

Open Peer Commentary

Speech and gesture are mediated by independent systems

Anna M. Barrett,^a Anne L. Foundas,^b and Kenneth M. Heilman^c

^aDepartment of Neurology and Neural and Behavioral Sciences, Pennsylvania State University College of Medicine, Hershey, PA 17033;

^bDepartment of Psychiatry and Neurology, Tulane University School of Medicine and the Veterans Affairs Medical Center, New Orleans, LA 70115;

^cCollege of Medicine, Gainesville, FL 32610, and North Florida South Georgia Veterans Affairs Medical Center, Gainesville, FL 32610.

ambarrett@psu.edu foundas@tulane.edu

heilman@neurology.ufl.edu

<http://www.som.tulane.edu/neurograd/foundahm.htm#Foundas>

Abstract: Arbib suggests that language emerged in direct relation to manual gestural communication, that Broca's area participates in producing and imitating gestures, and that emotional facial expressions contributed to gesture-language coevolution. We discuss functional and structural evidence supporting localization of the neuronal modules controlling limb praxis, speech and language, and emotional communication. Current evidence supports completely independent limb praxis and speech/language systems.

Arbib suggests that language coevolved with a neuroanatomic-behavioral system specialized for manual gesture production and imitation, relying upon a shared anatomic substrate. He suggests that the region adapted for this purpose corresponds to Broca's area in the human brain and area F5 in the monkey. Although this is an interesting hypothesis, there are behavioral and anatomic studies inconsistent with this theory.

First, the target article treats four forelimb gesture classes interchangeably: volitional meaningless movements or movements unrelated to tool use (e.g., reaching or grasping movements), movements standing for objects (emblems, e.g., flapping hands for "bird"), tool use pantomimes (limb praxis), and movements carrying linguistic meaning (signs). A comprehensive discussion of the literature supporting separable systems controlling these movement classes is beyond the scope of this commentary. The term "gesture" here refers to limb praxis unless otherwise noted.

The anatomic evidence is weak regarding Broca's area as a shared neuronal substrate for human gesture and language. Although abnormal skilled learned purposive movements (limb apraxia) may co-occur with speech and language disorders (aphasia), these two conditions are anatomically and functionally dissociable (Kertesz et al. 1984; Liepmann 1920; Papagno et al. 1993; see Heilman & Rothi 2003 for a review). Indeed, in left-handed subjects, brain regions supporting limb praxis may be localized in the right hemisphere, and those supporting speech and language, in the left hemisphere (Heilman et al. 1973; Valenstein & Heilman 1979). Right-handed subjects with limb apraxia and gestural production and imitation deficits are hypothesized to have lost function in brain regions supporting time-space-movement representations. Most have damage to posterior, rather than anterior, cortical areas, usually the inferior parietal lobe (Rothi et al. 1994). A report of an isolated disorder of gesture imitation was associated with posterior, not frontal, cortical injury (Ochipa et al. 1994). Premotor lesions have been associated with limb apraxia, but these lesions were in the supplementary motor area (Watson et al. 1986) or the convexity premotor cortex (superior to Broca's area; Barrett et al. 1998). Lesions in Broca's area can cause oral or buccofacial apraxia and apraxia of speech. However, although limb apraxia is associated with Broca's aphasia (Albert et al. 1981), we could find no reports of patients with limb apraxia for whom there was

anatomic evidence of a brain lesion restricted to Broca's area. Patients with nonfluent aphasias (e.g., Broca's aphasia) can acquire simplified signs and gestural emblems (Amerind; not a signed language) as an alternative method of communication (Albert et al. 1981; Skelly 1974) and are preferred candidates over patients with aphasia from posterior brain injury. If coactivity of the speech-language and praxis systems led evolutionarily to refinement of these functional systems, one might expect coactivity to be associated with competent gestural and linguistic performance. However, in normal subjects, spontaneously speaking while gesturing is associated with poorer limb praxis performance (Barrett et al. 2002). Thus, anatomic-behavioral evidence in humans does not support Broca's area as necessary for programming skilled learned forelimb gestures.

Although Arbib posits a relationship between primate orofacial gestures, vocalizations communicating emotional state, and human language, human speech-language functions may be localized quite differently from emotionally communicative functions. In most humans, the left hemisphere may be dominant in the control of vocalization associated with propositional speech, but the right hemisphere often controls vocalization associated with emotional prosody (Heilman et al. 2000; Ross 1981; Tucker et al. 1977), automatic speech (e.g., the Lord's Prayer; Speedie et al. 1993) and singing (Wildgruber et al. 1996). In patients with aphasia after left-hemisphere injury, comprehension and production of affective vocal prosody and emotional facial expressions may also be relatively spared (Barrett et al. 1999; Kanter et al. 1986). This double dissociation argues against left-hemisphere dominance for comprehending, imitating, or producing emotional facial expression or prosody.

Lastly, there is evidence that brain anatomical regions mediating language functions are not truly homologues to the equivalent regions in primates. In humans, the classic Broca's area is not one functional unit but is comprised of two anatomical subregions differing in gross morphology, microscopic cytoarchitectonic features, and functionally (Foundas et al. 1998). The human Broca's area includes the pars triangularis (PTR) and pars opercularis (POP), which may mediate different functions. The PTR is comprised of higher-order heteromodal-association cortex suited to complex cross-modal associations typical of higher-order linguistic functions (e.g., syntax, lexical semantics), whereas the POP is comprised of motor-association cortex suited to articulatory and motor speech functions. In a lesion study, Tonkonogy and Goodglass (1981) reported differences between two patients with delimited Broca subregion lesions. A patient with a lesion restricted to the PTR had linguistic deficits, whereas the patient with a lesion to the POP had disordered articulation and fluency, but preserved higher-order language functions. Using functional MRI to study phonemic versus semantic fluency in the inferior frontal gyrus, Paulesu et al. (1997) also found functional heterogeneity within Broca's area. Whereas phonemic and semantic fluency tasks activated the PTR, only phonemic fluency tasks activated the POP. Hence, the anterior triangular portion of Broca's area (PTR) and the POP may be functionally dissociated, with the PTR functioning more critically in lexical retrieval, and the POP selectively subserving articulatory motor speech functions. Nonhuman primates may not have a homologue to the PTR (Brodmann's area 45). Thus, rather than being a PTR homologue, area F5 in the monkey may represent the human POP (Brodmann's area 44), more directly linked to vocalization.

Arbib discusses the possibility that some forms of limb praxis and speech/language function may draw upon the ability to represent symbolic meaning. Although not all people with limb apraxia or aphasia demonstrate a loss of action semantics or linguistic semantics, and Liepmann (1905) rejected asymbolia as a sole explanation for limb apraxia, some patients with limb apraxia (e.g., conceptual apraxia; Heilman & Rothi 2003; Raymer & Ochipa 1997) demonstrate abnormal action-meaning systems apart from gestural tasks. Whether a separation of action semantics and linguistic semantics in the modern human brain devel-

oped over the course of human evolution, however, unfortunately cannot be determined at our present level of knowledge.

Beyond the mirror neuron – the smoke neuron?

Derek Bickerton

Department of Linguistics, University of Hawaii, Honolulu, HI 96822.
derbick@hawaii.rr.com

Abstract: Mirror neurons form a poor basis for Arbib's account of language evolution, failing to explain the creativity that must precede imitation, and requiring capacities (improbable in hominids) for categorizing situations and unambiguously miming them. They also commit Arbib to an implausible holophrastic protolanguage. His model is further vitiated by failure to address the origins of symbolization and the real nature of syntax.

Mirror-neuron theory is the second-latest (*FOXP2* is the latest) in a series of magic-bullet solutions to the problems of language evolution. To his credit, Arbib realizes it could not account for all of language. Unfortunately, his attempts to go beyond it fall far short of adequacy.

Even as a significant component of language, mirror neurons are dubious. There cannot be imitation unless someone has first created something to imitate, and mirror neurons offer no clue as to how totally novel sequences – complex ones, at that – could have been created *ab ovo*. Moreover, when someone addresses you, you don't just imitate what they said (unless you want to be thought an idiot); you say something equally novel.

Arbib treats as wholly unproblematic both the category “frequently occurring situation” and the capacity of pantomime to represent such situations. Situations, frequent or otherwise, do not come with labels attached; indeed, it is questionable whether any species could isolate “a situation” from the unbroken, ongoing stream of experience unless it already had a language with which to do so. For this task requires abstracting away from a potentially infinite number of irrelevant features – place, weather, time of day, number and identity of participants, and on and on. How, short of mind-reading powers that would leave professional clairvoyants gasping, could our alingual ancestors have known which features seemed relevant to the sender of the message, and which did not?

If Arbib answers “through pantomime,” one assumes he has never played charades. Those who have, know that even with the help of a large set of “disambiguating signs” – stereotypic gestures for “film title,” “book title,” and so on, elaborate routines of finger-counting to provide numbers of words and syllables – participants with all the resources of modern human language and cognition find it often difficult and sometimes impossible to guess what the pantomimer is trying to represent. When what is to be represented is not a monosyllabic word but something as complex as “The alpha male has killed a meat animal and now the tribe has a chance to feast together. Yum, yum!” or “Take your spear and go round the other side of that animal and we will have a better chance of being able to kill it” (Arbib's own examples, sect. 7, para. 2), the likelihood of successful guessing becomes infinitesimally small.

Arbib does see what I pointed out more than a decade ago (Bickerton 1990, pp. 97–98),¹ that any espousal of mirror neurons commits one to a holistic (Wray 1998; 2000) rather than a synthetic protolanguage – one that would have to represent “bird flying” with one symbol, rather than two (“bird” and “flying”) as all contemporary languages do (see Bickerton [2003] and especially Tallerman [2004] for discussion). True language is then supposed to develop straightforwardly through the “fractionation” of this protolanguage. Arbib asks us to “imagine that a tribe has two unitary utterances concerning fire which, by chance, contain similar

substrings” (sect. 7, para. 3). But won't similar substrings also occur in unitary utterances that have nothing to do with fire? Here he is on the horns of a dilemma. If he thinks they will not, he has smuggled in a ready-made word, and if all “similar substrings” behave similarly, a holistic stage becomes superfluous – all the separate words of a synthetic language are already present, clumsily disguised. If he thinks they will – and given the limited number of possible syllables even in modern languages, they will probably occur more often in sequences that have nothing to do with fire – why should they be taken as meaning “fire” in the rarer cases, and what will similar strings in other contexts be assumed to mean? And even before this dilemma can be addressed, Arbib must specify what would count as “similar enough” and explain why phonetic or gestural similarities would not be eroded by natural change processes long before hominids could correlate them with similarities of meaning. Moreover, to extract a symbol meaning “fire” from a holistic utterance, our ancestors must first have had the semantic concept of fire, and it becomes wholly unclear why, instead of going the roundabout holistic route, they could not immediately have given that concept a (signed or spoken) label. Real-world objects can be ostensively defined; events and situations cannot.

Two substantive issues lie at the heart of language evolution: how symbolism emerged, and how syntax emerged. No treatment that fails to deal with both can be taken seriously. Indeed, symbolism (as distinct from iconic or indexical reference, distinctions that Arbib nowhere makes) has seemed to some (e.g., Deacon 1997) to be the Rubicon between our species and others. Arbib mentions it several times, hypothesizing it as a “support” for protolanguage and noting the necessity for its “increasing sophistication” as true language emerges. But at no point does he even acknowledge the problem of how such an evolutionary novelty could have developed.

Syntax makes an even better candidate for a human apomorphy, since even with explicit instruction our nearest relatives fail to acquire the rudiments of it (Givon 2004). Arbib's dismissal of syntax as a “historical phenomenon” makes such uniqueness hard to explain. According to him, “Words as we know them then coevolved culturally with syntax through fractionation” (sect. 2, para. 2). Even if syntax meant only the most frequent word-order in simple affirmative sentences, this claim might be tricky to defend. In fact, syntax depends on a wide variety of relationships within complex hierarchical structures. Where do these structures and relationships come from? Arbib, ignoring the half-century of linguistic research that has revealed (if not explained) them, remains silent on this.

Arbib's treatment claims to go “beyond the mirror.” However, what he offers is only a smoke-and-mirrors version of language evolution, one in which all the real issues are obscured. His flow-charts and neurological references may look impressive, but they tell us nothing about the central problems of the field.

NOTE

1. It is surely worth reminding readers that all the features of mirror neurons (except for their catchy title) were described by David Perrett and his associates (Perrett et al. 1982; 1985) more than two decades ago – a fact seldom acknowledged in contemporary accounts, including Arbib's.