

# Is the evo-devo approach paradigmatic for language motor control emergence?

Virginie DUCEY-KAUFMANN

Institut de la Communication Parlée, UMR CNRS (INPG-Stendhal), Université Grenoble 3,  
ducey@icp.inpg.fr

## Keywords

babbling, pointing, evo-devo approach

## Abstract

The general problem that children encounter in different areas of development is in the control of their body parts. The same situation is found for the Speech Frame, which is characterized by the auditori-motor cyclic control of the jaw (*Frame/Content* Theory of MacNeilage, 1998) found in canonical babbling around the age of seven months, well before the appearance of the independent and coordinated control of the distal articulators carried by the jaw (the lips and tongue). This difference between proximal and distal control has been demonstrated in another study (Munhall & Jones, 1998), which compared measurements of the upper and lower lips in eight month old children and adults. Green et al. (2002) have also shown that throughout development, the upper lip becomes less and less compliant to the pushing up of the lower lip carried by the jaw, hence showing finally evidence of an autonomous *contact* control. Bickley et al. (1986) have shown that the frequency of this Speech Frame is about 3 Hz. This corresponds to the proto-syllable, and more generally to children's rhythmic activities (Thelen, 1981).

Our working hypothesis is based on the idea that there is a “developmental rendez-vous” between what we call the “*Sign Frame*” and what we call the “*Speech Frame*”. The Sign Frame is characterized by the visual-motor control of the “carrier arm”, leading to so-called “imperative pointing” with the index finger around the age of nine months. To study this frame, we adopt the notion of objecthood, which provides a basis for language learning. Agent, object, and event indexing provides a theory for establishing the ability to track one to three targets during scene analysis, as Leslie et al. (1998) *inter alia* have proposed. The mental index of agent, object, and event indexing theory is first achieved by the eye in a shared deictic attentional device (Baron-Cohen, 1995), then by imperative pointing, and finally by declarative pointing around the age of 12 months. Declarative pointing appears along with the first words, while the Speech Frame allows the child at that stage to coproduce (coarticulate) a vowel and a consonant (Sussman et al., 1999). Pointing with the index finger instantiates the indicated object (the argument) and the predicate corresponds to an utterance produced by oral-laryngeal articulation accompanied by facial expressions and intonation. The relative importance of the elements of this developmental rendez-vous at the time of the emergence of the first words remains to be explored.

Thus, we think that when babies produce their first words, these “prosodic” words can vary from one to two syllables. In order for the child to acquire a sign-word in speech, she must be able to integrate two jaw cycles – canonical babbling at around 3 Hz or 300-350 ms – in a foot at 1.5 Hz or 600-700 ms. Our idea is thus that what hearing children do with the jaw (Speech Frame), they integrate in the pointing gesture of the arm and the index finger of the extended hand (Sign Frame), in coordination with vocal prosody, i.e., the word. In order to account for the existence of a harmonic relationship between the Speech Frame and the Sign Frame, we examined an audio-visual corpus, recorded from six French children aged from 6 to 18 months. We studied the distribution of the babbling frequencies, and of the durations of the

pointing or “stroke” gestures. Our results show that for these six children, the average pointing duration was around 775.87 ms (std. dev.:  $\pm 267.36$  ms), i.e., around 1.3 Hz. Moreover, the babbling data of one of the children showed that the syllable duration was about 364.59 ms ( $\pm 211.84$  ms), i.e., around 2.7 Hz. We thus found a ratio between gesture and speech of around 2, which could account for the first words template. The harmonic relationship found in children could also lead us to postulate a bisyllabic model for the dominant word pattern in the world’s spoken languages (Rousset, 2004). Following MacNeilage (1998), in this paper, we ask questions about speech ontogenesis, acknowledging that ontogenesis at least in part recapitulates phylogenesis. This allows us to shed light on the evolutionary path of human speech, within the framework of new developments in evo-devo theory.

## References

- S. Baron-Cohen. *Mindblindness*, Cambridge, Massachusetts: MIT Press, 1995.
- C. Bickley, B. Lindblom, L. Roug. Acoustic measures of rhythm in infants’ babbling, or “All god’s children got rhythm”. Proceedings of the 12<sup>th</sup> International Congress on Acoustics, Toronto, 1986.
- R. Green, C.A. Moore, K.J. Reilly. The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research*, vol. 45, 66-79, 2002.
- A.M. Leslie, F. Xu, P. Tremoulet, B. Scholl. Indexing and the object concept: Developing “What” and “Where” systems. *Trends in Cognitive Sciences*, vol.2, n°1, 10-18, 1998.
- P. MacNeilage. The Frame/Content theory of evolution of speech production. *Behavioral and Brain Sciences*, 21/4, pp. 499-511, 1998.
- K.G. Munhall, J.A. Jones. Articulatory evidence for syllabic structure. *Behavioral and Brain Sciences*, 21/4, pp. 524-525, 1998.
- I. Rousset. *Structures syllabiques et lexicales des langues du monde. Données, typologies, tendances universelles et contraintes substantielles*. Thèse soutenue en Sciences du Langage, Université Grenoble III, Institut de la Communication Parlée, 10 juin 2004.
- H. Sussman, C. Duder, E. Dalston, A. Cacciatore. An Acoustic analysis of the development of CV coarticulation: a case study. *Journal of Speech, Language, and Hearing Research*, 42, pp.1080-1096, 1999.
- E. Thelen. Rhythmical behavior in infancy: an ethological perspective. *Developmental Psychology*, 17, pp. 237-257, 1981.