
The effect of voice familiarity via training on voice cue sensitivity and listening effort with vocoder degraded speech

Ada Biçer*^{†1,2}, Thomas Koelewijn^{1,2}, and Deniz Başkent^{1,2}

¹Department of Otorhinolaryngology / Head and Neck Surgery, University Medical Center Groningen, University of Groningen – Netherlands

²Graduate School of Medical Sciences, Research School of Behavioural and Cognitive Neurosciences, University of Groningen – Netherlands

Abstract

Understanding speech in real life can be challenging and effortful, such as in multiple-talker listening conditions. Fundamental frequency (F0) and vocal-tract length (VTL) voice cues can help listeners segregate between talkers, enhancing speech perception in adverse listening conditions. Previous research showed that degradations of cochlear implant (CI) hearing reduce sensitivity to F0+VTL voice cues compared to normal hearing (NH), contributing to difficulties in understanding speech in adverse listening. Nevertheless, when multiple talkers are present, familiarity with a talker could provide a speech intelligibility benefit. In this study, we investigated how voice familiarity could affect perceptual discrimination of voice cues, as well as listening effort, with or without vocoder degradations.

To establish voice familiarity, we implemented an implicit short-term voice training. NH participants listened to a recording of a book segment that was presented for approximately 30 minutes, and to ensure engagement, they had to answer text-related questions. Following voice training, just-noticeable-differences (JNDs) for F0+VTL were measured with an odd-one-out task implemented as a 3 alternative forced choice adaptive paradigm. During the procedure, the reference voice either belonged to the trained voice or an unfamiliar voice, presented in both unprocessed and vocoder-degraded (12-band with low spread of excitation) versions. Effects of voice familiarity (trained and untrained voice), vocoding (non-vocoded and vocoded) and item variability (fixed or variable consonant-vowel triplets presented across three items) on voice cue sensitivity (F0+VTL JNDs) and listening effort (pupillometry measurements) were analyzed.

Results showed that, voice training did not have a significant effect on voice cue discrimination. As expected, F0+VTL JNDs were significantly larger for vocoded conditions than for non-vocoded conditions and with variable item presentations than fixed item presentations. GAMM analysis over the time course of stimulus presentation showed that pupil dilation was significantly larger during F0+VTL discrimination while listening to unfamiliar voices compared to trained voices, but only for vocoded speech. Peak Pupil Dilation was significantly larger for vocoded conditions compared to non-vocoded conditions and variable items increased the pupil baseline relative to fixed items, which could suggest a higher anticipated

*Speaker

[†]Corresponding author: a.bicer@rug.nl

task difficulty. In this study, even though short training for voice familiarity did not improve sensitivity to small F0+VTL voice cue differences at the threshold level, voice familiarity still provided a benefit in listening effort for voice discrimination among vocoded voices.

Funding: VICI grant 918-17-603 from the Netherlands Organization for Scientific Research (NWO) and the Netherlands Organization for Health Research and Development (ZonMw), the Heinsius Houbolt Foundation, and a Rosalind Franklin Fellowship.