Although the view of Falk puts language emergence in a continuum that is closer to primate vocal communication than the framework of Michael Arbib, both models involve a progression atop the systems already preexisting in nonhuman primates. Arbib's work gives the first detailed account of putative evolutionary stages in the emergence of human language from a cognitive viewpoint. It therefore could be used as a framework to test specific links between cognitive human language and communicative human language emergence hypotheses, such as the one recently proposed by Falk.

Auditory object processing and primate biological evolution

Barry Horwitz,[a] Fatima T. Husain,[a] and Frank H. Guenther[b]

[a]*Brain Imaging and Modeling Section, National Institute on Deafness and Other Communications Disorders, National Institutes of Health, Bethesda, MD 20892;* [b]*Department of Cognitive and Neural Systems, Boston University, Boston, MA 02215.* **horwitz@helix.nih.gov** **husainf@nidcd.nih.gov**
http://www.nidcd.nih.gov/research/scientists/horwitzb.asp
guenther@cns.bu.edu **http://www.cns.bu.edu/~guenther/**

Abstract: This commentary focuses on the importance of auditory object processing for producing and comprehending human language, the relative lack of development of this capability in nonhuman primates, and the consequent need for hominid neurobiological evolution to enhance this capability in making the transition from protosign to protospeech to language.

The target article by Arbib provides a cogent but highly speculative proposal concerning the crucial steps in recent primate evolution that led to the development of human language. Generally, much of what Arbib proposes concerning the transition from the mirror neuron system to protosign seems plausible, and he makes numerous points that are important when thinking about language evolution. We especially applaud his use of neural modeling to implement specific hypotheses about the neural mechanisms mediating the mirror neuron system. We also think his discussion in section 6 of the necessity to use protosign as scaffolding upon which to ground symbolic auditory gestures in protospeech is a significant insight. However, the relatively brief attention Arbib devotes to the perception side of language, and specifically to the auditory aspects of this perception, seems to us to be a critical oversight. The explicit assumption that protosign developed before protospeech, reinforced by the existence of sign language as a fully developed language, allows Arbib (and others) to ignore some of the crucial features that both the productive and receptive aspects of speech require in terms of a newly evolved neurobiological architecture.

One aspect of auditory processing that merits attention, but is not examined by Arbib, has to do with auditory object processing. By auditory object, we mean a delimited acoustic pattern that is subject to figure-ground separation (Kubovy & Van Valkenburg 2001). Humans are interested in a huge number of such objects (in the form of words, melodic fragments, important environmental sounds), perhaps numbering on the order of 10^5 in an individual. However, it is difficult to train monkeys on auditory object tasks, and the number of auditory objects that interest them, compared to visual objects, seems small, numbering perhaps in the hundreds (e.g., some species-specific calls, some important environmental sounds). For example, Mishkin and collaborators (Fritz et al. 1999; Saunders et al. 1998) have showed that monkeys with lesions in the medial temporal lobe (i.e., entorhinal and perirhinal cortex) are impaired relative to unlesioned monkeys in their ability to perform correctly a visual delayed match-to-sample task when the delay period is long, whereas both lesioned and unlesioned monkeys are equally unable to perform such a task using auditory stimuli.

These results implicate differences in monkeys between vision and audition in the use of long-term memory for objects. Our view is that a significant change occurred in biological evolution allowing hominids to develop the ability to discriminate auditory objects, to categorize them, to retain them in long-term memory, to manipulate them in working memory, and to relate them to articulatory gestures. It is only the last of these features that Arbib discusses. In our view, the neural basis of auditory object processing will prove to be central to understanding human language evolution. We have begun a systematic approach combining neural modeling with neurophysiological and functional brain imaging data to explore the neural substrates for this type of processing (Husain et al. 2004).

Concerning language production, Arbib's model of the mirror-neuron system (MNS) may require considerable modification, especially when the focus shifts to the auditory modality. For instance, there is no treatment of babbling, which occurs in the development of both spoken and sign languages (Petitto & Marentette 1991). Underscoring the importance of auditory processing in human evolution, hearing-impaired infants exhibit vocal babbling that declines with time (Stoel-Gammon & Otomo 1986).

However, there has been work in developing biologically plausible models of speech acquisition and production. In one such model (Guenther 1995), a role for the MNS in learning motor commands for producing speech sounds has been posited. Prior to developing the ability to generate speech sounds, an infant must learn what sounds to produce by processing sound examples from the native language. That is, he or she must learn an auditory target for each native language sound. This occurs in the model via a MNS involving speech sound-map cells hypothesized to correspond to mirror neurons (Guenther & Ghosh 2003). Only after learning this auditory target can the model learn the appropriate motor commands for producing the sound via a combination of feedback and feed-forward control subsystems. After the commands are learned, the same speech sound-map cell can be activated to read out the motor commands for producing the sound. In this way, mirror neurons in the model play an important role in both the acquisition of speaking skills and in subsequent speech production in the tuned system. This role of mirror neurons in development of new motor skills differs from Arbib's MNS model, which "makes the crucial assumption that the grasps that the mirror system comes to recognize are already in the (monkey or human) infant's repertoire" (sect. 3.2, para. 7).

Our efforts to comprehend the biological basis of language evolution will, by necessity, depend on understanding the neural substrates for human language processing, which in turn will rely heavily on comparative analyses with nonhuman primate neurobiology. All these points are found in Arbib's target article. A crucial aspect, which Arbib invokes, is the necessary reliance on neurobiologically realistic neural modeling to generate actual implementations of neurally based hypotheses that can be tested by comparing simulated data to human and nonhuman primate experimental data (Horwitz 2005). It seems to us that the fact that humans use audition as the primary medium for language expression means that auditory neurobiology is a crucial component that must be incorporated into hypotheses about how we must go beyond the mirror-neuron system.

On the neural grounding for metaphor and projection

Bipin Indurkhya

International Institute of Information Technology, Hyderabad 500 019, India.
bipin@iiit.net

Abstract: Focusing on the mirror system and imitation, I examine the role of metaphor and projection in evolutionary neurolinguistics. I suggest that the key to language evolution in hominid might be an ability to project one's thoughts and feelings onto another agent or object, to see and feel things from another perspective, and to be able to empathize with another agent.