

## In the beginning was the song: The complex multimodal timing of mother-infant musical interaction

Elena Longhi and Annette Karmiloff-Smith

*Neurocognitive Development Unit, Institute of Child Health, University College London, London WC1N 1EH, United Kingdom.*

**E.Longhi@ich.ucl.ac.uk**    **A.Karmiloff-Smith@ich.ucl.ac.uk**

**[http://www.ich.ucl.ac.uk/ich/html/academicunits/neurocog\\_dev/n\\_d\\_unit.html](http://www.ich.ucl.ac.uk/ich/html/academicunits/neurocog_dev/n_d_unit.html)**

**Abstract:** In this commentary we raise three issues: (1) Is it motherese or song that sets the stage for very early mother-infant interaction? (2) Does the infant play a pivotal role in the complex temporal structure of social interaction? (3) Is the vocal channel primordial or do other modalities play an equally important role in social interaction?

In her target article, Dean Falk focuses predominantly on one side of mother-infant interaction. She eloquently illustrates the contribution that mothers make to the development of communication and speculates in fascinating ways about the evolutionary roots of motherese, or infant-directed speech. In this commentary, we raise three issues: (1) Is it language or an even more fundamental

form of communication, song, that sets the stage for very early mother-infant interaction? (2) Does the infant also play a pivotal role in the complex temporal structure of social interaction? (3) Is the vocal channel primordial or do other modalities play an equally important role in social interaction?

First, we consider whether motherese or song might lie at the basis of very early infant-mother interaction. We argue that one crucial difference between motherese and song is that motherese tends to stress meaning alongside social interaction, whereas the actual semantic content of songs is often completely irrelevant. What is crucial about songs is the rhythmic and segmental characteristics of the vocal message, and this may make it primordial over early linguistic interaction. Indeed, in a longitudinal, microdevelopmental study of infant-mother dyads at 3 and 7 months of age, Longhi (2003) made an in-depth analysis of the temporal structure of the songs mothers sing to their infants as well as of the temporal structure of the infants' responses. Her analysis included several modalities – vocal, visual, kinaesthetic, and tactile – and focused on the spontaneous segmentation of songs into hierarchically organized units. She showed not only that songs are used to regulate the infant's emotions, but, critically, that they also serve a structural purpose in helping the infant anticipate the segmental units of the mother's songs. While songs bear a clear-cut relationship to language due to their hierarchical structure, the more regular musical syntax of songs makes it possible for simpler segmental units to emerge: a song can be broken down into units of three or four notes, phrases, and larger units, and mothers convey these units to their infants. The fact that songs are very repetitive also plays a role.

Another important aspect of musical interaction turns out to be tempo, which again differentiates song from language. Longhi found that mothers vary the tempo of their songs according to the infant's behavioral state. In particular, they sing at a fast allegro tempo for attentional purposes – that is, to attract the infant's attention. On the other hand, mothers sing at a slower andante tempo for interactional purposes – that is, to maintain the infant's attention. Tempo also helps to create regularities that are crucial in the interaction. When analyzing the duration of the phrases of the song, Longhi (2003) found that the mothers also differentiate the two tempos in terms of which phrase is longer: for allegro, the second phrase of the song is significantly longer, whereas for andante it is the fourth phrase of the song that is significantly longer. Thus, infants can use these different tempos to anticipate the structure of the musical interaction. Mothers also stress the duration of upbeats significantly longer than downbeats, creating a gap effect, and thereby marking the segmental units of the song. Therefore, it could be that, alongside motherese, song plays a far more important role than hitherto realized. This might be worth considering from an evolutionary point of view, as well. Prior to the use of lexical terms, it is possible that mothers used structured melodic outputs to reassure their infants that they were close by – any time they had to put them down during foraging.

It is becoming increasingly clear that infants do not listen passively to their mother's songs. Rather, they actively participate in the synchronous timing of movements, thereby anticipating the song's segmental units. Thus, the temporal structure of mother-infant interaction via song does not only emanate from the mother; infants also play a pivotal role. Infants react not only by emotional responses such as smiling and cooing, but they also coordinate their movements with the temporal structure of their mother's song. Already at age 3 months, they seem to have a mental representation of the musical interaction, displaying synchronous behaviors with the stressed segments of the song significantly more often than with the nonstressed segments (Longhi 2003). However, there are also developmental differences. At the age of 3 months, infants tend to respond by clustering their behaviors around the main parts of the song (the beginning, middle, and end), whereas by 7 months they organize their synchronous behaviors more smoothly in time with the beat across the different phrases of the song. Moreover, with time, infants learn to antici-

pate the stressed elements, thereby increasingly demonstrating their sensitivity to the hierarchical structure of songs. It is possible that this helps them in their efforts to structure linguistic interaction also.

The importance of the mother's use of multimodal channels has been stressed with respect to speech (e.g., Gogate et al. 2000; Papousek & Papousek 1981), but the use of multiple modalities during singing is even more striking (Longhi 2003). Longhi carried out a musical analysis of the mother's beat, a frame-by-frame behavioral analysis of the mother's and infant's behaviors, and a musical plus behavioral analysis examining the synchronization of the partners' behaviors with the musical beat. She found that when mothers are singing, they mark boundaries between phrases while simultaneously shaking their heads, rocking their bodies, and patting their infant's bodies in an intricately timed fashion, to which their babies respond with similarly tightly synchronized movements. More interesting are the developmental changes that mothers introduce when singing to their infants. Longhi's study showed that when infants are around 3 months of age, their mothers use a great deal of synchronous physical contact with their infant while singing. By the time infants are 7 months of age, mothers produce more synchronous actions in contact with toys and other external objects while singing, significantly reducing those in direct contact with their infant. This highlights a change in the terms of the interaction from dyadic to prelinguistic triadic interaction through song.

In sum, it is worth considering the possibility that, with its simpler, regular structure, musical interaction in general, and song in particular, may have played a greater role in evolution and ontogeny than Falk's target article would suggest.