

Phonology grounded in sensorimotor speech: Elements of a morphogenesis theory

Jean-Luc Schwartz

Speech Communication Institute – Gipsa-Lab / ICP, UMR 5216 – CNRS – Grenoble Univ.

Where does language come from? This old question – forbidden during many years by the Paris Linguistic Society – now arouses a number of proposals and models, that can basically be divided between “pull” and “push” approaches. The first ones intend to determine what could be the evolutionary advantages of the “invention” of language, leading to its selection as a key component of modern humans. The second ones focus on the search for continuities, that are cognitive abilities already existing in non-human primates and able to have “pushed” towards the birth of language. This is the focus of the present talk.

I shall present and discuss a theoretical framework enabling to ground language in cognitive abilities surrounding it and pre-existing to it, around three questions.

Firstly, where does reference – that is the ability to refer to something in the world thanks to the production of an adequate communicative action – come from? I will show how the “Vocalize to Localize” framework developed in Grenoble (Abry et al., 2004) provides a possible bootstrap to reference, coupling the hand and the mouth in a deictic function (how to show things around us by both pointing and vocalizing).

Secondly, where does parity – that is the property enabling a speaker and a listener to be “pairs”, sharing the same sign system and able to both adequately produce and perceive these signs – come from? I will discard both “auditory” and “motor” theories of speech perception, and introduce the perceptuo-motor theory that we have proposed instead, called the “Perception-for-Action-Control Theory (PACT) (Schwartz et al., 2002, 2007), coupling the mouth and the ear and eye in the selection of adequate gestures for communication.

Thirdly, where does compositionality – that is the ability to combine units into larger items both at a phonological and syntactic level – come from? I will describe the solution proposed by our colleagues MacNeilage and Davis (2000) for compositionality in phonology through their “Frame-then-Content” theory of speech production, coupling the jaw and the other orofacial articulators in a principled way through development and phylogeny.

In this framework, I will present a number of experimental and simulation results showing how phonemes, syllables and words could emerge from these dynamic systems, inside a global complex system in which two speaking partners are coupled in face-to-face interaction through a double connection: a perceptuo-motor audio-visuo-orofacial loop and a mutual attention system.

Abry, C., Vilain, A. & Schwartz, J.-L. (2004). (Eds) 'Vocalize to Localize'? A call for better crosstalk between auditory and visual communication systems researchers. "Vocalize to Vocalize", *Interaction Studies*, 5, 313-324.

MacNeilage, P. F. & Davis, B. L. (2000) On the origin of internal structure of word forms. *Science*, 288, 527--531.

Schwartz, J.L., Abry, C., Boë, L.J., & Cathiard, M. (2002). Phonology in a theory of perception-for-action-control. In J. Durand, B. Laks (eds.) *Phonology : from Phonetics to Cognition* (pp. 255-280). Oxford: Oxford University Press.

Schwartz, J.L., Boë, L.J., & Abry, C. (2007). Linking the Dispersion-Focalization Theory (DFT) and the Maximum Utilization of the Available Distinctive Features (MUAF) principle in a Perception-for-Action-Control Theory (PACT). In M.J. Solé, P. Beddor & M. Ohala (eds.) *Experimental Approaches to Phonology*. OUP (to appear).