Daniela Lenti Boero

Les Echos - Society for alpine culture and environment research
Association pour l'étude des cultures et des environnements alpins
Associazione per lo studio delle culture e degli ambienti alpini
Università della Valle d'Aosta
d.lentiboero@icloud.com

Sounds of protolanguages:
Some preliminary insights
from developmental psychology

Abstract. In present paper I examine the transition from crying to word in human infants and propose that it might be a treasure acoustic box helping us to conceive how our ancestor's sounds might have been, and to outline breakpoints that might help us in making hypotheses about possible subsequent phases of language evolution.

Keywords: protolanguage, onomatopoeia, animal communication, speech, protophones, vocalization, origin of language

1. Introduction

We know that language in *Homo sapiens* is unique among mammals, due to: a) peculiar vocal apparatus allowing the existence of an articulatory space more similar to singing birds' than to mammals' (Marler 1970; Bolhuis *et al.* 2010), b) specialized brain areas allowing acoustic analysis of input speech sounds and nervous motor control of output speech sounds (Broca and Wernicke), c) a specialized cognitive equipment allowing the mapping of any acoustic label to objects in the world (Lenti Boero and Bottoni 2006). Although language includes syntax and grammar, I believe that words (and their referents) are the ground of linguistic communication, thus in present

paper I will adopt an ontogenetic view addressing the features described in points a), b), c). Those features develop in the first year of human life, when the human infant proceeds "from cry to words", i.e. from an early species-specific communication (see below), analogue to other mammalian infants (Peters 1980; Lingle 2012), towards a semantic holophrastic communication introducing syntactic discourse organization (Kagan 1984). It is acknowledged that evolutionary studies may profit from developmental ones and vice versa (Piaget 1972; Butterworth *et al.* 1985): even if we should be cautious in not adopting a recapitulationistic point of view, language evolution theories might very well profit from ontogenetic studies, especially from those aimed at identifying constraints common to both development and evolution (Costall 1985; MacNeilage 2008). Phylogeny is not a sequence of adults but a sequence of entire ontogenies (Striedter 2005), thus including the immature forms. For this reason, the path of the infant child toward language might be a treasure box because in their first year of life human infants go through a sequence of subsequent anatomical and neurophysiological changes ending in words. Those changes should be considered as steps travelled along the hominid phylogenetic line from the utterance of relatively simple sounds, as our nearest living species (great apes) do, to the complex articulate sounds that compose language (Mithen 2005; Tattersall 1995).

The aim of present paper is to propose that the analysis of the human infant acoustic outputs might help us in defining a logical sequence of sound making that hominids underwent from one species to another. By logical I mean the fact that more simple anatomical features and nervous motor controls might have themselves possessed simple communication systems, and generated the more complex ones. I will try to sketch the communicative boundaries of those systems, and to sketch the possible selective pressures that acted upon those "improvements" from one stage to the next. My models will be animal and early human communication. In our ancestors' path toward language natural, parental, and/or sexual selection might have acted in order to maintain traits relative to point 1 and 2. In the paragraphs to follow, I will try to hypothesize which selective forces might have been at work.

2. Acoustic output of the human infant: crying, laughter and low intensity sounds

Crying is the first sound uttered by a human being. It is an alarm signal conveying the information that homeostatic imbalance is endangered (Lenti

Boero *et al.* 1998). Crying has striking characteristics: loudness that can be perceived at least at about 100 m around the emitter – but can also be intensity modulated, and long lasting duration: up to ten minutes (Lenti Boero, personal observation). According to Newmann (1985), it has no homologous in infant primates. Crying can communicate individuality (Cismaresco *et al.* 1990; Rocca and Lenti Boero 2005), and possibly sex of the caller (Rocca and Lenti Boero 2005). In addition, the infant cry communicates urgency to a recipient (Lenti Boero *et al.* 2008) and we know that contemporary mothers perceive their own 4/8 months old infant's cry as "intentional" according to background variables, maternal emotional adjustment, and maternal interactive style (Feldman and Reznick 1996).

Some acoustic content of the cry can be extracted and paralleled to speech sounds, for instance, Italian vocalic sounds "a" and "e" and nasal sounds were heard by Italian musical trained listeners of cries uttered by less than one week old children (Lenti Boero *et al.* 2008). Interestingly, in middle age, the same prevalent vowels "a" and "ε/e" in newborns' cry were reported by the Florentine poet Antonio Pucci (1310/1388), though we might not agree with his interpretation: "infant males cry emission include the "a" in memory of Father Adam, while infant females include "ε/e" in memory of Mother Eve, thus carrying on those cry all world's troubles" (Giallongo 1997). Nervous motor controls for cry emission are the same as the ones reported in monkey models (Jurgens 1990, 1992; Jurgens and Ploog 1988; Lenti Boero 2009; Lester and Boukydis 1992), and the infant cry must be considered thoroughly an "animal" signal.

Beyond cry, infants utter vegetative sounds, such as cough and sobs that might have some non-intentional communication value poorly explored (Papoushek and Papoushek 1981), but possibly informing the care-giver about the infant's physiological state. In addition, in positive situational contexts (Papoushek and Papoushek 1981), infants produce low intensity sounds with distinctive timbre, intonation and F_0 contour from three days of age; and, though we have a poor knowledge about the ontogeny and development of laughter, preliminary findings show that a fully fledged laughter may be observed at three and half month (Decurti and Lenti Boero, personal observation). A key characteristic of early laughter is that it is a low intensity sound that can reach few meters around the emitter. Thus a human infant is endowed since her/his early months of life with a basic communication system signaling alarm, well being and basic physiological yet not very alarming states, the only long distant traveling signal being the cry.

We might thus think of a simple communication system based on the above described early infant sounds put under volitional control, as is the

case for some animal species (Seyfarth and Cheney 1980; Lenti Boero 1992). In a hominid social group endowed with the above described communication competencies, an individual could communicate alarm and urgency to group mates from a distance, or from a vantage point not reachable to predators, such as a tree, this communication might have included basic referentiality, as is the case in many living mammalian species: alpine marmots, (Lenti Boero 1992), dwarf mongoose (Rasa, 1986), vervet and Diana monkeys (Seyfarth and Cheney 1980; Zuberbühler 2005) among others. Eventually, low intensity sounds (Oller 2000) and laughter might have been used as a kind of acoustic grooming, *sensu* Dunbar (1993).

3. The human infant as sound analyzer

We know that during early development infant perception of surrounding sounds, including language, is much more advanced than motor competence (Lenti Boero 2014): infants discriminate language phonemes, sharing this capacity with many animal species: rhesus macaques, dogs, chinchilla, quails, and parrots (Adams *et al.* 1987; Bottoni *et al.* 2003, 2009; Dewson 1964; Kluender *et al.* 1987; Kuhl and Miller 1975; Miller 1977; Morse and Snowdon 1975; Pepperberg 2007). This is a key point: why both human infants and many animal species are competent in phonemic discrimination? I believe that this ability is a subcomponent of a more general competence in acoustic spectra analysis, and that it must be widespread at least among vertebrates. Animals, similarly to us, must be updated with the auditory scene, they must distinguish between abiotic sounds from the environment, and biotic sounds uttered by preys or dangerous predators (according to the view-point), between familiar and unfamiliar sounds, and this competence is only achieved by analyzing acoustic spectra (Bottoni *et al.* 2003).

Infants recognize the melodic contour of maternal language, and detect substitution of musical notes (Cooper and Aslin 1989; Mehler *et al.* 1988; Tervaniemi and Huotilainen 2003), they are also able to connect input and output sounds: infants shape their cries' melodic contours on their native language (Mampe *et al.* 2009). In addition, at three months of age, infants are able to imitate musical pitch (Kessen *et al.* 1979).

What does this means for language evolution? I believe that the competences outlined above, and common to other vertebrates, are fundamental in language evolution, that should be considered the product of coevolution between receiver and emitter, analogously as other communication devices (Shannon and Weaver 1949). This is evident in the

specialized adult language brain areas: Wernicke's and Broca's, and in the way infant language develops further on.

4. The ontogeny of musical babbling and of articulated sounds

A key breakpoint in the path towards language is the shift from relatively simple sounds as described above to the articulated ones. This phase starts approximately by the time of 2 months when infants start producing very low intensity quasi-musical sounds or protophones, cooing or gooing, those sounds are mostly, but not only, vowel-like, having more complex melodic contours paralleling vocal emotional expression as the infant grows (Hsu and Fogel 2001). In addition, the infant improves her/his control of the intensity of the acoustic output, with high volume correlating with high excitement both positive or negative (Decurti and Lenti Boero, personal observation) (Papoushek and Papoushek 1981; Oller 2000; Ruzza *et al*. 2003). I defined those protophones as musibabbling (Lenti Boero 2009).

Then, from the third month on, sequences of vowels (voiced sounds) are progressively accompanied by protoconsonant sounds, progressively articulated in the retropharynx, in the dental/alveolar, velar and glottal space and bilabial (Oller and Eilers 1992; Roug *et al*. 1989). This phase might have an internal positive reinforcement, because we know that infants spend much time in exploring their new possibilities, and that this practice gives great pleasure to them (de Boysson-Bardie 2001; Papoushek and Papoushek 1981). Those steps are allowed by a better control of the respiratory cycle, which lengthens the exhaling phase, in favor of vocal emission (Papoushek and Papoushek 1981); and implies a more developed nervous motor control of the entire vocal apparatus (mouth, lips, nose, throat) than the one needed for crying: the acoustic energy can be canalized through the nose producing nasal sounds, or kept lower in the throat, thus producing harsh sounds, and lips can be used as sound makers independently from the mouth (Lenti Boero personal observation).

According to MacNeilage (1998, 2008), this early articulatory component derives from an original precursor in early mammals related to lower jaw oscillation in the purpose of ingestive movements (chewing, sucking, licking). Davis and MacNeilage (2002) notice that the utterance of consonant and vowels requires incompatible mouth movements: depression of the mandible (mouth opening) is required for vowels, while lower jaw elevation is required for consonants. The novelty in this aspect is the total contrast of the two functions: ingestion requires the closing of the air flow and consequently

of the vocal apparatus (if food enters the respiratory apparatus death might occur), oppositely sound making requires an open vocal apparatus enabling air to enter in order to fuel the acoustic energy necessary for vocalize.

We know that those sounds are very much appreciated by parents and are preferred in front of cry (Lenti Boero and Bottoni 2009), probably because they lack the aversive component of cry (Barr 2004; Frodi 1985; Frodi and Senchack 1990; Levisky and Cooper 2000; Zeskind and Lester 1978; Seifritz *et al.* 2003).

4.1. Selective pressures

The above could point to the selective pressures that might have been at work in favoring articulated sounds in front of the fixed relatively simple sounds described above. In fact, evolution is not directive, and natural selection "should operate at all stages of development" (Hogan 1988): this is particularly true for language (Locke and Bogin 2006). Indeed, those selective pressures favored the exaptation of the ingestion apparatus toward the emission of articulated sounds. Locke (2006) was the first to consider parental selection as an agent for preference for early articulated sounds over cry due to the stressfulness of the infant cry. The "Intrinsic Musicality Hypothesis" (Lenti Boero and Bottoni 2008) could confirm Locke's proposal. This concept refers to the predisposition, shaped along hominid evolution, of our auditory system to generate conscious, "aesthetic" and/or emotional responses to all heard sounds (abiotic, biotic including human such as cry). Apparently, the aperiodicity of some segments of the infant cry, and its loudness, are aversive to listeners. In a pilot experiment Lenti Boero *et al.* (2009) showed that the aesthetic quality of musical protophones is rated higher than cry, and produces less stress and anxiety.

In adults, the musical protophone competence might have conferred higher sexual attraction to the carriers. Think of an individual endowed with a vocal apparatus allowing musical protophones, cooing and gooing sounds under volitional control. Because those sounds are nice, and have the possibility to be modulated in their fundamental frequency, they might be used in courtship and preferred over other less modulated signals. Darwin (1871) was the first to propose that musical ability might have been selected by sexual selection: "it is probable that the progenitors of man, either the males or females, or both sexes, before acquiring the power of expressing mutual love in articulate language, endeavoured to charm each other with musical notes and rhythm". Mithen (2006, 2009) proposes that until the appearance of *Homo sapiens*, hominids utilized a musical, modulated, holophrastic way

of communication. Beyond musicality, more complex melodic contours might have been chosen as advertisers of good motor control, an important trait conferring advantage to individuals in the challenges of everyday savannah life and thus indicator of "good genes".

4.2. Animal models

Vallet *et al.* (1997) propose an interesting model in canaries, where the ability to control singing is sexually selected as an advertiser of overall motor ability. As regards as sound imitation, we know that mammals are not able to imitate surrounding sounds; however, many birds are able to imitate surrounding sounds, including the Grey Parrot (*Psytthacus erithacus*) that was proposed as model for both music and language evolution (Bottoni *et al.* 2003, 2009; Pepperberg 2007).

To what extent do protophone articulated sounds allow a communication system? Mithen (2009) suggests that early communication might have included both emotional as well as mimetic (imitation from environmental sounds) signals. I agree with this position: sound imitation is possible with a good control of articulated sound emission, as early sound making shows, and a competence for acoustic spectra analysis (that the human infant demonstrate very early). By means of modulation of the vocal portion of sound emission our ancestors might have been able to introduce a subjective component by adding emotional information to the imitation of surrounding sounds.

The ability for sound mimicry might have been a great help in hunting: this might be confirmed by the hunting strategies of present-day hunters-gatherers and modern western hunters, who still use sound mimics – artificial or natural – in order to lure small prey in their surrounding for kills. Hunting success might have been an additional selective advantage because it enhances individual fitness and attracts mates.

But sound mimicry might have added another important advantage to individuals in groups: for group hunters, as in hominids (Tattersall 1995), it might be useful to share information with group mates about presence of prey or of an incoming danger. Indeed, sound mimicry has this double aspect: it lures the prey, but can also communicate its presence to fellow humans. This property could have raised the communication "entropy" in the social group by adding acoustical referential signs coupled with vocal emotional information to the gestural ones, thus freeing the emitter from being in the visual field of the audience, and might have been primarily shared within the kin hunting group (kin selection).

Is sound mimicry a full communication system? Emphatically not. Abiotic entities, such as stones and places in the environment are silent, a forest sounds only when moved by the wind, a stream has a sound, but a river or a lake might not, animals produce different sounds, and most of all, different actions, which are difficult or impossible to mimic by sounds. However, this might have been the heritage of earlier hominids to *Homo sapiens*.

5. Further development in infant communication

From 6 months of age infants enter the so-called *canonical stage* (de Boysson Bardie 2001; Oller 2000). Canonical syllables are composed by a "nucleus" of acoustic energy (a vowel), and at least a "margin" (a consonant), together those sounds last 100–550 milliseconds, the syllabic period guaranteeing the distinct perception of the syllable (Oller and Eilers 1992; Oller 2000). For the canonical stage to appear a crucial maturation is needed: the time for nucleus and margin emission are under nervous motor control for time length, in addition canonical stage is tied with acoustic automonitoring: deaf infants do not enter the canonical stage until the tenth month of life with rare exceptions (Oller and Eilers 1988). The canonical syllables phase is paralleled by many neuroanatomical changes that imaging studies have recently revealed (Matzusawa *et al.* 2001; de Graaf-Peters and Hadders-Algra 2005). In particular MEG results show that at sixth month of age there is an activation of the left inferior frontal portion of the infant brain (Imada *et al.* 2006).

5.1. Social scaffolding in language learning

Most important aspect of language acquisition is the role of social influences on language learning (Kuhl 2004).

After six months, infants show a perceptual magnet effect from their native language category and reflect sensitivity to the distributional properties of sounds in their language. This effect is uniquely human and requires linguistic experience, in fact by this time infants focalize on linguistic sounds from their native language (Kuhl 2004). Also, it is of interest that by this time (5/6 months) human infants lose their encyclopedic hearing (de Boysson-Bardie 2001) that is the ability to discriminate the phonetic contrasts of all languages (Kuhl 2004).

5.2. The role of phylogenetic inheritance

In modern humans canonical babbling is scaffolded by motherese, a specific melodic contour that activates the right part of the infant brain (Homae *et al.* 2006). By means of motherese infants are naturally guided towards the right pronunciation of syllables of the native language, to put different syllables in a sequence of different vowel/consonant or consonant/vowel string, thus overcoming the first phases of babbling when sounds are often repeated (mama, dada etc.), and subsequently conducted towards the acquisition of the meaning of syllables strings i.e. words, (socially shared acoustic icons for items (objects or subjects) or meaningful motor sequences (actions) (Kuhl 2007). Motherese has a peculiar prosody: higher mean, maximum and minimum frequency, higher Fo frequency range, shorter sentences and longer pauses (Fernald and Simons 1989; Fernald 1992; Fernald and Kuhl 1987). Motherese is transcultural: ascending intonation contours for encouragement and descending ones for sedation were similarly used by Chinese, German, US mothers; another study showed that French, English, US and Italian mothers used the same intonation contours for attention seeking, approval, prohibition and comfort (Papoushek 1992; Fernald 1992).

This specific intonation contour was already noted by Darwin while observing his own infants, and Mithen (2009) believes that *Homo sapiens*' motherese was inherited by earlier hominid melodic communication forms. Indeed, the prosodic aspects of linguistic communication are processed in the right part of the adult human brain (Zatorre *et al.* 1992) and it was demonstrated that infants perceive prosody even when sleeping (Homae *et al.* 2006).

It is universally acknowledged that language is a communicative tool socially shared, and all the studies of language acquisition confirm that (Kuhl 2004) language has been defined as non referent arbitrary acoustic (or gestural) significant for objects or actions. Thus, in a start-up situation of a human (*Homo sapiens*) social group sharing a complete language apparatus the linguistic code had to be built step by step.

5.3. Onomatopoeia: a transitory phase?

The first idea that earlier forms of language were onomatopoeic was proposed by Darwin in his *Descent of Man* (1871), and was inspired by his cousin and linguist Hensleigh Wedgwood, who had written a book on language five years before (Richards 1987). Very soon those ideas were discarded by the great Oxford linguist Friedrich Max Müller (1823–1900),

who derisively called them the "pooh-pooh" and "bow-wow" theories of language formation (Richards 1987), and this proposal was discarded. However, a very interesting and up-to-now unique study by Hashimoto and coll. (2006), might shed new light on those old theories: by means of an event related fMRI study those authors found that the brain regions involved in processing onomatopoeic sounds – i.e. sounds imitating the sound or action of objects, animals, and humans (e.g., "buzz", "whirr", "bow", "hoot", "meow", "squeak", "coo", "hush", and "boom") – include the regions that process both nouns and animal sounds. Onomatopoeias and nouns are human speech sounds of similar frequency components, while onomatopoeias and animal sounds are often repetitive and acoustically simple sounds with frequency modulation, and Hashimoto *et al.* (2006) conclude that onomatopoeic sounds can serve as a bridge between nouns and animal sounds, and postulate that the onomatopoeic sounds contain both verbal and nonverbal sound components. As MacNeilage (2008) asserts, meaning should not arrive out of nowhere: the meaning of words is socially shared; thus, how did early human speakers build their shared vocabulary? I believe that there must have been a transitory phase between sound mimicry (simple repetition of animal sounds) inherited from earlier hominid forms and later, more modern "full human" forms of linguistic communication. And a phase of onomatopoeia (repetition and also re-creation of animal sounds) must have been in the middle, between sound mimicry and full-fledged language. Hashimoto & coll. (2006) indicate that, but it is impossible to say by now if this phase was already in the *Homo sapiens* domain or in earlier hominid forms. As a matter of fact, when the task of pronouncing the first words is accomplished, children frequently use onomatopoeic sounds in order to indicate surrounding objects. We also know that many, if not all, cultures have a repertoire of onomatopoeic words/sounds used with and by children (for Japanese, Hashimoto 2006; for Italian, Lenti Boero and Habegger, personal observation).

6. Discussion and conclusions

The maturation of the neurophysiological and anatomical apparatus allowing the first words production is a constrained path for human infants and must have been so along the hominid line. In present paper I try to describe that "logical" sequences underwent by hominids in their path towards language. I sketched three benchmarks: 1) a fixed (mammalian like) phase of sound production, allowing both alarm sounds with basic referents, and prosocial

low intensity sounds; 2) a second phase, when low intensity, music-like sounds start to be articulated paralleled by vocal contours. It is noticeable that in this phase infants possess a wide and open range acoustic competence, defined as encyclopedic hearing (de Boysson-Bardie 2001); 3) a third phase when consonant and vowel sounds are uttered and refined in the dyadic context, and are basically helped by motherese (Khul 2004). In this phase infants lose their encyclopedic hearing and focus only on the sounds of their native language (Khul 2007).

This series cannot be reversed in the sense that canonical babbling could not precede cry nor musilanguage. Because natural selection acts on present features and not for future potential benefits (Fagen 1981), I try to figure out the selective pressures that might have been at work along this path. This aspect is seldom considered in language evolution studies, but see Locke and Bogin (2006) and Lenti Boero (2014).

In addition, I try to examine in more detail the characteristics of protocommunication that might have antedated the appearance of the first words: socially shared sounds indexing objects by means of acoustic arbitrary labels.

Was the transition from one stage to the subsequent the product of chance mutations regarding hominid communication tools? Or was it a byproduct of changes due to other aspects of hominid evolution? Or was the transition from one stage to another driven in an autocatalytic way? Those are among the still open questions in the complex puzzle of language evolution.

Acknowledgements

References

Adams, C. L., D. L. Molfese and J. C. Betz. (1987). Electrophysiological correlates of categorical speech perception for voicing contrasts in dogs. *Developmental Neuropsychology* 3.3–4:175–89.

Barr, R. (2004). Early infant crying as a behavioral state rather than a signal. *Behavioral and Brain Sciences* 27.4: 460–46.

Bolhuis, J. J., K. Okanoya and C. Scharff. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience* 11, 747–759.

Bottoni, L., R. Massa and D. Lenti Boero. (2003). The grey parrot (Psittacus erithacus) as musician: an experiment with the Temperate Scale. *Ethology, Ecology & Evolution* 15.2: 133–141.

Bottoni, L., S. Masin and D. Lenti Boero. (2009). Vowel-like sound structure in an African Grey Parrot (Psittacus erithacus) vocal production. *The Open Behavioural Science Journal* 3, 53–68.

Butterworth, G., J. Rutkowska and M. Scaife (eds.) (1985). *Evolution and developmental psychology*. Brighton: John Spiers, The Harvester Press Limited.

Cismaresco, A.S. and H. Montagner. (1990). Mother's discrimination of their neonates' cry in relation to cry acoustics: The first week of life. *Early Child Development & Care* 65: 3–13.

Costall, A. (1985). Specious origins? Darwinism and developmental theory. In G. Butterworth, J. Rutkowska and M. Scaife (eds.). *Evolution and developmental psychology*, 30–44. Brighton: John Spiers, The Harvester Press Limited.

de Boysson-Bardie, B. (2001). *How language comes to children*. Cambridge: MIT Press.

De Graaf-Peters, V.B. and M. Hadders-Algra. (2006). Ontogeny of the human central nervous system: what is happening when? *Early human development* 82: 257–266.

Cooper, R.P. and R. N. Aslin. (1989). The language environment of the young infant: implications for early perceptual development. *Canadian Journal of Psychology* 43: 247–265.

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: Murray.

Davis, B. L. and P. F. MacNeilage. (2002). The internal structure of the syllable. In T. Givòn and B. F. Malle (eds.), *The evolution of language out of pre-language*, 135–54. John Benjamins.

Dewson, J. H. (1964). Speech sound discrimination by cats. *Science* 144: 555–56.

Dunbar, R. I. M. (1993). Coevolution of Neocortical Size, Group Size and Language in Humans. *Behavioral and Brain Sciences* 16: 681–694.

Eimas P. D., E. R. Siqueland, P. Jusczyk and J. Vigorito. (1971). Speech perception in infants. Science 171: 303–306.

Fagen, R. (1981). *Animal play behavior*. New York: Oxford University Press.

Feldman, R. and J. S. Reznick. (1996). Maternal perception of infant intentionality at 4 and 8 months. *Infant Behavior and Development* 19.4, 483–496.

Fernald, A., (1992). Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens and M. Papousek (eds.), *Nonverbal vocal communication*, 262–282. Cambridge University Press..

Fernald, A. and P. K. Kuhl. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development* 10: 279–293.

Fernald, A. and T. Simons. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20: 104–113.

Frodi, A. (1985). When empathy fails: aversive infant crying and child abuse. In B. M. Lester and C.F.Z. Boukydis (eds.) *Infant crying: Theoretical and research perspective*, 263–278. New York and London: Plenum Press.

Frodi, A. and M. Senchack. (1990). Verbal and behavioral responsiveness to the cries of atypical infants. *Child Development* 61: 76–84.

Giallongo, A. (1997). *Il bambino medievale* [The medieval child]. Bari: Dedalo.

Homae F., H. Watanabe, T. Nakano, K. Asakawa and G. Taga. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neuroscience Research* 54.4: 276–280.

Hogan, J. A. (1988). Cause and Function in the Development of Behavior Systems. In E. M. Blass (ed.), *Handbook of Behavioral Neurobiology 9*, 63–106. New York: Plenum.

Hopkins, B. and T. von Wulfften Palthe. (1987). The Development of the Crying State during Early Infancy. *Developmental Psychobiology* 20: 165–175.

Hsu, H. C. and A. Fogel. (2001). Infant Vocal Development in a Dynamic Mother-Infant Communication System. *Infancy* 2.1: 87–109.

Imada, T., Y. Zhang, M. D. Cheour, S. Taulu, A. Ahonen and P. Kuhl. (2006). Infant speech perception activates Broca's area: a developmental magnetoencephalography study. *Neuroreport* 17.10: 957–962.

Janicke, T., S. Hahn, M. S. Ritz and H. Peter. (2008). Vocal performance reflects individual quality in a nonpasserine. *Animal Behaviour* 75: 91–98.

Jurgens, U. and D. Ploog. (1988) On the motor control of monkey calls. In J. D. Newman (ed.), *The physiological control of mammalian vocalization*, 7–19. Plenum Press.

Jurgens, U. (1990). Vocal communication in primates. In R. P. Kesner and D. S. Olton (eds.), *Neurobiology of comparative cognition*, 51–76. Hillsdale, New Jersey: Lawrence Erlbaum Associates Publishers.

Jurgens, U. (1992). On the neurobiology of vocal communication. In H. Papoushek, U. Jurgens and M. Papoushek (eds.), *Nonverbal vocal communication. Comparative and developmental approaches*, 31–42. Cambridge: Cambridge University Press.

Kagan, J. (1984). *The nature of the Child*. New York: Basic Books Publishers.

Kluender, K. R., R. L. Diehl and P. R. Killeen. (1987). Japanese quail can learn phonetic categories. *Science* 237: 1195–97.

Kessen, W. and J. Levine. (1979). The imitation of pitch in infants. *Infant Behavior and Development* 2: 93–99.

Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience* 5: 831–843.

Kuhl, P.K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science* 10.1: 110–120.

Kuhl, P. K. and J. D. Miller. (1975) Speech perception by chinchilla: Voiced voiceless distinction in alveolar plosive consonants. *Science* 190: 69–72.

Lenti Boero, D. (1992). Alarm calls in marmots: evidence for semantic communication. *Ethology, Ecology, Evolution* 4.2: 125–138.

Lenti Boero, D. (1997). Dal pianto alla parola: una revisione della letteratura in una prospettiva comparata [From crying to words: a revision of literature in a comparative perspective]. *Ricerche di Psicologia* 4: 33–79.

Lenti Boero, D. (2009). Neurofunctional spectrographic analysis of the cry of brain injured asphyxiated infants: a physioacoustic and clinical study. In C. Manfredi (ed.), *Models and analysis of vocal emissions for biomedical applications*, 3–6. Firenze: Università di Firenze-Firenze University Press.

Lenti Boero, D. (2013). Sounds of protolanguage: some insights from developmental psychology. In K. Buczek, A. R. Knapik, J. Mianowski and P. P. Chruszczewski (eds.), *Ways to Protolanguage3, Book of abstracts*, 30–31. Wroclaw: Granice Nauki.

Lenti Boero, D. (2014). Early human communication helps in understanding language evolution. *Behavioral and Brain Sciences* 37(06), 560-561.

Lenti Boero, D. and L. Bottoni. (2006). From crying to words: Unique or multilevel selective pressures? *Behavioral and Brain Sciences* 29.3: 292–293.

Lenti Boero, D. and L. Bottoni. (2008). Why we experience musical emotions: Intrinsic musicality in an evolutionary perspective. *Behavioral and Brain Sciences* 31.5: 585–586.

Lenti Boero, D. and L. Bottoni, L. (2009). Contrasting early cry and early babbling: results from a pilot study. http://www.protolang.umk.pl/2009/public/files/book_of_abstracts.pdf.

Lenti Boero, D., C. Lenti and C. Volpe. (1993). How and why preterm cries are aversive? *Proceedings of the XXIII International Ethological Conference*, 138. Wien: University of Wien.

Lenti Boero, D., C. Volpe, A. Marcello, C. Bianchi and C. Lenti. (1998). Newborns crying in different situational contexts: discrete or graded signals? *Perceptual and Motor Skills* 86: 1123–1140.

Lenti Boero, D., S. Miraglia, F. Ortalda, G. Nuti, L. Bottoni and C. Lenti. (2008). Biomusicological approach in infant cry listening. *Second European Conference on Developmental Psychology of Music*, 162–165. Hull: Hull Publishing.

Lenti Boero, D., F. Ortalda, L. Bottoni, S. Miraglia and M. Filippa. (2009). Listening to biotic sounds: a pilot study. In A. R. Addessi and S. Young (eds.), *Proceedings of the 4th Conference of the European Network of Music Educators and researchers of Young Children*, 173–178. Bologna: Bologna University Press.

Lester, B. M. and C. F. Z. Boukydis. (1992). No language but a cry. In H. Papoushek, U. Jurgens and M. Papoushek (eds.), *Nonverbal vocal communication. Comparative and developmental approaches*, 145–173. Cambridge: Cambridge University Press.

Levitzky, S. and R. Cooper. (2000). Infant Colic Syndrome Maternal Fantasies of Aggression and Infanticide. *Clinical Pediatrics* 39.7: 395–400.

Lingle S., M. T. Wyman, R. Kotrba, L. J. Teichroebi and C. A. Romanow. (2012). What makes a cry a cry? A review of infant distress vocalizations. *Current Zoology* 58.5: 698–726.

Locke, J. L. (2006). Parental selection of vocal behavior. Crying, Cooing, Babbling, and the Evolution of Language. *Human Nature* 17.2: 155–168.

Locke, J. L. and B. Bogin. (2006). Language and life history: a new perspective on the development and evolution of human language. *Behavioral and Brain Sciences* 29.3: 301–311.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21.4: 499–511.

MacNeilage, P. F. (2008). *The origin of speech*. Oxford: Oxford University Press.

Mampe, B., A. D. Friederici, A. Christophe and K. Wermke. (2009). Newborns' Cry Melody Is Shaped by Their Native Language. *Current biology* 19.23: 1994–1997.

Marler, P. (1970). Birdsong and speech development: could there be parallels? *American Scientist* 58: 669–673.

Matzusawa, J., M. Matsui, T. Konishi, K. Noguchi, R. C. Gur, W. Bilker and T. Miyawaki. (2001). Age-related volumetric changes of brain gray and white matter in healthy infants and children. *Cerebral Cortex* 11: 335–342.

Mehler, J., P. Juskzyc, G. Lamberz, N. Halsted, J. Bertoncini and C. Amiel-Tison. (1988). A precursor of language acquisition in young infants. *Cognition* 29: 143–178.

Miller, J. D. (1977). Perception of speech sounds in animals: evidence for speech processing by mammalian auditory mechanisms. In T. Bullock (ed.), *Dahlem workshop on recognition of complex acoustic signals. Life Sciences report, Vol. 5*, 49–58. Abakon.

Mithen, S. (2006). *The Singing Neanderthals. The origin of music, language, mind and body*. London: Orion Books Ltd.

Mithen, S. (2009). The music instinct: the evolutionary basis of musicality. In S. Dalla Bella, N. Kraus, K. Overy, C. Pantev, J. S. Snyder, M. Tervaniemi, B. Tillman and G. Schlaug (eds.), *The neuroscience and music III. Disorders and plasticity*, *Annals of the New York Academy of Sciences* 1169: 3–12.

Morse, P. A. and C. T. Snowdon. (1975). An investigation of categorical speech discrimination by rhesus monkeys. *Perception and Psychophysics* 17: 9–16.

Newman, J. D. (1985). The infant cry of primates: an evolutionary perspective. In B. M. Lester and C. F. Z. Boukydis (eds.), *Infant crying. Theoretical and research perspectives*, 307–324. New York and London: Plenum Press.

Oller, D. K. (2000). *The Emergence of the Speech Capacity*. Mahwah, New Jersey: Erlbaum.

Oiler, D. K. and R. E. Eilers. (1988). The role of audition in infant babbling. *Child Development* 59: 441–449.

Oller, D.K. and R.E. Eilers. (1992). Development of vocal signaling in human infants: toward a methodology for cross-species vocalizations. In H. Papoushek, U. Jurgens and M. Papoushek (eds.), *Nonverbal vocal communication*, 174–191. Cambridge: Cambridge University Press.

Papoushek, M. and H. Papoushek. (1981). Musical elements in the infant's vocalization: their significance for communication, cognition, and creativity. In L.P. Lipsitt (ed.), *Advances in infancy research. Vol 1*. Norwood, New Jersey: Ablex.

Papoushek, M. (1992). Early ontogeny of vocal and verbal development in human infants. In H. Papoushek, U. Jurgens and M. Papoushek (eds.), *Nonverbal vocal communication*, 230–261. Cambridge: Cambridge University Press.

Pepperberg, I. M. (2007). Grey parrots do not always 'parrot': The roles of imitation and phonological awareness in the creation of new labels from existing vocalizations. *Language Science* 29: 1–13.

Peters, R. (1980). *Mammalian communication. A behavioral analysis of meaning*. Monterey: Brooks/Cole.

Piaget, J. (1972). *Principles of Genetic Epistemology*. London: Routledge & Kegan Paul.

Rasa, O. A. E. (1986). Coordinated Vigilance in Dwarf Mongoose Family Groups: The 'Watchman's Song' Hypothesis and the Costs of Guarding. *Ethology* 71.4: 340–344.

Rocca, F. and D. Lenti Boero. (2005). Sex differences in human infant cry: A comparative view. In R. Sàndor (ed.) *Abstracts of the XXIX International Ethological Conference*, 186. Budapest: Késult a Codex Print Nyondàbam.

Richards, R. J. (1987). *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*. Chicago: University of Chicago Press.

Roug, L., I. Landberg and L. J. Lundberg. (1989). Phonetic development in early infancy: A study of 4 Swedish children during the first 18 months of life. *Journal of Child Language* 16: 19–40.

Ruzza, B., F. Rocca, D. Lenti Boero and C. Lenti. (2003). Investigating the musical qualities of early infant sounds. In G. Avanzini, C. Faienza, D. Minciacchi, L. Lopez and L. Majno (eds.), *The Neuroscience and Music*, 527–530. New York, New York: The New York Academy of Sciences.

Seyfarth, R. M. and D. L. Cheney. (1980). Monkey Responses to Three Different Alarm Calls: Evidence of predator classification and semantic communication. *Science* 210: 801–803.

Seifritz, E., F. Esposito, J. G. Neuhoff, A. Luthi, H. Mustovic, G. Dammann, U. von Bardeleben, E. W. Radue, C. Sossio, G. Tedeschi and F. Di Salle. (2003). Differential Sex-Independent Amygdala Response to Infant Crying and Laughing in Parents versus Nonparents. *Biological Psychiatry* 54: 1367–1375.

Shannon, C. E. and W. Weaver. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Striedter, G. F. (2005). *Principles of brain evolution*. Sinauer Associates.

Tattersall, I. (1995). *The Fossil Trail: How We Know What We Think We Know About Human Evolution*. New York: Oxford University Press.

Tervaniemi, M. and M. Huotilainen. (2003). The promises of change-related brain potentials in cognitive neuro science of music. In G. Avanzini, C. Faienza, D. Minciacchi, L. Lopezmac and L. Majno. (eds.), *The Neuroscience and Music*, 527–530. New York, New York: The New York Academy of Sciences.

Vallet, E. M., M. L. Kreutzer, I. Beme and L. Kiosseva. (1997). Sexy syllables in male canary songs: honest signals of motor constraints on male vocal production? *Advances in Ethology* 32: 132.

Vorperian, H. K., S. Wang, M. K. Chung, E. M. Schimek, R. B. Durtschi, R. D. Kent, A. J. Ziegert and L. R. Gentry. (2009). Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of Acoustic Society of America* 125: 1666–1678.

Wedgwood, H. (1866). *On the Origin of Language*. London: Trübner.

Zatorre, R.J., A. C. Evans, E. Meyer and A. Gjedde. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256: 846–849.

Zeskind P. S. and B. M. Lester. (1978). Acoustic features and auditory perceptions of the cries of newborns with prenatal and perinatal complications. *Child Development* 49: 580–589.

Zuberbühler, K. (2000). Referential labelling in Diana monkeys. *Animal Behaviour* 59.5: 917–927.