Abstract

Three experiments were conducted to provide a better understanding about the fundamental processes involved in Chinese speech recognition. Specifically, we intended to answer three questions. First, are subsyllabic units like individual phonemes or whole syllables the basic encoding units in Chinese speech recognition? Second, does tone play a significant role in generating candidate words before correct identification? Third, how can the different meanings of homophones be resolved? In Experiment 1, we used the gating paradigm to explore the three issues. Results suggested that both subsyllabic (onset) and syllabic representations were important in recognizing Chinese monosyllables. Tonal constraints emerged only when context was available. And context also facilitated homophone recognition. In Experiment 2, the visual-world paradigm was used to verify the major findings in gating. While the salience of syllable and the absence of tonal constraints without context were replicated, the onset effect was greatly diminished. Further analyses suggested that acoustic similarity might also play a role in speech recognition. Experiment 3 also employed the visual-world paradigm. The resolution of Chinese homophones was found to be influenced by relative meaning frequency and context position. Based on these findings and those from related studies, we proposed a model of Chinese speech perception, in which initially, segmental and suprasegmental types of information were processed in separate but interacting pathways. Outputs from the two pathways were then combined at a later time point and jointly activated the corresponding morpheme. Implications of the model and its relations to previous findings are discussed.