Classification of stop consonant place of articulation: Combining acoustic attributes. Atiwong Suchato, Speech Communication Group, RLE, MIT, Cambridge, MA, USA. [Full Paper Available on CD]

This study evaluates the classification of stop consonant place of articulation in running speech using knowledge-based cues. Acoustic attributes are chosen to capture four categories of cues: amplitude and energy of burst, formant movement of adjacent vowels, spectrum of noise after the release, and some temporal cues. Correlation analysis shows no redundant information among cross-category attributes, although a few attributes in the same category show high correlation. Combinations of non-redundant attributes are used for the place classifications based on Mahalanobis distance. When stops contain release bursts, the classification accuracies are better than 90%. When bursts are absent, the classification can only rely on formant structures, and this results in reduced accuracy. It is also shown that voicing and vowel frontness contexts lead to a better classification accuracy of stops in CV context. However, there is no evidence of significant improvement for stops in VC context. When stops are located between two vowels, information on the formant structures in the vowels on both sides can be used. The probabilities of each place of articulation from two classifiers, each of which uses information on one side of the stops, are combined in order to obtain final decisions. Such combination yields the best classification accuracy of 95.5%. By using appropriate methods for stops in different contexts, an overall classification accuracy of 92.1% is achieved. The overall accuracy could be improved significantly if stops in VC context with no release burst were classified more accurately.