

Effect of Different Frequency Mappings on Speech Intelligibility for Cochlear Implant Listeners

Matthew J. Goupell¹, Bernhard Laback¹, Piotr Majdak¹, Wolf-Dieter Baumgartner²

¹ Acoustics Research Institute, 1010-Vienna, Austria, Email: goupell@kfs.oeaw.ac.at

² ENT-Department, University Hospital, 1090-Vienna, Austria

Introduction

There have been several studies that have tested the number of channels necessary for good speech understanding in cochlear implant (CI) listeners or normal hearing (NH) listeners using a CI simulation (e.g. [1]). Studies in this area show a saturation in performance between three to eight channels in quiet. In noise, a larger number of channels has been shown to be beneficial for NH listeners but not necessarily for CI listeners (e.g. [2]).

There have also been studies that varied the mapping of frequency-to-tonotopic place. The conclusion of these studies has been that a well-matched frequency-to-place mapping (e.g. [3]) was the best. However, there appears to be a benefit of not removing lower frequencies for the sake of a well-matched frequency-to-place mapping (e.g. [4]).

Methods

Seven CI listeners and six NH listeners using a CI simulation listened to Oldenburg sentences [5] processed by the CIS strategy. Two parameters were varied to change the frequency-to-place mapping, the number of channels or electrodes (N) and the spectral content (M). The variable M controls the amount of spectral content by varying the upper frequency boundary. The values of M used in this experiment were $M = 12, 10, 8, 6,$ and 4 which correspond to an upper-frequency boundary $f = 8.5, 4.9, 2.8, 1.6,$ and 0.9 kHz respectively. The values of N used in this experiment were $N = 12, 10, 8, 6,$ and 4 channels. For convenience the conditions were labeled by, for example, $M_{12}N_{12}$ for the baseline twelve-channel condition with frequencies up to 8.5 kHz ($M = 12$). CIs typically use this condition. By testing all combinations of $M = 12 - 6$ and $N = 12 - 6$, the condition M_4N_4 , and an extended frequency condition, $M_{14}N_{12}$, with an upper-frequency boundary $f = 16$ kHz, 18 conditions were tested in total.

Listeners verbally repeated words from the Oldenburg sentences and an experimenter recorded the number of correct words. Listeners listened to 90 sentence blocks for one frequency-to-place mapping. The first 10 sentences were warm-up stimuli, not counted in the experiment. The percentage of correct words (P_c) was found from the next 80 sentences that were presented at four different signal-to-noise ratios (SNR): $0, +5, +10$ dB and in quiet. There were 20 sentences for each SNR and the different SNRs were presented in random order.

Results

Figure 1 shows the results for the matched conditions, where $M = N$. Not all the CI listeners were included because there were two populations: high-performance and low-performance. The high-performance CIs had $P_c > 90\%$ for $M_{12}N_{12}$ in quiet like the NHs. The low-performance CIs had $P_c \approx 75\%$ for $M_{12}N_{12}$ in quiet, had highly variable data between conditions, and thus were omitted from the plot. The performance improves with increasing N for both the CI and NH listeners. Helmert contrasts were calculated to determine the level of the factor N above which no further improvement in performance occurs. For all seven CI listeners, increasing the number of channels results in a significant improvement ($p < 0.05$) of performance up to $N = 8$ for all four SNRs. For the NH listeners, increasing the number of channels results in a significant improvement of performance up to $N = 8$ in quiet and up to $N = 10$ for all conditions with noise.

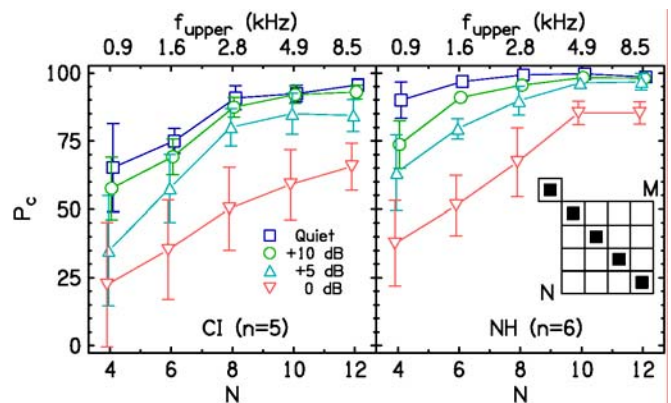


Figure 1: Percent correct scores as a function of matched channels $M = N$ for five high-performance CI listeners (left panel) and six NH listeners (right panel). Two low-performance CI listeners were not included. Data points have error bars that are two standard deviations in length.

Figure 2 shows the results for the unmatched conditions ($M \neq N$). Figure 2 shows ΔP_c , the change in P_c from the matched conditions, as a function of number of channels in each panel. Therefore, the matched condition, which is indicated by the dotted vertical line in each panel, has exactly zero ΔP_c and no error bars. The data are plotted as a difference to better identify relative trends in the data and to combine the high-performance and low-performance CI listeners. Note that on this plot the points to the left of the dotted line represent conditions with a decreased frequency resolution, namely less channels, for a constant M . Points to the right of the dotted line represent conditions with increased frequency resolution, namely more channels.

In Fig. 2, the matched case was often the best listening condition. Significant decreases from the matched condition (solid symbols) are seen for mostly conditions that differ by more than two channels from the matched case. The same occurs if the data is plotted as a function of M , instead of N . One notable exception is the extended frequency range condition $M_{14}N_{12}$ (not shown). It was found that CIs had significantly worse speech understanding for this condition compared to $M_{12}N_{12}$, but not NHs.

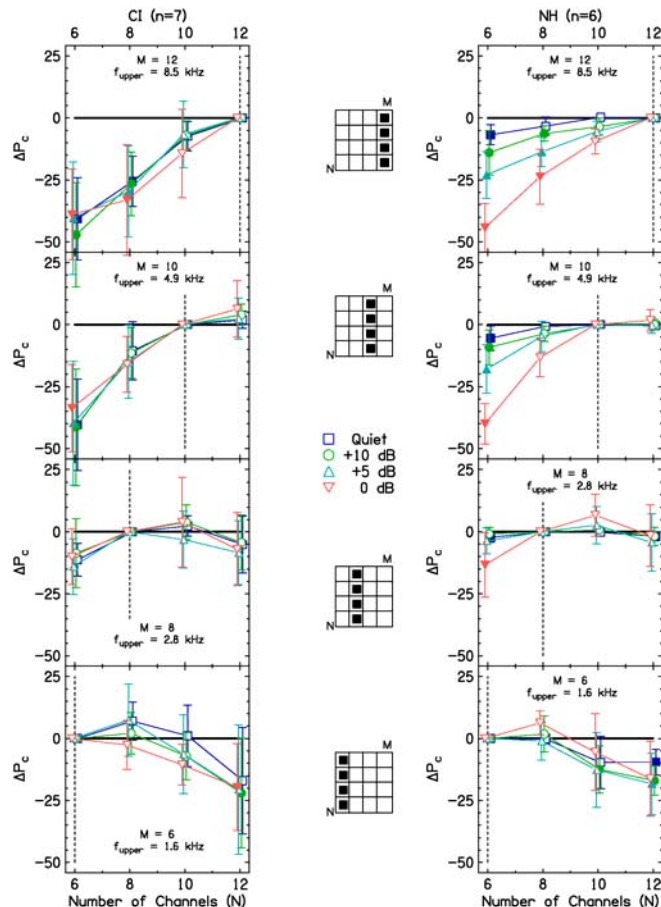


Figure 2: Percent correct scores relative to the matched conditions (shown by the vertical dotted line) vs the number of channels (N) for fixed spectral content (M). Closed symbols show a significant difference from the matched case ($p < 0.05$). The matched case was often the best case. Significant decreases were seen mostly for changes of more than two channels from the matched case.

An analysis to find the number of conditions not different from the baseline condition, $M_{12}N_{12}$, was performed. It was found that the conditions where speech understanding is equal to the $M_{12}N_{12}$ case are: $M_{10}N_{12}$, $M_{12}N_{10}$, $M_{10}N_{10}$, and M_8N_8 for both CI and NH listeners. An additional condition, M_8N_{10} , was also not different from the baseline for CI listeners only.

Discussion and Conclusion

Speech understanding is found to improve significantly when the number of matched channels increases up to 8 in case of the CI listeners for all SNRs, and up to 8 or 10 channels in case of the NH listeners, depending on the SNR. A different way of interpreting the results of the CI listeners is that speech understanding does not decrease significantly

when the number of matched channels is reduced to 8. These values coincide with those reported in the literature [1,2].

Our results for both the CI and NH groups show that there is approximately a “plus-or-minus two rule”, where increasing or decreasing either spectral range (M) or number of channels (N) by two causes no significant decrease in speech understanding performance compared to the nearest matched condition, which is consistent with other studies [1,3].

When comparing our results with previous studies, it is important to recall that we varied the number of channels by changing the upper frequency boundary and held the lower constant, while all cited studies (except for [4] who varied the upper frequency boundary) held the frequency range constant and varied the number of channels within this fixed range.

The purpose of this study was to see if a decrease in the number of channels or a decrease in the amount of spectral information presented to CI listeners causes a decrease in speech understanding. Since there are conditions that show no decrease compared to the baseline, it may be possible to use a CI for mapping spectral localization cues without affecting the speech understanding. The most important spectral cues used for vertical plane localization occur between 4 and 16 kHz. Current processing strategies that use spectral information up to 8.5 or 10 kHz (depending on processor type) would already be able to implement peaks and notches in the lower half of the important spectral region. However, a direct mapping of frequency information may not be possible because the extended frequency range condition showed a significant decrease in speech understanding from the baseline for CI listeners.

Literature

- [1] Dorman, M. F., Loizou, P. C., Fitzke, J., and Tu, Z. (1998). The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels. *J. Acoust. Soc. Am.* 104, 3583-3585.
- [2] Garnham, C., O'Driscoll, M., Ramsden, R., and Saeed, S. (2002). Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets. *Ear Hear.* 23, 540-552.
- [3] Baskent, D., and Shannon, R. V. (2004). Frequency-to-place compression and expansion in cochlear implant listeners. *J. Acoust. Soc. Am.* 116, 3130-3140.
- [4] Baskent, D., and Shannon, R. V. (2005). Interactions between cochlear implant electrode insertion depth and frequency-to-place mapping. *J. Acoust. Soc. Am.* 117, 1405-1416.
- [5] Wagener, K., Brand, T., Kühnel, V., and Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache. I. Design des Oldenburger Satztests. *Zeitschrift für Audiologie/Audiological Acoustics* 38, 4-14.

This study was funded by the Austrian Science Fund, FWF, project number P18401-B15.