

McNeill, D., Bertenthal, B., Cole, J. and Gallagher, S. 2005. Gesture-first, but no gestures? Commentary on Michael A. Arbib. *Behavioral and Brain Sciences* 28: 138-39.

“Gesture-first, but no gestures?”

David McNeill, Department of Psychology, University of Chicago, 5848 S. University Ave., Chicago, IL 60637 USA. Email: dmcneill@uchicago.edu. Corresponding author phone number: 773-702-8833.

Bennett Bertenthal, Department of Psychology, University of Chicago, 5848 S. University Ave., Chicago, IL 60637 USA. Email: bbertent@uchicago.edu.

Jonathan Cole, Clinical Neurological Sciences, University of Southampton, UK. Email: jonathan@cofamily.org.uk.

Shaun Gallagher, Department of Philosophy, University of Central Florida, Colbourn Hall 411, University of Central Florida, Orlando, FL 32816-1352 USA. Email: gallaghr@mail.ucf.edu.

Abstract

While Arbib’s extension of the mirror system hypothesis neatly sidesteps one problem with the ‘gesture-first’ theory of language origins, it overlooks the importance of gestures that occur in current day human linguistic performance, and this lands it with another problem. We argue that, instead of gesture-first, a system of combined vocalization and gestures would have been a more natural evolutionary unit.

Michael Arbib’s extension of the mirror system hypothesis for explaining the origin of language elegantly sets the stage for further discussion, but we think it overlooks a crucial source of data – the kinds of gestures that actually occur in current human linguistic performance. These data lead us to doubt a basic claim of the ‘gesture-first’ theory, that language started as a gesture language that was gradually supplanted by speech. Arbib has modified this theory with his concept of an expanding spiral, but this new model does not go far enough in representing a speech-gesture system that evolved together.

Classic gesture-first

The enduring popularity of ‘gesture-first’ seems to presuppose that gestures are simple and that as we, and language, became more complex speech evolved and to an extent supplanted gesture, a belief that emerged as part of the Enlightenment quest for the natural state of man and is credited to Condillac, and which has continued since, e.g., Hewes (1973), Armstrong, et al. (1995), and Corballis (2002). However, contrary to the traditional view, we contend that gesture and language, as they currently exist, belong to a single system of verbalized thinking and communication, and neither can be called the simple twin of the other. It is this *system*, in which both speech and gesture are crucial, that we should be explaining. It makes little sense to ask which part of an unbroken system is ‘simpler’ – a better question is how the parts work together.

McNeill, D., Bertenthal, B., Cole, J. and Gallagher, S. 2005. Gesture-first, but no gestures? Commentary on Michael A. Arbib. *Behavioral and Brain Sciences* 28: 138-39.

In this system, we find synchrony and co-expressiveness – gesture and speech convey the *same idea unit, at the same time*. Gesture and speech exhibit what Wundt described long ago as the “simultaneous” and “sequential” sides of the sentence (Blumenthal 1970, p. 21) and Saussure, in notes recently discovered, termed “l’essence double du langage” (Harris 2002). Double essence, not enhancement, is the relationship, and we do not see how it could have evolved from the *supplanting* of gestures by speech. In the remainder of this commentary, we summarize three sources of evidence to support this assertion.



1. Consider the attached drawing.¹ The speaker was describing a cartoon episode in which one character tries to reach another character by climbing up inside a drainpipe. The speaker is saying, “and he tries going **up through** it this time”, with the gesture occurring during the boldfaced portion (the illustration captures the moment when the speaker says the vowel of “through”). Co-expressively with “up” her hand rose upward and co-expressively with “through” her fingers spread outward to create an interior space. These took place together, and were synchronized with “up through”, the linguistic package that combines the same meanings.

The effect is a uniquely gestural way of packaging meaning – something like ‘rising hollowness’, which does not exist as a semantic package of English at all. Speech and gesture, at the moment of their synchronization, were co-expressive. The very fact there is shared reference to the character’s climbing up inside the pipe makes clear that it is being represented by the speaker in two ways simultaneously – analytic/combinatoric in speech and global/synthetic in gesture. We suggest it was this very simultaneous combination of opposites that evolution seized upon.

2. When signs and speech do combine, in contemporary human performance, they do not synchronize. Kendon (1988) observed sign languages employed by Aboriginal women – full languages developed culturally for (rather frequent) speech taboos – which they sometimes combine with speech. The relevant point is that in producing these combinations speech and sign start out synchronously but then, as the utterance proceeds, speech outruns the semantically equivalent signs. The speaker stops speaking until the signs catch up and then starts over, only for speech and signs to pull apart again. If, in the evolution of language, there had been a similar doubling up of signs and speech, as the supplanting scenario implies, they too would have been driven apart rather than into synchrony, and for this reason too we doubt the replacement hypothesis.

3. The Wundt/Saussure “double essence” of gesture and language appears to be carried by a dedicated thought-hand-language circuit in the brain. This circuit strikes us as a prime candidate for an evolutionary selection at the foundation of language. It implies

¹ Computer drawing from video by Fey Parrill, University of Chicago.

McNeill, D., Bertenthal, B., Cole, J. and Gallagher, S. 2005. Gesture-first, but no gestures? Commentary on Michael A. Arbib. *Behavioral and Brain Sciences* 28: 138-39.

that the aforementioned *combinations* of speech and gesture were the selected units, not gesture first with speech supplanting it. We observe this circuit in the unique neurological case of IW, who lost all proprioception and spatial position sense from the neck down at age 19, and has since taught himself to move using vision and cognition. The thought-language-hand link, located presumably in Broca's Area, ties together language and gesture, and, in IW, survives and is partly dissociable from instrumental action. We can address Arbib's pantomime model by observing the kinds of gestures the dedicated link sustains in IW's performance, in the absence of vision: his gestures are a) co-expressive and synchronous with speech, b) not supplemental, and c) not derivable from pantomime. IW is unable to perform instrumental actions without vision but continues to perform speech-synchronized, co-expressive gestures that are virtually indistinguishable from normal (topokinetic accuracy is reduced but morphokinetic accuracy is preserved) (Cole, Gallagher, and McNeill 2002). His gestures without vision, moreover, minimize the one quality that could be derived from pantomime, a so-called 'first-person' or 'character' viewpoint, in which a gesture replicates an action of a character (cf. McNeill 1992).

More generally, an abundance of evidence demonstrates that spontaneous, speech-synchronized gestures should be counted as part of language (McNeill 1992). Gestures are frequent (accompanying up to 90% of utterances in narrations). They synchronize exactly with co-expressive speech segments, implying that gesture and related linguistic content are co-active in time and jointly convey what is newsworthy in context. Gesture adds cohesion, gluing together potentially temporally separated but thematically related segments of discourse. Speech and gesture develop jointly in children, and decline jointly after brain injury. In contrast to cultural emblems, like the "O.K." sign, speech-synchronized gestures occur in all languages, so far as is known. Finally, gestures are not 'signs', with an independent linguistic code. Gestures exist only in combination with speech, and are not themselves a coded system.

Arbib's gesture-first

Arbib's concept of an expanding spiral may avoid some of the problems of the supplanting mechanism. He speaks of scaffolding and spiral expansion, which appear to mean, in both cases, that one thing is preparing the ground for or propping up further developments of the other thing – speech to gesture, gesture to speech, and so on. This spiral, as now described, brings speech and gesture into temporal alignment (see Arbib's Fig. 6), but also implies, rather than the evolution of a single 'thing', with a double essence, two things juxtaposed. Modification to produce a dialectic of speech and gesture, beyond scaffolding, does not seem impossible. However, the theory is still focused on gestures of the wrong kind for this dialectic – in terms of 'Kendon's Continuum (see McNeill 2000 for two versions), signs, emblems, and pantomime. Because it regards all gestures as simplified and meaning-poor, it is difficult to see how the expanding spiral can expand to include the remaining point on the Continuum, 'gesticulations' – the kind of speech-synchronized co-expressive gesture illustrated above. A compromise is that pantomime was the initial proto-language, but was replaced by speech *plus gesture*, leading to the thought-language-hand link that we have described. This hypothesis has the interesting implication that different evolutionary trajectories landed at different points along Kendon's Continuum. One path led to pantomime,

McNeill, D., Bertenthal, B., Cole, J. and Gallagher, S. 2005. Gesture-first, but no gestures? Commentary on Michael A. Arbib. *Behavioral and Brain Sciences* 28: 138-39.

another to co-expressive and speech-synchronized gesticulation, etc. These different evolutions are reflected today in distinct ways of combining movements with speech. While we do not question the importance of extending the mirror system hypothesis, we have concerns about a theory that predicts, as far as gesture goes, the evolution of what did not evolve instead of what did.

References

- Armstrong, David F., Stokoe, William C., & Wilcox, Sherman E. 1995. *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.
- Blumenthal, Arthur (ed. and trans.). 1970. *Language and Psychology: Historical aspects of psycholinguistics*. New York: John Wiley & Sons Ltd.
- Cole, J., Gallagher, S., and McNeill, D. 2002. Gesture following deafferentation: A phenomenologically informed experimental study. *Phenomenology and the Cognitive Sciences* 1: 49-67.
- Corballis, Michael C. 2002. *From Hand to Mouth: The origins of language*. Princeton, NJ: Princeton University Press.
- Harris, Roy. 2002. *Times Literary Supplement* 26 July, 5182: 30.
- Hewes, Gordon W. 1973. Primate communication and the gestural origins of language. *Current Anthropology* 14:5-24.
- Kendon, Adam. 1988. *Sign Languages of Aboriginal Australia: Cultural, semiotic and communicative perspectives*. Cambridge: Cambridge University Press.
- McNeill, David. 1992. *Hand and Mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- McNeill, David. 2000. Introduction. In D. McNeill (ed.), *Language and Gesture*, pp. 1-10. Cambridge: Cambridge University Press.