

DOCTORAL THESIS ABSTRACT

Multidimensional speech analysis methods for evaluation of articulatory features of sibilant sounds

Author: mgr inż. Zuzanna Miodońska

Supervisor: dr hab. inż. Paweł Badura

Auxiliary supervisor: dr inż. Joanna Czajkowska

The presented doctoral thesis is focused on speech signal processing in computer-aided speech diagnosis. The main objective was to develop methods of children speech analysis for evaluation of sibilant sounds' pronunciation. Such tools could provide additional possibilities of acoustic description of sigmatism – a speech disorder, in which sibilant sounds are misarticulated. The main thought of this work has been formulated as ***Speech signal features and statistical models of sibilants /s/ and /z/ can be successfully employed for detection of selected non-normative articulatory features of these sounds in preschool children's speech.***

The study was performed on a dedicated speech corpus. The corpus was registered in a group of 5 and 6 year-olds and recordings were annotated by speech pathologists. One of the main tasks was to propose a speech signal feature set that would reflect the specifics of analyzed phones best. The set contained different subsets of parameters typically used in speech analysis, as well as novel features describing high-frequency noise characteristic to sibilant sounds. The parameters were subjected to statistical analysis. The aim of these experiments was to verify, whether selected features are distinctive for 4 types of pronunciation found in the speech corpus: normative, addental, interdental and dental articulation.

The feature set was subsequently employed as a basis for statistical models of articulation, which were built using Hidden Markov Models (HMM). The most important element of the thesis was a method of articulatory features evaluation. The proposed approach was based on hybrid use of HMM models of articulation and Support Vector Machines. Two strategies were presented: first one for binary case (selected non-normative feature vs. norm) and the second one for multiclass case.

Experimental results confirmed, that signal parameters describing high-frequency noise, as well as proposed methods of classification based on statistical models of articulation allowed for more accurate pathology detection than state-of-the-art methods.