A Computational Model of Infant Speech Development

Ian Spencer Howard1 & Piers Messum2
1 Biological & Machine Learning Lab, Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK
2Phonetics & Linguistics, University College London, Gower Street, London WC1E 6BT, UK
Correspondence should be sent to ish22@cam.ac.uk

ABSTRACT

Almost all theories of infant speech acquisition assume an infant learns speech sounds by direct imitation, performing an acoustic matching of an adult's utterances and some postulate an innate link between perception and production.

We present a computer model which treats speech production and perception as separate processes with no innate link and has no requirement for acoustic matching on the part of the infant. Instead, learning to speak makes use of associative mechanisms and reinforcement learning.

Initially the infant explores its speech apparatus and reinforces its own action on the basis of sensory salience, developing local motor schemes. As the infant's production improves, it will start to generate speech like utterances, some of which will start to generate a response from its mother. A second phase of development then arises from this interaction. Because the mother is a learned speaker, her experienced perceptive system can evaluate the infants output within the phonological system of the ambient language L1.

The mother's response to the infant's vocal output can reinforce it in a number of ways. Simply generating a response will tend to encourage and reinforce the infant's production of a given utterance. However, most significantly, during imitative exchanges in which the mother reformulates the infant's speech, this enables the infant to learns equivalence relations between its motor activity and its mother's acoustic output, and thus to solve the correspondence problem. Notice that the infant does not learn equivalence relations between its own acoustic output and that of its mother based on acoustic similarity. On the contrary, this similarity based matching is performed by its mother.

We show that this model is able to progress through distinct stages of speech development. It begins by generating simple sounds and ends up producing word-like utterances which it can associate with its mother's spoken input.