



Comment

A modality-independent, neurobiological grounding for the  
combinatory capacity of the language-ready brain  
Comment on “Towards a Computational Comparative  
Neuroprimatology: Framing the language-ready brain” by  
Michael A. Arbib

Ina Bornkessel-Schlesewsky\*, Phillip M. Alday, Matthias Schlesewsky

*Cognitive Neuroscience Laboratory, School of Psychology, Social Work and Social Policy, University of South Australia, Adelaide, Australia*

Received 7 January 2016; accepted 21 January 2016

Available online 22 January 2016

Communicated by L. Perlovsky

In this comprehensive review of his past and current work on language evolution, Arbib [1] argues that “the capability for protosign – rather than elaborations intrinsic to the core vocalization systems – may [...] have provided the essential scaffolding for protospeech and evolution of the human language-ready brain” (p. 25). He hypothesises that this evolutionary trajectory is based on the mirror system and mechanisms of complex imitation that developed by drawing on systems “beyond the mirror”. As Arbib himself discusses in detail, the claim that gestural combinatorics of increasing complexity and symbolisation formed a prerequisite for the evolution of auditory speech and language is rather controversial. Though, in our own previous work, we have emphasised the importance of the computational properties of the auditory system in defining the language-ready brain [2], we would like to focus on a somewhat different, and perhaps even more foundational issue for the purposes of this commentary: are there basic neurobiological mechanisms that underlie combinatorial processing irrespective of modality?

The answer to this question may be surprisingly simple. Internal (forward and inverse) models, currently one of the “hot topics” in cognitive neuroscience (e.g. [3]), indeed provide a basic, biologically plausible and – perhaps even more importantly – unifying mechanism for combinatorial processing. Originally proposed for the domain of motor control and sensorimotor integration [4], internal models can serve as models of the body and its surrounding environment [5], and they have been used to account for non-motor processing in domains such as one’s own sense of agency [6], social interaction [7], theory of mind [8] and language [2,9,10]. The common denominator between forward models in all of these areas is that they serve to encode predictions about the environment that can be compared against sensory information, thereby either confirming the model or leading to its adaptation. This type of mechanism provides a natural foundation for combinatorial processing, as it (a) implements the (predictive) processing of previously encountered sequences; and (b) allows for the interpretation of novel sequences, by enabling an input item

DOI of original article: <http://dx.doi.org/10.1016/j.plrev.2015.09.003>.

\* Corresponding author at: School of Psychology, Social Work and Social Policy, University of South Australia, GPO Box 2471, Adelaide, SA 5001, Australia. Tel.: +61 (0)8 8302 2545; fax: +61 (0)8 8302 4377.

E-mail address: [Ina.Bornkessel-Schlesewsky@unisa.edu.au](mailto:Ina.Bornkessel-Schlesewsky@unisa.edu.au) (I. Bornkessel-Schlesewsky).

encountered at position  $n + 1$  in a sequence to be understood in the context of the predictions generated at the position of input item  $n$ . Furthermore, the hierarchical (nested) organisation of internal models [11] allows for differing degrees of abstraction, with the lowest level of the hierarchy serving to predict the concrete sensory input, and higher levels predicting more abstract features or feature combinations (e.g. the category of a word rather than a specific word) [2].

Given the ubiquity of internal models across modalities and the resulting availability of combinatorial mechanisms to various brain systems, it is not clear why combinatorics in the manual–visual system need have been primary in leading to protolanguage – and the evolution of speech and language in the auditory domain. An alternative possibility is that combinatorial aspects of communication emerged in multiple modalities, thereby allowing for modality-specific restrictions to be overcome both in the manual–visual and in the auditory domain: in contrast to auditory signals, manual–visual signals have the advantage of being three-dimensional (as in human sign language); conversely, auditory signals have the advantage of not occupying the hands and the opposable thumbs, and of allowing for communication over distances and without direct line of sight (cf. [12] for the importance of this property, particularly in difficult terrain). From this perspective, the considerable anatomical overlap between the neural circuits comprising the mirror system and action observation network [13] and those of the dorsal auditory stream [14] is not surprising. Consequently, protolanguage may have evolved from the combinatorial processing of sequences in time in both the auditory and visual systems – that is, from proto symbol sequences bearing characteristics of both domains, rather than solely from pantomime and protosign.

An account along these lines opens up an interesting possible solution to two remaining problems: how fully-fledged languages evolved from protolanguage, and why the auditory system came to dominate speech and language processing as well as language acquisition. (Note that we do not deny the importance of co-speech gestures. However, the primacy of the auditory modality in language acquisition appears indisputable.) Notably, even in macaques and thereby likely in the last common ancestor of macaque, chimpanzee and human (LCA-m), the auditory system, and specifically the ventral auditory stream, demonstrably supports auditory object recognition [14,15]. While we agree with Arbib that it is questionable whether innate monkey calls should be analysed as symbols, the complexity of at least certain monkey vocalisation systems certainly appears to require a symbolic analysis. As described in detail by Collier and colleagues in a recent review of combinatorics in animal communication [16], Campbell monkeys (*Cercopithecus campbelli campbelli*) employ vocalisation sequences which encode at least rudimentary compositional semantics and thus require some level of symbolic encoding. Like macaques, Campbell monkeys are Old World Monkeys and thus evolved from LCA-m. From this type of compositional combinatorics, only a small step is required to yield one of the key characteristics of human language: the ability to produce and interpret combinations of non-adjacent elements (i.e. non-sequence-based, associative interpretation). Thus, the key advantage of the auditory system for the evolution from protolanguage to fully-fledged languages may have lain in its existing neural infrastructure of a dorsal pathway that overlapped strongly with the domain-general combinatorial circuits described above, combined with a ventral pathway that allowed for sound-to-meaning mapping in the recognition of auditory objects [14,17] and provided a basis for the interpretation of sequence-independent, non-adjacent dependencies [2].

To conclude, we strongly agree with Arbib that comparative neuroprimatology provides a highly promising approach towards language evolution and the neurobiology of language (i.e. to understanding both the language-ready brain and the brain that supports language in use). We also share his conviction that no further genetic modification was required in order for language to evolve from protolanguage and the language-ready brain. We suggest, however, that it may be fruitful to consider the notion of modality-independent “protocombinatorics” based on basic neurobiological mechanisms as opposed to a primacy of the visual–manual system in the evolution of language.

## References

- [1] Arbib MA. Towards a computational comparative neuroprimatology: framing the language-ready brain. *Phys Life Rev* 2016;16:1–54. <http://dx.doi.org/10.1016/j.plrev.2015.09.003> [in this issue].
- [2] Bornkessel-Schlesewsky I, Schlesewsky M, Small SL, Rauschecker JP. Neurobiological roots of language in primate audition: common computational properties. *Trends Cogn Sci* 2015;19:142–50. <http://dx.doi.org/10.1016/j.tics.2014.12.008>.
- [3] Friston K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 2010;11:127–38. <http://dx.doi.org/10.1038/nrn2787>.
- [4] Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. *Science* 1995;269:1880–2.
- [5] Grush R. The emulation theory of representation: motor control, imagery, and perception. *Behav Brain Sci* 2004;27:377–96.
- [6] Blakemore S-J, Wolpert DM, Frith CD. Central cancellation of self-produced tickle sensation. *Nat Neurosci* 1998;1:635–40.
- [7] Wolpert DM, Doya K, Kawato M. A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond B, Biol Sci* 2003;358:593–602. <http://dx.doi.org/10.1098/rstb.2002.1238>.

- [8] Koster-Hale J, Saxe R. Theory of mind: a neural prediction problem. *Neuron* 2013;79:836–48. <http://dx.doi.org/10.1016/j.neuron.2013.08.020>.
- [9] Pickering MJ, Garrod S. Do people use language production to make predictions during comprehension? *Trends Cogn Sci* 2007;11:105–10. <http://dx.doi.org/10.1016/j.tics.2006.12.002>.
- [10] Pickering MJ, Garrod S. An integrated theory of language production and comprehension. *Behav Brain Sci* 2013;36:329–47. <http://dx.doi.org/10.1017/S0140525X12001495>.
- [11] Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 1999;2:79–87.
- [12] Carreiras M, Lopez J, Rivero F, Corina D. Linguistic perception: neural processing of a whistled language. *Nature* 2005;433:31–2.
- [13] Grafton ST. Embodied cognition and the simulation of action to understand others. *Ann NY Acad Sci* 2009;1156:97–117. <http://dx.doi.org/10.1111/j.1749-6632.2009.04425.x>.
- [14] Rauschecker JP. Cortical processing of complex sounds. *Curr Opin Neurobiol* 1998;8:516–21.
- [15] Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 1995;268:111–4.
- [16] Collier K, Bickel B, van Schaik CP, Manser MB, Townsend SW. Language evolution: syntax before phonology? *Proc R Soc Lond B, Biol Sci* 2014;281:20140263. <http://dx.doi.org/10.1098/rspb.2014.0263>.
- [17] Rauschecker JP, Tian B. Mechanisms and streams for processing of what and where in auditory cortex. *Proc Natl Acad Sci USA* 2000;97:11800–6.