Frequency-Place Mapping and Speech Intelligibility: Implications for a Cochlear-Implant Localization Strategy

Piotr Majdak[1], Matthew Goupell[1], Bernhard Laback[1], and W-D. Baumgartner[2]

[1] Acoustics Research Institute, Austrian Academy of Sciences, Austria

[2] ENT-Department, Vienna University Hospital, Austria

ARO 31st MidWinter Meeting

Association for Research in Otolaryngology

February 16-21, 2008, Phoenix, Arizona, USA

INTRODUCTION

Localization of sounds in the vertical plane is primarily based on spectral cues. The information in these cues is related to peaks and notches in the head related transfer functions and is most salient at higher frequencies (above 4 kHz). For the speech, the most important frequency region is below 4 kHz.

The primary function of a cochlear implant (CI) is to provide speech understanding to profoundly deaf people. Thus, for the implementation of a spatial localization strategy in CIs, it is important to combine the speech and spectral cues. However, combining the speech information and spectral cues can cause a degradation in speech understanding. This is because of a low frequency resolution in CI listeners (up to 20 frequency channels only) and a reduced sensitivity to spectral peaks and notches.

A potential way to include the spectral cues in a processing strategy could be to present the speech information and spectral cues on different electrodes. This even may include a change in the frequency-to-place mapping resulting in a frequency warping. Of course, with such mappings, good speech understanding in quiet and noise without spectral cues is a prerequisite for further investigations.

In this study we changed the upper-frequency boundary of speech information and used frequency warping to determine the number of electrodes required for good speech understanding. This allows us to determine if an inclusion of spectral cues is possible in future CI processing strategies.

EXPERIMENT I: ACUTE TESTS

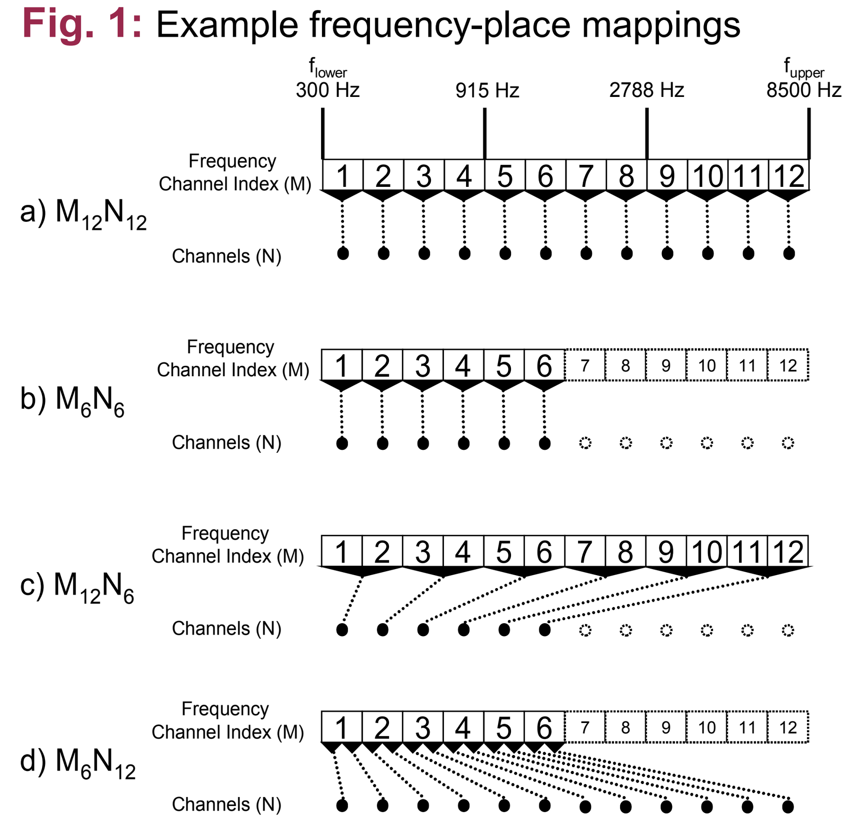
1. Subjects

- 7 CI listeners using direct electric stimulation
- 6 NH listeners using an acoustic CI simulation (noise-vocoder)

2. Stimuli

- OLSA Sentences: 5-word German nonsense sentences
- The frequency-to-place mapping (Fig. 1) was altered by:
  - ➢ varying the upper-frequency boundary (M) while holding the lower-frequency boundary fixed at 300 Hz, and
  - ➢ varying the number of electrodes (N).

Fig. 1: Example frequency-place mappings



3. Conditions (Fig. 2)

- 18 combinations of M and N presented in quiet and in noise (SNR: +10 dB, +5 dB, and 0 dB).

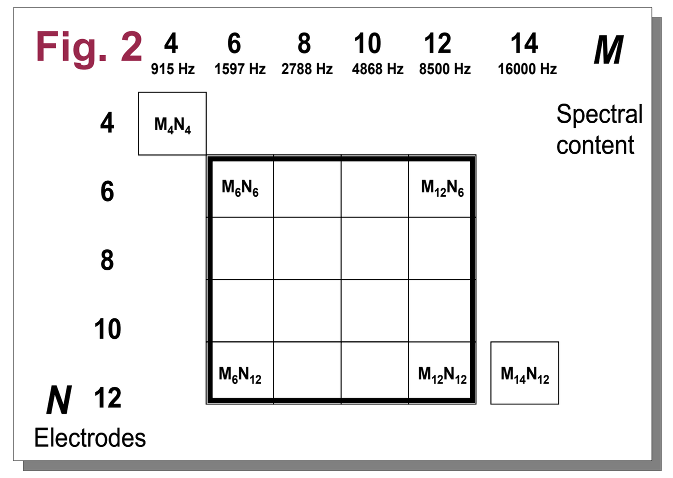
4. Procedure

- Verbal repetition of the words, the number of correct words was recorded.
- One block contained one combination of M and N with 10 presentations as warm-up (condition: quiet) and 20 presentation of each SNR (random order).
- Training with $M_{12}N_{12}$ (4 to 14 blocks) until stable P_C
- Remaining 17 blocks were presented in balanced order.

3. Statistical Analysis

- Factor effects: RM ANOVA
- Saturation tests: Helmert contrasts

Fig. 2



4. Results: Matched Conditions (M = N = N'; Fig. 3):

- Saturation in performance with increasing N' up to:
  - ➢ NH, in quiet: N' = 8 (N'_6 vs. $N'_{>6}$: $p < 0.05$; N'_8 vs. $N'_{>8}$: $p = 0.89$).
  - ➢ NH, in noise: N' = 10 (N'_8 vs. $N'_{>8}$: $p < 0.05$; N'_{10} vs. $N'_{>10}$: $p > 0.63$).
  - ➢ CI, in quiet and 5 dB SNR: N' = 8 (N'_6 vs. $N'_{>6}$: $p < 0.05$; N'_8 vs. $N'_{>8}$: $p > 0.061$).
  - ➢ CI, in 0 and 10 dB SNR: N' = 10 (N'_8 vs. $N'_{>8}$: $p < 0.05$; N'_{10} vs. $N'_{>10}$: $p > 0.44$).
- Agrees with Garnham et al. (2002): Eight channels required (lower and upper frequency boundaries constant).
- Agrees with Başkent and Shannon (2005): Eight channels required (upper-frequency boundary constant).
- Number of required channels appears to be independent of type of spectral manipulation when no frequency warping is involved.

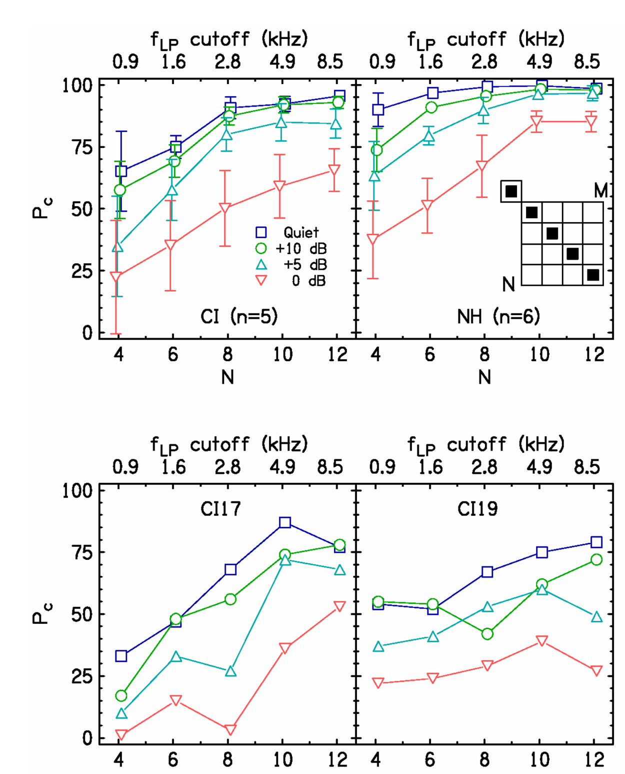


Fig. 3: Results for matched conditions (M = N = N').

Upper-right panel: NH listeners

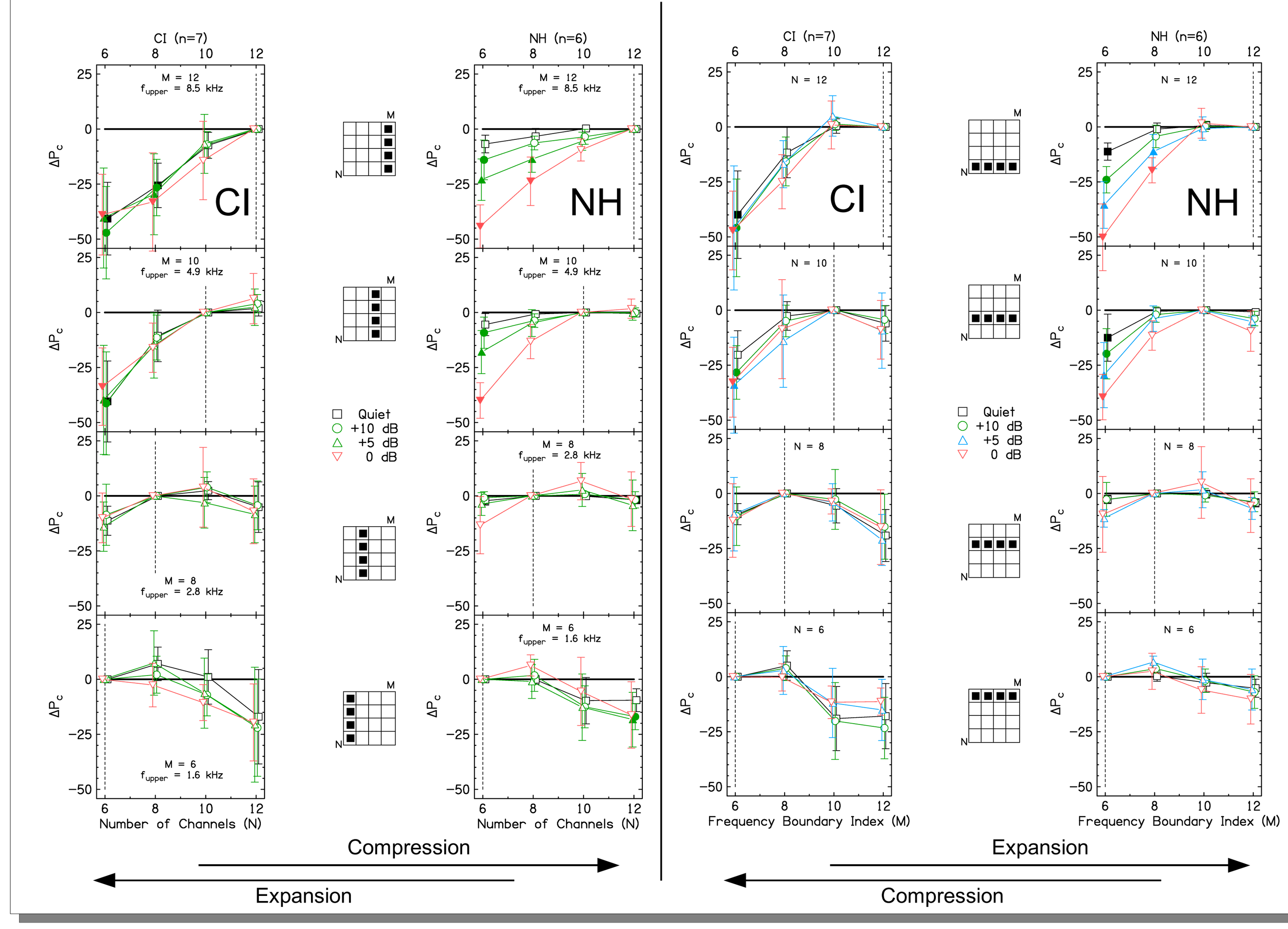
Upper-left panel: High-performance CI

Lower panels: Low-performance CI

2. Results: Unmatched Conditions (M ≠ N; Fig. 4):

- Conditions not different from baseline ($M_{12}N_{12}$; $p > 0.05$):
  - ➢ NH: $M_{10}N_{12}$, $M_{12}N_{10}$, $M_{10}N_{10}$, M_8N_8.
  - ➢ CI: $M_{10}N_{12}$, $M_{12}N_{10}$, $M_{10}N_{10}$, $M_{10}N_8$, M_8N_8.
- Changes greater than ± 2 N (electrodes) from the matched conditions result in sig. decrease in P_C.
- Changes greater than ± 2 M (upper-frequency boundary) from the matched conditions result in significant decrease in P_C. This corresponds to a shift of ± 0.77 octaves of the most basal electrode and a decreasing shift towards the most apical electrode. This shift is greater than other studies that do not hold the lower frequency boundary fixed (e.g. Başkent and Shannon, 2004).
- Asymmetric decrease in ΔP_C: Removing spectral content or electrodes is more detrimental than adding spectral content or electrodes.

Fig. 4: Results for the unmatched conditions: P_C scores relative to the matched conditions. Left panels: Fixed spectral content (M). Right panels: Fixed number of channels (N). Vertical dotted lines show the matched conditions. Filled symbols show significant differences (p < 0.05).



3. Extended frequency range ($M_{14}N_{12}$)

- Contains frequencies up to 16 kHz and results in a frequency shift of -0.88 octave for the most basal electrode.
- Significant decreases for CIs in conditions quiet and 10 dB SNR (see Tab. 1).
- The shift is a little bit higher than the shift in $M_{12}N_{10}$ (-0.77 oct.), for which the performance does not significantly decrease. Upper frequency boundary of 14.5 kHz (-0.74 oct.) may be appropriate in the future.
- Difference between NH and CI: possible ceiling effects in NH listeners' results.

Tab. 1: Results for extended frequency range conditions.

CI SNR (dB)	$M_{12}N_{12}$ Average	$M_{12}N_{12}$ Stand. Dev.	$M_{14}N_{12}$ Average	$M_{14}N_{12}$ Stand. Dev.	Difference	p
Quiet	90.57	8.79	79.29	13.50	-11.29	**0.002**
+10	87.86	9.19	72.14	15.14	-15.71	**0.004**
+5	76.86	15.94	61.86	22.14	-15.00	0.064
0	56.86	16.22	36.43	26.44	-20.43	0.054

NH SNR (dB)	$M_{12}N_{12}$ Average	$M_{12}N_{12}$ Stand. Dev.	$M_{14}N_{12}$ Average	$M_{14}N_{12}$ Stand. Dev.	Difference	p
Quiet	98.50	1.76	98.67	1.03	0.17	0.72
+10	98.00	1.55	97.83	1.17	-0.17	0.41
+5	96.67	3.01	95.17	4.45	-1.50	0.43
0	85.33	4.08	88.33	3.98	3.00	0.19

EXPERIMENT II: TRAINING

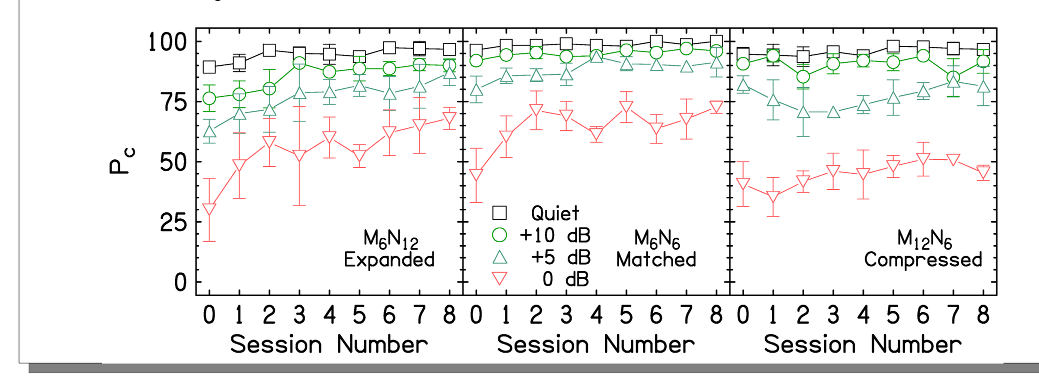
1. Methods (differences to Exp. I):

- One session s_i: feedback training with subsequent testing.
- Two tests: (M_6N_6 and M_6N_{12}) and (M_6N_6 and $M_{12}N_6$).
- Eight sessions (s_1 to s_8) per test. Three subjects, NH listeners only.

2. Results (Fig. 5):

- M_6N_6: no sig. learning after 2nd session (s_1 vs. $s_{>1}$: $p = 0.002$, s_2 vs. $s_{>2}$: $p=0.36$), performance saturated within 8 sessions ($s_{6,7,8\ \text{test 1}}$ vs. $s_{6,7,8\ \text{test 2}}$: $p = 0.51$).
- M_6N_{12}: no sig. learning after 3rd session (s_2 vs. $s_{>2}$: $p = 0.007$, s_3 vs. $s_{>3}$: $p=0.17$).
- $M_{12}N_6$: no sig. learning at all (s_1 vs. $s_{6,7,8}$: $p = 0.1$).
- All three conditions before and after training are significantly different ($p < 0.001$) to the untrained baseline condition $M_{12}N_{12}$. Thus, the training did not change the relative differences between the conditions.

Fig. 5: Learning experiment results for 3 NHs over 8 sessions. Session s_0 is from the unmatched conditions in Fig. 4.



CONCLUSIONS

1. Increasing the number of matched channels beyond **eight to ten** does not improve speech understanding substantially, depending on SNR.
2. Slight changes in frequency-place mapping do not cause significant decreases:
   - Conditions: $M_{10}N_{12}$, $M_{12}N_{10}$, $M_{10}N_{10}$, M_8N_8, or
   - Generally, **±2 electrodes** or **±0.77 octaves** (for the most basal electrode, with fixed lower-frequency boundary).
3. The results indicate that eight electrodes might be used as "**speech electrodes**" and the remaining four could be used as "**spatial electrodes**" in a sound localization strategy. However, more than a straight forward frequency mapping, as was done for $M_{14}N_{12}$, is required in a future CI sound localization strategy.

REFERENCES

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.

Başkent, D., and Shannon, R. V. **(2004).** "Frequency-to-place compression and expansion in cochlear implant listeners," J. Acoust. Soc. Am. **116**, 3130-3140.

Başkent, 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.

Goupell, M. J., Laback, B., Majdak, P., and Baumgartner, W-D. **(2008).** "Effects of upper-frequency boundary and spectral warping on speech intelligibility in electrical stimulation," J. Acoust. Soc. Am. **123** (in press).

Corresp. Author: Piotr Majdak, Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, A-1040 Wien, Austria

E-Mail: piotr@majdak.com

http://www.kfs.oeaw.ac.at

We thank Mr. M. Mihocic for running experiments.
We also thank the CI listeners for their enthusiastic participation in this study.
The equipment for direct electrical stimulation was provided by MED-EL Corp.
This research was supported by the Austrian Science Fund (Project P18401-B15) and the Austrian Academy of Sciences.