Falk, in addressing the origins of a form of speaking directed at infants that is typically referred to as *motherese,* likewise presents the case that this attribute is unique to humans, based on its apparent absence in the genus *Pan.* However, Falk fails to look beyond *Pan* to other animals. Had she done so, she would have discovered that mother-infant communication is common in mammals (and also present in some birds). This includes vocal behavior that is clearly directed at the infant, which is the criterion she uses to describe and define the characteristics of motherese. Even in human mothers, one form of infant-directed (ID) vocalization is humming. Is this so different from the purring that everyone associates with felines (but is also found in other mammals, such as the squirrel monkey, when an infant is nursing and that is given by both mother and infant)? Even where evidence for ID vocalization by mothers or other caregivers is lacking, mammalian infants universally make vocalizations that are structurally and functionally similar to the crying of human infants (cf. Newman 1985b; 2003a), and are clearly directed at their mother or other caregiver. In some species, evidence exists for mother-infant communication that takes the form of a dialogue. In squirrel monkeys, for example, mothers make vocalizations (mutters, chucks, err chucks, caregiver calls) and infants respond with peeps, oinks, and chucks (cf. Biben 1992; Newman 1985a). Thus, ID vocalization in nonhominin mammals may be the evolutionary precursors of the motherese found in humans. I have argued elsewhere (Newman 2003b) that the brain circuits underlying mother-infant communication arose early in mammalian evolution and have had a conservative evolutionary history. In primates, a region of the rostral midline (limbic) cortex is important for the expression of "crying" sounds (isolation calls) (Newman 2003a; 2003b). Recently, evidence has emerged indicating that this same region is activated in human mothers upon hearing infant cries (Lorberbaum et al. 2002). Thus, as MacLean suggested many years ago (1985), there appears to be brain circuitry dedicated to mother-infant communication that is part of his "thalamo-cingulate" division of the limbic system, and part of the "paleomammalian" component of his "triune brain" model of mammalian brain organization (MacLean 1990).

None of this is to detract from the basic outlines of Falk's proposal that motherese served an important function in hominin females – namely, to quiet infants that had been "parked" so that the mother could forage. The main issue of contention presented here is how ancient the antecedents to present-day motherese really are. Falk argues that motherese arose with our hominin ancestors, originally as ID-affective vocalizations from earlier hominins. As I argue above, ID-affective vocalizations are widespread in mammals, hence their apparent absence in *Pan* is the real mystery. Perhaps further, careful observation of free-ranging chimp or bonobo groups, with the specific focus of identifying ID vocalizations, would reveal that they are more common than might be concluded from studying single mother-infant pairs in captivity. How humans came to use hyperarticulated vowels embedded in words to communicate with preverbal infants is also unknown. The origins of spoken motherese are as clouded with mystery as are the origins of speech. Falk's hypothesis that selection for vocal language occurred after early hominin mothers began engaging in routine affective vocalization toward their infants may not be far off the mark. Many hypotheses of the origins of spoken language have become discredited; one hypothesis which suggests that spoken language arose out of ID vocalizations by hominin mothers, although not formally presented in Falk's target article, has some merit worth further consideration. It seems likely that female hominins, like female mammals in general, had evolved for some functions more advanced brains than their male counterparts. Certainly this would seem likely to include infant care, including ID vocalization. Brain circuitry for ID vocalization would already have been in place due to the likely presence of ID vocalization over a long period of mammalian evolution. Given the importance of the rostral limbic cortex (anterior cingulate gyrus) in speech (cf. Brown 1988), the fact that this same region is involved in mother-infant communication perhaps presents the neural circuitry required to explain the ID incorporation of speech sounds as affective stimuli. The fact that "fatherese" appears to be rare also suggests that the brains of mothers (and other female caregivers) have specializations for ID communication that evolved in females but not in males.

NOTE

The author of this commentary is employed by a government agency and as such this commentary is considered a work of the U.S. government and not subject to copyright within the United States.

Walkie-talkie evolution: Bipedalism and vocal production

Robert R. Provine

Department of Psychology, University of Maryland Baltimore County, Baltimore, MD 21250. **provine@umbc.edu**

Abstract: A converging pattern of evidence from laughter, tickling, and motherese suggests that bipedal locomotion plays a critical and unanticipated role in vocal evolution. Bipedalism frees the thorax of its support role during quadrupedal locomotion, which permits the uncoupling of breathing and striding necessary for the subsequent selection for vocal virtuosity and speech.

Dean Falk reviews several new, converging themes in vocal evolution and offers a synthesis of this material based on motherese and related infant-mother interactions. I will focus on the critical role of bipedality, a secondary but pervasive theme in Falk's presentation, because its contribution is both unappreciated and central to many aspects of primate and human vocal behavior. Bipedalism, as Falk notes, is the definitive trait of hominins. The proposed benefits of walking upright on two legs range from efficiency in locomotion, predator detection, genital display and cooling, to freeing the hands to throw, manipulate objects, make and use tools, gesture, forage, and carry infants or objects. Missing from this list is the essential role that bipedality plays in a completely different realm, uncoupling the acts of breathing and locomotion, setting the stage for the subsequent evolution of human speech. Falk cites research (Provine 2000) that describes the important role that bipedality and resultant breath control plays in the emergence of vocal competence, using laughter as an example. I will provide additional details about the bipedal ("walkie-talkie") theory of speech evolution (Provine 2000), focusing on *why* bipedalism affects vocalization. Essential insights come from the comparative analysis of laughter.

A point of agreement in the often-contentious area of ape language is that chimpanzees cannot learn to produce human speech. Although they may have formidable skills in manual signing and recognizing human speech, they cannot talk. Laughter is an ideal subject for the study of primate vocal evolution because, unlike the case of spoken English, laughter is in the vocal repertoire of both chimpanzees and humans, which permits comparative analyses. In important details, chimpanzee laughter differs from that of humans (Provine 1996). Laughing humans chop an outward breath into a series of short (1/15 second), vowel-like blasts ("ha") that repeat about every 1/5 second ("ha-ha"). The sounds have a marked harmonic structure consisting of a fundamental frequency and its overtones. Chimpanzee laughter, however, is breathy, lacking the vowel-like sounds, clear harmonic structure, and clean onset and termination of the human "ha." Chimpanzee laughter ranges from a low amplitude panting sound ("pant-pant") to more intense and guttural "ah-grunting." Unlike the chopped exhalations of human laughter, chimpanzee utterances have only one sound produced per inward and outward breath. As in speech, humans modulate an outward breath to produce the sounds of laughter. Chimpanzees and probably other great apes lack such vocal control, being more confined in syllable production per exhala-

https://doi.org/10.1017/S0140525X04410115 Published online by Cambridge University Press

tion. Although many investigators from Darwin to the present have reported laughter in great apes, they seldom described the distinct sound of ape laughter, and never noted the nonhuman means of laugh production or its ramifications for speaking.

The characteristics of laughter in chimpanzees point to a critical constraint on the evolution of speech and language in the great apes. Chimpanzees are captive to an inflexible neuromuscular system that is still tied to the ancient and essential labor of breathing. They share this characteristic with other quadrupeds whose breathing and running are closely synchronized (one stride per breath) so as to brace the thorax for forelimb impacts. (Without full lungs and breath-holding at the instant of impact, the thorax would be a weak air-filled bag, and the running animal might collapse headfirst into the dirt.) This respiratory pattern may be more limiting to the evolution of speech than the more publicized structures of the tongue, larynx, and vocal tract. The evolution of bipedalism in human ancestors permitted natural selection for greater flexibility in the coordination of breathing and locomotion because the thorax no longer was required to absorb forelimb impacts during running. A human runner, for example, may employ a variety of patterns. According to Bramble and Currier (1983), the ratio of strides per breath in humans can be 4:1, 3:1, 5:2, 2:1, 3:2, or 1:1, with 2:1 being the most common. Breaking the rigid link between stride and breath allowed the subsequent natural selection for utterances in which individual sounds were no longer tied to single breaths, permitting speech as we know it and, incidentally, the characteristic human laugh.

Although a powerful probe into the mechanisms of vocal production, laughter also provides one of the best cases of how a specific human vocalization evolved. Literally, *laughter is the sound of play.* The rhythmic panting of chimpanzee laughter is the sound of the labored breathing during tickle games between mother and infant or rough-and-tumble play with peers. The panting sounds of chimpanzee laughter later came to symbolize ("ritualize") the labored panting of the physical play even in the absence of physical exertion. The primal "pant-pant" of our chimpanzee cousins evolved into the human "ha-ha." Along the way, the range of stimuli for laughter expanded from the physical contact and tickle of chimpanzees to the more subtle symbolic play of humor. My candidate for the most ancient joke, and the only one that works as well with chimpanzees as human babies, is the "I'm going to get you game" of feigned tickle. Mothers playing this game with their babies are engaging in one of our species most ancient rituals.

Prosody as an intermediary evolutionary stage between a manual communication system and a fully developed language faculty

Andreas Rogalewski[a], Caterina Breitenstein[a], Agnes Floel[b], and Stefan Knecht[a]

[a]*Department of Neurology, University of Muenster, 48129 Muenster, Germany;* [b]*Human Cortical Physiology Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892.* **rogalewski@uni-muenster.de**
caterina.breitenstein@uni-muenster.de **floela@ninds.nih.gov**
knecht@uni-muenster.de
http://neurologie.unimuenster.de/ger/mitarbeiter/rogalewski

Abstract: Based on the motor theory of language, which asserts an evolution from gestures along several stages to today's speech and language, we suggest that speech ontogeny may partly reflect speech phylogeny, in that perception of prosodic contours is an intermediary stage between a manual communication system and a fully developed language faculty.

The precise mechanism of speech evolution is still unclear to a large extent. It has been proposed that Broca's region mediates higher order forelimb movement control in humans, which resembles the neuronal mechanisms subserving communication (Binkofski et al. 2000). These data – together with the finding of "mirror neurons" in area F5 of the monkey's brain (Di Pellegrino et al. 1992; Galaburda et al. 1982; Gallese et al. 1996; Preuss et al. 1996; Rizzolatti et al. 1996) – support the idea that this observation-evoked motor resonance could be the functional expression of a brain mechanism involved in "understanding" actions made by others. Thus, human language could have evolved from the ability to recognize abstract motor behavior relevant for communication (Liberman & Mattingly 1985). Within this action-perception network, communication may have arisen from a mirror system for grasping to a manual-based communication system, and then moved on to articulatory gestures (protospeech) and from there to today's elaborate language system (Corballis 1992; Hewes 1973; Lieberman 1979).

Rizzolatti speculated that communication occurred through oro-facial gestures with a variety of different concepts being communicated (Rizzolatti & Arbib 1998). Oro-facial movements are also used for communication by other primates. Dean Falk describes the role of facial expressions in the mother-infant communication in chimpanzees and humans. The vocalization system evolved into the originally manual- and facial-based communication system. The combination of orofacial communication with vocalization might lead to a higher level of expression. Prosody served to broaden the initially limited spectrum of vocalization by adding another dimension to communication and to compensate for the reduction in sustained mother-infant physical contact. For instance, Falk notes that by around 3 months of age, human infants develop the ability to modulate their cries to express different emotions.

We previously examined whether, and to what extent, language activates the hand motor system. Using transcranial magnetic stimulation (TMS) we showed that speaking, covert reading, and listening to speech bilaterally activate the hand motor system (Floel et al. 2003). This indicates a direct link between the language and the manual/facial action systems, a link far more extensive than previously thought and one that may still be functionally relevant in humans. If language gradually evolved from gestures along several stages, listening to prosodic utterances without semantic content should activate the extended action-perception system. In a subsequent neurophysiological study using TMS, we found that whether subjects are listening to variable prosodic contours without meaning or to sentences with meaning and grammar, the hand motor cortex is bilaterally activated to a similar extent. Furthermore, the pitch variability of the sentences correlates selectively with the size of the MEP amplitudes after right motor cortex stimulation (Rogalewski et al. 2003). In addition, pilot data from our lab using functional magnetic resonance imaging provided converging evidence for this preactivation of the hand motor cortex during perception of prosodic features and normal sentences (Rogalewski et al., submitted). This shows that the prosodic aspect of speech is sufficient to activate the manual gesture system.

Behavioral studies in infants show that children learn the phonology and prosodic intonation contours of their mother tongue during the first year, before they learn the meaning of single words and the rules governing how to combine single words into sentences (Jusczyk & Hohne 1997). For example, 4-day-old neonates and 2-month-old infants discriminate sentences spoken in their native language from sentences in a foreign language. When the prosodic contour is not available, because the stimuli are played backwards, the preference for the mother tongue is no longer observed. Other studies showed that cotton-top tamarin monkeys discriminate between languages just as humans newborns do (Ramus et al. 2000). This may reflect the parallelism of ontogeny and phylogeny.

Our data provide strong evidence for a joint neural system that processes manual as well as prosodic gestures. We observed that both listening to sentences and listening to prosodic contours activate an extended cortical network. Our data cannot prove the

https://doi.org/10.1017/S0140525X04410115 Published online by Cambridge University Press