Barycentre spectral et perception de la hauteur de la voix et de la longueur de tractus vocal chez les normo-entendants et les utilisateurs d’implants cochléaires

Etienne Gaudrain

To cite this version:


HAL Id: hal-01558903
http://hal.univ-brest.fr/hal-01558903

Submitted on 19 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Introduction

Voice differences are useful for listeners to discriminate competing sentences (Brungart, 2001, Darwin et al., 2003, Baskent and Gaudrain, 2016; see Figure 1). The perception and discrimination of voices relies on two principal cues: the fundamental frequency (F0) and vocal-tract length (VTL). F0 is related to voice pitch. VTL is related to the size of the speaker and affects the spectral envelope such that all the formants are shifted by the same ratio when VTL is changed (Figure 2).

While normal-hearing (NH) listeners have good access to these cues, cochlear-implant (CI) users show much enlarged discrimination thresholds along these dimensions (Figure 3). This deficit, especially along the VTL dimension, is likely responsible for difficulties in identifying speaker gender (Massida et al., 2013, Fuller et al., 2014) and in taking advantage of voice differences for speech-on-speech intelligibility (adapted from Başkent and Gaudrain, 2016).

The objective of the present study is to determine whether spectral centroid could be a plausible cue for VTL perception, first in NH listeners, and then in CI listeners. While VTL undoubtedly affects the spectral centroid, the very nature of the speech signal also warrants that the spectral centroid is unlikely to contribute to the observed JNDs in NH listeners. However, the cue could be used by CI listeners for both F0 and VTL discrimination tasks.

Methods

The spectral centroid was calculated using 7 different methods. Spectral centroid is defined by the formula on the right. The weights w and the frequency f can be expressed in various units that lead to different centroid values. For each method, distributions were obtained over the corpus of 6 Dutch CV-syllables used to obtain the JNDs shown in Figure 3.

The d' estimates indicate the spectral centroid is unlikely to be used by CI listeners. However, the cue could be used by CI listeners for both F0 and VTL discrimination tasks.

References


