Factors limiting vocal-tract-length perception in cochlear-implants
Etienne Gaudrain, Deniz Başkent
University Medical Center Groningen, Department of Otorhinolaryngology | University of Groningen, School of Behavioural, Cognitive Neuroscience | Groningen, Netherlands

Introduction
To categorize speakers as ‘male’ or ‘female’, or to track a voice in a crowded environment, normal-hearing (NH) listeners rely on two vocal characteristics: glottal-pulse rate (GPR), which defines F0, and vocal-tract length (VTL), related to the size of the vocal tract. Previous studies showed that normal-hearing (NH) listeners are extremely sensitive to these two dimensions: with vowels, the JND for F0 is about 2%, and between 5 and 7% for VTL (Smith and Patterson, 2005).

Figure 1 shows gender categorization by NH listeners and cochlear implant (CI) users, as a function of F0 and VTL from Gaudrain et al. (2013). This study shows that gender categorization is abnormal in CI users because their perception of VTL is very limited.

The purpose of the present study is to test various possible causes for this lack of VTL sensitivity using different types of vocoders.

Figure 2 shows gender categorization by NH listeners and cochlear implant (CI) users, as a function of F0 and VTL from Gaudrain et al. (2013). This study shows that gender categorization is abnormal in CI users because their perception of VTL is very limited.

14 NH participants listened to triplets of Dutch CV syllables in an adaptive 3AFC task tracking the JND in various directions of the F0-VTL plane. The original utterances were recorded from a female speaker and manipulated with Sinusoid to effect changes in F0 and/or VTL. All JNDs were measured relative to the original voice.

The stimuli were then used to modulate the amplitude of a sine-wave centered on the frequency band. The temporal envelope was extracted by half wave rectification and low pass filtering (filtering below 300 Hz zero-phase 4th order Butterworth filter). The envelope was then used to modulate the amplitude of a sine-wave centered on the frequency band.

Figure 4 shows the average F0 and VTL JNDs for the sinewave, noise, harmonic noise and phase-spreading harmonic complexes (PSPHC, see MacKenzie et al., poster P21). Individual results are shown as light symbols. The dotted lines represent the F0 and VTL difference between the male and female talkers from Gaudrain et al. (2013).

4th order Butterworth filter (-72 dB/oct.) are very sharp. Only the VTL JNDs for 4th and 12th order filters (-72 dB/oct.) are very shallow and produce a fair amount of channel interaction. 8th order filters (-48 dB/oct.) are reasonably sharp and 12th order filters (-72 dB/oct.) are very sharp. Only the VTL JNDs were measured in this experiment.

Another characteristic of CIs is the fact that the frequency band allocated to a specific electrode may not correspond to frequency to which the place is tuned. This place-frequency mismatch can be quantified in millimeters along the basilar membrane and can be simulated by shifting the synthesis filter in the vocoder by a fixed distance. Twelve bands were used, with two filter orders and two vocoder settings.

Figure 8 shows VTL JNDs for 4th and 6th order filters, as a function of place-frequency shift. The light symbols represent individual data for the 12 participants.

Exp. 1: Number of bands

Sinewave vocoder does not allow simulation of the spread of excitation that happens in actual CIs. It also provides spectral pitch cues which are not present in real implants. Here we test various carriers that do not provide spectral cues, and which have more or less flat envelopes, providing more or less clear F0-modulation information. The main objective is to evaluate the effect on VTL JNDs.

Exp. 2: Type of carrier

Sinewave vocoder does not allow simulation of the spread of excitation that happens in actual CIs. It also provides spectral pitch cues which are not present in real implants. Here we test various carriers that do not provide spectral cues, and which have more or less flat envelopes, providing more or less clear F0-modulation information. The main objective is to evaluate the effect on VTL JNDs.

Exp. 3: Spread of excitation

Using a noise carrier, spread of excitation can be simulated by changing the order of the bandpass filters. 4th order filters (-48 dB/oct.) are very shallow and 12th order filters (-72 dB/oct.) are very sharp. Only the VTL JNDs were measured in this experiment.

Exp. 4: Place-frequency shift

Another characteristic of CIs is the fact that the frequency band allocated to a specific electrode may not correspond to frequency to which the place is tuned. This place-frequency mismatch can be quantified in millimeters along the basilar membrane and can be simulated by shifting the synthesis filter in the vocoder by a fixed distance. Twelve bands were used, with two filter orders and two vocoder settings.

Figure 8 shows VTL JNDs for 4th and 6th order filters, as a function of place-frequency shift. The light symbols represent individual data for the 12 participants.