberty. Fitch suggests that the latter was driven, in evolution, by body size exaggeration, and concludes that the same selective pressure triggered the primary laryngeal lowering in early hominins. But there's a highly problematic inference here. First, there is no logical link between the initial laryngeal descent in human PHYLOGENY and the secondary developmental lowering, occurring in boys alone, in human ONTOGENY. Since the former occurred in all humans, it seems highly unlikely to have been driven by size exaggeration; human females have no need to exaggerate body size. Fitch notes that both male and female big cats have a reconfigured vocal tract, again driven by size exaggeration, occurring because both sexes defend territories – but this hardly applies to humans, male or female, because hominin individuals do not defend territories. Nor, unlike red deer, do humans practise the system of polygyny in which only the dominant male within a group reproduces; body size exaggeration is clearly adaptive for these deer, but not so clearly applicable to hominin evolution.

Moreover, Fitch himself points out that even before the genus *Homo* appears (over two million years ago), sexual dimorphism in hominins was greatly reduced; in other words, males had stopped needing to be – or to appear – significantly larger than females. This alone makes the size exaggeration hypothesis implausible as an account of the initial stage in human laryngeal descent. Finally, humans did not evolve in a habitat in which they would typically be heard long before they were seen – while modern chimps still inhabit dense forest, our ancestors occupied much more varied terrain, including mosaic woodland and grassland. As for the secondary laryngeal descent in boys, this may simply be down to sexual

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

selection for voice quality, as indicated by the fact that it appears at puberty: all it takes is for generations of females to prefer males with deeper voices (in fact, still the case today; Locke 2009).

Still on the question of adaptations for speech, we come to volitionality of vocalization. It is hard to discover when our ancestors might have gained volitional control of the vocal tract, which is not found in other great apes but is clearly required for speech; almost certainly, neural changes are essential to handle this. As Fitch notes (335), interesting evidence concerning the timing of speech comes from the enlargement of the thoracic canal (MacLarnon & Hewitt 1999, 2004), indicative of highly specific adaptations for the enhanced breathing control that the production of speech requires. These adaptations are not found in earlier *Homo erectus* fossils (*c.* 1.5 million years ago) but are present both in anatomically modern humans and in Neanderthals; thus, also, in their last common ancestor, some half a million years ago (Endicott, Ho & Stringer 2010). Since SPEECH depends on this fine breath control, the most likely explanation for these adaptations is that some form of spoken LANGUAGE had already emerged by this time, with selection pressures for better speech driving further vocal adaptations. A reasonable conclusion would thus be that a spoken protolanguage existed by half a million years ago. In a similar vein, Philip Lieberman (e.g. Lieberman 1984/2006) argues that *erectus* must have had speech of some kind before the restructuring of the SVT, since only the selective advantage of improved communication provides the necessary pressure for the initial vocal tract modifications to be retained, indeed amplified, in later hominins. Fitch's view of the initial use hominins made of their new vocal abilities is, however, radically different. He proposes that 'true meaningful speech was preceded phylogenetically by a song-like system' (335). I turn next to an examination of this strange claim.

4. A MUSICAL PROTOLANGUAGE?

The concept of 'protolanguage' in the language evolution literature connotes a hypothetical biological stage (or stages) before the emergence of full language; note the usage distinct from that of historical linguistics, which indicates an extinct language reconstructed on the basis of existing forms. It was Bickerton's (1990) book that popularized this usage of protolanguage to refer to putative ancestral pre-language, though (and I learned this from the book under review) the term was introduced by Hewes (1973). We have no direct evidence for protolanguage, and all indirect evidence is highly controversial, including extrapolation from language acquisition in children, or from emergent languages such as new sign languages, or pidgins and creoles. In the Minimalist literature there is much rhetoric against the very concept of protolanguage. For instance, Berwick (2011b: 99) claims that 'there is no possibility of an intermediate language between a

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

non-combinatorial syntax and full natural language syntax – one either has Merge in all its generative glory, or one has effectively no combinatorial syntax at all'; similar sentiments are expressed by Berwick & Chomsky (2011) and Chomsky (2010).[9] Of course, it should be noted that there is no evidence for the alternative Minimalist picture either, in which full language emerged *ex nihilo* as a saltation.

Fitch, in any case, parts company with adherents of Minimalism once again by adopting the starting point prevailing in language evolution studies, namely that protolanguage is a valid concept. So what might an initial protolanguage be like? Fitch's views will definitely startle many linguists, and, I think, give a quite inaccurate impression of the careful argumentation surrounding protolanguage found in the work of linguists such as Jackendoff (2002), Bickerton (2009) or Hurford (2012). As we'll see directly, Fitch believes that protolanguage started life as music, more specifically as non-lyrical SONG. To see how Fitch got there we need to understand, as mentioned earlier, that his central concern lies not with the evolution of syntax, symbolism or phonology, but rather, with the appearance of a system of learned vocalization.

For many readers, the emergence of the speech modality will be somewhat tangential to what they really want to know about language evolution, which typically centres on the cognitive features of the language faculty. As noted above, syntax in particular receives little attention from Fitch, but so also does phonology: the focus on speech involves its mechanics and the selection pressures leading to it, not on the cognitive phonological systems. Tellingly, neither phonetics nor phonology have index entries. Given Fitch's extensive research on the biology of vocalization in animals, it's not too surprising that this topic figures prominently here too. Much emphasis is placed on the evolution of vocal learning, complex vocal imitation ('the ability to imitate complex multisyllabic vocalizations', 340) and vocal control, all of which are rare among mammals generally, and, at best, extremely limited in non-human primates. One of Fitch's key objectives is thus to account for the emergence of these traits, which he regards as a major discontinuity in the hominin lineage, and he is critical of existing work on protolanguage which, in his view, downplays the significance of learned vocalization.

So, if we want to find out about the evolution of vocal learning, imitation and control, where to look? 'Data on repeated convergent evolution of "song" in animals … provide a powerful source of potential insight into a critical aspect of the evolution of spoken language' (341). Specifically, Fitch adopts an idea of Darwin's: 'that the first protolanguage was musical, and that this stage was driven by sexual selection (by analogy with learned birdsong)' (399). Moreover, reprising his point (discussed in Section 2 above)

[9] Extensive arguments against these views appear in Heine & Kuteva (2007), Hurford (2012: Chapter 8) and Tallerman (2013a).

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

that we should still find elements of protolanguage in modern humans, Fitch sees evidence for musical protolanguage in the continued existence of music in modern cultures: 'music is seen as an ongoing exemplar of an earlier protolanguage' (433); 'music … is the best example so far of a hypothetical protolanguage that is "alive and kicking"' (470).

Just in case the word 'musical', as applied to protolanguage above, seems like a metaphor for something else, here's a representative sample of what Fitch is claiming:

> The acquisition of complex vocal learning occurred during an initial song-like stage of communication that lacked propositional meaning. … This system of "bare phonology" provided a learned, complex, generative vocal communication system, with multiple units being combined into a hierarchical, but propositionally meaningless, signaling system. (503)
>
> To the extent that there was a lexicon, it was a simple list of tunes or "riffs" – complex, multi-unit phrases linked to whole, context-bound events. (476)
>
> [T]he central insight of musical protolanguage theories [is] their ability to account for phonological flexibility and generativity. (485)
>
> [M]usical protolanguage provided a preadaptation for spoken phonology and some aspects of syntax. (481)

By 'music', Fitch means 'bare phonology', or 'non-lyrical song' – 'song that lacks meaningful words' (468); he notes that the term 'protomusic' would be just as valid as 'protolanguage' (479). In an astonishing understatement, Fitch asserts that 'the main difference between spoken language and non-lyrical song is SIMPLY that the latter lacks specific, propositional meaning' (469, emphasis added). The inference is clearly that it's but a small step from song to language; indeed, he claims, 'song possesses the characteristics of openness and generativity, as well as cultural transmission' needed for language (469). Thus, at a stroke, just about all the complexity that keeps linguists in gainful employment is swept aside.

Let's unpack all this and see where it takes us. Fitch's use of the term 'phonology' (note: not 'phonetics') is wholly deliberate; he really believes that: 'the generative aspect of phonology … emerged before it was put to meaningful use' (471). The first problem is that there's no possible phonology in 'bare phonology'. Fitch envisages musical protolanguage as containing 'phonological "words" which are just chunks of syllables, but not morphemes' (499). As evidence for such a system, he cites pseudowords – possible but non-existing words in extant languages: '[t]he very existence of a category of pseudowords … encourages us to consider phonology as an independent generative system' (468) which may have a meaningless output, meanings being attached later. But of course, pseudowords are dependent on the prior existence of MEANINGFUL words, which establish the phonotactic

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

constraints of a language in the first place.[10] These are cognitive facts; phonology is linguistic, and cannot exist in isolation from language. Crucially, Fitch's idea of 'phonology first, semantics later' (475) is entirely the wrong way round. As Studdert-Kennedy & Goldstein have observed, we cannot assume the prior existence of a set of 'independent phonetic segments, devoid of meaning'; consonants and vowels are linguistic (cognitive) entities, 'precisely what an evolutionary account must explain' (2003: 238). Demonstrably in ontogeny, and putatively in phylogeny, an expanding phonology derives from a growing vocabulary, and not the other way round. The phylogenetic position is elegantly established, using computational modelling, by Björn Lindblom (e.g. Lindblom 1992, 1998; see also Zuidema & de Boer 2009, de Boer & Zuidema 2010, Studdert-Kennedy 2012). Lindblom models the emergence of discrete phonological gestures, syllables and phonemes by processes of self-organization. He demonstrates that 'phonological units are EMERGENTS of the developing lexical system rather than prespecified entities' [emphasis in original] and that 'phonemic coding is an automatic consequence of the accumulation of a large lexicon (absent in other species)' (Lindblom 1998: 255). Thus, Fitch's idea that the first stage in the evolution of protolanguage involved 'meaningless sung phrases of complex phonological structure' (496) is unsubstantiated.

At the heart of Fitch's argument is the claim that sexual selection in early hominins drove the evolution of a learned 'song-like stage of communication' (503). The analogy with learned birdsong, also driven by sexual selection, is repeatedly made.[11] This raises the question of how good an analogy learned birdsong or whalesong makes either with music in humans or with the acquisition of language. In fact, the musicologist Aniruddh Patel presents detailed arguments against the idea of a biological parallel between learned animal song and music (Patel 2008). Moreover, Fitch's belief (e.g. 490ff.) that the evolution of music itself was driven by sexual selection (following Darwin 1871 and Miller 2000, 2001) has come under sustained criticism in a great deal of recent work. Reviewing another book containing similar speculations about a musical protolanguage (viz. Mithen 2005), the music psychologist John Bispham notes that '[a] sexual selection model for

[10] A referee for *JL* suggests that phonotactics may simply be knowledge of transitional probabilities, as is indicated by the fact that subjects can learn inductively the distinction between possible and impossible words in artificial languages (Heinz 2007). Thus, supporting Fitch, phonotactics could exist independently of meaning. I'm not sure how this would be applicable to language evolution, however, since in language the phonotactic constraints are driven by vocabulary that does have meaning.

[11] Actually, ornithologists propose various possible explanations for the emergence of learned song. After discussing four potential selective pressures, Peter Slater, probably the world authority on birdsong, concludes that '[t]he reason why songbirds learn the songs that they sing therefore remains something of a mystery' (Slater 2012: 100).

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

music … has been widely and comprehensively rejected by a range of arguments' (2006: 591).

Let us then compare the biology of birdsong and other animal song with that of language, including some of Patel's (2008) arguments regarding music along the way.[12] Certainly, there is one notable biological parallel between birdsong and human speech (which oddly, Fitch does not discuss): 'there are significant similarities in the neural circuitry used for the production and learning of vocalizations' (Hurford 2012: 22). Though learned vocalization clearly evolved independently in humans and in vocal-learning bird groups, Jarvis (2004) suggests that the neural pathways involved in both vocalization systems are at least partially homologous. Interesting though this is, vocalization is not language, but merely a modality for language. Moreover, our ancestors weren't birds, but primates, and even focussing exclusively on vocalization, what we most need to know is how THEY gained vocal communication systems that were learned, not innate, and volitional, not affective (i.e. emotionally driven). Can animal song tell us anything about this? Can it tell us anything about language? Here is a brief survey of the potential parallels between animal song and language.

First, animal song is purely an acoustic display, and is closely tied with fitness in an evolutionary sense: birds sing because it enhances their chances of reproducing. Music, on the other hand, is created rather than being a display (Patel 2008), and while it may be linked in some way to fitness (Bispham 2006), it's not at all obvious how. Language in modern humans is not tied to fitness in any immediate sense. While we assume that the evolution of language was associated in SOME way with fitness, it's very hard to make a convincing case for this (though it's easy to propose a rather facile argument, since we can only too readily imagine the adaptive consequences of even a limited protolanguage).

Second, birdsong often involves a huge repertoire, but it has a very restricted message (Slater 2012), essentially concerning either the attraction of a mate, or defence of territory (and of course, these factors are linked); there is no productivity either in the message or, actually, in its form (Hurford 2012: Chapter 1 has an excellent summary). Though animal songs sometimes have 'dialects', in the sense that they differ in form from place to place, no distinctions in meaning attach to these different forms. Clearly, both language and music differ radically in this regard from animal song. Music has a massive variety of forms and culturally-specific 'messages', and language is

[12] There is, of course, a large literature on parallels and distinctions between the (contemporary) STRUCTURES of language and music, as well as work on their biology and possible neurological similarities; some central references are Fitch (2006), Lerdahl & Jackendoff (2006), Patel (2008), Jackendoff (2009, 2011), Katz & Pesetsky (2011). This topic is well beyond both my expertise and the scope of this article. David Pesetsky (p.c.) suggests that 'music is parasitic on the language faculty', which to me makes perfect sense; in the book under review, Fitch is apparently proposing the converse.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

fully productive with an unlimited set of meanings, and in form is again highly culture-specific.

Third, animal songs are mostly produced by males, and are hormonally driven, thus displaying seasonal peaks. Music exhibits none of these features; nor – crucially – does language. Fitch isn't terribly worried by these major discrepancies. He points out (342) that females of many bird species do sing, and suggests one scenario in which it is co-defence of territory that prompts the emergence of learned song in both sexes (491). But this must surely be discounted as a selection pressure in human evolution: as noted in Section 3 above, there is no evidence that hominin male–female pairs defended a territory; on the contrary, our ancestors are assumed to have lived in small social groups, possibly with the fission/fusion behaviour typical of modern chimpanzees (Wynn 2012). The other possibility Fitch suggests (491) is that mutual mate choice can lead to both sexes having learned song. If males become choosy about female partners (entirely possible in species where males contribute to raising offspring), then females as well as males may need vocal display. According to this story, thus could musical protolanguage arise in hominins of both sexes. The problem with this is that if it were true, the size exaggeration hypothesis for the lowering of the larynx (321ff.) goes out of the window too – as noted above, size exaggeration is about males getting access to as many females as possible, and there is no room in that picture for males to be choosy.[13]

A fourth factor is that – like language – learned song is culturally transmitted, and has a sensitive period for acquisition. Music, conversely, has no critical period for acquisition, and though it is culturally transmitted, this is by no means inevitable in humans – music often fails to be acquired, and around 4% of humans are tone deaf (Patel 2008). But here, surely, is the crucial comparison that Fitch needs between learned song and language – or so it appears at first glance: both have a critical period. However, Gibson (2012) points out that sensitive periods in animal behavioural systems are extremely widespread, and by no means confined to learned vocalization. Moreover, there are important differences between the learning that occurs in the sensitive periods for birdsong and language. As Hurford (2012: 23) notes: 'a remarkable difference between nestlings and babies growing up and learning their language is the fact that birds do not learn their song incrementally through a process of discourse with their parents (or other group members). Birds store the patterns they hear as nestlings, and then only later, sometimes as much as eight months later, start to produce their own songs'. It is as if human infants were to learn some stock phrases by overhearing

[13] Christina Behme suggests to me (p.c.) that Fitch's scenario could be salvaged if these two stages occurred in sequence: first, the hominin vocal tract lowers for size exaggeration; later, both sexes practise territorial defence singing. As noted above, though, evidence for either stage is lacking.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

adults, but say nothing themselves until they reached puberty, when they would start to reproduce these stored phrases verbatim. And as Fitch himself has indicated elsewhere (2004: 287), the sexual selection proposal for language is problematic because 'most sexual display characters in the animal kingdom arise at puberty', which is why only sexually mature birds sing; clearly, this has no parallel whatever to language.

Thus, the sexual selection model for the evolution of language, based on a proposed parallel between learned animal song and language, does not seem well supported by the evidence. In sum, there are two questions that it's crucial to ask regarding this putative analogy between vocal learning in animals and that of humans.

a. Can the convergent evolution of a learned song system in birds and other taxa tell us anything about the evolution of learned vocalization in humans?
b. Can the evolution of learned vocalization in humans tell us anything about the evolution of language AS A FACULTY?

Regarding question a., I'd say on the basis of the summary above, rather little. The parallels are mostly tenuous; though the similarities in neural pathways are appealing, the resulting systems are so massively different from each other that these findings remain an intriguing observation rather than a basis for productive research. Regarding question b., I'd say that the answer is: only in an extremely limited sense. It's definitely of interest to explore, as Fitch does, the question of how our ape ancestors gained a learned communication system involving vocal imitation. But language itself is a cognitive entity (this includes many of the features that make up FLB, such as memory, storage and retrieval of lexical items), and though speech serves it well for externalization, speech is not the essence of the language faculty. To remind the reader, the book title is *The evolution of language*.

5. PROTOLANGUAGE BEYOND MUSIC

So far, we've explored Fitch's claims that linguistic phonology originates as 'bare phonology' (phonology without meaning) in a musical or 'prosodic' protolanguage (his preferred term). In fact, Fitch envisages the scope of this idea as being significantly more extensive: 'this model delivers not just the generative aspects of phonology, but also some important parts of the syntactic interface' (475). Again, lest this sounds improbable, I offer some illustrative quotes (475):

> [T]he "bare phonology" postulated by this model provides sequences of units, arranged into phrases, and also rules for attributing phrase boundaries … [It thus] goes far beyond simple phonetic segments to include some aspects of syntax.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

> This model of protolanguage has other virtues … Regarding syntax, prosodic protolanguage provides meaningless but hierarchically structured signals, which include phrases but lack many other syntactic complexities.

And later, we learn that 'bare phonology' provides 'simple "phonological" aspects of syntax (sequencing, hierarchy, and phrase structure)' (503).

Insofar as I understand these proposals, Fitch seems to be saying that phrases already existed in the earliest forms of protolanguage, just waiting to be filled by meaningful words that came along later. 'Later' is pictured by Fitch as a (spoken) holistic protolanguage stage, which would follow the prosodic protolanguage phase. I'll digress briefly to outline what this means. Holistic protolanguage (like 'bare phonology') has no words as such; it is not compositional, but contains 'holistic' signals, each with its own propositional meaning: 'whole phrases [are] linked to whole meanings' (497) and 'broad, holistic meanings ('stay away' etc.) were attached to … complex signals' (496). Over time, words are purported to fractionate out of these unanalysable signals, as early humans notice chance similarities between chunks of sound and pieces of meaning. Such ideas often seem appealing to non-linguists (e.g. Mithen 2005), since they apparently offer a way to get from no language to language without positing too much new stuff in primate cognition and/or vocalization. But non-human primates have far less cortical control over vocalization than humans; moreover, their innate vocalizations – and indeed our own innate calls, such as crying – are driven by distinct neural circuits from those controlling spoken language (Ploog 2002). So newly evolved mechanisms must be accounted for anyway. Moreover, modern primates lack the 'fractionation' skills required to form words from unanalysable sound strings,[14] so even if such skills existed in our ancestors, they clearly aren't phylogenetically ancient: thus, proponents of holistic protolanguage need to explain how these abilities arose too. Finally, holistic protolanguage is a cul-de-sac since words are ultimately going to be needed in any case.

However, as mentioned earlier, Fitch believes that '[h]olistic protolanguage is still with us today' (504) in the form of unanalysed formulaic utterances such as *How do you do?*, and he assumes that this constitutes evidence in favour of a holistic pre-language stage. But actually, formulae of this kind start off life as fully analysable and compositional, and become unanalysed through historical processes, whereas the putative holistic utterances of pre-language are just the opposite, starting off as unanalysable and

[14] A referee for *JL* suggests that Hauser, Newport & Aslin (2001) reach the opposite conclusion in their study of cotton-top tamarins. In fact, the authors conclude that the monkeys are 'capable of computing the same rather complex statistical regularities on fluent and rapid streams of synthetic human speech that have been demonstrated … in human infants' (2001: 62). This is not remotely the same as forming new words.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

subsequently getting analysed. For these processes, modern analogues are limited to back formations like *edit* from *editor* and reanalyses like *beef+burger* from *hamburg+er*, which are in any case totally reliant on pre-existing meaningful morphemes (/-ər/ for *edit*; *ham*).[15]

Returning to the main story, Fitch's proposal is that phrases are exapted – taken over to fulfil a new function – by syntax when the holistic stage is finally resolved into words.[16] But phrases in syntax are not pre-existing hollow vessels, either in ontogeny or phylogeny. Crucially, syntactic phrases are HEADED, and what they are headed by is a lexical item of a specific syntactic/semantic class. Let's assume, reasonably, that the earliest words (proto-heads?) were purely semantic, with no lexical requirements. I think it's also reasonable to assume, as argued by Hurford (2012: Chapter 9.3), that the earliest semantic distinction was the Topic/Comment contrast. Gradually, semantic heads accrue dependents, again at first purely semantic, and entirely optional, since there is no syntax – these are just protowords that are sometimes or typically used with heads of a certain meaning. Over time, relationships develop between a semantic head and its dependents; some of the latter start to be used consistently, and their precise properties are selected by the semantic head. Note that distinct LEXICAL classes of head can't exist until there's a behavioural contrast between at least two of them. When a 'phrase' consists of a head and a specific, obligatory dependent, we have syntax; now, a head is syntactic, rather than merely semantic. Since different head types select different dependent types, we then have distinct word classes; putatively, just nouns and verbs at first (Heine & Kuteva 2007). Now compare phrases in music or birdsong. These have radically different building blocks (Katz & Pesetsky 2011). They are not built up of lexical items, or anything analogous; words have no non-linguistic counterpart in music or animal song. Put simply, in syntax words must come first; phrases are built up around single words, as indeed they are in child language acquisition today. Thus, suggesting that phrases evolved in protolanguage BEFORE there were word classes is once again entirely the wrong way round.

Unfortunately, Fitch's casual use of terms such as 'phrase structure', 'hierarchical structure' and 'phrase boundaries' is likely to misinform non-linguists, who may conclude that there's nothing special at all about phrases

[15] Holistic proposals are generally not well supported by linguists; see Tallerman (2007, 2010) and other papers in Arbib & Bickerton (2010), Hurford (2012: Chapter 8.5). Conceptual and empirical evidence against holistic models is now overwhelming, and I will not rehearse all the arguments again here.

[16] Worryingly, Fitch comments that '[h]olistic utterances … bear more resemblance to "words" in highly polysynthetic languages …, in that each is composed of many syllables, but before they have been successfully analyzed into their component morphemes by the child' (498). This comparison does serious injustice to polysynthesis, where of course each morpheme has a specific position in a template, and is systematically attached as a prefix or suffix to a stem. In authors' sketches of the putative holistic utterances of pre-language, there is nothing systematic whatever, and there are no stems or grammatical affixes.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

in language. The problem is that Chapter 14, which is the final chapter apart from a brief concluding chapter, gives the impression that pretty well everything syntactic was already in place long before any meaningful language had emerged, so that little had to evolve to give us the kind of syntax we see in language today. I believe that impression to be a potentially damaging one for the wider field, even likely to stifle further progress in language evolution if taken seriously by biologists, neurologists and so on.

6. Cheapness and honesty: A problem for language evolution?

Above, I focus on what Fitch regards as one of the major problems in the evolutionary debate: the emergence of learned vocalization. A second 'fundamental problem' in language evolution (194) is the development of a system of cheap, honest signals, unexpected in animal communication (Maynard Smith & Harper 2003). These terms are likely to require some gloss for the present audience. Fitch notes that externalized language is a biologically 'cheap' signal, in the sense that speech – metabolically speaking – 'costs' nothing to produce (presumably, sign is slightly more costly, but not much). This is the prevailing view, but I think there is more to say. SPEECH may well be cheap, but that cannot be taken to imply that the emergence of LANGUAGE was without cost: specifically, the gradual evolution of the large human brain (today, approximately three times the size of other great ape brains) was certainly not low cost. Metabolically, as Fitch notes (288), our brains are highly expensive, in modern humans consuming up to 25% of our calorific intake but accounting for just 2% of body weight (see Aiello & Wheeler 1995). While the relationship between brain size and language is unknown (did brain growth facilitate language? Did an evolving language faculty drive brain growth? How large a brain do we need for the full language faculty?), most observers assume that there must be SOME causal link here (see Bickerton 1990: 133–136 for discussion). It's true that once a linguistic system is in place, linguistic signals themselves are cheap to produce. But the brain still has to be maintained, and this may be the real cost associated with language.

Honesty appears more problematic. 'Honest' signals (not necessarily vocal signals) in animal communication provide reliable information to those that perceive the signal. But the evolution of a stable system of reliable signals is surprising, because if an animal gains some advantage over conspecifics by deceptive signalling, then it's adaptive to deceive. In time, however, others realize that those signals are unreliable, and ignore them. Then there's no point in producing unreliable signals, since they have no further effect, but producing reliable signals may be maladaptive – so signalling just ceases entirely. Thus, Fitch states that accounting for the cooperation that would be necessary to share 'detailed, truthful information' is

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

'a core issue in the evolution of language' (414), and that 'detailed, propositional information exchange … is one of the central oddities of our species from an evolutionary viewpoint' (425). So how could such a system evolve?

Various evolutionary 'options' keep animal signals reliable (Hurford 2007: Chapter 8 and Scott-Phillips 2008 have good summaries). Some signals are necessarily honest because they cannot be faked: an example is the conspicuous sexual swellings occurring in many female primates when in oestrus. Handicaps, such as a deer stag's antlers or the peacock's tail, are honest signals because they 'cost' the possessor (often by literally weighing him down, so that only strong specimens can bear the cost); thus, biological quality is reliably signalled both to male competitors and potential mates. But these conditions don't apply to language, and no handicaps (apart from brain growth as noted) seem to be associated with it. Fitch sums up the problem as he sees it:

> Athough humans freely and continuously use language to share accurate information with unrelated others, we appear to have avoided any costly or "handicap" requirement for honest signaling. Indeed, human language appears … to have evaded ALL known theoretical routes to honest signaling. (195; emphasis in original)

Before turning to Fitch's proposed solution, let's examine the premise – that the primary (in the sense of both 'original' and 'major') function of language was the sharing of information between hominins. Such behaviour, as Fitch rightly notes, would be unexpected in evolutionary terms. True altruism is rare in the natural world, since helping a conspecific leaves an animal with fewer resources to handle its own daily survival, thus potentially reducing its biological fitness (reproductive success). So why would our ancestors give away useful information gratuitously? Yet here, surely, we have a Fallacy of First Use problem (Bickerton 2002: 222): the focus on 'information sharing' mistakes current utility for 'original' purpose. We don't know what the earliest users of externalized language did with it. Perhaps the 'original' function of (proto)language was phatic social communication with no 'information' content whatever – quite an appealing idea, actually, given that one of the major biological distinctions between externalized language and many animal signals is that humans DELIBERATELY communicate with each other.[17] The realization that conspecifics are trying to communicate something is essential, and must surely precede any kind of protolanguage.[18]

[17] In fairness, Fitch (428) also considers phatic communication and suggests it may be an adaptation.

[18] The interesting question of how the first protolanguage users even understood that others were trying to communicate with them (how to 'signal signalhood'?) is addressed in recent work on the ostensive-inferential nature of communication (e.g. Scott-Phillips et al. 2009).

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

Somewhere very early on in hominin evolution, the crucial ability to engage in a dialogue with conspecifics must have evolved, part of a suite of cognitive social abilities that are particularly well developed in humans (shared intentionality; Tomasello et al. 2005, Tomasello 2008). 'Dialogue' here should not be taken to mean any kind of linguistic communication; for instance, at first it might simply mean one person smiling and the other nodding. Ultimately, communication about a third entity is required (joint attention – which perhaps starts with pointing, another characteristic not found in wild apes). Talking about external stuff is central to linguistic communication, since it buys the property of REFERENCE (Hurford 2007: Chapter 7 has a full survey). But maybe the earliest pre-linguistic dialogues were 'about' nothing at all, but nonetheless served to build trust and helped to create that very human mutuality. This is pure speculation. More concrete is the observation that externalized (proto)language must have been very well established before it had the structural and semantic resources to convey 'detailed, truthful information' (414). Keeping this thought in mind, let's examine Fitch's solution to the honest signalling problem.

Though true altruism is unlikely in nature, many animals are strongly predisposed to help close relatives; in the case of mammals and birds, this particularly means helping young offspring, for instance by provisioning them. This is known as KIN SELECTION. Helping offspring survive increases an animal's DIRECT FITNESS: its genetic contribution to the next generation. Honest signalling to offspring is thus predictable. Other close relatives also share many of one's genes, and helping them as well increases an animal's INCLUSIVE FITNESS – the overall contribution that a single genotype makes to the next generation (Hamilton 1963, 1964). Textbook examples include the predator alarm calls that occur in various monkey species, where members of a troop are closely related. So the answer to the otherwise surprising evolution of apparently altruistic acts typically lies in kin selection. For Fitch, kin selection is relevant to the very earliest stage of language evolution: 'information sharing evolved to help closely related individuals, and particularly dependent young' (425).[19]

Fitch's views are by no means unusual in the recent literature; kin selection/language evolution links of one sort or another are also proposed by other authors, notably Locke & Bogin (2006) and Falk (2004). Tallerman (2013b) argues that these proposals are highly problematic. Here, I consider some of the arguments against the kin selection idea for language. First, let's

[19] Fitch even suggests that kin selection operates at a pre-semantic stage 'as the selective force behind musical protolanguage: aiding relatives' (492). Since little or no information could be transmitted at such a stage, I regard this as an entirely separate proposal. Arguments against it can be found in Tallerman (2013b).

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

consider what could be transmitted at Fitch's 'Stage 1: kin selection for information exchange' (425):

> Adults who could inform young relatives about cryptic or seasonal sources of food, water, or danger, the characteristics of materials for tools, the techniques for food processing and hunting, and a host of other valuable facts would be increasing their inclusive fitness with every proposition conveyed. (426–427)

The holistic protolanguage that Fitch supports, putatively in use in the *Homo erectus* era, is intended to function largely to manipulate others and in this sense 'is continuous with nonhuman primate communication systems' (497). This doesn't suggest to me that it was indeed capable of conveying 'a host of valuable facts'. Fitch himself describes holistic protolanguage as a 'limited, functional system' which 'sufficed for the communicative needs' of *erectus* (498). Logically, there seem to be two possibilities here. The first is that *erectus* actually didn't have much to communicate to their kids, in which case we might save holistic protolanguage, with its limited linguistic resources and limited function, but we'd need to throw out the kin selection proposal for protolanguage: if you can't communicate productively about pressing survival matters such as 'seasonal sources of food' then you aren't going to help your offspring much by talking to them. Alternatively – and far more realistically – *erectus* had a lot of useful things to communicate to offspring, in which case we must throw out holistic protolanguage in favour of a more promisingly productive kind of protolanguage (lexical protolanguage, such as Jackendoff's 2002 or Bickerton's 2009).

Second, however, is the problem that explicit TEACHING about 'the characteristics of materials for tools, the techniques for food processing and hunting' etc. (426) is not at all how even modern humans, with full linguistic resources, go about acquiring practical skills. Though we might imagine that language is indispensable for this procedural knowledge, an extensive ethnographic literature clearly shows just the opposite: both in industrialized and non-industrialized societies, direct verbal instruction in practical training is extremely limited. Instead, skills of all kinds, from tool-making and stone-knapping through to silversmithing, weaving, farming, fishing and even dentistry and surgical skills are learned predominantly via observation and imitation (Tallerman 2013b has a short overview). Experimental research, involving teaching students to make stone tools either with or without verbal instruction (Ohnuma et al. 1997), also reached the same conclusion: the 'with language' group didn't learn faster. Not only is language unnecessary for teaching practical skills, it's apparently not even particularly helpful. I certainly don't doubt that teaching and learning can be adaptive (Thornton & Raihani 2008). However, for the 'language as kin selection' argument to hold water, our ancestors' limited protolinguistic resources would need to have been so successful in teaching their offspring

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

that this was adaptive – enhancing fitness in both parent and child. For procedural knowledge (learning practical skills) this clearly appears not to be the case.

Moreover, it seems that the most important factor in the supremely human talent for learning from others lies not in teaching, but in learning itself. For a start, humans have evolved a far greater tendency to learn in a social setting than other great apes; the term used for this disposition is DOCILITY, which describes 'persons who are adept at social learning, who accept well the instruction society provides them' (Simon 1990: 1666). Since learning critical survival skills is inherently likely to increase one's reproductive fitness, it's adaptive, and thus docility, the propensity to learn from others, is also adaptive. In addition, since languages rely on TRADITIONAL TRANSMISSION (being passed on from one generation to the next), (proto)language itself must have required docility from its earliest users. Conversely, of course, language transmission does not require any teaching skills. Thus, by stressing the importance of social learning rather than teaching, we build in an ability that is needed independently in language evolution.[20]

Third, using the comparative method favoured by Fitch, we find that the other great apes are not in the habit of direct teaching. Thornton & Raihani (2008: 1828) argue that '[a]ll reported examples of teaching among apes are qualitative, anecdotal and open to alternative interpretation'. Even reports of ape teaching are rare.[21] This is not too surprising, since, as Thornton & Raihani (2008) argue, teaching is a form of cooperative behaviour, in which humans but not other apes excel. The ape evidence is important, because if, as Fitch says, '[k]in communication … drive[s] the evolution of propositional information sharing' (426), then the tendency to transmit useful information NON-VERBALLY must already be in place in our ancestors before any kind of protolanguage exists. If protolanguage enables better teaching, then teaching must exist first, since evolution cannot plan ahead and see what might be useful. But the comparative method shows that teaching is not phylogenetically ancient; rather, we are the innovators, while the other great apes engage little in direct social transmission of skills, and instead their infants mainly learn by trial and error (Tomasello 1999a).

Since teaching is not phylogenetically ancient, I suggest that what is more likely is that protolanguage evolved BEFORE teaching, and was then subsequently used for transmitting declarative knowledge (teaching facts, such as

[20] It's also clear that human learning procedures are very different from those of other great apes (Whiten et al. 2004). Humans acquiring a skill pay more attention to the actions and techniques of the proficient user, thus focussing on the process itself, whereas apes concentrate on the desired outcome (Tomasello 1999b, 2008: 208f.).

[21] One paper always cited in this regard is Boesch (1991), which documents chimpanzee mothers in the Taï National Park helping their offspring to learn the difficult skill of nut-cracking using a hammer and anvil. However, Boesch also notes that teaching hasn't been observed in other wild chimp populations, which lack the nut-cracking culture of Taï.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

information about seasonal food sources and so on). From this point, I have no arguments with claims that language can be adaptive in the context of teaching; what I cannot endorse is the idea that language evolved 'for' teaching kin, and that this was the major selection pressure in language evolution.

I have examined a variety of evidence suggesting that 'kin selection for information exchange' (425) is not well supported. However, even if the earliest protolanguages were used for information-sharing, as Fitch assumes, a straightforward resolution to the biological problem of honesty need not involve kin selection at all. Some solutions that seem highly promising already exist. Scott-Phillips (2008: 278) proposes the idea of 'reputation as deterrent'. Here, costs in the form of social exclusion are incurred by unreliable signallers; large-brained primates are well able to recall instances of bad behaviour in conspecifics, and if the perpetrators are isolated, this acts as a strong deterrent to antisocial behaviour in a highly social species. Moreover, as protolanguage develops, it can also be used to bolster this arrangement, since information can be exchanged about the reliability of others, 'and that exchange would itself be kept reliable by the very same mechanisms' (Scott-Phillips 2008: 279). Crucially, no altruism is involved.

A slightly different scenario is outlined by Tomasello (2008: Chapter 5). For instance, '[o]ur proposal is that human cooperative communication was adaptive initially because it arose in the context of mutualistic collaborative activities in which individuals helping others were simultaneously helping themselves' (2008: 170). Tomasello also explicitly invokes the need for individuals to preserve a good reputation in order to maintain social acceptance. A similar solution – again not involving any altruism – is also proposed by Bickerton (2009). For Bickerton, the specific collaborative activity engaged in by early hominins was power scavenging: this involves recruiting aid from as many people as possible to drive away dangerous animal scavengers from the carcasses of dead megafauna, so that puny hominins can butcher the meat in safety. The idea is that one person spots the carcass, and uses signals (some early form of protolanguage) to recruit other people. It pays the signaller to be 'honest', since it's far too dangerous for one person alone to attempt to exploit the carcass; co-operation thus benefits all parties. Reputation again plays a part here; faking information about a distant carcass would quickly result in social isolation (and no access to food). Bickerton also points out that co-operation between non-kin is vital in this scenario. All these ideas provide biologically plausible solutions to the problem of honest signalling – none involve any problematic altruism – and they additionally build in the specifically human type of co-operation that is otherwise so hard to account for.[22]

[22] Many of the issues surrounding the evolution of an honest signalling system are explored in greater detail by Clark (2011) in an insightful review of Bickerton (2009) in this journal.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

7. Should we adopt *The Evolution of Language*?

Since courses in language evolution are a growth area in higher education, it is worth briefly considering what book might make a good set text for such modules. It's true that there are, as yet, few alternative options. Christiansen & Kirby (2003) was a great survey of the field a decade ago, but the field moves very quickly; moreover, it's not actually a textbook. Johansson (2005) is very readable and would be (mostly) an excellent choice if it were updated and perhaps retitled. Though it covers much of the same ground as Fitch, I cannot recommend the very non-linguistic (and highly anti-generative) Beaken (2011). Anderson's (2004) excellent text comparing animal communication systems and language also overlaps quite a bit with Fitch in content, but it's not intended to centre on language evolution. Tallerman & Gibson (2012) is not a textbook, and is not intended primarily for undergraduate use. Hurford's two volumes (2007, 2012), both of which have been referred to extensively in this review, are engaging, comprehensive and very well written. However, they're BIG books and too expensive for use as texts. So yes, I do mention *The Evolution of Language* as a possible purchase (though not a set text) for my own course, because there is indeed much valuable material in there – but the parts on protolanguage must be used with extreme caution, and most linguists would want the emphasis to be on language as a cognitive property, and not on vocalization.

REFERENCES

Aiello, Leslie C. & Peter Wheeler. 1995. The expensive-tissue hypothesis – the brain and the digestive system in human and primate evolution. *Current Anthropology* 36, 199–221.

Anderson, Stephen R. 2004. *Doctor Doolittle's delusion: Animals and the uniqueness of human language*. New Haven, CT & London: Yale University Press.

Arbib, Michael A. & Derek Bickerton (eds.). 2010. *The emergence of protolanguage: Holophrasis vs compositionality*. Amsterdam & Philadelphia, PA: John Benjamins.

Beaken, Mike. 2011. *The making of language*, 2nd edn. Edinburgh: Dunedin Academic Press.

Berwick, Robert. 2011a. All you need is Merge: Biology, computation, and language from the bottom up. In Di Sciullo & Boeckx (eds.), 461–491.

Berwick, Robert. 2011b. Syntax facit saltum redux: Biolinguistics and the leap to syntax. In Di Sciullo & Boeckx (eds.), 65–99.

Berwick, Robert & Noam Chomsky. 2011. The biolinguistic program: The current state of its development. In Di Sciullo & Boeckx (eds.), 19–41.

Bickerton, Derek. 1990. *Language and species*. Chicago, IL: University of Chicago Press.

Bickerton, Derek. 2000. How protolanguage became language. In Chris Knight, Michael Studdert-Kennedy & James R. Hurford (eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form*, 264–284. Cambridge: Cambridge University Press.

Bickerton, Derek. 2002. Foraging versus social intelligence in the evolution of protolanguage. In Alison Wray (ed.), *The transition to language*, 207–225. Oxford: Oxford University Press.

Bickerton, Derek. 2009. *Adam's tongue: How humans made language, how language made humans*. New York: Hill & Wang.

Bickerton, Derek & Eörs Szathmáry (eds.). 2009. *Biological foundations and origins of syntax*. Cambridge, MA & London: MIT Press.

Bispham, John. 2006. "Music" means nothing if we don't know what it means. *Journal of Human Evolution* 50, 587–593.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

Boesch, Christophe. 1991. Teaching among wild chimpanzees. *Animal Behaviour* 41, 530–532.
Byrne, Richard W. & Anne E. Russon. 1998. Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences* 21, 667–721.
Cheney, Dorothy L. & Robert M. Seyfarth. 2005. Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review* 22, 135–159.
Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2010. Some simple evo devo theses: How true might they be for language? In Larson et al. (eds.), 45–62.
Chomsky, Noam. 2011. On the poverty of the stimulus. Lecture presented at UCL, 10 October 2011. Available at http://www.ucl.ac.uk/psychlangsci/research/linguistics/news-events/latest-news/n_chomsky.
Christiansen, Morten H. & Simon Kirby (eds.). 2003. *Language evolution*. Oxford: Oxford University Press.
Clark, Brady. 2011. Scavenging, the stag hunt, and the evolution of language. *Journal of Linguistics* 47, 447–480.
Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.
Darwin, Charles. 1871. *The descent of man, and selection in relation to sex*. London: John Murray.
de Boer, Bart & Willem Zuidema. 2010. Multi-agent simulations of the evolution of combinatorial phonology. *Adaptive Behavior* 18, 141–154.
Di Sciullo, Anna Maria & Cedric Boeckx (eds.). 2011. *The Biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*. Oxford: Oxford University Press.
Diller, Karl & Rebecca L. Cann. 2009. Evidence against a genetic-based revolution in language 50,000 years ago. In Rudolf P. Botha & Chris Knight (eds.), *The cradle of language*, 135–149. Oxford: Oxford University Press.
Endicott, Phillip, Simon Y. W. Ho & Chris B. Stringer. 2010. Using genetic evidence to evaluate four palaeoanthropological hypotheses for the timing of Neanderthal and modern human origins. *Journal of Human Evolution* 59, 87–95.
Falk, Dean. 2004. Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences* 27, 491–541.
Fitch, W. Tecumseh. 2004. Kin selection and 'mother tongues': A neglected component in language evolution. In D. Kimbrough Oller & Ulrike Griebel (eds.), *Evolution of communication systems: A comparative approach*. Cambridge, MA & London: MIT Press, 275–296.
Fitch, W. Tecumseh. 2006. The biology and evolution of music: A comparative perspective. *Cognition* 100, 173–215.
Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179–210.
Gibson, Kathleen R. 1993. Tool use, language and social behavior in relationship to information processing capacities. In Kathleen R. Gibson & Tim Ingold (eds.), *Tools, language and cognition in human evolution*, 251–269. Cambridge: Cambridge University Press.
Gibson, Kathleen R. 1996. The ontogeny and evolution of the brain, cognition and language. In Andrew Lock & Charles R. Peters (eds.), *Handbook of human symbolic evolution*, 409–433. New York: Clarendon Press.
Gibson, Kathleen R. 2012. Are other animals as smart as great apes? Do others provide better models for the evolution of speech or language? In Tallerman & Gibson (eds.), 120–130.
Greenfield, Patricia M. 1991. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behaviour. *Behavioral and Brain Sciences* 14, 531–595 (with continuing commentary and a new response (1998), 21, 154–163).
Hamilton, William D. 1963. The evolution of altruistic behavior. *American Naturalist* 97, 354–356.
Hamilton, William D. 1964. The genetical evolution of social behavior, parts I and II. *Journal of Theoretical Biology* 7, 1–52.
Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it and how did it evolve? *Science* 298, 1569–1579.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

Hauser, Marc D., Elissa L. Newport & Richard N. Aslin. 2001. Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition* 78, B53–B64.
Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar: A reconstruction*. Oxford: Oxford University Press.
Heinz, Jeffrey. 2007. *Inductive learning of phonotactic patterns*. Ph.D. dissertation, UCLA.
Hewes, Gordon. 1973. Primate communication and the gestural origin of language. *Current Anthropology* 14, 5–24.
Hurford, James R. 2003. The language mosaic and its evolution. In Christiansen & Kirby (eds.), 38–57.
Hurford, James R. 2007. *The origins of meaning: Language in the light of evolution*. Oxford: Oxford University Press.
Hurford, James R. 2012. *The origins of grammar: Language in the light of evolution*. Oxford: Oxford University Press.
Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar and evolution*. Oxford: Oxford University Press.
Jackendoff, Ray. 2009. Parallels and non-parallels between language and music. *Music Perception* 26, 195–204.
Jackendoff, Ray. 2010. Your theory of language evolution depends on your theory of language. In Larson et al. (eds.), 63–72.
Jackendoff, Ray. 2011. What is the human language faculty? Two views. *Language* 87, 586–624.
Jarvis, Erich D. 2004. Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences* 1016, 749–777.
Johansson, Sverker. 2005. *Origins of language: Constraints on hypotheses*. Amsterdam: John Benjamins.
Katz, Jonah & David Pesetsky. 2011. The identity thesis for language and music. http://ling.auf.net/lingbuzz/000959.
Larson, Richard K., Viviane Déprez & Hiroko Yamakido (eds.). 2010. *The evolution of human language: Biolinguistic Perspective*. Cambridge: Cambridge University Press.
Lerdahl, Fred & Ray Jackendoff. 2006. The capacity for music: What's special about it? *Cognition* 100, 33–72.
Lieberman, Philip. 1984/2006. *Toward an evolutionary biology of language*. Cambridge, MA: Harvard University Press.
Lindblom, Björn. 1992. Phonological units as adaptive emergents of lexical development. In Charles A. Ferguson, Lisa Menn & Carol Stoel-Gammon (eds.), *Phonological development: Models, research, implications*, 131–163. Timonium, MD: York Press.
Lindblom, Björn. 1998. Systemic constraints and adaptive changes in the formation of sound structure. In James R. Hurford, Michael Studdert-Kennedy & Chris Knight (eds.), *Approaches to the evolution of language: Social and cognitive bases*, 242–264. Cambridge: Cambridge University Press.
Locke, John L. 2009. Evolutionary development linguistics: Naturalization of the faculty of language. *Language Sciences* 31, 33–59.
Locke, John L. & Barry Bogin. 2006. Language and life history: A new perspective on the development and evolution of human language. *Behavioral and Brain Sciences* 29, 259–325.
MacLarnon, Ann. 2012. The anatomical and physiological basis of human speech production: Adaptations and exaptations. In Tallerman & Gibson (eds.), 224–235.
MacLarnon, Ann & Gwen Hewitt. 1999. The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology* 109, 341–363.
MacLarnon, Ann & Gwen Hewitt. 2004. Increased breathing control: Another factor in the evolution of human language. *Evolutionary Anthropology* 13, 181–197.
Marler, Peter. 1998. Animal communication and human language. In Nina Jablonski & Leslie Aiello (eds.), *The origin and diversification of language*, 1–19. San Francisco, CA: California Academy of Sciences.
Maynard Smith, John & David Harper. 2003. *Animal signals*. Oxford: Oxford University Press.
Miller, Geoffrey. 2000. Evolution of music through sexual selection. In Nils L. Wallin, Björn Merker & Steven Brown (eds.), *The origins of music*, 329–360. Cambridge, MA & London: MIT Press.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

Miller, Geoffrey. 2001. *The mating mind: How sexual choice shaped the evolution of human nature*. London: Vintage.
Mithen, Steven. 2005. *The singing Neanderthals: The origins of music, language, mind and body*. London: Weidenfeld & Nicolson.
Ohnuma, Katsuhiko, Kenichi Aoki & Takeru Akazawa. 1997. Transmission of tool-making through verbal and non-verbal communication: Preliminary experiments in Levallois flake production. *Anthropological Science* 105, 159–168.
Patel, Aniruddh D. 2008. *Music, language, and the brain*. Oxford: Oxford University Press.
Pepperberg, Irene M. 2000. *The Alex studies: Cognitive and communicative abilities of Grey parrots*. Cambridge, MA: Harvard University Press.
Piattelli-Palmarini, Massimo & Juan Uriagereka. 2011. A geneticist's dream, a linguist's nightmare: The case of FOXP2. In Di Sciullo & Boeckx (eds.), 101–125.
Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13, 707–784.
Ploog, Detlev. 2002. Is the neural basis of vocalisation different in non-human primates and *Homo sapiens*? In Timothy J. Crow (ed.), *The speciation of modern Homo sapiens* (Proceedings of The British Academy 106), 121–135. Oxford: Oxford University Press.
Savage-Rumbaugh, Sue & Duane Rumbaugh. 1978. Symbolization, language, and chimpanzees: A theoretical reevaluation based on initial language acquisition processes in four young *Pan troglodytes*. *Brain and Language* 6, 265–300.
Savage-Rumbaugh, Sue, Stuart G. Shanker & Talbot J. Taylor. 1998. *Apes, language, and the human mind*. New York & Oxford: Oxford University Press.
Scott-Phillips, Thomas C. 2008. On the correct application of animal signalling theory to human communication. In Andrew D. M. Smith, Kenny Smith & Ramon Ferrer i Cancho (eds.), *The evolution of language: Proceedings of the 7th International Conference (EVOLANG7)*, 275–282. New Jersey: World Scientific.
Scott-Phillips, Thomas C., Simon Kirby & Graham R. S. Ritchie. 2009. Signalling signalhood and the emergence of communication. *Cognition* 113, 226–233.
Simon, Herbert A. 1990. A mechanism for social selection and successful altruism. *Science* 250, 1665–1668.
Slater, Peter. 2012. Bird song and language. In Tallerman & Gibson (eds.), 96–101.
Studdert-Kennedy, Michael. 2012. The emergence of phonetic form. In Tallerman & Gibson (eds.), 417–422.
Studdert-Kennedy, Michael & Louis Goldstein. 2003. Launching language: The gestural origin of discrete infinity. In Christiansen & Kirby (eds.), 235–254.
Tallerman, Maggie. 2007. Did our ancestors speak a holistic protolanguage? *Lingua* 117, 579–604.
Tallerman, Maggie. 2010. Holophrastic protolanguage: Planning, processing, storage, and retrieval. In Arbib & Bickerton (eds.), 83–97.
Tallerman, Maggie. 2013a. The evolutionary origins of syntax. In Andrew Carnie, Yosuke Sato & Dan Siddiqi (eds.), *The Routledge handbook of syntax*. London: Routledge.
Tallerman, Maggie. 2013b. Kin selection, pedagogy, and linguistic complexity: Whence protolanguage? In Rudie Botha & Martin Everaert (eds.), *The evolutionary emergence of human language*, 77–96. Oxford: Oxford University Press.
Tallerman, Maggie & Kathleen R. Gibson (eds.). 2012. *The Oxford handbook of language evolution*. Oxford: Oxford University Press.
Thornton, Alex & Nichola J. Raihani. 2008. The evolution of teaching. *Animal Behaviour* 75, 1823–1836.
Tomasello, Michael. 1999a. The human adaptation for culture. *Annual Review of Anthropology* 28, 509–529.
Tomasello, Michael. 1999b. *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
Tomasello, Michael. 2008. *Origins of human communication*. Cambridge, MA: MIT Press.
Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne & Henrike Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28, 675–735.

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press

Whiten, Andrew, Victoria Horner, Carla A. Litchfield & Sarah Marshall-Pescini. 2004. How do apes ape? *Learning and Behavior* 32, 36–52.
Wynn, Thomas. 2012. The Palaeolithic record. In Tallerman & Gibson (eds.), 282–295.
Zuidema, Willem & Bart de Boer. 2009. The evolution of combinatorial phonology. *Journal of Phonetics* 37, 125–144.

Author's address : Linguistics Section, SELLL, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
maggie.tallerman@ncl.ac.uk

https://doi.org/10.1017/S0022226713000017 Published online by Cambridge University Press