Title:

Longitudinal investigation of sibilant articulation development in children: a statistical modeling approach based on instrumental evidence and data mining methods

Acquisition of the phonological system of one's language is a process that takes years. Child speech researchers, e.g., speech and language pathologists, pediatricians, psycholinguists, assume that the process of speech acquisition proceeds in stages, and the child masters the correct (similar to the adult norm) pronunciation around the age of six. Deviations from adult pronunciation may be related to insufficient articulatory proficiency or poor mastery of the language phonology. Speech sounds that require more precise articulatory movements are initially omitted or pronounced as simpler sounds (sound substitution). This applies primarily to mastering sounds considered to be more difficult, e.g., most sibilants, in written Polish spelled with the letters "s, z, c, dz, sz, ż, cz, dż".

Although the topic of speech acquisition in children has been addressed by many researchers, relatively little is known about the changes in the speech signal that occur within this process. We know that language development does not happen in leaps but gradually. However, there is no data showing whether the gradual character of this process can be observed, measured, and mathematically described based on children's recordings in the timeline of their development.

The project aims to develop a statistical model describing the nature and pace of changes in the improvement of sibilant articulation based on parameters determined from audio recordings and video images of the face. We will invite 50 typically developing children, who, from the fourth to the sixth birthday, will undergo a speech and hearing examination every six months; we will also record samples of their speech and video data showing how their articulators work while speaking. Having data on the development of their speech described by speech therapists, we will be able to trace how the audible and visible phenomena, e.g., better differentiation of sounds pronounced by children, are reflected in instrumental data, recorded with measuring devices and analyzed using modern methods data analysis.

The result of the project will be a mathematical model describing the development of sibilant articulation using numerical parameters. The use of acoustic and image data analysis tools to study child speech in the long term will allow us to verify, clarify, and enrich our understanding of the process of speech acquisition and articulation improvement in children. This, in turn, will be valuable in the field of language research in both theoretical and application terms – in the future, such knowledge will help to objectify and improve the process of speech therapy diagnosis by providing speech therapists with tools for computer assessment of children's pronunciation.