COMMENTARY

The role of rhythmic cyclicities in infant action development

Peter F. MacNeilage and Barbara L. Davis

The University of Texas at Austin, USA

These four papers raise a number of issues concerning the ontogeny of the action component of language production, providing a structured opportunity to consider this question in general terms. Koopmans-van Beinum *et al.* compare prelinguistic vocal development in hearing and hearing-impaired infants and raise the question of the role of auditory input in infants' vocal development. Oller *et al.* are concerned with the parent's awareness of the first truly speech-like stage of development, the babbling stage. They provide a springboard for consideration of this critical first stage. Takei is concerned with the question of manual babbling in infants learning sign language. She raises the general question of the relation between pre-lexical vocal and manual output. Ejiri and Masataka add to a growing body of information suggesting a relation between the rhythmic property of babbling and other general rhythmic action by analyzing the on-line interaction between these two aspects of rhythmic development.

One theme that unifies the four papers, though it is not readily apparent in every paper, is the central role of rhythmic cyclicities in the development of language-related action, and, indeed, infant action in general. The parents in Oller *et al.*'s study are cueing on auditory correlates of the open–close rhythmic cycle of the mandible, which generates a series of acoustic energy peaks in the open vowel-like phase and acoustic minima during the consonant-like closing phase. Lack of coordination between the phonatory and articulatory system in the hearing-impaired infants studied by Koopmans-van Beinum *et al.* is (judging by the low incidence of two or more articulatory movements after 7 months in the hearing-impaired infants) primarily due to the lack of normal development of the mouth close–open alternation. Perhaps Takei's most important result is the confirmation of an earlier finding by Meier and Willerman (1995) that hearing-impaired infants who have access to adult signing produce a surprisingly large number of rhythmic cyclicities using various subcomponents of the manual apparatus. Ejiri and Masataka conclude that concurrent body rhythmicities may influence the form of babbling episodes, which show a higher tempo than babbling episodes occurring alone.

The important role of rhythmic cyclicities indicated by these papers was anticipated by Thelen (1981), who noted the presence of a wide variety of such cyclicities in the first few months of life – 'kicking, rocking, waving, bouncing, banging, rubbing, scratching, swaying' (p. 238). She suggested that they might play an important role in both ontogeny and phylogeny of action. This theme has subsequently been examined by a number of researchers. Thelen noted that this behavior 'stands out not only for its frequency but for the peculiar exuberance and seemingly pleasurable absorption often seen in infants moving in this manner' (p. 238). She views these stereotypies as transitional behavior between uncoordinated behavior and complex, coordinated motor control.

The importance of rhythmic cyclicities in infant action development is not surprising as the biphasic cycle is the main way in which living forms produce time-extended action sequences (MacNeilage, 1998a). Thelen considers these cyclicities to be 'phylogenetically available to the immature infant'. We have suggested a fundamental role of the cyclicity of the mandible in particular, as the 'frame' in a 'frame/content' theory of evolution of speech. The frame is considered to have a parallel role of early dominance in phylogeny and ontogeny (MacNeilage, 1998a; MacNeilage & Davis, 2000). The mandibular frame could have originated in ingestive cyclicities of early mammals, and then appeared in an intermediate stage as visuofacial communicative cyclicities (e.g. lipsmacks) before phonation was added, ultimately resulting in the protosyllable. The ease with which parents can correctly report whether their infant is babbling or not, shown by Oller *et al.*, suggests that the

Address for correspondence: Peter F. MacNeilage, Department of Psychology, University of Texas at Austin, Austin, TX 78712, USA; e-mail: macneilage@mail.utexas.edu

mandibular cyclicity emerges more or less fully formed rather than being put together in the uncoordinated manner characteristic of many infant action sequences (e.g. using culinary utensils). Oller *et al.*'s findings constitute a validation of the findings of Koopmans-van Beinum on babbling onset shown in their Figure 1. These results were obtained by telling parents of prebabbling infants about babbling and asking them 'to report the day on which they recognized it for the first time' (van der Stelt & Koopmans-van Beinum, 1986, p. 164).

The sudden onset of well-formed rhythmic babbling in infants has important evolutionary implications if this rhythmic cyclicity provided a frame for the earliest speech of hominids, just as it does for infants. This view includes the implication that subsequent stages of both phylogeny and ontogeny primarily involved *differentiation* of an initial frame into 'content' elements (MacNeilage, 1998a), namely independently controllable aspects of consonants and vowels. The view can perhaps be profitably contrasted with the opposite view of Ohala (1998) that the origin of speech in hominids may have instead involved *integration* – adding together of individual vowel and consonant productions. Although infants may produce steady-state vowels before babbling onset, they do not mirror the phylogenetic sequence suggested by Ohala by systematically producing stop consonants – the main consonants of babbling – in isolation before combining them with vowels.

If babbling has a deep-seated motor basis, why do Koopmans-van Beinum *et al.* and others find it so elusive in hearing-impaired infants. Understanding this absence may be enhanced by consideration of another putative emergent property of hominids, beyond the frame itself: the evolution of a generalized capacity for 'mimesis' – the capacity for representation and simulation of the actions of others (Donald, 1991). Donald asks how we can explain the centrality, and cultural specificity, in hominids, of a wide variety of behaviors such as tribal ritual, dance, opera, singing and music, games and sport, none of which have close analogs in our closest living relatives. For vocal communication in particular, there is little evidence that monkeys learn their calls (Hauser, 1997) although some species such as marmosets may be able to learn to modulate specific aspects of their calls according to social context (Snowdon, 1999). From Donald's perspective it can be argued that part of the phylogenetic and ontogenetic basis for the origin of speech is the ability to simulate language-related actions of the community. However, the ability of infants to imitate varied types of non-speech environmental sounds, many of which are culture-specific, is consistent with Donald's view that the ability to learn speech is not speech-specific.

According to the frame/content theory (MacNeilage, 1998a) phonation and mandibular oscillation had different functional histories, and therefore needed to be put together phylogenetically for speech. Consistent with this view of the two subsystems, hearing-impaired infants have a problem in coordinating phonatory and articulatory actions. However, more than the auditory simulation capacity is apparently involved in this coordination. Meier, McGarvin, Zakia and Willerman (1997) found, in both hearing and hearing-impaired infants, a phenomenon they call 'jaw wags' – bouts of silent mandibular oscillation. The relevance of this phenomenon to speech development is shown by some instances in which a series of these wags was intermittently accompanied by phonation. However, their occurrence without phonation is consistent with Donald's thesis of a generalized mimetic capacity, not specific to speech. No linguistically based definition of speech includes time-extended syllable-like productions without phonation.

It has been argued here that an endogenous motor cyclicity and an emergent mimetic capacity are jointly responsible for the onset of vocal babbling. As a prelude to a comparison of manual and vocal prelinguistic development, it is necessary to note additional motor constraints that determine the detailed form of the vocal babbling repertoire. While a high level of rhythmicity characterizes early vocal simulations of speech by hearing infants, they are otherwise rather gross. There is wide overall variation in the extent to which infants approximate specific adult models. Babbling features simple consonant–vowel alternations, but many adult languages are not so simple. While infants tend to repeat the same syllable within an utterance, syllable repetition occurs at below-chance levels in many languages (MacNeilage, Davis, Kinney & Matyear, 2000). This suggests that simulation during babbling is rather global in hearing infants, dominated by the basic motor frame capacity, and perhaps triggered by the hearing of rhythmic alternation between high and low acoustic amplitudes, but capitalizing on visual as well as acoustic information from adult utterances.

Recent research on intracyclical structure in the vocal modality has shown that both babbling and adult language have in common three kinds of consonant– vowel co-occurrence preferences – coronal (front) consonants with front vowels, dorsal (back) consonants with back vowels, and labial (lip) consonants with central vowels. These regularities suggest that infants might simulate adult intracyclical structure on the basis of auditory information. However, infants have comparable co-occurrences between the vowel and the following consonant, while languages do not (MacNeilage *et al.*,

2000). Thus the infant co-occurrence constraints appear to be due to basic biomechanical constraints on intracyclical diversity, which adult languages supercede at the vowel–consonant boundary, perhaps as part of the origin of explicit syllabification, which mandates syllable boundaries, which most typically occur after vowels.

With respect to sounds, babbling infants favor stop consonants, nasals, glides (Locke, 1983) and low and mid front and central vowels (MacNeilage & Davis, 1990). They disfavor fricatives and liquids, and high and back vowels. This overall pattern of difference between infants and languages, which usually contain all these sounds at high frequencies, is presumably a result of motor constraints rather than any differential ability to apprehend the perceptual substance of sounds. Fricatives and liquids are often considered difficult to produce as they make high demands on precision of articulator placement (fricatives) and tongue configuration (liquids) (Kent, 1992). As the tongue's primary anchor is in the low front part of the oral cavity, low frequencies of high and back vowels might initially be expected on biomechanical grounds.

How then does vocal babbling compare with manual babbling? The nonreferential signs studied by Takei, like vocal babbling, might be explained in terms of a combination of basic motor capabilities of the manual apparatus and general-purpose mimetic abilities not specific to language. Takei's finding of a high incidence of rhythmically repeated nonreferential manual actions in two hearing-impaired infants with signing parents is a confirmation of Meier and Willerman's (1995) findings. However, Meier and Willerman found much higher levels of repetition in the hearing-impaired offspring of signing parents than in hearing infants. In other ways, contrary to a well-known earlier study by Petitto and Marentette (1991) in which a common innate basis for speech and sign language was postulated, hearing and hearing-impaired infants were very similar. (See MacNeilage (1998b) for a methodological and conceptual critique of Petitto and Marentette's study.)

The finding of both Meier and Willerman, and Takei, of high levels of multicycle manual movements in hearing-impaired infants, which were not found in Meier and Willerman's hearing group, is evidence for the role of a general-purpose mimetic factor. One can assume that only the hearing-impaired infants were regularly exposed to time-extended series of manual gestures, many of which are repeated, due to the nature of sign language. Multicyclical movements of the hearing-impaired infants may then be a simulation of adult sign language similar to the way in which babbling is a simulation of adult speech. As Thelen pointed out, both make use of readily available cyclical capacities.

However, there is an important difference between the vocal and manual cyclical simulations. A single cyclical form – the mandibular cycle – is a simulation that is appropriate for virtually every syllable in any language. In contrast, while many signs involve repetitions of a movement cycle, this is far from typical, and the sign repetitions can be made with various parts of the system – arms, hands, fingers etc. To the degree that repetitions are not typical in sign language, repetitions by infants are not appropriate simulations. It is difficult to see why Takei would take multicycle movement as indicative of the syllable in sign language, when such movement is not typical of sign. It is not entirely clear that there is a sign language analog of the syllable, and if so, what its properties are (Coulter, 1993). Consequently there can at present be no straightforward claim that speech and sign language consist of series of syllables, so no appeal is possible at present to this unit as a common linguistic entity in speech and sign language.

As in vocal babbling, there is a strong element of motor determination of early manual gestures. Meier, Mauk, Mirus and Conlin (1998) found evidence of motor difficulty effects in early lexical signs, many of which are similar to prelinguistic gestures. While the location of the sign is typically correct, there are limitations on the ability to make more complex movements and handshapes. Handshape difficulty was reflected in more use of gross proximal limb configurations relative to distil (hand-internal) configurations. Many of these motor constraints observed by Meier *et al.* are sufficiently fundamental that they also show up in the 'home sign' of 3–4-year-old infants – signs invented for language-like communication by hearing-impaired infants in the absence of an adult model (Goldin-Meadow, 1999). While, as in speech, these constraints are also typical of prelinguistic nonreferential gestures, Meier *et al.* found no difference in these gestures between infants exposed and not exposed to sign.

Takei takes two findings as the primary evidence that nonreferential gestures and vocal babbling are components of a single amodal system. She notes that 'nonreferential gestures (like babbling) were observed 3 or 4 months before the onset of first signs', and 'nonreferential gestures were produced at the same age as hearing infants started vocal babbling'. Like Petitto and Marentette, who also claimed that onsets of vocal and manual babbling were identical, Takei makes this claim on the basis of results from only two infants. Consider for a moment what data would be needed to support this claim. Koopmans-van Beinum *et al.* presented (Figure 1) a relatively normal distribution of babbling onset times for hearing infants over a range of 30 weeks. If the mean and range of manual babbling

onset times were comparable, which they should be if both distributions reflect a common amodal system, then a relatively large number of infants would need to be sampled in order for an appropriate comparison of mean time of onset to be made.

Even if a large number of infants were available, there remains the problem of what criteria should be used to determine a characteristic time of manual babbling onset. This problem apparently has no counterpart in the vocal realm, where onset is easily agreed upon, and apparently results from a systematic pairing of frames with phonation. The only behavior so far found to be specific to hearing-impaired infants exposed to sign language is the high frequency of multicyclical gestures. We have argued that this behavior is not globally appropriate as a prelinguistic simulation of sign language, as most signs are not repeated. It may simply be a reflection of a general-purpose mimetic ability and motor system rhythmicity, rather than a language-specific capacity to babble in the manual modality. If other nonreferential gestures that do not presently seem to be affected by sign language experience are to be used, and it is found that their mean onset times are similar in hearing and hearing-impaired infants, can this be taken to indicate the onset of manual babbling?

An additional cautionary note on the question of manual babbling onset comes from the work of Rönnquist and von Hofsten (1994) who analyzed 2530 spontaneous finger movements of 10 newborn infants over two 5 min periods: 'Slightly less than half of these were opening and closings of the whole hand. The next most numerous kind of movement only involved the thumb and the index finger (22%). Almost all of these movements were thumb–index finger oppositions. Finally there were a number of less frequent types of movement like flexion of the two central digits of the hand, an increased flexion or increased extension from Digit 2 to Digit 5, and extension of the index finger, as in pointing' (p. 130). These findings suggest that some manual phenomena, which might be called nonreferential gestures, are present from birth.

Ejiri and Masataka provide evidence of an actual functional relationship between vocal and other bodily rhythmicities in on-line behavior. It is important to note that as vocal cyclicities involving the mandible are more appropriate approximations to spoken language than manual cyclicities are of sign language, concurrent interactions between vocal and manual cyclicities cannot be evidence of the existence of a common language-related controller. Ejiri and Masataka's data might better be regarded as evidence of the early existence of a capacity to interrelate vocal and manual activity for what could be called panlinguistic purposes. This capacity reaches a high level of development in adults in the types of concurrent language–gesture relationships noted by McNeill (e.g. McNeill, 1985).

In conclusion, much of the information in these four papers is consistent with the presence of two main causal factors in the development of vocal and manual language. One is endogenous rhythmic motor cyclicities, in a context of other motor constraints. The other is a generalized mimetic capacity. No language-specific factors seem to be involved in the infant's actions at these early stages in either modality. A concluding theme is available from Donald (1991). He emphasizes the *integrative* capacity of mimesis, perhaps revealed in early form in the results of Ejiri and Masataka. Donald considers that an emerging capacity to use rhythm is the main contribution to the mimetic integration capacity. From this standpoint, which echoes Thelen (1981) and which we endorse, the harnessing of endogenous rhythmic capacities of infant hominids for mimetic purposes should be a central issue in the developmental science of action.

Acknowledgement

References

Coulter, G.R. (Ed.) (1993). *Phonetics and phonology*, Vol. 3: *Current issues in ASL phonology*. San Diego, CA: Academic Press.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge, MA: Harvard University Press.

Goldin-Meadow, S. (1999). The form of gestures in deaf children's invented homesign systems. Paper presented at the VIIIth International Congress for the Study of Child Language, San Sebastian, Spain, July.

Hauser, M.D. (1997). *The evolution of communication*. Cambridge, MA: MIT Press.

Kent, R.D. (1992). The biology of phonological development. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (Eds), *Phonological development: Models, research, implications* (pp. 65–90). Timonium, MD: York Press.

Locke, J.L. (1983). *Phonological acquisition and change*. New York: Academic Press.

MacNeilage, P.F. (1998a). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, **21**, 499–511.

MacNeilage, P.F. (1998b). Evolution of the mechanism of language output: comparative neurobiology of vocal and manual communication. In J.R. Hurford, C. Knight & M.G.

Studdert-Kennedy (Eds), *Approaches to the evolution of language: Social and cognitive bases* (pp. 222–241). Cambridge: Cambridge University Press.

MacNeilage, P.F., & Davis, B.L. (1990). Acquisition of speech production: the achievement of segmental independence. In W.J. Hardcastle & A. Marchal (Eds), *Speech production and speech modelling*. Dordrecht: Kluwer.

MacNeilage, P.F., & Davis, B.L. (2000). Evolution of speech: the relation between phylogeny and ontogeny. In J.R. Hurford, C. Knight & M.G. Studdert-Kennedy (Eds), *The evolutionary emergence of language* (pp. 146–160). Cambridge: Cambridge University Press.

MacNeilage, P.F., Davis, B.L., Kinney, A., & Matyear, C.L. (2000). The motor core of speech: a comparison of serial organization patterns in infants and languages. *Child Development*, **71**, 153–163.

McNeill, D. (1985). So you think gestures are nonverbal. *Psychological Review*, **92**, 350–371.

Meier, R., & Willerman, R. (1995). Prelinguistic gesture in deaf and hearing infants. In K. Emmorey & J.S. Reilly (Eds), *Language, gesture and space*. Hillsdale, NJ: Erlbaum.

Meier, R., McGarvin, L., Zakia, R.A.E., & Willerman, R. (1997). Silent mandibular oscillations in vocal babbling. *Phonetica*, **54**, 153–171.

Meier, R.P., Mauk, C., Mirus, G.R., & Conlin, K.E. (1998). Motoric constraints on early sign acquisition. In E. Clark (Ed.), *Proceedings of the Child Language Research Forum*, Vol. 29. Stanford, CA: CSLI Press.

Ohala, J. (1998). Content first, frame later. *Behavioral and Brain Sciences*, **21**, 525–526.

Petitto, L., & Marentette, P. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science*, **251**, 1493–1496.

Rönnquist, L., & von Hofsten, C. (1994). Varieties and determinants of finger movements in neonates. *Early Development and Parenting*, **3**, 81–94.

Snowdon, C.T. (1999). An empiricist view of language evolution and development. In B.J. King (Ed.), *The origins of language: What nonhuman primates can tell us* (pp. 79–114). Santa Fe, NM: School of American Research Press.

van der Stelt, J.M., & Koopmans-van Beinum, F.J. (1986). The onset of babbling related to gross motor development. In B. Lindblom & R. Zetterstrom (Eds), *Precursors of early speech* (pp. 163–174). New York: Stockton Press.

Thelen, E. (1981). Rhythmic behavior in infancy: an ethological perspective. *Developmental Psychology*, **17**, 237–257.