



Topic: Speech Coding Strategies
2nd Topic: Sound Processing
3rd Topic: Sound Coding

Title: Gender identification and Intelligibility of whispered Speech in Cochlear Implant Users: Evaluation and Analysis

Author(s): Hazrati, Oldooz¹, Ali, Hussnain¹, Hansen, John H.L.¹, Tobey, Emily A.²

Institute(s): ¹The University of Texas at Dallas, Electrical Engineering, Richardson, United States, ²The University of Texas at Dallas, Behavioral and Brain Sciences, Richardson, United States

Abstract: Cochlear implant (CI) devices enable profoundly deafened individuals to identify clean and neutral speech to a great extent. CI users understand spectrally reduced speech (SRS) quite well as long as no distortion is introduced to the signal. These distortions may be associated with the listening environment due to room reverberation and background noise, or be due to speaking style of speakers such as in shouted or whispered speech. Because of the limited temporal and spectral resolution of speech signal processed by CIs, CI users' ability to recognize speech drops substantially in challenging listening scenarios. Whispering, in particular, changes temporal fine structure of speech and consequently results in deteriorated speech perception in CI listeners. In this study, we evaluate the effect of whispering on speech intelligibility in CIs as our first goal.

In addition to speech understanding, which is the primary goal of speech recognition for human listeners, speaker and/or gender identification is a secondary aim in recovering linguistic information. It has been shown that fundamental frequency and formant frequencies are important factors in speaker gender classification. Despite the reduced temporal and spectral cues of processed speech, CI users can correctly identify speaker gender to a great extent for quiet and neutral speech. The perceived pitch of whispered vowels corresponds to the first and second formant values. Therefore, whispered speech gender identification task becomes challenging for CI listeners with SRS. As a secondary goal of this study, we assess the effect of whispering on speaker gender identification in CI



users.

CI listeners are tested with the speech sentences from TIMIT database spoken by native speakers of American English recorded under neutral and whispered vocal efforts. For the intelligibility listening tests, 20 sentences (10 spoken by a female and 10 spoken by a male speaker) are used for each vocal effort condition (neutral and whisper). Sentences spoken by female speakers with $F_0 = 229$ Hz and 220 Hz are used for neutral and whispered listening conditions, respectively. Speech from male speakers with $F_0 = 137$ Hz and 146 Hz were used for the neutral and whispered intelligibility tests, respectively. For gender classification under each condition (neutral or whisper), twenty sentences are used from 10 female speakers ($174 \text{ Hz} < F_{0N} < 242 \text{ Hz}$ and $171 \text{ Hz} < F_{0W} < 240 \text{ Hz}$) and 10 male speaker ($111 \text{ Hz} < F_{0N} < 157 \text{ Hz}$ and $110 \text{ Hz} < F_{0W} < 166 \text{ Hz}$). The results are evaluated in terms of significance of difference between neutral and whispered speech both for intelligibility and speaker gender identification tasks. The intelligibility and speaker gender identification scores are also evaluated for each class of speakers (male and female) separately, in order to better assess the conditions in which CI users' performance drops more significantly.