no less necessary to early talk than infant design (Spurrett & Cowley 2004). Third, not only may babies lack genetic propensities for word production but persons, not brains, seem to sustain early speech. As neural systems self-organize, infants come to control action and perception in ways that prompt vocally mediated interaction. Generally, then, Falk's argument is weakened by the unsupported claim that word-based protolanguage emerged from a genetic propensity. Other problems also arise. Above all, Falk links infant-directed speech to conventional form-based meanings rather than to interpersonal, affective events. By making prehistoric talk sign-based, protolanguage becomes a matter of producing and recognizing speech acts. However, unless communication draws on interpersonal events, syllabic invariants are likely to be products of an individual's recurrent affective states. In modern infants, this is not what occurs. Rather, words arise from iconic-indexical events that integrate activity between persons and across modalities (Cowley et al. 2004). Finally, Falk's appeal to ontogenetic and phylogenetic parallels is often not persuasive. If, say, phonological and semantic bootstrapping occur in ontogenesis, they rely on producing formally consistent meanings. By definition, however, form-based processes cannot *precede* protolanguage.

Many reject the view that species differences depend on words. Neither Chomsky's recent work (Hauser et al. 2002) nor that based on Wittgenstein invoked genetic propensities to explain verbalizations. Whereas Taylor (1997) and Shanker (2001) posited no inner linguistic mechanisms, Hauser et al. (2002) has hypothesized that "most, if not all" verbal aspects of language use "mechanisms shared with nonhuman animals" (p. 1573). For both sets of theorists, what sets language apart is a human capacity for off-line modification of utterance-activity. Hauser et al. (2002) appealed to a neurally based mechanism for "recursion" and Taylor (2000) emphasized our capacity to talk about talk, or "linguistic reflexivity." Remarkably, both sets of theorists agree that what matters is that, in the course of speaking, we modify what is uttered. It follows therefore that (nonverbal) Ur-language emerged as hominins extended bodily expression. Wittgensteinians and Chomskyans concur that no specialized genetic propensities are needed to sustain simple vocal-production. While disagreeing about how to explain off-line modification, they agree that nonhumans share social mechanisms used in language. In defending a continuity view, Falk addresses the wrong target. The folk mislead us: Even if words are unique, they are not the taproot of language.

Given emphasis on multimodality, Falk's argument can be reframed in terms of the origins of utterance-activity or Hauser et al.'s (2002) "language faculty-broad sense." *Putting the baby down* changed ecology in line with both bipedalism and neonates' enlarging brains. The thesis, then, sustains the view that joint behavior is shaped by mother-infant interaction. In phylogeny, as Wray (1998) argued, this may have used holistic vocal (and, presumably, other) patterns. Like social grooming (Dunbar 1996), utterance-activity may have come to dominate social coregulation. Then, as now, in Fernald's (1993) terms it may have "engaged and persuaded" infants by inducing "subtle changes in emotions and intentions" (p. 80). If so, instead of appealing to ontogenetic and phylogenetic parallels, we can ask how interactional events give rise to cognitive outcomes. With Laland et al. (2000), *putting the baby down* may have led to "choices, activity, and metabolic processes" (p. 132) that influenced natural selection through "niche construction." The newly created niche altered both maternal vigilance and the epigenetic processes that affect how infants attend and respond to multimodal expression. As infants became sensitive to the mother's appraisal of circumstances, there would have been a partial decoupling of expression from affect. Real-time feedback could shape the mother-infant relationship and, by extension, the evolution of development. With Owings and Morton (1998), "assessment" would drive an arms race which ensured that increasingly more differentiated expression was being used to "manage" infants. Utterance-activity began to exploit Ekman (1972) and Fernald's (1993) invariants as well as the micro-temporal dynamics of infant-caregiver play (Bateson 1979; Stern 1977). As joint events became affectively coregulated, vocal power and sensitivity increased. In this view, the ability to use words depends not on genes but on mutual adjusting that is supported by neurodevelopmental change.

Niche construction allows *putting the baby down* to be seen as helping prosody and gesture take on new affective, cognitive, and practical roles. Social learning may have used behavioral ecology to reshape both intrinsic motive formation (see Trevarthen et al. 1999) and perception-action systems (Preston & de Waal 2001). Study of this natural history can throw light on, say, coregulation (Fogel 1993), interactional synchrony (Condon & Sander 1974), emotional contagion (Hatfield et al. 1994), accommodation (Giles et al. 1991), and real-time understanding (Cowley 1998; Gumperz 1996). Reframed in terms of niche construction, Falk's argument can promote new thinking about language. Not only does it allow for skepticism about the role of words in Ur-language, but it prompts us to ask how joint behavior induces *belief* in verbal entities. Beyond that, there lies a harder question: Is consilience possible between seeking the taproot of language in neural capacities for recursion and viewing reflexivity as the product of how infants participate in – and talk about – utterance activity?

Continuity, displaced reference, and deception

Lee Cronk

Department of Anthropology, Rutgers University, New Brunswick, NJ 08901.
lcronk@anthropology.rutgers.edu
http://anthro.rutgers.edu/faculty/cronk.shtml

Abstract: Falk's contribution to a continuity theory of the origins of language would be complemented by an account of the origins of displaced reference, a key characteristic distinguishing human language from animal signaling systems. Because deception is one situation in which nonhumans may use signals in the absence of their referents, deception may have been the starting point for displaced reference.

Falk's interesting and persuasive argument that human language was built, at least in part, upon a substrate of infant-directed communication is framed in terms of the contrast between continuity and discontinuity theories of the origin of language. However, unless we resort to saltationism, a choice between continuity and discontinuity is as false in the study of language origins as it is in any evolutionary scenario. Although examination of the end points of any episode of divergence will create the appearance of discontinuity, gradual change is the only plausible scenario within a Darwinian framework.

This is not to say that evolution's gradual, continuous, and incremental nature means that "differences between human language and nonhuman primate communication are only quantitative" (King 1996, p. 193). Even a gradual process can result in important qualitative differences over time. Human language differs from nonhuman signaling systems in a variety of ways. Falk shows that infant-directed communication is likely to have had a role in bridging that gap, and King (1996) has provided a similarly plausible gradualist account of the origins of syntax. Another key difference between nonhuman signaling systems and human language is displaced reference – that is, the ability to refer to things and to understand references to things that are absent. Unlike humans, nonhumans can use their signaling systems to discuss only things that are currently in evidence: "There is a predator nearby," "Here is a food source," and so on. Although they can signal the presence of, say, a snake, they cannot use that signal as the starting point for a discussion about snakes or as a way to teach their young about the dangers of snakes. They can express their own hunger, but they cannot have a conversation about the problem of hunger while their own bellies are full.

https://doi.org/10.1017/S0140525X04300117 Published online by Cambridge University Press

A gradualist account of the origins of displaced reference might start with the observation that the only circumstance in which nonhumans send signals in the absence of the referent is when they are engaging in deception, such as when birds send false alarm signals in order to frighten competitors from a food source (Munn 1986). Of course, in order for our ancestors to have been able to discuss things not in evidence, the receiver of the signal would have had to be clued into the trick, which would preclude actual deception. Perhaps the line was crossed when two individuals formed a coalition to deceive another, enabling the coalition members to share an understanding that a signal was to be used independent of its referent. Once it was established that a signal could be used without its referent being present, it would have been a relatively short step to real displaced reference, unconnected to deception. Although it is a very long way from coalition-based deception using signals to human language as we now know it, perhaps this was how the transition from an animal signaling system to human language began (see Wray [2002] for more on the evolution of displaced reference). As Knight (1998a; 1998b; see also Knight et al. 1995) has pointed out, such a scenario would require high levels of trust among coalition members. This might have been facilitated by kinship and, in line with Falk's scenario, a signaling system rooted in the trustworthy soil of motherese and its precursors.

Whether displaced reference has its origin in coalitional deception or somewhere else, one thing is certain: Only a continuity theory of the origin of human language can account for this or any other discontinuity between it and nonhuman signaling systems.

ACKNOWLEDGMENTS

I thank William Irons, Beth Leech, and Lars Rodseth for their comments on a draft of this commentary.

Syntax: An evolutionary stepchild

Danielle Dilkes and Steven M. Platek

Department of Psychology, Drexel University, Philadelphia, PA 19102.
Danielle.Dilkes@drexel.edu steven.m.platek@drexel.edu
http://psychology.drexel.edu/platek.htm

Abstract: Dean Falk has strategically explored "mother-infant gestural and vocal interactions . . . in chimpanzees and humans" in order to offer hypotheses "about the evolutionary underpinnings that preceded the first glimmerings of language." Though she offers compelling evidence for many interesting hypotheses as to the epigenesis of language, other possibilities have yet to be explored. Here we explore the role of gestural communication among deaf signers and the neural correlates associated with this type of communication.

In her article *Prelinguistic evolution in early hominins: Whence motherese?*, Dean Falk strategically explores "mother-infant gestural and vocal interactions . . . in chimpanzees and humans" in order to offer hypotheses "about the evolutionary underpinnings that preceded the first glimmerings of language." Though she offers compelling evidence for many interesting hypotheses as to the epigenesis of language, other possibilities have yet to be explored.

One such possibility is whether the structure/syntax of the languages we use today was molded to best fit a preestablished cortical organization for linguistics and the related tasks, and, if so, is this organization modality dependent? Is linguistic structure/syntax a function of the organization of the left-hemisphere? Is language innate; can it be evolutionarily traced? If so, what implications does this have in the ever-present question of the evolution of language?

We know from existing literature and in vivo studies that nonhuman primates communicate using gestures, a type of "signed language," and that humans for the most part communicate using a spoken language. The primary difference between signed and spoken language is that sign relies "on spatial contrasts while speech is linear and non-spatial" (Goldin-Meadow 1999). In verbal communicators, a lesion to the left hemisphere usually produces deficits on linguistic tasks, whereas damage to the right hemisphere usually produces deficits in spatial tasks. Similarly, when human nonverbal communicators sustain damage to the left hemisphere, they perform more poorly on linguistic tasks but do not exhibit the same spatial deficits that signers with right-hemisphere damage do. The implications of these findings are that in humans, sign seems to be processed as linguistic information rather than spatial information; therefore implicating the left hemisphere in linguistics, regardless of transmission (Goldin-Meadow 1999).

When deaf children of nondeaf parents are not taught to sign and have not acquired speech because of their hearing impairment, they independently create a system of gestural communication that takes on a structure similar to that of spoken language and is consistent across cultures (Goldin-Meadow 1999; Goldin-Meadow & Mylander 1998). A possible explanation for why deaf children create linguistically oriented gestures and hearing children do not, may relate to the notion that gesture needs to take on grammatical properties only when it has to carry the full burden of communication. When used in conjunction with speech, gesture does not have to convey (all) meaning; therefore, it does not assume a language-like form (Goldin-Meadow 1999).

A cortical region implicated in nonverbal communication is the superior temporal sulcus (STS). When congenitally deaf signers and hearing expert signers are presented both with sign language and with nonmeaningful gestures, activation of the STS was noted (Allison et al. 2000). Furthermore, while viewing American Sign Language sentences, those who are unfamiliar with the language showed no activation of the STS. These results are indicative of the STS's role in the perception of ASL. Further support of this hypothesis can be seen when studying monkeys. "In monkeys, responsiveness of STS cells was greater to a hand making a movement than to a bar of the same size making the same movement, demonstrating that the cells are preferentially responsive to biological motion" (Allison et al. 2000; Rizzolatti & Arbib1998; 1999; Rizzolatti et al. 1996; 2002). This applies to humans in that the cellular organization of the STS may provide a predisposition for the perception of communicative or meaningful hand gestures, but not for meaningless hand movements.

The cortical response to the observation of action in both human and nonhuman primates is very similar and supports the above findings. In humans, PET studies revealed that the observation of an action, such as grasping, activated the STS, the inferior parietal lobule, and the inferior frontal gyrus (area 45); all activation sites were limited to the left hemisphere (Rizzolatti & Arbib 1998). The activation found in humans parallels that found in nonhuman primates on similar tasks, thereby indicating "that, in primates, there is a fundamental mechanism for action recognition" (Rizzolatti & Arbib 1998). This is very interesting because the stimuli used in these experiments were not tied to linguistics; however, the findings may implicate "that this action-recognition mechanism has been the basis for language development" (Rizzolatti & Arbib 1998).

These findings suggest that the left hemisphere may not simply be responsible for language tasks, but for all linguistic tasks, including the recognition and processing of multiple modalities of communication – one of these modalities being gestural communication, from which it may be possible that language as we now know it has evolved. In Rizzolatti and Arbib (1998), a notion is put forth that the nonhuman primate homolog to the human cortical area known as Broca's is area F5 (the rostral part of the monkey ventral premotor cortex). "The reasons for this view are that both F5 and Broca's area are parts of inferior area 6 and their location within the agranular frontal cortex is similar; and cytoarchitectonically, there are strong similarities between area 44 (the caudal part of Broca's area) and F5" (Rizzolatti & Arbib 1998).

The major difference in conceptualization of these two areas is that Broca's is commonly associated with speech, F5 with hand

https://doi.org/10.1017/S0140525X04300117 Published online by Cambridge University Press