Voice discrimination by cochlear implant users.

Verena Pyschny[1], Martin Walger[2], Julia Weber, Hasso von Wedel[2], and Hartmut Meister[1]

[1] Jean-Uhrmacher Institute for Clinical ENT-Research, University of Cologne
[2] University ENT-Hospital Cologne, University of Cologne

Introduction

The voice of an individual is an important attribute, since vocal characteristics contribute to differences that help to distinguish one speaker from another. The fundamental frequency (f0), which corresponds to perceived voice pitch, and the vocal-tract (VT), which can be described in terms of formant frequencies, are two important measures for voice classification.

Due to the limited processing of spectral cues, voice discrimination might be restricted in cochlear implant (CI) users. A very basic representation of different voices is given by the speaker gender.

In order to examine the underlying cues, f0 and the length of the vocal-tract (VTL) respectively were manipulated and speaker gender discrimination and identification were tested in CI-users. These manipulations were performed using a script from Darwin et al (2003) who examined the role that changes in these both parameters had on the ability to attend to one of two competing talkers in normal listeners.

In our study the range of the VT ratios was interpolated from those given in Darwin et al and exceeds the average formant-frequency ratio between female and male voices. Thus, it was ensured that not only normal hearing listeners (NL) but also CI-users are able to recognize differences.

Some hypotheses could be made based on a study of Fu et al. (2005) who examined voice gender identification by CI-users. It is hypothesized that speaker gender identification is generally possible in CI-recipients. Following Darwin et al. (2003) changes in f0 should be easier to recognize than changes in VTL at least in NL. Thus, speakers of similar gender should be more difficult to discriminate and identify. Changes of both parameters affect the results additively.

Subjects and Methods

Stimuli were generated from an original utterance (male voice: f0 = 156 Hz; VTL = 13.5 cm). A short phrase (“Die Krankenschwester ißt einen Apfel.”) was manipulated in terms of lowering or increasing f0 and/or modifying the length of the vocal-tract and thus changing the formant frequencies. Manipulations were done with the software package ‘Praat’ (Boersma et al., 1996) using the above mentioned “change gender script” by Darwin et al. (2003).

Two experiments were conducted. The first one was a gender discrimination experiment using a three alternative forced choice (3-AFC) procedure the subjects were asked to indicate the stimulus which differed from the other two presented.

The second experiment was dedicated to gender identification. Using a 2-AFC procedure the subjects’ task was to indicate if the stimulus was spoken by a male or a female talker. Tab. 1 shows the set of nine different stimuli (labelled 2 to -2) used with the gender identification paradigm.

Signals were presented via a single loudspeaker with 0 azimuth at 70 dB SPL in a sound proofed booth.

Subjects were six postlingually deafened adult CI-recipients with an average age of 58 years (39 - 71 years) and a control group of six normal hearing listeners (NL) with an average age of 28 years (23 - 38 years).

Stimuli	f0 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]
2.0 (female)	199.96	422.86	1609.5	2911.9
1.5	187.65	419.73	1597.4	2875.4
1.0	174.75	410.83	1548.7	2790.2
0.5	166.13	404.52	1524.3	2753.7
0	156.7	392.35	1487.8	2680.7
-0.5	135.27	380.18	1439.1	2595.5
-1.0	127.20	378.34	1421.1	2571.1
-1.5	114.54	365.33	1390.4	2498.1
-2.0 (male)	110.98	362.72	1378.3	2473.7

Tab. 1: Parameter changes (in Hertz) for the 9 different stimuli presented with the gender identification task

Results

Experiment 1:

For the discrimination experiment data were collapsed across subjects and displayed in relation to f0 and the formant frequencies (Fig. 1). CI-recipients needed about two times higher changes in f0 compared to the original stimulus than the control group of normal listeners (NL).

Results for VTL-based changes (i.e. formant frequencies) were undetermined, since some CI- recipients unfortunately needed much higher differences than had been created for the experiment. This exceeded discrimination thresholds beyond 25% differences. NL needed 4.4% differences.

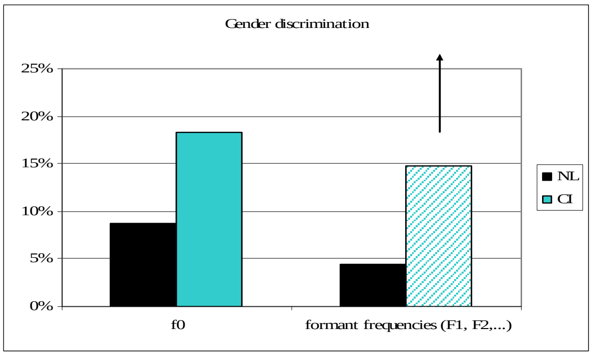

Fig. 1: Results for the gender discrimination task for just noticeable differences in f0 or formant frequencies based on VTL-changes (in %). Baseline values: f0: 134 Hz; F1: 736 Hz; F2: 1310 Hz; F3: 2479 Hz

Experiment 2:

Fig. 2 shows the results for the gender identification experiment based on changes in f0, VTL or both parameters combined. The steep increase of the curves with changes in f0 and changes in f0 and VTL (upper left and right panel) revealed good identification, since virtually every stimulus is consistently assigned to either a male or a female voice.

Results of the separated VTL changes (bottom middle panel) were near chance level (p=0.5). Thus, subjects of both groups were not able to identify the speaker gender based on VTL-changes alone.

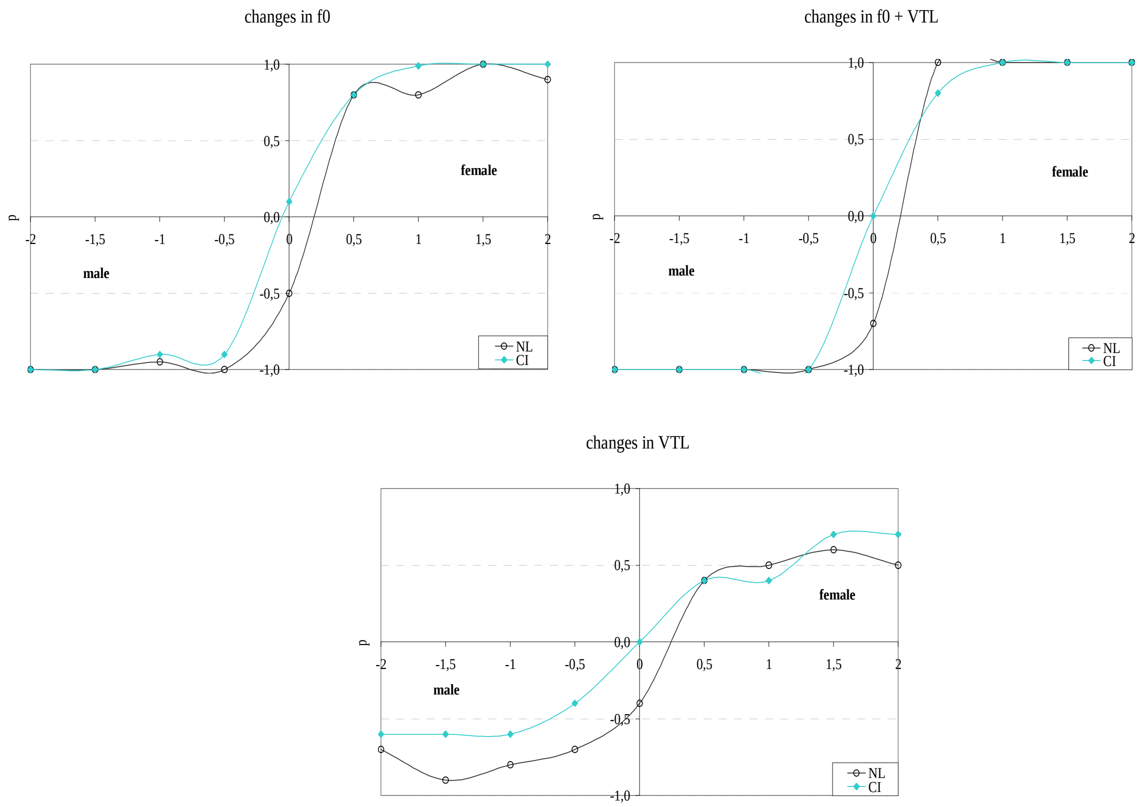

Fig. 2: Results for the gender identification task (chance response p = 0.5). Upper left panel: changes in f0, upper right panel: changes in f0 and VTL, bottom middle panel: changes in VTL.

Discussion

The aim of the present study was the investigation of CI-listeners' ability to discriminate and identify different voices depending on changes of the two characteristic voice parameters, namely f0 and formant frequencies.

Results revealed that f0 is the most dominant factor to discriminate voices for both, NL and CI-recipients. However, as can be seen in Fig. 1 just noticeable differences of CI-users exceed that of NL by a factor 2. Changes in formant frequencies alone frequently couldn't be perceived correctly by CI-users even with the greatest dissimilarity between the stimuli. This might be due to the restricted spectral resolution in Cochlear Implants.

The speaker gender identification experiment did not show large differences between NL and CI-recipients (see Fig. 2). For both listener groups, comparison across conditions revealed that the identification of male and female voices is possible when manipulating f0 alone or the combination of f0 and VTL.

In contrast, separated VTL-changes (i.e. formant frequencies) are not sufficient to identify the speaker gender. This shows that the VT alone is no sufficient parameter for identifying speaker gender. However, adding VTL-changes to alterations in f0 seems to improve gender identification at least for the NL whereas CI-recipients don't have an additional benefit.

In conclusion, voice gender identification based on differences in fundamental frequency is basically possible in CI-users. However, for voice discrimination in general more difficulties can occur since e.g. differences in voices of the same sex are mainly based upon changes in formant frequencies.

Acknowledgements

We would like to thank our subjects for their time and dedication. This research was supported by Cochlear ltd, Germany.

References

Boersma P, Weenink D (1996): Praat, a system for doing Phonetics by Computer, version .4.22. Institute of Phonetic Sciences, University of Amsterdam; www.praat.org.

Darwin C J, Brungart D S, Simpson B D (2003): Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. J. Acoust.

Soc. Am. 114 (5), 2913 – 2922.

Fu Q-J, Chinchilla S, Nogaki G, Galvin III J J (2005): Voice Gender identification by cochlear implant users: The role of spectral and temporal resolution. J. Acoust. Soc.

Am. 118 (3) , 1711–1718.