NOTE

1. When chimpanzee infants, however, are nursed by human caretakers and sensitized to face-to-face interaction with humans, they appear able, as Bard (1998) has shown, to imitate human facial models of certain gestures. I have a video record of a chimpanzee infant (age 39 days) engaging in a sort of turn-taking vocal interplay with his foster parent, but I have never observed this in infant-adult chimpanzee interaction. Further, although adult males in the wild may interact relatively little with infants (see target article, Note 1), the captive males I have observed sometimes do. For example, a Beta male is sometimes used as a baby-sitter by one of the mothers when she goes off in search of food (Braten 2000, p. 282). In any event, never have I observed prolonged facial eye-to-eye contact between infants and adults, males or females.

Prosody does not equal language

Robbins Burling

Department of Linguistics, University of Michigan, Ann Arbor, MI 48109.
rburling@umich.edu http://www-personal.umich.edu/~rburling

Abstract: Prosody, in motherese as in all forms of language, has a very different form and a very different use than the central lexical, phonological, and syntactic components of language. Whereas the prosodic aspects of motherese probably derive from primate vocalization, this does not help us to understand how the more distinctive parts of language emerged.

Dean Falk makes a strong argument that human motherese began with affective vocalization and that "*the use of prosody in human maternal speech is similar to the use of vocal signals by some nonhuman primates*" (Fernald 1994; quoted and given emphasis by Falk). Even though, as Falk makes clear, neither chimpanzee nor bonobo mothers engage in much infant-directed vocalization, I am sympathetic with her argument that the prosodic component of human infant-directed (ID) speech shows continuity with primate communication and I find it plausible that ID vocalization could have formed an important bridge between primate and human communication. Among other things, the early development of ID vocalization in phylogeny could help to solve the puzzle of how vocal/auditory language, rather than a manual/visible rival, became dominant. Most primates have much better voluntary control over their hands and arms than over their mouths and tongues. This should have given a decisive head start to a manual language. If voluntary control over the vocal organs had already been achieved with the help of such things as ID vocalization, a vocal language might have been viable from the very start.

Most of Falk's article is concerned with very early forms of parent-infant communication, and I am in general agreement with her discussion. I feel less comfortable with the sections of the paper, starting with 3.2, where Falk seeks to relate motherese to language. If Falk is right, the earliest forms of human ID vocalization had none of the specifically linguistic features that have been so difficult to account for in an evolutionary framework. Present-day motherese makes use of the same words, combinatorial phonology, and heirarchical syntax that we find in other linguistic styles; it is set apart primarily by its characteristic prosodic features. Falk's hypothesized prelinguistic ID vocalization has prosodic features of the sort found in modern human motherese but lacks its lexical, phonological, and syntactic features. To say "Over time, words would have emerged in hominins from prelinguistic melody (Fernald 1994, p. 65) and become conventionalized" (sect. 3.2.1) seems to beg the question. Just *how* would this emergence have come about?

Tone of voice, the ability of the voice to convey such emotions as joy, excitement, and anger, and the soothing tones of motherese are important uses of prosody, and I find it reasonable to see them as emerging from (and still, I believe, belonging to) a primate call system. However, this prosody lacks the system of contrastive phonology that is characteristic of language. As with other kinds of human and animals calls, the referential potential of prosody is more limited than that of words. Prosody is better at conveying emotions, whereas words are better at reference. To be sure, prosody has become deeply entangled with contrastive phonology in modern languages, but they do remain easily distinguishable. Parents have no trouble extracting their infant's first words from the abundant primate vocalizations that they have been listening to since the baby's birth. Some features of that vocalization will forever accompany their child's language in the form of prosody. In other words, prosody has both a different form and different functions than phonology or the lexicon, and it is the new form and functions of language that need to be accounted for if we are ever to understand how it emerged in phylogeny. I continue to think that the best way to understand what happened is to conclude that "tone of voice [along with the other aspects of prosody] amounts to an invasion of language by something that is fundamentally different" (Burling 1993, p. 30). We ignore the most interesting and difficult parts of the puzzle if we take for granted that all of language somehow emerged from prosody.

I seem to have failed to make myself clear in my 1993 article, for Falk is not the first person to conclude that I believe in the sudden emergence of language. In that article, I did express deep skepticism about finding the origins of language in a call system, but such skepticism need not imply that language emerged suddenly. One could believe, and I do believe, that language emerged very gradually from something other than a primate call system. Human cries, laughs, and screams, after all, constitute a fine primate call system – the call system of the human primate – and neither our own calls nor the calls of other primates show the degree of continuity with language that we might expect if language had emerged from a call system. Falk is right that the phonetic aspects of motherese are derived from primate vocalizations. Sadly, this tells us very little about the origin of the most distinctive parts of language: contrastive phonology, syntax, and the lexicon.

Early hominins, utterance-activity, and niche construction

Stephen J. Cowley

Department of Social Sciences and Humanities, University of Bradford, Bradford BD7 1DP, United Kingdom; Psychology, University of KwaZulu-Natal, South Africa. **s.j.cowley@bradford.ac.uk**

Abstract: Falk's argument takes for granted that "protolanguage" used a genetic propensity for producing word-forms. Using developmental evidence, I dispute this assumption and, instead, reframe the argument in terms of behavioral ecology. Viewed as niche-construction, *putting the baby down* can help clarify not only the origins of talk but also the capacity to modify what we are saying as we speak.

Invoking "protolanguage," Falk uses cross-primate comparisons to speculate on how hominins set off toward full-fledged language. *Putting the baby down,* she suggests, prompted words to arise in response to alterations in mother-infant interaction. Use of a comparative method allows due weight to be given to the multimodality of this "utterance-activity." Instead of emphasizing the prosodic, however, Falk's argument stresses conventionalized events. Rejecting this focus on "words" and protolanguage, I use behavioral ecology to reframe the thesis. Stronger arguments arise if caregiver-infant interaction is seen in terms of "niche construction" (Laland et al. 1999).

Taking the folk view that words distinguish us from chimps and bonobos, Falk posits a "genetically driven propensity to produce natural protolanguage" (sect. 3.2.1, para. 7). Did this exist? First, as no other species exploits simple language, words may owe more to brain-culture coevolution than to genes (Deacon 1997). Second, intention attribution is crucial in learning to talk because, without sympathetic others, infant vocalizations make little sense. In Dennett's terms (1987), taking an "intentional stance" may be

https://doi.org/10.1017/S0140525X04290112 Published online by Cambridge University Press

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/S0140525X04290112 Published online by Cambridge University Press