A Diversity of Divides in Devising System Models for Language Evolution

Chomsky (1956) showed that certain classes of grammars align with certain classes of automata – most famously, the alignment between context-free grammars and push-down automata. Over the years, Chomsky has moved on from this highly formal framework to seek to characterize a class of grammars broad enough to encompass the syntax of all human languages, with the resultant framework changing quite drastically every decade or so. Chomsky’s writings introduce another automaton, the Language Acquisition Device (LAD), but this time without a precise formal framework. It is posited to be innately specified, and present in (almost) every human infant. Its task of is to accept strings from a single language L as input and to infer a grammar G(L) which generates all and only the strings of L. Most of Chomsky’s writing discounts the notion that there can be a useful account of the evolution of language (see Hauser et al., 2002, for an apparent exception -- but one that seems inconsistent with much that has gone before). But for those who wish to study language evolution in the Chomskian framework (e.g., Pinker and Bloom, 1990), the study of evolution reduces to breaking Universal Grammar into several components that can be placed within an evolutionary succession. On this view, the available data are simply well-formedness judgments for each language under study. However, working backwards in time, we can see that different approaches to language evolution may consider language in a broader sense of performance that includes:

· Data on language use in communication and thought, including use in conversation, including a whole range of speech acts.
· 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.
· Animal communication systems in general, seeking commonalities with and differences from language.
· Animal behavior in general, seeking commonalities with and differences from communication systems.
· Data on brain function.

Our general concern, then, is: how did evolution equip Homo sapiens to enable human societies to develop a range of languages; and an individual to acquire the language of the community in which it is raised? The evolutionary challenge is both biological and cultural and there are many divisions over what the nature of this challenge might be:

**Biological: How did biological evolution endow Homo sapiens with brains and bodies that can acquire and use language?**

- **Modeling Divide 1:** Our ancestors evolved a capability for protolanguage – which had an open lexicon but little if any syntax – before they developed language versus no intermediary was involved.
- **Modeling Divide 2:** Models must be speech-centered, equating speech with language, versus we seek to understand the shared mechanisms that support signed languages as well as spoken languages.
- In any case, the importance of spoken language asks us to understand the evolution of the vocal apparatus which made it able to produce speech, and the related mechanisms for perception, and for the neural control of perception and production.
- **Modeling Divide 3:** Biological evolution of language must make crucial use of data on the brain versus the primary data are those of linguistics alone.

**Cultural: What aspects of language are innate, and what are the fruits of historical change?**

- **Modeling Divide 4:** Biological evolution endowed us with a Universal Grammar versus it gave us mechanisms which made the eventual invention of language possible.
- In either case, one needs to understand what the child learns during language acquisition, and what biological evolution provided to make such learning possible.
• In either case, one needs to understand what happens during historical language change, and what biological evolution provided to make such processes possible.

**Modeling Divide 5:** Initially, much of protolanguage was holophrastic (with a protoword describing a frequently occurring or significant situation) versus protolanguage started with words akin in scope to modern words (such as nouns and verbs)

**Modeling Divide 6:** Language evolution is to be understood solely in terms of adaptive pressures for communication or thought versus language evolution rests in part on the exaptation of adaptations that are not directly related to communication.

These six modeling divides (and, of course, there are others) define 64 overall approaches to language evolution. Thus any general framework must 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 mechanisms responsible for some specific set of data, 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.

Returning to our general discussion, then, our choice of behavioral data determines whether we look at Performance Systems, Developmental Systems, Historical Systems, and Evolutionary Systems. The full talk will outline the different modeling challenges these pose, and how each relates to the later types.

**The choices made in developing the Mirror System Hypothesis**

The Mirror System Hypothesis (Arbib, 2005; Arbib and Rizzolatti, 1997; Rizzolatti and Arbib, 1998) offers an approach to language evolution rooted in the data of primatology, comparative neurobiology, and neurolinguistics. In terms of the six Modeling Divides (MDs) presented in the previous section, the Mirror System Hypothesis makes the following choices:

• MD1: Our ancestors evolved a capability for protolanguage before they developed language.
• MD2: Models must seek to understand the shared mechanisms that support signed languages as well as spoken languages.
• The importance of spoken language asks us to understand the evolution of the vocal apparatus which made it able to produce speech, and the related mechanisms for perception, and for the neural control of perception and production.
• MD3: Biological evolution of language must make crucial use of data on the brain.
• MD4: Biological evolution gave us mechanisms which made the eventual invention of language possible.
• One needs to understand what the child learns during language acquisition, and what biological evolution provided to make such learning possible.
• One needs to understand what happens during historical language change, and what biological evolution provided to make it possible.
• MD5: Initially, much of protolanguage was holophrastic.
• MD6: Language evolution rests in part on the exaptation of adaptations that are not directly related to communication.
• What is the social structure that brings the child into the community using a specific language, and what neural capabilities are needed in the brains of the child and its caregivers to support this process?

The full talk will summarize where the Mirror System Hypothesis has offered a computational model or a conceptual model for a subtopic within the framework, outlining the modeling approach used and what data have been addressed in each case. In particular, it will discuss how the modeling of brain mechanisms in the monkey may suggest new ways of approaching the modeling of language mechanisms, complemented by analysis of models of
visual scene perception in relation to descriptions of the scene observed (Arbib, 2010; Arbib and Bonaiuto, 2008; Arbib and Lee, 2008).

Acknowledgement: The present research was supported in part by the National Science Foundation Program in Perception, Action, and Cognition under Grant BCS-0924674 (Michael A. Arbib, Principal Investigator).

References