Baby-talk words (e.g., "mama," "bow-wow," "wee-wee") have a number of characteristics relevant to the question of the nature of first words. They tend to be phonetically simple because most are produced by infants as well as adults. Presumably, simple forms are favored in the first words. Also, many baby-talk words are probably old, a prerequisite for their being potential early words of language. Ferguson (1964) cites evidence of "the persistence of baby talk words for food, drink, and sleep for two thousand years in the Mediterranean area" (p. 104).

In addition, baby-talk words provide two specific phonetic sources of evidence of their relevance to the origin of first words. This evidence comes from Ferguson's (1964) sample of 213 baby-talk words derived from six language communities: Arabic, Comanche, English, Gilyak, Marathi, and Spanish. The words were grouped into four categories: kinship terms (for "mother," "father," and "infant"), words for body parts and functions, "qualities" (descriptive words), and words for animals and games. The first evidence is that all the words in Ferguson's corpus for female parent contained a nasal sound, whereas none of the words for male parent did. This not only supports Falk's suggested linkage between nasal sounds and the word for female parent but also suggests the development of something that might be expected in a true linguistic system. There is a clear phonetic *contrast* between the words for female parent and the words for male parent: the former are linked with the nasal airway (one of the two basic speech airways) and the latter are linked with the oral airway.

The second source of evidence comes from our analysis of the overall phonetic structure of 80 consonant-vowel-consonant-vowel (CVCV) words in the Ferguson baby-talk corpus. We show that they fit the pattern suggested for the first words in the Frame/Content theory of the evolution of speech (MacNeilage l998). In particular, the baby-talk words show three CV co-occurrence patterns that have now been repeatedly found in babbling and first-word corpora of infants (Davis & MacNeilage 2000) and are also widespread in languages (MacNeilage et al. 2000). They are coronal (tongue-front) consonants with front vowels; dorsal (tongue-back) consonants with back vowels; and labial (lip) consonants with central vowels. These patterns suggest biomechanical constraints against tongue movement from consonant to vowel that were probably present in the first attempts at words. In the baby-talk corpus, we obtained the following observed-to-expected ratios of the three patterns: coronal-front, 1.30; dorsal-back, 2.20; labial-central, 1.33. (Ratios above 1.0 indicate above-chance occurrence.)

In addition, as in the babbling and first words of infants, we also found three similar patterns in vowels and *following* consonants (VC) in Ferguson's baby-talk words (1.50, 1.64, and 2.38, respectively), consistent with the fact that first and second syllables in all these samples tended to be reduplicated (repeated), as we assume they tended to be in the earliest words. In contrast, in modern languages, there are no such VC co-occurrences, which is consistent with the fact that successive syllables in evolving languages tend to be different, presumably associated with the evolution of the syllable as a separate functional entity, in the service of increases in the size of the message set.

One could contend that baby-talk words are not real words, and therefore could not have been first words. However, it is likely that parental terms now considered legitimate may have typically derived from baby-talk words, as suggested for English by Jakobson (1960). He noted that the intimate, emotional, childishly tinged words of baby talk coexist today with more general and abstract, exclusively adult parental terms. Thus, in English, various forms of "mama" and "papa" or "dada" differ in use from the higher terms "mother" and "father." Jakobson then suggested that "in Indo-European, the intellectualized parental designations *mater* and *pater* were built from the nursery forms with the help of the suffix *-ter* used for various kin terms"(1960, p. 60). Then, as he described it, a sound change occurred in *pater* that made the initial stop [p] and the intervocalic stop [t] into fricatives (hence "father"), while another such change had the same effect on the [t] of "mater" ("mother").

In addition, there is direct evidence that the use of the nasal airway for maternal terms and the oral airway for paternal terms is present today in languages proper, as would be expected if they were derived from baby-talk terms. In a study of kinship terms in 474 languages, Murdock (1959) found that 78% of the words for "mother" begin with a nasal consonant and 66% of the words for "father" begin with an oral consonant. He also presented data on vowel frequencies, using a slightly different categorization of vowels than our own. Using these data, we found that the first syllables of Murdock's words showed three patterns of consonant-vowel co-occurrence constraints similar to the ones we have found (coronal, high front 1.34; dorsal, high back, 2.27; labial, low vowel, 1.12). (Most low vowels in language are central [Maddieson l984].) Murdock did not present data on the second syllables of these words, but we predict that, given their probable origin in first words, these terms, though now in languages proper, are more likely to involve syllable reduplication than are other typical words of modern languages.

How exactly were the phonetic forms originally linked with the concepts of male parent and female parent? Falk suggests, with respect to the maternal term, that the infant's nasal sound put a name to the face of the caregiver. However, it was necessary for the *mother* to make the connection (i.e., "this stands for me") and to use the result of this pairing in communication with others in order for the nasal form to function as a true word. Then, perhaps by default, the infant's *oral* vocalizations were deemed to signal "male parent." Conscious reflections on these pairings might eventually have led to the "naming insight" – the insight that things could be given names (McShane l979).

ACKNOWLEDGMENT

This paper was prepared with support from research grant No. HD 27733–10 from the Public Health Service.

Affective prosody: Whence motherese

Marilee Monnot[a], Robert Foley[b], and Elliott Ross[a]

[a]*Department of Neurology, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104;* [b]*Leverhulme Centre for Human Evolutionary Studies, University of Cambridge, Cambridge CB2 3DZ, United Kingdom.* **marilee-monnot@ouhsc.edu r.foley@human.evol.cam.ac.uk elliott-ross@ouhsc.edu http://www.human-evol.cam.ac.uk/**

Abstract: Motherese is a form of affective prosody injected automatically into speech during caregiving solicitude. Affective prosody is the aspect of language that conveys emotion by changes in tone, rhythm, and emphasis during speech. It is a neocortical function that allows graded, highly varied vocal emotional expression. Other mammals have only rigid, species-specific, limbic vocalizations. Thus, encephalization with corticalization is necessary for the evolution of progressively complex vocal emotional displays.

Falk's contribution to the discussion of the evolution of language is a complex review of communication and infant-rearing characteristics in which the continuity hypothesis is invoked. It is clear that motherese speech is an important element of modern human communication and must have its own evolutionary origin and a connection to language. Falk's model raises important issues addressed in this commentary (mechanisms/evolution of modern motherese; contemporary infant parking, neurology of affective prosody, and species-specific emotion displays in relation to evolution of language).

Motherese characteristics may be more important to a discussion of the evolution of language than Falk's emphasis on chimpanzee/primate rearing behavior. Motherese, or infant-directed speech (IDS), compared to adult-directed speech (ADS), is characterized by simpler phrases, longer pauses, more repetition, wider tone range, and higher overall tone when addressing an infant. Higher overall tone and wider tone range are speech char-

https://doi.org/10.1017/S0140525X04390114 Published online by Cambridge University Press

acteristics of happiness, and are the most salient cues in IDS (Frick 1985; Scherer 1974; Scherer & Oshinsky 1977). These changes are *injected into speech automatically and unconsciously.* IDS is thought to be a universal human trait (Fernald 1992), and has been documented in diverse language groups including click, tonal (e.g., Mandarin), nontonal (e.g., English), and even sign languages (Ferguson 1964; 1977; Fernald 1991; Fernald & Mazzie 1991; Fernald et al. 1989; Masataka 1992; Papousek & Papousek 1992). A nonhuman primate homologue of motherese has been found in which squirrel monkey adults addressed infants with an acoustically different utterance (Biben et al. 1988). Infants prefer IDS in any language (Cooper & Aslin 1989; Colombo et al. 1995; Cooper et al. 1997). *Premature* infants have been shown to respond to IDS with calm awareness, in contrast to other forms of stimulation, such as stroking, which overwhelm their immature nervous systems (Eckerman et al. 1995). Infants under 6 months of age have been shown to grow more rapidly if their primary caregiver speaks high quality/quantity IDS (Monnot 1999). Automatic prosodic changes, when addressing anyone who elicits caregiving concern, reveal the speaker's emotion, so IDS appears to be a form of affective prosody.

The use of motherese to explain the transition from emotional vocalizations common to all extant primates to human language is plausible but complicated, given the neurology of paralinguistic attributes. Spoken language conveys both semantic messages and emotion information. Affective prosody (AP), a paralinguistic feature of language with acoustic features such as pitch/tone, intonation patterns, stress, timing, rhythm, and differential pausing, gives the listener information about the speaker's emotional and attitudinal states (Bolinger 1980; Crystal 1975; Kent & Read 1992; Monrad-Krohn 1963; Ross 2000). Modulation of both comprehension and expression of affective prosody appears to be a dominant function of the *right* hemisphere (Blonder et al. 1991; Bowers et al. 1993; Meyer et al. 2002; Mitchell et al. 2003; Ross 1997). Right hemisphere lesions produce aprosodias (syndromatic deficits in comprehension or production of affective prosody, or both) that may be accompanied by severe communication and social/emotional difficulties (Ross 1982; 1997; Ross & Mesulam 1979) based on the patient's inability to impart affect into speech and/or to understand affect in the speech of others (Carton et al. 1999). If the linguistic message is at odds with the AP content, the affective prosodic message usually takes precedence (Ackerman 1983; Bolinger 1972). Honest signals are hypothesized to be very important to human communication, especially because words and syntax can convey other messages; human survival and reproduction have been hypothesized to be dependent on vocal communication in general and on language specifically (Foley 1995).

Falk's review of modern motherese infers that the automatic modulation of vocalizations when addressing infants is very similar to the automatic changes in voice that occur when a vocalizer of *any species* experiences emotion such as fear or joy. However, affective prosody is a right hemisphere, *neocortical,* function that allows graded and highly varied emotional expressions that are under considerable conscious control. Other mammals have only rigid, automatic species-specific emotion displays that are predominantly *limbic* in origin (Fulton 1941; Harlow 1936; Jerison 1985). Progressively complex emotion displays may have become true prosodic utterances but this required encephalization, with corticalization allowing the change from primitive rigid responses to complex variable responses specific to each species (Jerison 1985; Jurgens 1979).

Because chimpanzee/bonobo vocalizations to infants are both *rare and simplified* compared to humans, Falk insists that some mechanism must be found to explain the transition to IDS during hominin evolution (hominin parents put undeveloped altricial hominin infants down while foraging, necessitating a new form of reassurance to the infant). "Occam's razor" caution must be applied here. Modern apes rarely put infants down, and the infants themselves resist this vigorously and vociferously.[1] Were hominin infants, subject to the same pressures of natural selection as their mothers, really as inadequate as Falk suggests? Perhaps these infants became cognitively precocious before birth so that they elicited more communication cues from mother to compensate for physical prematurity. If modern infants still retain a strong grasping reflex, how long ago did they really lose the ability to cling to mother's fur/hair? And when did hominins lose the body hair so important to infant grasp and thus survival? Clothing may be a recent innovation, implying that abundant fur/hair was present until quite recently (Kittler et al. 2003).

A more parsimonious explanation is that the bipedal mother foraged with her infant clinging to her back as she strode upright, necessitating complex enticing vocalizations to the infant when the dyad was not face-to-face. Because motherese speech is the most effective way to stimulate *premature* neonates to "calm awareness" (Eckerman et al. 1995) and to encourage faster infant growth (Monnot 1998; 1999), stimulating emotional utterances to a hominin child resulted simply from the necessity to ensure *infant growth and survival* (Monnot 1999). Emotional/prosodic vocalizations certainly seem to have laid the foundation for propositional language (Ploog 1992), but infant "parking" seems to be a less-than-efficient explanation for the evolution and encephalization of IDS, affective prosody, and human language.

ACKNOWLEDGMENT

Monnot is indebted to Janette Wallis, Ph.D., for many stimulating and enlightening discussions about primate behavior and vocalizations.

NOTE

1. Personal communication from Janette Wallis, Ph.D., based on years of observing chimpanzees at Gombe Stream and other chimpanzee settings (wild and captive populations).

Motherese by any other name: Mother-infant communication in non-hominin mammals

John D. Newman

Laboratory of Comparative Ethology, National Institute of Clinical Health and Human Development, National Institutes of Health, Department of Health and Human Services, Poolesville, MD 20837. **jdnewman@helix.nih.gov**
http://gpp.nih.gov/researchers/viewbook/Newman_John.html

Abstract: The definition of motherese is extended to infant-directed vocalizations in non-hominin mammals. In many species, vocal interactions between mothers and their infants are common. The neural substrates mediating these interactions include the rostral limbic cortex of the frontal lobe. Spoken language may have arisen from hominin females vocalizing to their infants.

Because behavior doesn't leave a fossil record, statements regarding the evolution of behavior are inevitably inferential and based largely on comparisons between related groups of living species. In the case of humans, this is particularly difficult as we have no close living ancestors. Behavioral adaptations thought to be specific to the tribe Homini (*Australopithecus, Ardipithecus, Paranthropus,* and *Homo*) within the Homininae are generally assumed to have arisen after the time period during which the Homini shared a common ancestor with the genus *Pan* (chimpanzee and bonobo), around 5 to 7 million years ago. We can only speculate at which point in the evolution of the Homini the various behavioral attributes that are exclusively found in *Homo sapiens,* our species, actually arose. Speech and language, as typically defined, are unique to humans. Because they do not occur in the genus *Pan,* it may reasonably be concluded that these attributes arose within the Homini. That speech and language are unique to humans, among existing animals, can be assumed based on numerous studies of nonhuman primates and other animals. (Whether the antecedents of speech exist in other animals is a separate issue, for which there is compelling, although controversial, evidence.)

https://doi.org/10.1017/S0140525X04390114 Published online by Cambridge University Press