Features of Normative and Pathological Realizations of Sibilant Sounds for Computer-Aided Pronunciation Evaluation in Children

Authors : Zuzanna Miodonska, Michal Krecichwost, Pawel Badura

Abstract : Sigmatism (lisping) is a speech disorder in which sibilant consonants are mispronounced. The diagnosis of this phenomenon is usually based on the auditory assessment. However, the progress in speech analysis techniques creates a possibility of developing computer-aided sigmatism diagnosis tools. The aim of the study is to statistically verify whether specific acoustic features of sibilant sounds may be related to pronunciation correctness. Such knowledge can be of great importance while implementing classifiers and designing novel tools for automatic sibilants pronunciation evaluation. The study covers analysis of various speech signal measures, including features proposed in the literature for the description of normative sibilants realization. Amplitudes and frequencies of three fricative formants (FF) are extracted based on local spectral maxima of the friction noise. Skewness, kurtosis, four normalized spectral moments (SM) and 13 mel-frequency cepstral coefficients (MFCC) with their 1st and 2nd derivatives (13 Delta and 13 Delta-Delta MFCC) are included in the analysis as well. The resulting feature vector contains 51 measures. The experiments are performed on the speech corpus containing words with selected sibilant sounds (/ʃ, ʒ/) pronounced by 60 preschool children with proper pronunciation or with natural pathologies. In total, 224 /ʃ/ segments and 191 /ʒ/ segments are employed in the study. The Mann-Whitney U test is employed for the analysis of stigmatism and normative pronunciation. Statistically, significant differences are obtained in most of the proposed features in children divided into these two groups at p < 0.05. All spectral moments and fricative formants appear to be distinctive between pathology and proper pronunciation. These metrics describe the friction noise characteristic for sibilants, which makes them particularly promising for the use in sibilants evaluation tools. Correspondences found between phoneme feature values and an expert evaluation of the pronunciation correctness encourage to involve speech analysis tools in diagnosis and therapy of sigmatism. Proposed feature extraction methods could be used in a computer-assisted stigmatism diagnosis or therapy systems.