



Learning phonetic categories by tracking movements

Bruno Gauthier a,*, Rushen Shi a, Yi Xu b

We explore in this study how infants may derive phonetic categories from adult input that are highly variable. Neural networks in the form of self-organizing maps (SOMs; Kohonen, 1989, 1995) were used to simulate unsupervised learning of Mandarin tones. In Simulation 1, we trained the SOMs with syllable-sized continuous F0 contours, produced by multiple speakers in connected speech, and with the corresponding velocity profiles (D1). No attempt was made to reduce the large amount of variability in the input or to add to the input any abstract features such as height and slope of the F0 contours. In the testing phase, reasonably high categorization rate was achieved with F0 profiles, but D1 profiles yielded almost perfect categorization of the four tones. Close inspection of the learned prototypical D1 profile clusters revealed that they had effectively eliminated surface variability and directly reflected articulatory movements toward the underlying targets of the four tones as proposed by Xu and Wang (2001). Additional simulations indicated that a further learning step was possible through which D1 prototypes with one-to-one correspondence to the tones were derived from the prototype clusters learned in Simulation 1. Implications of these findings for theories of language acquisition, speech perception and speech production are discussed.

全文: [2008_31.pdf](#)

[上一章](#) [目录](#) [下一章](#)