

1 Introduction: rewards and challenges of multi-perspectival work on the evolution of language and speech

RUDOLF BOTHA

The chapters in this volume on the prehistory of language are meant to serve two general purposes. First, they are representative of work of substance currently being done in an area where quality is not consistently dissociated from quantity. Which, of course, is not to say that the accounts offered in these chapters are therefore without limitations. The point, rather, is that in presenting work by scholars of repute, these chapters give a good idea of the depth of understanding that can be achieved at present.

Second, collectively, the chapters provide a striking illustration of how such understanding is achieved: by the adoption of a diversity of perspectives and approaches. Thus, in attempting to account for facets of the evolution of language and speech, these chapters represent a variety of perspectives: social, cultural, archeological, paleoanthropological, musicological, anatomical, neurobiological, primatological, and linguistic, to mention only some. The authors, moreover, adopt a variety of approaches for unraveling the evolution of language and speech, including ones involving comparison, correlation, simulation, and theoretical analysis. Since various of the perspectives and approaches at issue are associated with different disciplines, the chapters in this volume also clarify the sense in which present-day work on the evolution of language and speech can be said to be “multidisciplinary.”

Taken individually, then, what are these chapters about? In what follows, I identify the main perspective—often there are ancillary ones as well—from which each chapter elucidates the one or more facets on which it focusses. And I indicate how the chapter relates in terms of perspective to chapters that precede or follow it. In outlining each chapter’s central argument, I am forced by limitations of space to abstract away from specifics.

As a consequence, I am unable to show just how richly varied most of the chapters are in what they claim about the evolution of language or speech, in how they argue for their claims, and in how they engage with relevant literature. Readers will discover this for themselves, and to their delight, I believe.

In Chapter 2, Robin Dunbar, characterizing himself as an “evolutionary biologist,” takes up the general questions “Why did language evolve?,” “When did language evolve?” and “Why do only humans have language?” He argues that language evolved primarily not to facilitate the exchange of factual information but rather to serve social bonds by providing a substitute for social grooming, which he considers to be the main mechanism that our fellow primates use for bonding social relations. This hypothesis allows him to account for unrelated facts about brain size, group size, grooming patterns in primates, how we use language, what we talk about, conversation group sizes, and so on. From this hypothesis, moreover, follows naturally why, among the primates, only humans are likely to have evolved language: no other species evolved group sizes large enough to require more than grooming for social bonding. Addressing the question of when language evolved, Dunbar hypothesizes that this happened relatively late, probably with the appearance of anatomically modern humans about 200,000 years ago. And he argues that language evolved out of non-verbal forms of music-like vocalization, so-called chorusing, which served for the purpose of social bonding from about the appearance of archaic humans approximately 500,000 years ago. Among the perspectives offered by Dunbar on the evolution of language, the social one is clearly the central one.

In Chapter 3, Luc Steels provides another social perspective on the evolution of language, his main hypothesis being that sociality is a crucial prerequisite for the emergence of language. He argues in essence that human language, being symbol-based, can easily be used to cheat at the linguistic level by pretending that something is called or expressed in one way whereas you know that it is not, and at the factual level by lying. Steels assumes that, for language to be able to develop, individuals should (i) be willing to adopt linguistic conventions and categorizations introduced by others; (ii) align their conceptual and linguistic inventories as much as possible to those of others, and (iii) use language in an honest way, being maximally cooperative in doing all of this. If this sociality assumption is not adopted, language-like communication does not get off the ground

or is much less successful. This is argued by Steels on the basis of results of agent-based language games developed by him and his colleagues. The argument is consonant with a scenario—discussed by Chris Knight in Chapter 17 of the companion volume *The Cradle of Language* (henceforth *The Cradle*)—in terms of which a “social revolution” begun in Africa enabled early hominins to lift themselves out of a Darwinian world into a social and cultural one.

Two other chapters link up with Dunbar’s in positing an evolutionary link between language and music: the first by Steven Mithen, the second by Ian Cross and Eliot Woodruff. Mithen in Chapter 4 argues, to begin with, that we should return to ideas about the relationship between language and music advocated by scholars such as Rousseau, Darwin, and Jespersen. Next, Mithen further articulates the view that language and music co-evolved, a view which he ties in with recent arguments to the effect that protolanguage was holistic. According to Mithen, the proposal of a music-like protolanguage enables us not only to explain certain continuities between human speech and primate vocal communication but also to explain the seeming alacrity with which newborn infants respond to language and music alike, and the significant overlaps of the respective brain regions recruited for language and music. In addition, he cites reasons of different kinds for assuming that protolanguage used holistic phrases, not compositional ones. And he discusses a number of reasons why so-called hominin holistic phrase communication would have had a degree of musicality. In interweaving various strands of evidence in support of his views, Mithen gives an illustration of the extent to which work on language evolution has become in his view an interdisciplinary endeavor.

The idea that language and music may have co-evolved is entertained also in Chapter 5, with Cross and Woodruff proposing that language and music constitute complementary components of what they refer to as the “human communicative toolkit.” Drawing on ethnomusical, cognitive, and neuroscientific evidence, they suggest that music is a communicative medium with features that are optimally adapted for the management of situations of social uncertainty. They propose that music achieves this by presenting the characteristics of an honest signal, while underspecifying goals in a way that permits individuals to interact even while holding personal interpretations of goals and meanings that may actually be in conflict. In support of their proposal, Cross and Woodruff adduce a theory

of meaning in music: on this theory, the experience of music is accounted for in specific ways by reference to principles that are said to underlie both animal communication in general and human communicative interaction in particular. Exploring the implications of this theory for the evolution of language, Cross and Woodruff argue that, as complementary components of the “modern human communicative toolkit,” music and language are best thought of as having co-evolved from a precursor communicative system that embodied features of both.

The perspective on the evolution of language provided in Chapter 6 by John Odling-Smee and Kevin Laland is partly similar to those of earlier chapters in assigning a central role to social and cultural factors. Their main objective, though, is not to comment on what the evolution of language might have involved by way of specifics but rather to provide an alternative evolutionary framework within which that evolution can be explored. This framework is based on two reciprocal causal processes that feature in evolution: natural selection and niche construction, the latter being used by organisms to choose, regulate, construct, and destroy their environments. To Odling-Smee and Laland, cultural niche construction is particularly pertinent, since in their view the need for better ways of transmitting information is crucial to the cultural niche in which language may have evolved. More specifically, they suggest that language may have co-evolved with human cultural niche construction, with language serving as a means of facilitating and advancing the social transmission of life-skills to young hominids, particularly in our own species.

Sonia Ragir and Sue Savage-Rumbaugh, in Chapter 7, develop yet another social perspective on the evolution of language, the one afforded by social play. Social play and language use, they maintain, are similar in a number of ways. In both cases, participants negotiate hierarchically ordered moves and exchanges that can be modified and rearranged through repetitive actions and shared goals into normative, rule-governed behavior. Such similarities, they propose, make social play a “proper model” for understanding the emergence of language. Support for this comes, they claim, from data which on their analysis show that social play is indeed a “profoundly normative and, thus, self-organising communicative activity.” They rely for evidence on a fine-grained analysis of a period of social play among apes in an outdoor enclosure at the Language Research Center, Georgia State University—thereby adding to their chapter a primatological perspective on the evolution of language.

Chapters 8 and 9 present in some detail two further primatological perspectives on the evolution of language. Thus in the former chapter, David Leavens, Timothy Racine, and William Hopkins review evidence, accumulated over the last hundred years or so, for deixis in great apes. Some of this evidence suggests that great apes easily develop deictic repertoires in the complete absence of any explicit attempt to train them. Leavens, Racine, and Hopkins accordingly conclude that deixis—in the sense of the ability to direct the attention of another to a specific locus—is a capacity shared by great apes and humans. And assuming that deixis in great apes cannot ultimately derive from bipedalism or other adaptations, they conclude that our hominin ancestors were pre-adapted for joint attention, which makes deixis a component of the faculty of language in the broad sense of Hauser, Chomsky, and Fitch (2002).

In their review of the evidence for deixis in great apes, Leavens, Racine, and Hopkins refer to a particular deictic gesture which has been called the “directed scratch” by Simone Pika and John Mitani. In Chapter 9, Pika and Mitani give a description of how this gesture is used referentially by chimpanzees in the wild. It involves one chimpanzee male making a relatively loud and exaggerated scratching movement on a part of its body such that the movement can be seen and heard by his grooming partner. This gesture is shown by Pika and Mitani to be used communicatively to indicate a precise spot on the body and to request future action, namely grooming. The directed scratch appears to Pika and Mitani to be similar in form and in function to homesigns—the signs used communicatively by deaf children who have not been exposed to a manual sign language. Pika and Mitani hold that, like homesigns, directed scratches involve a form of reference and are therefore able to specify a distinct action—a property which, in turn, qualifies them as “characterizing signs.” Pika and Mitani accordingly conclude that directed scratches may constitute the first step towards symbolic gestures. In addition, they consider their findings to be consistent with the hypothesis that certain gestures of a sort still used today by our closest living relatives may have been crucial in providing the modality within which the precursors of symbolic communication evolved.

In Chapter 10—the first that looks at the evolution of language from a linguistic perspective—Maggie Tallerman investigates the origins of some basic features of the human lexicon. She proposes that a word-based lexicon evolved by a process of building on ancient conceptual categories that

were probably shared by all primates. This process, she argues, furthermore involved the use of the hierarchical structure that was already in place in primate cognition. In terms of Tallerman's proposal, a cognitive continuity is established between early humans—possibly *Homo ergaster*—and pre-human primates. That continuity is manifested, on her account, in other ways as well: she argues that the learning of categories is aided by labels both in humans and in non-human primates, and she argues that word-learning is aided by a set of innate learning biases. In support of her various hypotheses, Tallerman draws evidence from psycholinguistic studies, from work on category-specific brain defects, and from the study of pre-linguistic infants and non-human primates—making her chapter one of those that are richly textured from an evidentiary point of view.

In the next two chapters, aspects of the evolution of syntax are considered from the perspective of syntactic theory and diachronic change. Thus, subscribing to Noam Chomsky's Minimalist Program, Eric Reuland argues in Chapter 11 that it is "too simplistic" to view language as primarily a symbolic system used for communication. This view, he maintains, leads to an interpretation of the archeological record that is "too naïve." Central to Reuland's argument is the assumption that natural language is a computational system by which linguistic form and semantic interpretation are mapped systematically on to each other. The mapping is based on an inventory of lexical items and a combinatory system that includes the process known as "recursion" which, roughly, has the capacity to form infinitely long sentences by embedding phrases within phrases. The introduction of this process, Reuland argues, altered the nature of linguistic signs, severing the direct connection between form and interpretation. This gave rise to desymbolization, which he considers to be the "most characteristic" property of language. If his view is correct, evidence of symbolic activity by itself would not be a proper diagnostic of the presence of language, Reuland concludes.

Elly van Gelderen in Chapter 12 argues explicitly that Chomsky's biolinguistic approach has much to contribute to the study of language evolution, a view implicit to Reuland's chapter. To develop her argument, she pursues the question of what historical syntax can reveal about the "shape of original language"—with the question couched now in terms of this biolinguistic framework. Her position, in essence, is that the emergence of syntax followed the path followed by diachronic language change, a path also taken by children in acquiring language. She provides for two

steps along this path, the first being the organizing of the thematic layer of language through Merge, a syntactic principle by which two expressions are combined into a composite one. Grammaticalization, she argues, is the other step that is responsible for markings in the grammatical layer. As typical examples of grammaticalization, she cites instances where prepositions take on the function of case markers, verbs that of auxiliaries and affixes, and pronouns that of agreement morphemes. These diachronic changes, van Gelderen maintains, occurred in early language too and continue to occur in contemporary languages. And, she argues, they can be captured in terms of cognitive economy of syntactic derivation, as provided for in the Minimalist Program.

Noting that recursion is considered to be the hallmark of modern language, Frederick Coolidge and Thomas Wynn address in Chapter 13 two fundamental questions about its evolutionary emergence: “What is the relationship of recursion to modern language and thinking?” and “What might be the mechanism or subspecies of recursion that bestows its advantages to cognition?” In addressing these questions, they cite empirical evidence which in their opinion shows that recursion requires not only greater working memory capacity but also greater phonological storage capacity. And they propose that recursion arose as a function of an increase in phonological storage capacity and/or working memory capacity. In their view, these capacities were enhanced by a genetic neural mutation that occurred sometime between 150,000 and 30,000 years ago. That change made possible longer recursive and canonical utterances and a consequent increase in the complexity and information content of sentences. Considering the question of how enhanced working memory, by way of recursion, may have enabled modern thinking, Coolidge and Wynn speculate that it (i) may have given the speaker the ability to “hold in mind” a greater number of options, giving him more behavioral flexibility and even creativity; (ii) may have aided the rapid evolution of culture through “thought experiments”; and (iii) may have been required for fully symbolic thought, as reflected in therianthropoc art such as the *Löwenmensch* of Hohlenstein-Stadel and Hohle Fels cave.

In Chapter 14, Bart de Boer investigates the effect of the lowering of the larynx in humans, providing an articulatory/acoustic perspective on the evolution of speech. For his investigation, he uses Mermelstein’s model of the geometry of the human male vocal tract, a model in which the contours correspond to the actions of the muscles involved in speech. In

the experiment run by de Boer with this model, the area of the acoustic space that is accessible by a model of the male vocal tract—a space similar to the maximum vowel space—was compared with the accessible area of the female vocal tract. Since these vocal tracts differ not only with respect to the position of the larynx (lower for the male), he included in the comparison an artificial model on which the larynx has a female shape but is located in the male position. On de Boer's interpretation of the simulation results, the female vocal tract is better than the male tract for producing distinctive speech sounds. All in all, this indicates to de Boer that there is an evolutionary advantage to a vocal tract that has a pharyngeal and an oral cavity of equal length, as in the case of the female tract. He accordingly concludes that a different evolutionary explanation for the lower position of the male larynx needs to be found, the theory of size exaggeration as proposed by Tecumseh Fitch and his colleagues being a likely candidate.

In Chapter 15, the final one, Wendy Wilkins sets out a strategy for investigating the evolutionary biology of language. Central here is the following thesis: In order to understand the emergence of linguistic capacity as an innovation in the hominid line, it is necessary to work backwards from language-relevant anatomy. The assumption is that each piece of the anatomical mosaic will have a different evolutionary story, and that each story will be more or less evident in ancestral species, depending on the availability of biological evidence in the fossil record. Wilkins illustrates the use of this strategy by discussing the evolution of Broca's area and the parietal-occipital-temporal junction (POT) plus Wernicke's area, areas of the brain taken by Wilkins to be "necessary, if not sufficient, for language." In the view of Wilkins and her research associate Jenny Wakefield, the complex comprising Broca's area and the POT was evolutionarily shaped to improve the neurological control of the hand and thumb, and became available for exaptation after the divergence of the hominid and pongid lineages. This position, Wilkins argues, gains further support from recent work on primate neuroanatomy. She argues, too, that certain evidence from primate neuroanatomy indicates that there is a particular aspect of conceptual structure which is specific to humans and, moreover, specific to language—these properties making this a candidate for inclusion in the faculty of language narrowly construed (FLN).

As the chapters in this volume demonstrate, multi-perspectival work on the evolution of language and speech has generated a wealth of ideas

about what may have been involved in the relevant processes. Multi-perspectivalism, thus, clearly has its rewards. But it also brings with it new challenges for the future, most notably the challenges that are raised by the need to integrate ideas which, on the face of it, seem hard to reconcile in unifying and internally coherent theories of the evolution of language and speech. To illustrate what challenges of this sort are about, I sketch below three instances drawn from chapters in this volume. This is not meant to detract in any way from the merit of these chapters, all rated highly in peer reviews. The concern here is with something else altogether: namely, a fundamental question that arises in multi-perspectival work on the evolution of language and speech, the question of how to deal with the divergence of ideas.

Consider as the first instance two of the notions of language that feature in this volume:

- (1) Language is open-ended and fluid, particular languages lacking any clear or definite underlying system (Steels, Chapter 3).
- (2) Language is a computational system that embodies a systematic mapping between form and interpretation (Reuland, Chapter 11).

How, if at all, can these two notions of language be reconciled in a coherent, unifying theory of language evolution? On what basis can their relative merit be appraised? Is there sufficient agreement about the conditions that a conception of language should meet? (Botha, *The Cradle*: ch. 5). Or, to consider one more alternative, are Steels and Reuland actually concerned with the evolution of two different entities related terminologically only—by being called “language?” Incidentally, the notions of language indicated above are but two from among a large number that have been adopted in work on the evolution of language (Botha 2003: 13–15).

As a second instance, here are two of our contributors’ ideas about the primary function for which language emerged:

- (3) Language evolved primarily as a substitute for social grooming (Dunbar, Chapter 2).
- (4) Language evolved as a means of facilitating and advancing the social transmission of life-skills to young hominids, particularly in our own species (Odling-Smee and Laland, Chapter 6).

How can these two ideas—intriguing as they are—be made to cohere within an inclusive framework? For instance, could they be subsumed under a more general, unifying conception? And how do they relate to other ideas about what the primary function of language might have been—for instance, to aid thinking or to express thought? More generally, as for the merit of an idea about the primary function of emerging language—on what basis is it to be judged? Could the nature of that basis be other than empirical? This question arises if Fitch, Hauser, and Chomsky (2005: 185) are correct in believing that “from an empirical perspective, there are not and probably never will be data capable of discriminating among the many plausible speculations that have been offered for the original function(s) of language.” It also arises if “primary” and “original” are to mean the same thing. These are some of the questions that are bound to arise in attempts at knitting together the two ideas stated above within a coherent, unifying theory of the evolution of language.

The third instance is to do with similarities between language and some other phenomenon that are claimed to be evolutionarily significant. Again, the specimen claims below are drawn from the present volume:

- (5) Similarities between modern language and music indicate that language and music have a shared, music-like, precursor (e.g. Mithen, Chapter 4).
- (6) Similarities between language and social play indicate that social play is a “proper model” for understanding the emergence of language (Ragir and Savage-Rumbaugh, Chapter 7).

How would these two interesting ideas fit together, if at all, in a unifying theory of language evolution? What is to be the basis for judging that the similarities between language and a given phenomenon, say music, possess lesser or greater evolutionary significance than do the similarities between language and a given other phenomenon, say social play?

These three pairs of seemingly divergent ideas—and many more such pairs, triples, and so on are to hand in the literature—give some indication of the challenges that will have to be faced in future research aimed at incorporating into unifying theories the ideas thrown up by multi-perspectival work on the evolution of language and speech.

A first step in constructing such theories will be to subject truly incompatible ideas to comparative appraisal across disciplinary or sub-disciplinary boundaries. This will require some consensus among the participating disciplines or subdisciplines about the principles on which individual comparative appraisals are to be based. A second step will be to obtain clarity across disciplinary boundaries about what theoretical unification involves. Transcending these boundaries may well be the most daunting challenge to be faced in future work on the evolution of language and speech.