A neural network model of lexical segmentation and recognition

Shogo Makioka

Osaka Prefecture University

Abstract: A neural network models is presented for generating a representation of words from the input phoneme sequences. It uses an unsupervised learning algorithm that compares the current input with its memory of previous sequences, and generates a new representation of the common subsequence. Although the generated representation is quite noisy, the model can extract the consistent pairs of subsequences, which in high probability matches the words in the sentence. Simulations were conducted using child-directed speech in the CHILDES database (MacWhinney, 2000) as the training stimuli. The models performance of lexical segmentation was comparable with that of symbol-based models (Brent & Cartwright, 1996