Evolving a bridge from praxis to language

Michael A. Arbib
University Professor, University of Southern California

We first address diverse criteria on what a theory of language evolution should explain, focusing on six divides: evolution did/did not yield a Universal Grammar; brain evolution is/is not important; language is to be viewed as speech or multimodal communication; language evolution is/is not best understood solely with reference to tools for communication; we do/do not need a notion of protolanguage as a precursor to language; and protolanguage was/was not in great part holophrastic. We argue against a role for an innate Universal Grammar in language acquisition and language change, and then present a brief case study of the emergence of Nicaraguan Sign Language in a few decades. Finally, we present the mirror system hypothesis on the evolution of the language-ready brain locating it within the 6 divides and charting a path for biological evolution supporting mechanisms for simple and complex imitation, pantomime, protosign and protospeech in turn, claiming that this provided an adequate base for true languages to emerge through cultural evolution.

1. Introduction

Many species have rich communication systems, but only humans have the capability to master a language in the sense of the combination of an open-ended lexicon with a grammar that allows lexical elements to be combined hierarchically to create "utterances" whose meaning, encoded by the sender by means of *compositional semantics*, can be inferred (more or less accurately) by the receiver. The challenge of studying language evolution is twofold:

1. How did *biological evolution* (natural selection) yield a species that is "language-ready" in the sense that the young can acquire language if raised in a community in which language is used for communication? and
2. How did *cultural evolution* (processes of historical and social change) yield the diversity of languages that we know today?

Studies in Language Companion Series : New Perspectives on the Origins of Language. Amsterdam, NL: John Benjamins Publishing Company, 2013. ProQuest ebrary. Web. 4 November 2016.

The core of the present paper is a brief introduction (Section 4) to the mirror system hypothesis on the evolution of the language-ready brain. This hypothesis roots language parity (the hearer's ability to recognize [much of] the meaning of the speaker's utterance) in a property shared with monkeys, and thus presumably our last common ancestor – namely the possession of mirror neurons that are engaged in both the production and recognition of specific hand actions. Since monkeys (and, again, presumably our last common ancestor) do (did) not have language, the hypothesis seeks to delineate key extensions of the capabilities of our ancestors' brains that led to the human language-ready brain, with imitation grounding the path from pantomime to protosign and protolanguage.

As background to this, we first turn in Section 2 to an assessment of what a theory of language evolution should explain. It presents six dichotomies, which define 64 possible alternatives for the set of choices that must be made and justified in presenting an extended theory. This section invites the reader not only to assess the specific choices made in developing the mirror system hypothesis but also to make explicit the choices underlying other chapters in this volume, thus encouraging the development of the reader's own coherent perspective.

A number of researchers have taken the view that biological evolution has endowed the brain with an innate Universal Grammar, and that this biological endowment greatly simplifies the problems of language development for the human child. However, Section 3 argues against this view, in a discussion of language acquisition and language change. In particular, we complement the emphasis of most of this volume on the evolution of language-as-speech by looking at the emergence of novel sign languages, stressing the interaction of social patterns with the language-readiness of the individual brain. This look at cultural evolution of language provides a useful backdrop for the presentation of the mirror system hypothesis on the biological evolution of the language-ready brain in Section 4. Section 5 then places the hypothesis in perspective by responding to an alternative view that sees language as a by-product of a more general development of cognitive capabilities.

It is worth emphasizing that the mirror system hypothesis is a living hypothesis in that it is itself evolving as it expands to take account of more data from diverse disciplines.

2. What should a theory of language evolution explain?

Between them, studies of language evolution must address:

- Data on language use in communication and thought, including use in conversation;
- Data on language acquisition;

- Data from historical linguistics;
- Archeological data which seek to infer clues to the language of a culture from the material traces of that culture;
- Data on the communication systems of animals in general or primates in particular, seeking commonalities with and differences from language;
- Data on animal behavior in general or primate behavior in particular, seeking commonalities with and differences from communication systems;
- Data on brain function.

But no single paper or chapter can address more than a fraction of the available data. Instead, each theory will constrain the range of data that are examined, and will in turn be limited by restrictions of this search space. Before presenting my own theory and sketching the range of data that have shaped its development, I will start by charting a diversity of divides in approaching theories of language evolution.

Divide 1: Universal Grammar

Any satisfactory theory of language evolution must explain the human child's ability to acquire any human language if raised in interaction with a community that uses it. This yields our first divide:

(1a) Biological evolution endowed us with a Universal Grammar in the sense of a complex set of rules and constraints that must be inborn because they cannot be learned from the data available to the child.

versus

(1b) Biological evolution gave us a language-ready brain that made the eventual *invention* of language possible.

The general issue is to assess which aspects of language are innate, and which are the fruits of historical change. In either case, one needs to understand

- what the child learns during language acquisition, and what biological evolution provided to make such learning possible;
- what happens during historical language change, and what biological evolution provided to make such processes possible.

Divide 2: Brain

(2a) The study of language evolution must compare human brain imaging and lesion data with data on brain mechanisms for "language-related" functions in other species to ground an understanding of what has been conserved and what has been changed in human brain evolution.

versus

(2b) The primary data are those of linguistics alone.

Divide 3: Speech versus multimodal communication

(3a) Language is equated with speech.

versus

(3b) The shared mechanisms that support signed as well as spoken languages are stressed.

In either case, the importance of spoken language requires us to understand the evolution of the vocal apparatus and the neural control that made it able to produce speech, and the related mechanisms for perception.

Divide 4: The broader context

(4a) Language evolution is to be understood solely in terms of adaptive pressures for communication or thought.

versus

(4b) Language evolution rests in part on the exaptation of adaptations that are not directly related to communication.

Divide 5: Protolanguage

(5a) Our ancestors evolved a capability for *protolanguage* – which had an open lexicon but little if any syntax – before they developed language.

versus

(5b) No intermediary was involved.

Here, "protolanguage" is being used in the sense of "something intermediate between the communication systems of our last common ancestor with the apes and language, but which is not itself a language" (Bickerton 1995; Hewes 1973). This is to be distinguished from its use in historical linguistics (Dixon 1997; Lass 1997) as "an actual language posited to be 'ancestral' to a family of related languages," in the way that Latin is viewed as a protolanguage for the Romance languages.

Divide 6: Holophrasis

(6a) Initially, much of protolanguage was *holophrastic* (with protowords describing frequently occurring or significant situations).

versus

(6b) Protolanguage started with words akin in scope to modern words (such as nouns and verbs).

64 classes of frameworks

These six divides (and, of course, there are others) define 64 overall approaches to language evolution. Any *general* framework should justify (at least) which side of each of the six divides it lies on. By contrast, more focused models may ignore many of these issues to address the possible evolution of specific language-related mechanisms, such as how the human speech apparatus can produce the observed sounds of human languages, though even here assumptions from some general framework may play a crucial role.

In Section 4, I will introduce the Mirror System Hypothesis, whose general outline was laid down by Giacomo Rizzolatti and myself (Arbib & Rizzolatti 1997; Rizzolatti & Arbib 1998) but which I have refined and extended over the subsequent years (for an extensive synthesis, see Arbib 2012) in considering new data while responding to a range of critiques (such as those contained in the commentaries on the target article in Arbib 2005a, and the dozen commentaries on Arbib (2012) set forth in *Language and cognition* 5(2–3), 2013.). The choices made in this work are as follows:

Divide 1: Biological evolution gave us mechanisms which made the eventual *invention* of language possible.
Divide 2: We make crucial use of data on the brain.
Divide 3: We seek to understand the shared mechanisms that support signed languages as well as spoken languages.
Divide 4: Language evolution rests in part on the exaptation of adaptations that are not directly related to communication.
Divide 5: Our ancestors evolved a capability for protolanguage before they developed language.
Divide 6: Initially, much of protolanguage was holophrastic.

Let me briefly indicate what arguments justify the choices made:

Divide 1: Section 3, "Against Universal Grammar," outlines how an understanding of both language acquisition and historical change militate against the idea of a Universal Grammar, at least if that is taken to provide an innate encapsulation of the overall principles and parameters that span the syntax of all human languages.

Divide 2: The claim that "the study of language evolution must make crucial use of data on the brain" does not, of course, mean that *every* study of language evolution must make crucial use of data on the brain. As most chapters in this volume show, many insights are to be gained without invoking neural data. Nonetheless, if we are to address the biological evolution that made language possible, then an understanding of how the human brain and body have evolved is indispensable. This study must include the various learning mechanisms that give humans a language-ready brain, that is, a brain that endows human children with the ability to learn language, which is denied to all other species. A number of books offer a rich sense of the brain-based approach to the evolution of language (e.g. Corballis 2002; Deacon 1997; Lieberman 2000), and the Mirror System Hypothesis furthers this tradition.

Divide 3: Since there are intriguing data on the diverse calls of monkeys (Seyfarth et al. 1980b), and since the social relations of baboons offer complexities that in some way foreshadow certain structures evident in language (Cheney & Seyfarth 2005), it certainly seems plausible that spoken language evolved directly from the vocalization system of our common ancestor with monkeys. However, the Mirror System Hypothesis offers evidence for a gestural origins theory (Armstrong et al. 1995; Armstrong & Wilcox 2007; Corballis 2002; Hewes 1973) in the somewhat modified form presented later. Peter MacNeilage has vigorously defended a non-gestural theory (MacNeilage 1998; MacNeilage & Davis 2005; see also his contribution to this volume) and I have spelled out the counterargument in some detail (Arbib 2005b). Further support comes from the fact that the repertoire of calls of nonhuman primates is genetically specified, whereas different groups of apes of a given species can each develop and use some group-specific gestures (Arbib et al. 2008).

Divide 4: Whether one accepts a "speech only" or a "gesture first" view of language origin, it seems unlikely that the transition from monkey-like vocalizations or ape-like gestures to languages in the sense defined above was accomplished in one generation. Here, following Hewes (1973) and Bickerton (1995), we will use the term *protolanguage* to indicate any of the intermediate forms that preceded true languages in our lineage. (As noted earlier, this is not to be confused with the usage in historical linguistics, where a *protolanguage* for a family of languages is a full language posited to be ancestral to the whole family.)

Divide 5: Here, the debate is between two views. The *holophrastic* view (Arbib 2005a; Wray 1998) holds that in much of protolanguage, a complete communicative act involved a "unitary utterance" or "holophrase" whose parts had no independent meaning. For example, the innate leopard alarm call of vervet monkeys (Seyfarth et al. 1980a, 1980b) might be interpreted as "Danger! A leopard is nearby. Run up a tree to escape – and pass on the message." The key point is that the call cannot be broken into parts that correspond to words of an English translation, and – because of its imperative – it could not be used as a word for leopard to convey meanings like "There's a dead leopard. Let's scavenge it." Turning from innate calls to socially constructed holophrases, the holophrastic view holds that, as "protowords" were fractionated or elaborated to yield words for constituents of their original meaning, so were constructions developed to arrange the words to reconstitute those original meanings and many more besides. Opposing this, the *compositional* view (Bickerton 1995) hypothesizes that *Homo erectus* communicated by a protolanguage in which a communicative act comprised a few words in the current sense strung together without syntactic structure. In this view, the "protowords" (in the evolutionary sense) were so akin to the words of modern languages that languages evolved from protolanguages just by "adding syntax."

Before explaining the Mirror System Hypothesis, let me summarize why I reject the notion of an innate Universal Grammar.

3. Against Universal Grammar

Each human child enters a world of physical and social interactions in which adults use words that at first mean nothing to the child and combine them in many different ways. Over a few years, the child becomes a member of the language community, able to use ever more words in ever more subtle combinations to request help from others and to share interests, addressing an increasing range of physical and social needs, even as the child's understanding expands from the here and now to encompass the remembered past, the possible future, and the worlds of stories. Some linguists distinguish between E-language and I-language – between the wild variety of utterances "out there" and the knowledge of language inside the head of the individual speaker. The child joins the language community by constructing an I-language that approximates aspects of the ambient E-language. Yet the E-language is sufficiently heterogeneous and the child's sample sufficiently small that different children may end up with slightly different I-languages. As old speakers die and new speakers arise, the set of I-languages of the community changes. Their combined effect may preserve the E-language of the previous generation, change it in minor ways, or start changes that may cumulatively yield a new E-dialect or E-language. This leaves open (at least) two major questions:

Language acquisition: What is the nature of an I-language and how is it developed by the individual? A sub-question is how much of the I-language is acquired by the young child, and how much may an individual's I-language and the E-language of a community change over an individual's lifespan?

Language change: A language may change because of innovations invented and disseminated by young and older adults, or it may change because young language learners extract new patterns from the ambient fluctuations of E-language to create a new population of I-languages. Is this an either/or situation, or are both processes operative?

3.1 Language acquisition

Generative grammar (associated particularly with the work of Noam Chomsky 1965, 1981, 1992) sees the grammar of a language as comprising a relatively compact set of autonomous syntactic rules and principles that put words together in very general ways and without regard for the meaning of the result. The use of these rules to express meaning is then the problem of semantics, which is rather separate from syntax. Many proponents of generative semantics further hold that the basic rules of the syntax of *all* human languages are contained in a *Universal Grammar* which is genetically encoded in humans such that – at least in one version of the theory (Baker 2001; Chomsky & Lasnik 1993), known as *principles and parameters* – it can establish within the infant brain a range of parameters that enable children to acquire the syntax of their native languages by setting each parameter simply by hearing a few sentences to determine which value of the parameter is consistent with them. Lightfoot (2006)[1] states that "children are internally endowed with certain information, what linguists call Universal Grammar (UG), and, when exposed to primary linguistic data, they develop a grammar … or I-language. The essential properties of the eventual system are prescribed internally and are present from birth." By contrast, Evans and Levinson (2009) summarize decades of cross-linguistic work by typologists and descriptive linguists to demonstrate the radical variations in sound, meaning, and syntactic organization across the diversity of the world's languages. They argue that, although there are significant recurrent patterns in organization, these are better explained as stable engineering solutions satisfying multiple design constraints, reflecting both cultural-historical factors and the constraints of human cognition, rather than as the expression of an innate UG.

Construction grammar replaces a limited set of rules of autonomous syntax by a more or less language-specific set of constructions (Croft 2001; Fillmore et al. 1988; Goldberg 2003) which combine form (how to aggregate words) with meaning (how the meaning of the words constrains the meaning of the whole). The latter seems more hospitable to accounts (historical linguistics/cultural evolution) of how languages emerge and change over time (Croft 2000). In an example of language acquisition that can now be seen as related to construction grammar (though before this theory was introduced – see Tomasello 2003 for a more extended account), Jane Hill (1983) showed that the child may first acquire what the adult perceives as two-word utterances as *holophrases* – for example, saying "want-milk" without understanding that this has pieces with different meanings – prior to developing a more general construction (e.g. "want *x*") in which "*x*" can be replaced by the name of any "wantable thing." Further experience will yield more subtle constructions and the development of word classes such as "noun" defined by their syntactic roles in a range of constructions rather than their semantic role. Around 1985, I sent Noam Chomsky (with whom I had had many informative conversations while I was a graduate student at MIT) a draft of a paper (possibly Arbib & Hill 1984) based on Hill's work. She modeled data she had gathered from a 2-year-old child by a process in which the child interiorizes

1. This section is based in part on my review of Lightfoot's book (Arbib 2007).

fragments of what he or she hears to build constructions based on the child's current concerns and interests. Over time, some constructions merge with others, some fall into disuse, some gain complexity, and the lexicon grows and is categorized in tandem with the development of these constructions. Chomsky's reply was memorable: he said he would not read the paper, but then kindly typed a page or so explaining "the truth." This anecdote is completely neutral as to which approach to language acquisition is correct, but it does suggest a certain attitude to the relation between data and theory which is also apparent in Lightfoot's book.

Lightfoot and Chomsky share what I will refer to as the UG/LAD hypothesis – that there is an innate UG and that the child develops an I-language (in their view, the child's syntax) via a Language Acquisition Device (LAD), which scans the ambient E-language to determine which settings of the parameters of UG match the sentences the child hears. The poverty of the stimulus argument asserts that children receive too little explicit instruction in syntax to acquire their I-language unless the alternative rules are already present innately. However, if we leave the domain of autonomous syntax and assess how cognition and language can cooperate in mental development, as demonstrated by Hill and Tomasello, then the stimulus turns out not to be poor at all, and feedback is supplied by communicative efficacy. Further arguments within the linguistics literature also militate against the poverty of the stimulus argument (e.g. Pullum & Scholz 2002). Lightfoot argues convincingly that one approach to the UG/LAD hypothesis – setting parameters by scanning through the entire E-language (Clark 1992; Gibson & Wexler 1994) – must fail. His argument is based on the logic of a combinatorial explosion, not an analysis of developmental data. He then argues – again, I think, convincingly – that the child does not attend to the full complexity of adult sentences but instead attends to fragments, as indeed was a property of Hill's model. The child will not process complex sentences (at least at first) but will focus on simple sentences in the search for structure. Where Lightfoot and I differ is that he sees the process of scanning fragments as a search for "cues" in the form of items from UG that can structure the sentence fragments. The catch is that his cues are highly theory-laden. For example, the child is required to recognize an empty verb slot or a V-DP structure. No account is given of how young children recognize an empty slot or know a DP when they see one. However, once one accepts that children are not looking at the entire E-language to switch in a complete grammar from among those licensed by UG, but instead are slowly building up a repertoire of "generation strategies" for fragments of a simplified subset of the E-language, then the way is open to test the hypothesis that an innate UG is unnecessary after all. Instead one may posit that children build up a set of constructions that gradually extend their ability to use the ambient language both to get others to minister to their needs and to engage in a broadening range of social interactions. The key to work on language acquisition is to analyze data on how children acquire language. Here a useful compendium is Eve Clark's (2003)

First Language Acquisition. Clark integrates social and cognitive approaches to how children understand and produce sounds, words, and sentences within the context of learning to use language to cooperate and achieve goals.

3.2 Language change

I earlier raised the question of whether languages change because of innovations invented and disseminated by adults, or because of the way young language learners extract new patterns from E-language, or both. Lightfoot (2006) develops a child-centered theory of historical language change. He insists that languages change by switching UG parameters, rather than by the accretion of new constructions, and then asserts, without providing any supporting data, that children's formation of new I-languages is the key to how languages change. However, Croft (2000) offers a very different view of language change. He even devotes his Section 3.2 to a critique of "the child-based theory of language change," for example by citing a variety of studies which show that changes attested in historical language change are different from those found in language acquisition. Lightfoot does state that changes in E-language (e.g. due to the Viking invasions of Britain) must have triggered historical changes but offers no evidence that these changes were not made first by adults accommodating to the novel language flux, with children's language acquisition playing a secondary role. To take a contemporary example: the rapid changes in cell phone technology are driven in large part by the enthusiasm of teenagers for new devices and services, but these new devices and services are usually created by 30-year-olds (more or less), while the consumers provide selective pressure. In the same way, whereas language acquisition is the main cause of language change in Lightfoot's view, it seems strange to me to privilege the language learning of infants over the effects of adult innovation in response to a whole host of historical processes including language contact.

Heine (this volume) provides insight into the processes of grammaticalization whereby new grammatical structures may emerge without benefit of selection from the innate repertoire of a UG. Lefebvre (this volume) usefully stresses the interaction of substrate and superstrate languages in the formation of a creole, as against a Bioprogram Hypothesis (Bickerton 1984) or expression of "unmarked" parameters in UG. Here, I complement these insights by a quick look at the emergence, over the course of a few decades, of Nicaraguan Sign Language (NSL) – a useful "control case" because the people who created the language over the course of some 30 years had no extant sign languages to creolize. Although a number of the people involved in the cumulative emergence of NSL shared a knowledge of Spanish, neither the lexicon nor the constructions of this spoken language could provide either a substrate or superstrate language of the kind involved in the formation of creoles. Moreover, no existing sign language (such as American Sign Language, ASL) provided a basis for NSL. Rather, fragments of gestural communication were amalgamated; but, I argue, this process was influenced by the *idea* of language even in the absence of adaptation of specific lexical items or constructions from existing languages.

Before the 1980s, deaf Nicaraguans had little contact with each other. But when a vocational school for the deaf opened in 1981 in Managua, the children began to develop a new, gestural system for communicating with each other, in part by consolidating the different home signs (Goldin-Meadow 2003; Goldin-Meadow & Feldman 1975) each had developed for rudimentary communication with their families. The gestures soon expanded to form a rudimentary sign language. By 1986, the early collection of gestures developed into an expressive sign language, NSL, which is still in a state of flux, as the sign lexicon and constructions keep changing.

One study of such change (Senghas et al. 2004) compares the performance of the first, second, and third cohorts (deaf Nicaraguans who acquired NSL in the first, second, or third decade of its existence). Subjects were asked to use their hands to describe a segment of a Tweety Bird and Sylvester cartoon in which Sylvester, having swallowed a bowling ball, "rolls" down the hill. Manner (rolling) and path (downward) are expressed simultaneously in the co-speech gestures of Spanish speakers as well as in the signs of many early-cohort NSL signers. By contrast, third-cohort NSL signers tend to express manner and path sequentially, first signing 'rolling' and then 'downward'. Thus, NSL is not a copying of Spanish co-speech gestures. This separation of manner and path is a novel conventionalization. Moreover, many sign languages do express manner and path simultaneously, so NSL is developing its own signs, not simply copying other sign languages.

If manner and path are expressed separately, it may no longer be clear that the two aspects of movement occurred within a single event. *Roll* followed by *downward* might mean 'rolling, then descending'. Senghas et al. (2004) show that NSL developed a way to put the pieces back together: it now has the X-Y-X construction, such as *roll-descend-roll*, to express simultaneity. This string can serve as a structural unit within a larger expression like *cat [roll descend roll]*, or it can even be nested, as in *waddle [roll descend roll] waddle*. This construction never appeared in the gestures of the Spanish speakers and is unlike any construction of spoken Spanish.

Elsewhere (Arbib 2009), I have developed the hypothesis that NSL differs from home signs because the existence of a community provides more opportunities to use and choose signs, so that although some get lost to the community, many gain power by being widely shared. An "engine" for this process is provided by the fact that, since knowledge of another language (Spanish, in this case) *is* possessed by some members of the community, they seek to translate this knowledge into the new medium (as is proven for the lexicon which includes a sign for *Thursday*), and some of these attempts to capture a given property will become widespread – even though NSL need *not* reproduce the lexicon or constructions of these other languages. The foundational document for this analysis is Laura Polich's book analyzing the changing social matrix that supported the emergence of NSL, *The Emergence of the Nicaraguan Deaf Community in Nicaragua: "With Sign Language You Can Learn So Much"* (Polich 2005). She stresses that the turning point came when the vocational school was established that kept adolescents and young adults together "at a time when they were carving out their identities and craving a peer group in which to try out and enact their abilities to be social actors" (Polich 2005: 146). In its early stages, the community being formed included hearing people who spoke Spanish, while even those who could not speak had at least a small vocabulary of written Spanish. The impressive achievement of creating this new language, NSL, did not have to rest solely on innate capabilities of the human brain (which distinguish us from other primates, for example) but could indeed exploit the cultural innovations of existing language communities.

It has been argued that the brain of *Homo sapiens* was biologically ready for language perhaps 200,000 years ago but, if the increased complexity of artifacts like art and burial customs correlate with language of some subtlety, then human languages as we know them may have arisen at most 50,000 to 90,000 years ago (Noble & Davidson 1996). If we accept the idea that it took humans with modern-like brains 100,000 years or more to invent language as we know it, we must ask what advantage the community which developed NSL in a few decades had that early humans lacked. As mentioned above, I suggest that NSL differs from home signing because (i) the existence of a community provides more opportunities to use and choose signs, thus losing others; and (ii) since some community members possess knowledge of another language, they seek to translate their knowledge into the new medium. Early humans would have shared (i) but not (ii) with the deaf Nicaraguans. The very idea of language can have a catalytic effect. It seems almost inconceivable that the *idea* of language had to be invented, but we know that writing was only invented some 5,000 years ago. We believe that no genetically based changes in the *Homo sapiens* brain were required to support literacy; nevertheless, literacy can change the brain as it develops. For example, functional hemispheric specialization depends on both genetic and environmental factors (Petersson et al. 2000). Illiterate subjects are consistently more right-lateralized than literate controls even though the two groups show a similar degree of left–right differences in early speech-related regions of superior temporal cortex. Further, the influence of literacy on brain structures related to reading and verbal working memory affects large-scale brain connectivity more than gray matter. Moreover, many societies have lasted till modern times with no written form for their spoken language. Yet, once one has the idea of writing, it is a matter of years rather than millennia to invent a writing system. Around 1820, Sequoyah, a Cherokee who knew very little English and was illiterate, invented a Cherokee syllabary, with 86 characters to represent the sounds of the Cherokee language, inspired solely by the *idea* of writing (Walker & Sarbaugh 1993). Thus, the notion that languages emerged through cultural rather than biological evolution is indeed worthy of consideration.

In summary, I have argued that the natural selection/biological evolution that yielded the modern human genotype that supports development of a brain that can invent, learn, and use languages did *not* yield an innately specified Universal Grammar, but rather yielded a brain that could support mechanisms that allowed cultural evolution to support, over many tens of millennia, the transition from ape-like systems of limited, but mildly open, communication to protolanguages to full languages. We see here that a cultural factor, literacy, influences the functional hemispheric balance. The broader claim, then, is that basic language mechanisms of the human brain, not just the modern mechanisms supporting literacy, may reflect the effects of development within a certain kind of human community without requiring a genetic specification of anything like a UG.

With this, we turn to a brief outline of the mirror system hypothesis on the evolution of the genetic basis for the mechanisms which make the human brain language-ready.

4. The mirror system hypothesis on the evolution of the language-ready brain

The guiding hypotheses for the evolutionary approach to language offered here are:

1. Language did not evolve as a separate "faculty." Rather, brain mechanisms for the perceptual control of action provided the "evolutionary basis" for the language-ready brain.
2. Manual dexterity provides a key to understanding human speech: Birdsong provides an example of superb vocal control but has never become the basis for a language. By contrast, primates are precocious in their manual dexterity, and so we seek to establish that humans exploited this dexterity to become the only primates possessing language.[2] Indeed, speaking humans, even blind ones, make extensive use of manual co-speech gestures (McNeill 2005), and children raised appropriately can learn the sign languages of the deaf as readily as hearing children learn a spoken language.

To situate grammar in relation to praxis (practical manual actions) as systematizing strategies for achieving a communicative goal, consider the example of a restaurant

2. This view of the importance of dexterity and gesture to an understanding of language evolution is shared by Michael Corballis and Michael Tomasello (this volume), thus leading to one of the few general conclusions of the Summer School: "Michaels support the importance of gestures in the origins of language."

manager whose communicative goal is to get a waiter to serve a particular customer (Arbib 2006). He keeps replanning his sentence until he thinks that any ambiguity is resolved.

> *Serve the man on the left.*
> Still ambiguous?
> *Serve the young man on the left.*
> Still ambiguous?
> *Serve the young man on the left in the green sweater.*

Still ambiguous? Apparently not, and so he says it to the waiter. We see here a sentence planning strategy of repeating the construction *add adjective or relative clause until (you think) ambiguity is resolved,* adapting and "unfurling" a nested hierarchical structure to extract a set of actions to reach a communicative goal.

With this, let me turn to a fuller exposition of the Mirror System Hypothesis. A *mirror neuron,* as observed in macaque brains, is a neuron that fires vigorously both when the animal executes an action and when it observes an other execute a more or less similar action. Our theory of language evolution is grounded in evidence from brain imaging (e.g. Grafton et al. 1996) that there is a human *mirror system for grasping* – that is, a brain region activated for both grasping and observation of grasping – in or near Broca's area, an area traditionally viewed as involved in speech production. But why might a mirror system for grasping be associated with such an area? The fact that aphasia of signed, and not just spoken, languages may result from lesions based on Broca's area (Emmorey 2002; Poizner et al. 1987) supports the view that one should associate Broca's area with *multimodal* language production rather than with speech alone.

Different brain regions (not individual neurons) may be implicated in the human brain as mirror systems for different classes of actions, and many researchers have attributed high-level cognitive functions to human mirror regions, such as imitation (Buccino et al. 2004), intention attribution (Iacoboni et al. 2005), and language (Rizzolatti & Arbib 1998). However, monkeys do not imitate or learn language. Thus, any account of the role of human mirror systems in imitation and language must include an account of the evolution of mirror systems and their interaction with more extended systems within the human brain – *beyond the mirror system.*

The original *Mirror System Hypothesis* (Arbib & Rizzolatti 1997; Rizzolatti & Arbib 1998) argued that the basis for *language parity* evolved "atop" the mirror system for grasping, rooting speech in communication based on manual gesture. In other words, we trace a path from praxis to communication. Developing this basic premise, I have argued (Arbib 2002, 2005a) that a language-ready brain resulted from the evolution of a progression of mirror systems and linked brain regions "beyond the mirror" that made possible the full expression of their functionality.

4.1 Imitation

Monkeys have, at best, a very limited capacity for imitation (Visalberghi & Fragaszy 2001; Voelkl & Huber 2007) far overshadowed by what I call *simple imitation* as exhibited by apes. Myowa-Yamakoshi and Matsuzawa (1999) observed that chimpanzees took 12 or so trials to learn to "imitate" a behavior in a laboratory setting, focusing on bringing an object into a relationship with another object or the body, rather than the actual movements involved. Byrne and Byrne (1993) found that gorillas learn complex feeding strategies but may take months to do so. Consider eating nettle leaves. Skilled gorillas grasp the stem firmly, strip off the leaves, remove the petioles bimanually, fold the leaves over the thumb, pop the bundle into the mouth, and eat. The challenge of acquiring such skills is compounded because ape mothers seldom if ever correct and instruct their young (Tomasello 1999) and because the sequence of "atomic actions" varies greatly from trial to trial. Byrne (2003) posits that *imitation by behavior parsing*, a protracted form of statistical learning whereby certain *subgoals* (e.g. nettles folded over the thumb) become evident from repeated observation, is common to most performances. In his account, the young ape may acquire the skill over many months by coming to recognize the relevant subgoals and derive action strategies for achieving them by trial and error.

This ability to learn the overall structure of a specific feeding behavior over many, many observations is very different from the human ability to understand any sentence of an open-ended set as it is heard, and generate another novel sentence as an appropriate reply. In many cases of praxis (i.e. skilled interaction with objects), humans need just a few trials to make sense of a relatively complex behavior if the constituent actions are familiar and the subgoals these actions must achieve are readily discernible, and they can use this perception to repeat the behavior under changing circumstances. We call this ability *complex imitation* (extending the definition of Arbib 2002 to incorporate the goal-directed imitation of Wohlschläger et al. 2003).

A mirror system does not support imitation in itself. A monkey with an action in its repertoire may have mirror neurons active both when executing and observing that action, yet does not repeat the observed action. Nor, crucially, does it use observation of a novel action to add that action to its repertoire. Thus, the first step in the Mirror System Hypothesis is to hypothesize that evolution embeds a monkey-like mirror system in more powerful systems in two stages:

- **a simple imitation system for grasping**, shared with the common ancestor of human and apes; and
- **a complex imitation system for grasping**, which developed in the hominid line since that ancestor. This involves an increased ability for both perceptual and motor chunking.

Even though they involve praxis rather than explicit communication, both of these changes could create an evolutionary advantage in supporting the transfer of novel skills between the members of a community.

4.2 From pantomime to protosign and protolanguage

We now explore the stages whereby our distant ancestors made the transition to *protolanguage*, in the sense of a communication system that supports the ready addition of new utterances by a group through some combination of innovation and social learning. It is open to the addition of new "protowords," in contrast to the closed set of calls of a group of nonhuman primates, yet lacks any tools, beyond the mere juxtaposition of two or three protowords, to put protowords together to continually create novel utterances from occasion to occasion. Continuing, the Mirror System Hypothesis suggests that brain mechanisms for complex imitation evolved to support not only pantomime of manual actions but also

- **pantomime of actions outside the pantomimic's own behavioral repertoire** (e.g. flapping the arms to mime a flying bird). This provides the basis for the evolution of systems supporting protosign;
- **protosign:** conventional gestures used to formalize and disambiguate pantomime (e.g. to distinguish 'bird' from 'flying').

The transition from complex imitation and the small repertoires of ape gestures (perhaps 10 or so novel gestures shared by a group) to protosign involves, in more detail, first the pantomime of grasping and manual praxic actions then of non-manual actions, complemented by conventional gestures that simplify, disambiguate, or extend pantomime. Pantomime transcends the slow accretion of manual gestures by ontogenetic ritualization, providing an "open semantics" for a large set of novel meanings (Stokoe 2001). However, such pantomime is inefficient – both in the time taken to produce it, and in the likelihood of misunderstanding. Conventionalized signs extend and more efficiently exploit the semantic richness opened up by pantomime. Processes like ontogenetic ritualization can convert elaborate pantomimes into a conventionalized "shorthand," just as they do for praxic actions. This capability for protosign – rather than the elaborations intrinsic to the core vocalization systems – may then have provided the essential scaffolding for protospeech and evolution of the human language-ready brain:

- **protosign and Protospeech (an expanding spiral):** conventionalized manual, facial, and vocal communicative gestures ("protowords") emerge which are separate from (but initially grounded in) pantomime.

I then argue, controversially, that a brain with the above systems in place was language-ready, and that it was cultural evolution in *Homo sapiens* that yielded language incrementally. Getting from praxic and pantomimic actions to reduced, conventionalized, arbitrary ones is relatively straightforward, but what of the transition to predication? Elsewhere (Arbib 2008b), I have argued that the earliest protolanguages were in large part holophrastic, and that as they developed through time, each protolanguage retained holophrastic strategies while making increasing use of compositional strategies.[3] As protowords that represented frequently occurring or highly significant events were fractionated to yield words for actions and objects, constructions emerged to put the pieces back together again – but now with the added power that combinations could be formed for which no protowords existed. We have seen in the "want x" example the transition from holophrase to predicate in our discussion of language development in a modern human child. The claim is that complex imitation, when coupled with protosign, gave the language-ready brain the ability to both fractionate and recombine perceptual and motor structures. However, where modern children are exposed to a host of fractionated utterances that provide a ready basis for acquiring much of a language in a few years, early humans had to invent fractionations one by one, with cumulative generalizations emerging at first very slowly over tens of millennia to bridge the spectrum from protolanguages to languages.

We now come to the final stage, the transition from protolanguage to language, the development of syntax and compositional semantics. This might have involved grammatically specific biological evolution. Pinker and Bloom (1990) argue that Universal Grammar is innate, evolving through multiple stages, but their definition of Universal Grammar is incomplete, and some of their stages seem as amenable to cultural as to biological evolution. As I have argued earlier, it seems more plausible that the diversity of grammar should be captured in the history of different societies rather than in the diversity of the genes. The nature of the transition to language remains hotly debated and lies beyond the scope of this chapter (but the later sections of Arbib 2008a do offer some suggestions). Although the Mirror System Hypothesis posits that complex imitation evolved first to support the transfer of praxic skills and then came to support protolanguage, it is important to note its crucial relevance to modern-day language acquisition and adult language use. Complex imitation has two parts: (i) the ability to perceive that a novel action may be approximated by a composite of known actions associated with appropriate subgoals; and (ii) the ability to employ this perception to perform an approximation to the observed action, which may then be refined through

3. A recent book (Arbib & Bickerton 2010) includes the 2008 paper and presents both sides of a vigorous debate.

practice. Both parts come into play when the child is learning a language, whereas the former predominates in adult use of language as the emphasis shifts from mastering novel words and constructions to finding the appropriate way to continue a dialogue.

5. Discussion

In reviewing a draft of this paper, Stefano Nolfi offered the following critique (which I have edited slightly):

> I suggest considering the possibility that the expansion of individuals' behavioral and cognitive repertoires in general might have constituted the basis for the emergence of language rather than restricting the causal role to manual dexterity and imitation only. Moreover, the relation between the mirror neurons and the Broca's area might be the result of skill/neural reuse. This implies that this correlation does not necessarily demonstrate that a genetically driven brain change might had been at the basis of language evolution.
>
> In that respect, manual dexterity and imitation certainly represent two of the key elements for behavioral and cognitive expansions but not the only ones. Of course, the presence of mirror neurons in an area involved in speech production is an evidence in support of the author's hypothesis. On the other hand, such data might also be explained as a consequence of the fact that motor control and speech production are related processes. The development of language ability might thus have benefited from the reuse of such areas dedicated to related skills without necessarily implying that the brain structures subsuming those skills represent a necessary prerequisite for the acquisition of a language ability.
>
> On the basis of these considerations, it seems to me that related but more parsimonious theories can be hypothesized. These theories would postulate that language might have originated by a gradual expansion of individuals' behavioral repertoire that might have increased the adaptive need for cognitive and social skills (such as selective attention, memory, sequential control, social learning, and imitation), which in turn might have produced a further increase of the individuals' behavioral repertoire. This process might have created the adaptive need for the development of more sophisticated communication systems (compositional – at least in part – protolanguages) that had been originally developed through the reuse of previously developed cognitive and behavioral skills (and that are acquired during infant development by reusing related, previously acquired, skills).

This may well be true, but this hypothesis suggests that human cognitive abilities expanded first and that language was an automatic consequence. But is it plausible to imagine a great broadening of cognition that did not depend on the elaboration of an ability for communication to at least the level of protolanguage, as a driver rather than a mere consequence? How elaborate can one's conceptual repertoire be without the calibration provided by (proto)language, which enhances the ability to share useful concepts across a social group? Saying this does not prove Nolfi wrong, but it does suggest that any attempt to elaborate his theory might in the end converge with many of the ideas presented here. For example, Nolfi suggests that the presence of mirror neurons in an area involved in speech production might be explained as a consequence of the fact that motor control and speech production are related processes. But the fact is that no other creature has speech processes that support the expression of language, and so the issue is not "Are motor control and speech production related?" The fact that they are is a central tenet of the Mirror System Hypothesis. The question, rather, is "How did speech production arise as a form of motor control unique to humans?" A few more observations may suggest issues that will ground further research that can address Nolfi's challenges.

In putting parity at stage center in the mirror system hypothesis, we adhere to the view that the primary function of language is communication. Language is a shared medium, and thus parity is essential to it. No matter how useful a word may be as a tool for cognition, we must learn the word in the first place; and we must then engage in numerous conversations if, in concert with our own thoughts, we are to enrich our understanding of any associated concept and our ability to make fruitful use of it. Having said this, both the external social uses of language and the internal cognitive uses of language could have provided powerful and varied adaptive pressures for further evolution of such capacities as anticipation, working memory, and autobiographic memory as language enriched both our ability to plan ahead, to consider counterfactual possibilities, and to mull over past experience to extract general lessons (Suddendorf & Corballis 2007). Indeed, where we lay stress on parity in the evolution of the language-ready brain, Aboitiz et al. (Aboitiz & Garcia 1997; Aboitiz et al. 2006a, 2006b) lay primary stress on the evolution of working memory systems. I see such alternatives as complementary, rather than either excluding the other.

Jürgens (1979, 2002), working primarily with squirrel monkeys rather than macaques, found that voluntary control over the initiation and suppression of monkey vocalizations – initiation and suppression of calls from a small repertoire, not the dynamic assemblage and coarticulation of articulatory gestures that constitutes speech – relies on the mediofrontal cortex including the anterior cingulate gyrus. A major achievement of the Mirror System Hypothesis is to develop a plausible explanation as to why Broca's area in humans, which corresponds to the macaque F5 brain region, rather than the vocalization area of cingulate cortex, lies at the core of language production. Ferrari et al. (2003, 2005) found that F5 mirror neurons include some for orofacial gestures involved in feeding. Moreover, some of these gestures (such as lip smacking and teeth chattering) do have auditory side-effects which can be exploited for communication. This system has interesting implications for language evolution (Fogassi & Ferrari 2004, 2007), but is a long way from mirror neurons for speech.

Intriguingly, the squirrel monkey F5 region does have connections to the vocal folds (Jürgens, personal communication, 2006), but these are involved only in closing them and not in vocalization (but see Coudé et al. 2007). We thus hypothesize that the emergence of protospeech on the scaffolding of protosign involved expansion of the F5 projection to the vocal folds to allow vocalization to be controlled in coordination with the control of the use of tongue and lips as part of the ingestive system.

In short, the Mirror System Hypothesis presents a range of challenges that must be addressed in seeking to understand the role of the evolving brain in the emergence of language on phylogenetic, historical, and ontogenetic timescales.

References

Aboitiz, F., & Garcia, V.R. (1997). The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. *Brain Research. Brain Research Reviews, 25*, 381–396. ⟨http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9495565⟩

Aboitiz, F., Garcia, R.R., Bosman, C., & Brunetti, E. (2006a). Cortical memory mechanisms and language origins. *Brain and Language, 98*, 40–56. ⟨http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16481035⟩

Aboitiz, F., García, R., Brunetti, E., & Bosman, C. (2006b). The origin of Broca's area and its connections from an ancestral working memory network. In Y. Grodzinsky & K. Amunts (Eds.), *Broca's region* (pp. 3–16). New York, NY: Oxford University Press.

Arbib, M.A. (2002). The mirror system, imitation, and the evolution of language. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 229–280). Cambridge, MA: The MIT Press.

Arbib, M.A. (2005a). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics [with commentaries and author's response]. *Behavioral and Brain Sciences, 28*, 105–167.

Arbib, M.A. (2005b). Interweaving protosign and protospeech: Further developments beyond the mirror. *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems, 6*, 145–171.

Arbib, M.A. (2006). A sentence is to speech as what is to action? *Cortex, 42*, 507–514. ⟨http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16881260⟩

Arbib, M.A. (2007). How new languages emerge (Review of D. Lightfoot, 2006, *How new languages emerge*, Cambridge University Press). Linguist List 18–432, Feb. 8, 2007, ⟨http://linguistlist.org/issues/17/17–1250.html⟩

Arbib, M.A. (2008a). From grasp to language: Embodied concepts and the challenge of abstraction. *Journal of Physiology – Paris, 102*, 4–20. ⟨http://www.sciencedirect.com/science/article/B6VMC-4S32NK4-2/1/b625a008ff4ae8c5c5d94aa27abda4dc⟩

Arbib, M.A. (2008b). Holophrasis and the protolanguage spectrum. *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems, 9*, 151–165.

Arbib, M.A. (2009). Invention and community in the emergence of language: A perspective from new sign languages. In S. M. Platek & T. K. Shackelford (Eds.), *Foundations in evolutionary cognitive neuroscience: Introduction to the discipline* (pp. 117–152). Cambridge, UK: Cambridge University Press.

Arbib, M.A. (2012). *Beyond the mirror: Evolving language and our social selves*. Oxford: Oxford University Press.

Arbib, M.A., & Bickerton, D. (Eds.). (2010). *The emergence of protolanguage: Holophrasis vs compositionality*. Amsterdam: John Benjamins.

Arbib, M.A., & Hill, J.C. (1984). Schemas, computation, and language acquisition. *Human Development, 27*, 282–296.

Arbib, M.A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology, 49*, 1053–1076. (Follow the link ⟨http://www.journals.uchicago.edu/doi/full/1010.1086/593015#apa⟩ for access to supplementary materials and videos.)

Arbib, M.A., & Rizzolatti, G. (1997). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition, 29*, 393–424.

Armstrong, D.F., Stokoe, W.C., & Wilcox, S.E. (1995). *Gesture and the nature of language*. Cambridge, UK: Cambridge University Press.

Armstrong, D.F., & Wilcox, S.E. (2007). *The gestural origin of language*. Oxford: Oxford University Press.

Baker, M. (2001). *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences, 7*, 173–221.

Bickerton, D. (1995). *Language and human behavior*. Seattle, WA: University of Washington Press.

Buccino, G., Vogt, S., Ritzl, A., Fink, G., Zilles, K., Freund, H., et al. (2004). Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron, 42*, 323–333.

Byrne, R.W. (2003). Imitation as behavior parsing. *Philosophical Transactions of the Royal Society of London B, 358*, 529–536.

Byrne, R.W., & Byrne, J.M.E. (1993). Complex leaf-gathering skills of mountain gorillas (*Gorilla g. beringei*): Variability and standardization. *American Journal of Primatology, 31*, 241–261.

Cheney, D.L., & Seyfarth, R.M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review, 22*, 135–159.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1992). A minimalist program for linguistic theory. In K. Hale & S.J. Keyser (Eds.), *The view from building 20: essays in linguistics in honor of Sylvain Bromberger* (pp. 1–52). Cambridge, MA: The MIT Press.

Chomsky, N., & Lasnik, H. (1993). The theory of principles and parameters. In J. Jacobs, A. v. Stechow, W. Sternefeld & T. Vennemann (Eds.), *Syntax: An international handbook of contemporary research* (pp. 506–569). Berlin: De Gruyter.

Clark, E. (2003). *First language acquisition*. Cambridge, UK: Cambridge University Press.

Clark, R. (1992). The selection of syntactic knowledge. *Language Acquisition, 2*, 83–149.

Corballis, M.C. (2002). *From hand to mouth: The origins of language*. Princeton, NJ: Princeton University Press.

Coudé, G., Ferrari, P.F., Roda, F., Maranesi, M., Veroni, V., Monti, F., et al. (2007). Neuronal responses during vocalization in the ventral premotor cortex of macaque monkeys. *Society for Neuroscience Annual Meeting (San Diego California), Abstract 636.3*.

Croft, W. (2000). *Explaining language change: An evolutionary approach*. London: Longman.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Deacon, T.W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W. W. Norton & Co.

Dixon, R.M.W. (1997). *The rise and fall of languages*. Cambridge, UK: Cambridge University Press.

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Evans, N., & Levinson, S.C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429–448. ⟨http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=6427084&fulltextType=RA&fileId=S0140525X0999094X⟩

Ferrari, P.F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17*, 1703–1714. ⟨http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12752388⟩

Ferrari, P.F., Maiolini, C., Addessi, E., Fogassi, L., & Visalberghi, E. (2005). The observation and hearing of eating actions activates motor programs related to eating in macaque monkeys. *Behavioural Brain Research, 161*, 95–101. ⟨http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15904715⟩

Fillmore, C.J., Kay, P., & O'Connor, M.K. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language and Cognitive Processes, 64*, 501–538.

Fogassi, L., & Ferrari, P.F. (2004). Mirror neurons, gestures and language evolution. *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems, 5*, 345–363.

Fogassi, L., & Ferrari, P.F. (2007). Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science, 16*, 136–141.

Gibson, E., & Wexler, K. (1994). Triggers. *Linguistic Inquiry, 25*, 407–454.

Goldberg, A.E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science, 7*, 219–224.

Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York, NY: Psychology Press.

Goldin-Meadow, S., & Feldman, H. (1975). The creation of a communication system: A study of deaf children of hearing parents. *Sign Language Studies, 8*, 225–234.

Grafton, S.T., Arbib, M.A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by PET: 2. Observation compared with imagination. *Experimental Brain Research, 112*, 103–111.

Hewes, G.W. (1973). Primate communication and the gestural origin of language. *Current Anthropology, 12*, 5–24.

Hill, J.C. (1983). A computational model of language acquisition in the two-year-old. *Cognition and Brain Theory, 6*, 287–317.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology, 3*, e79. doi:10.1371/journal.pbio.0030079.

Jürgens, U. (1979). Neural control of vocalization in nonhuman primates. In H.D. Steklis & M. J. Raleigh (Eds.), *Neurobiology of social communication in primates* (pp. 11–44). New York, NY: Academic Press.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26*, 235–258. ⟨Go to ISI⟩://000174534600006.

Lass, R. (1997). *Historical linguistics and language change*. Cambridge, UK: Cambridge University Press.

Lieberman, P. (2000). *Human language and our reptilian brain: The subcortical bases of speech, syntax and thought*. Cambridge, MA: Harvard University Press.

Lightfoot, D.W. (2006). *How new languages emerge*. Cambridge, UK: Cambridge University Press.

MacNeilage, P.F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences, 21*, 499–546.

MacNeilage, P.F., & Davis, B.L. (2005). The frame/content theory of evolution of speech: A comparison with a gestural-origins alternative. *Interaction Studies, 6*, 173–199. ⟨http://www.sign-lang.uni-hamburg.de/BibWeb/LiDat.acgi?ID=63566⟩

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

Myowa-Yamakoshi, M., & Matsuzawa, T. (1999). Factors influencing imitation of manipulatory actions in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 113*, 128–136.

Noble, W., & Davidson, I. (1996). *Human evolution, language and mind: A psychological and archaeological inquiry*. Cambridge, UK: Cambridge University Press.

Petersson, K.M., Reis, A., Askelof, S., Castro-Caldas, A., & Ingvar, M. (2000). Language processing modulated by literacy: A network analysis of verbal repetition in literate and illiterate subjects. *Journal of Cognitive Neuroscience, 12*, 364–382. ⟨http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10931764⟩

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707–784.

Poizner, H., Klima, E., & Bellugi, U. (1987). *What the hands reveal about the brain*. Cambridge, MA: The MIT Press.

Polich, L. (2005). *The emergence of the deaf community in Nicaragua: "With sign language you can learn so much."* Washington, DC: Gallaudet University Press.

Pullum, G.K., & Scholz, B.C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review, 19*, 9–50.

Rizzolatti, G., & Arbib, M.A. (1998). Language within our grasp. *Trends in Neuroscience, 21*, 188–194.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science, 305*, 1779–1782.

Seyfarth, R.M., Cheney, D.L., & Marler, P. (1980a). Monkey responses to three different alarm calls: Evidence for predator classification and semantic communication. *Science, 210*, 801–803.

Seyfarth, R.M., Cheney, D.L., & Marler, P. (1980b). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour, 28*, 1070–1094.

Stokoe, W.C. (2001). *Language in hand: Why sign came before speech*. Washington, DC: Gallaudet University Press.

Suddendorf, T., & Corballis, M.C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences, 30*, 299–351.

Tomasello, M. (1999). Emulation learning and cultural learning. *Behavioral and Brain Sciences, 21*, 703–704.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Visalberghi, E., & Fragaszy, D. (2001). Do monkeys ape? Ten years after. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 471–499). Cambridge, MA: The MIT Press.

Voelkl, B., & Huber, L. (2007). Imitation as faithful copying of a novel technique in marmoset monkeys. *PLoS ONE, 2*, e611. (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17622356)

Walker, W., & Sarbaugh, J. (1993). The early history of the Cherokee syllabary. *Ethnohistory, 40*, 70–94.

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London B, 358*, 501–515.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication, 18*, 47–67.