The perception of prosodic cues in normal listeners and cochlear implant recipients

Hartmut Meister[1], Verena Pyschny[1], Markus Landwehr[1], Martin Walger[2], Hasso von Wedel[2],
[1] Jean-Uhrmacher Institute for Clinical ENT-Research, University of Cologne
[2] Clinic of Othorhinolaryngology, Head and Neck Surgery, University of Cologne

Introduction

Prosody reflects rhythmic and melodic aspects in speech and might carry information e.g. as with sentence stress or word stress and questions vs. statements. Prosodic cues are based on spectral, temporal, and amplitude-related changes in the signal. The most prominent factor is the course of the fundamental frequency (F0): Typically, the pitch of the voice is higher for stressed syllables or at the end of a question.

Cochlear implants (CIs) are limited in the processing of the above mentioned parameters, especially F0. Thus, it is interesting to ask to what extend CI recipients are able to perceive prosodic cues based on changes in F0. Therefore, a study was conducted which is based on defined modifications of F0 of a natural utterance and on the assessment of the resulting changes in perception. This might give further insight into the processing of prosodic cues in both, normal listeners and CI users.

Methods and Subjects

Stimuli were based on natural utterances with defined changes of F0-contours. Modifications were performed with the PSOLA (pitch-synchronous add and overlay) algorithm embedded in the signal processing software PRAAT©. Three different experiments addressing alterations of F0 were conducted, namely sentence stress, questions vs. statements, and a speaker gender experiment. Modifications of F0 within the different experiments are schematically depicted in Fig. 1. The tab. 1 shows the values of the corresponding changes in F0. Stimuli were processed within nine different steps labelled from -2 to +2.

The stimuli were presented to the subjects in a free field condition at 70 dB SPL. The task for the participants was to indicate, whether stress was on the subject or the verb in a short phrase (experiment 1), whether it was a statement or a question (experiment 2) and whether the stimulus was spoken by a male or a female speaker (experiment 3). Therefore, a 2 AFC-procedure was used.

Twelve normal listeners (NL) aged 34 to 68 years (mean 47 years) and 12 CI recipients aged 38 to 75 years (mean 58 years) participated in the study.

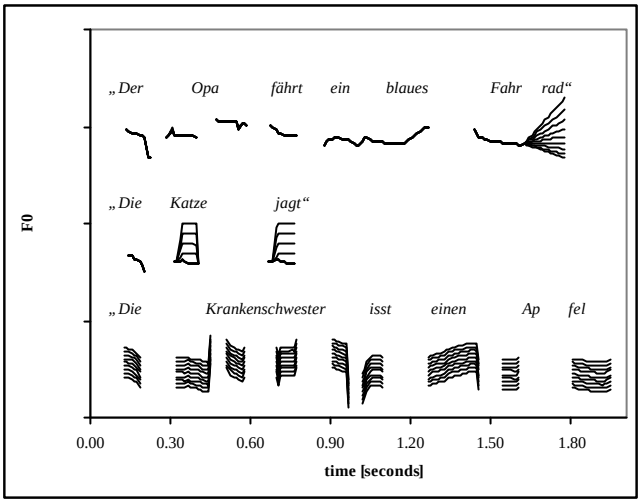


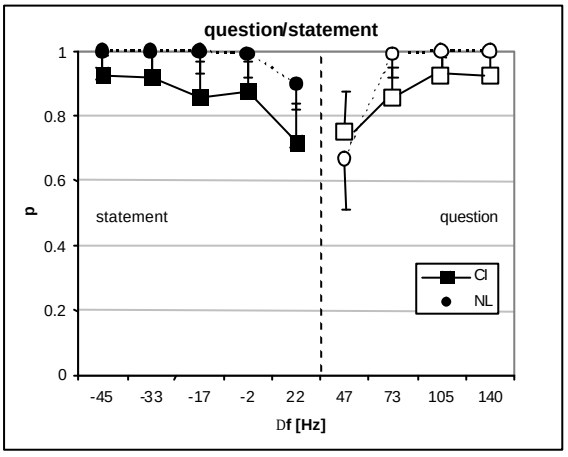
Figure 1: Schematically depicted changes in F0-contour for the three different experiments. Upper panel: question vs. statement, middle panel: sentence stress, lower panel: speaker gender

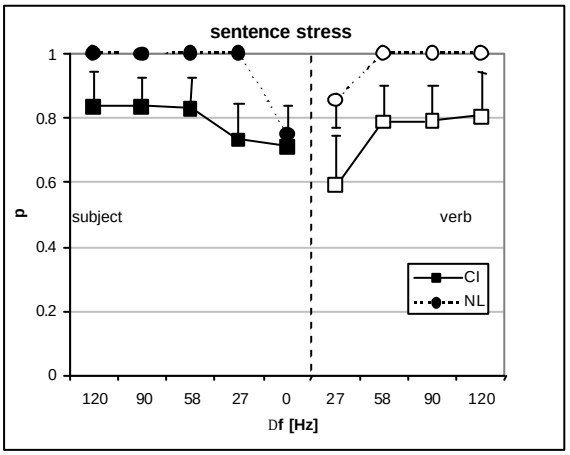
	Question vs. Statement	Sentence Stress	Speaker Gender
Category unit	ΔF0 [Hz]	ΔF0 [Hz]	F0 [Hz]
-2	140	120	120
-1.5	105	90	131
-1	73	58	142
-0.5	47	27	153
0	22	0	164
0.5	-2	27	175
1	-17	58	186
1.5	-33	90	197
2	-45	120	208

Tab. 1: Changes in F0 for the three different experiments.

Results

Figure 2 shows the results for the three experiments (upper left panel: questions vs. statements, upper right panel: sentence stress, middle panel: speaker gender). Percent correct scores are plotted against the changes in F0 corresponding to tab. 1. The questions/statements-paradigm reveals similar results for NL and CI users in that the break-over point is comparable between both study groups. A slight increase of F0 is still identified as a statement, since changes in the fundamental frequency serve also as cues for phrase boundaries. Higher changes (i.e. +47 Hz) yield the perception of a question. However, in contrast to the normal listeners who show scores of 100% correct even for small alterations in F0, the CI recipient's perception seems in general slightly restricted with values around 90-95%. Clearly worsened is the perception of sentence stress in the CI users. Whilst the break-over point is comparable again, the CI subjects reach correct responses of maximally 85% even for very high alterations in F0 of about 120 Hz. In contrast, the speaker gender results do not reveal large differences between NL and CI users with the latter showing slightly more uncertainty.





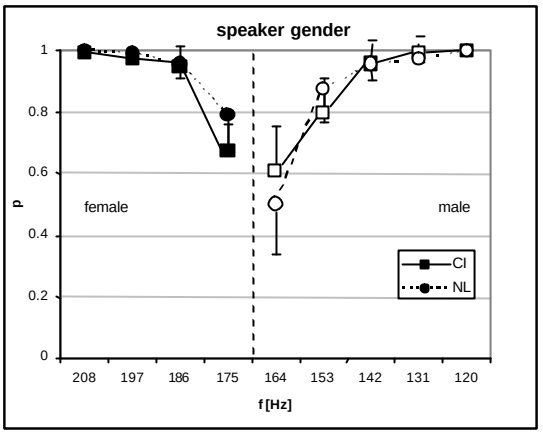


Figure 2: The amount of correct responses for both subject groups as a function of changes in f0.

Discussion

The study revealed that prosodic cues are similarly identified by normal listeners and cochlear implant recipients in that the break-over points between questions and statements, stress on different words in a phrase and between male and female voices, are comparable. Qualitatively, the CI users show worse results at least for the question vs. statement paradigm and the sentence stress procedure. Especially with the latter one, prosody perception is reduced even for very large alterations in F0. In contrast, problems with the speaker gender recognition based on F0 alterations are comparably low in CI users. This might be explained with the fact that with experiment 3 changes in the fundamental frequency were present over the entire length of the phrase (approx. 2 sec). The other two experiments were based on very short alterations of F0 of approx. 140 ms (questions vs. statements) and 100 ms (sentence stress), respectively.

Another interesting finding was that the CI recipients showed considerable interindividual scatter in their performance. Whilst some of them performed closely to the normal listeners others were hardly able to identify even the largest changes in F0 consistently. This effect could not only be attributed to the signal coding strategies

used by the subjects, since use of the same processors and implants yielded different performance. That is, besides technical aspects, auditive as well as cognitive factors might also play a role. However, since clinically used CI systems as a rule do not attempt to explicitly code F0, progress in coding strategies might also improve the perception of prosodic cues in future. It might also be speculated, that a combination of electric and acoustic hearing as with bimodal fitting and hybrid implants might improve prosody perception due to the transmission of fine structure in the low frequency area.

Acknowledgements

Parts of this study were supported by Medical Electronics, Med-El. Innsbruck.