Published online: 2019-04-23

Speech differences between CI users with pre- and postlingual onset of deafness detected by speech processing methods on voiceless to voice transitions

T. Arias-Vergara[1,2,3], S. Gollwitzer[1], J .R. Orozco-Arroyave[2,3], J. C. Vasquez-Correa[2,3], E. Nöth[2], C. Högerle[1], M. Schuster[1]

1. Department of Otorhinolaryngology, Head and Neck Surgery. Ludwig-Maximilians-University, Munich, Germany,
2. Pattern Recognition Lab. Friedrich-Alexander University, Erlangen-Nürnberg, Germany,
3. Faculty of engineering. Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellin, Colombia

Introduction

The onset of deafness affects the speech in different ways:

- Prelingually deafened people have difficulties learning how to speak intelligibly.
- Speech quality of postlingually deafened people decreases due to the lack of adequate auditory feedback.

Our hypothesis is that it is possible to differentiate between CI users with pre- and postlingual hearing loss considering acoustic features extracted automatically from speech recordings independently from age effects on speech.

Data

Speech recordings of 22 pre- and 22 postlingual CI users (10 male, 34 female) were considered for the tests. The age of the prelingual group ranges from 15 up to 71 years old (33,8 ± 18,5). In the postlingual group the age ranges from 15 up to 78 years old (53,5 ± 16,8). All of the patients read 97 words, which contain every phoneme of the German language in different positions within the words.

Methods

Speech was recorded using a Beyerdynamics microphone Opus 54.15-3 with a sampling frequency of 22.05 kHz.

Automatic speech analysis is performed considering speech segments with transitions from voiceless to voiced sounds. The method used in this work to identify these transitions is based on the presence of the fundamental frequency of speech in short-time frames with a length of 40 ms (Figure 1).

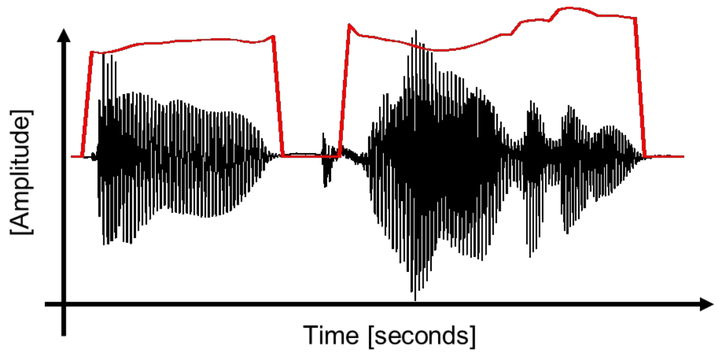

Figure. 1. Speech signal and pitch contour.

To extract the transitions, we search for the boundary between voiceless and voiced sounds using the fundamental frequency with a constant segment of 80 ms to the left and right (Figure 2).

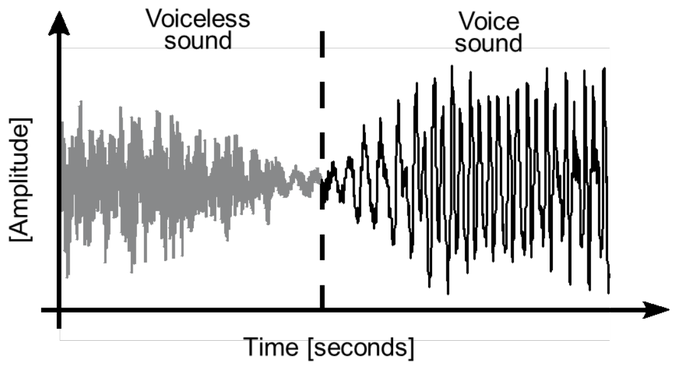

Figure. 2. Transition from voiceless to voice sound.

After segmenting the transition, 13 Mel-Frequency Cepstral Coefficients and the first and second derivatives are extracted from the segment. The mean, standard deviation, skewness and kurtosis are computed from the descriptors, forming a 156-dimensional feature vector per speaker.

Wilcoxon signed-rank test is used to find differences between the pre- and postlingual groups.

Additionally, a support vector regressor was trained to evaluate the age independence of the selected features.

Results

The Wilcoxon signed-rank test was performed for each descriptor, and significant differences between the pre- and postlingual groups ($\alpha<0.05$) were found in 8 of the 156 features.

Support vector regressor revealed not a strong correlation between the age of the speakers and the selected features ($\rho<0.40$).

Discussion

Speech patterns differ significantly between pre- and postlingual deafened CI users at the transitions of voiceless to voiced sounds. Altered transitions are also found in neurologic diseases such as Parkinson's disease and stand for an altered articulatory motor control. In patients with prelingual hearing loss, they are caused by congenital disturbed auditory feedback. Altered transitions might be correlated with the voice onset time.

Literature:

- Ruff S et al. Speech Production Quality of Cochlear Implant Users with Respect to Duration and Onset of Hearing Loss.
- Orozco-Arroyave J.R., Analysis of speech of people with Parkinson's disease.

Acknowledgment

The work is part of the Training Network on Automatic Processing of PAthological Speech (TAPAS) funded by the Horizon 2020 program of the European Commission.