Technical University of Denmark

Interaural Place-Mismatch Estimation With Two-Formant Vowels in Unilateral Cochlear-Implant Users

Guérit, François; Santurette, Sébastien; Chalupper, Josef; Arweiler, Iris

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Guérit, F., Santurette, S., Chalupper, J., & Arweiler, I. (2014). Interaural Place-Mismatch Estimation With Two-Formant Vowels in Unilateral Cochlear- Implant Users. Poster session presented at Association for Research in Otolaryngology 37th MidWinter Meeting, San Diego, CA, United States.

Interaural Place-Mismatch Estimation with Two-Formant Vowels in Unilateral Cochlear-Implant Users

PS-492

François Guérit[1], Sébastien Santurette[1], Josef Chalupper[2], Iris Arweiler[2], Torsten Dau[1]

[1]*Centre for Applied Hearing Research, Technical University of Denmark, Denmark*, [2]*Advanced Bionics European Research Center GmbH, Germany*

Introduction

For cochlear implant (CI) users with residual hearing in the contralateral ear, usually a default frequency-to-electrode map is used in the CI. This assumes that the human brain can adapt to interaural place-pitch mismatches.

This "one-size-fits-all" method might be partly responsible for the large variability of the individual bimodal benefit. Therefore, knowledge about the location of the electrode array is an important prerequisite for optimal fitting. Theoretically, the electrode location can be determined from CT-scans. However, these are often not available in audiological practice.

Behavioral pitch matching between the two ears has also been suggested, but has been shown to be tedious and unreliable (Carlyon et al., 2010). Here, an alternative method using two-formant vowels was developed and tested.

Research question: Can we use the second formant (F2) of a two-formant vowel as a pitch matching stimulus by presenting it either on the aided/normal side or on the implanted side?

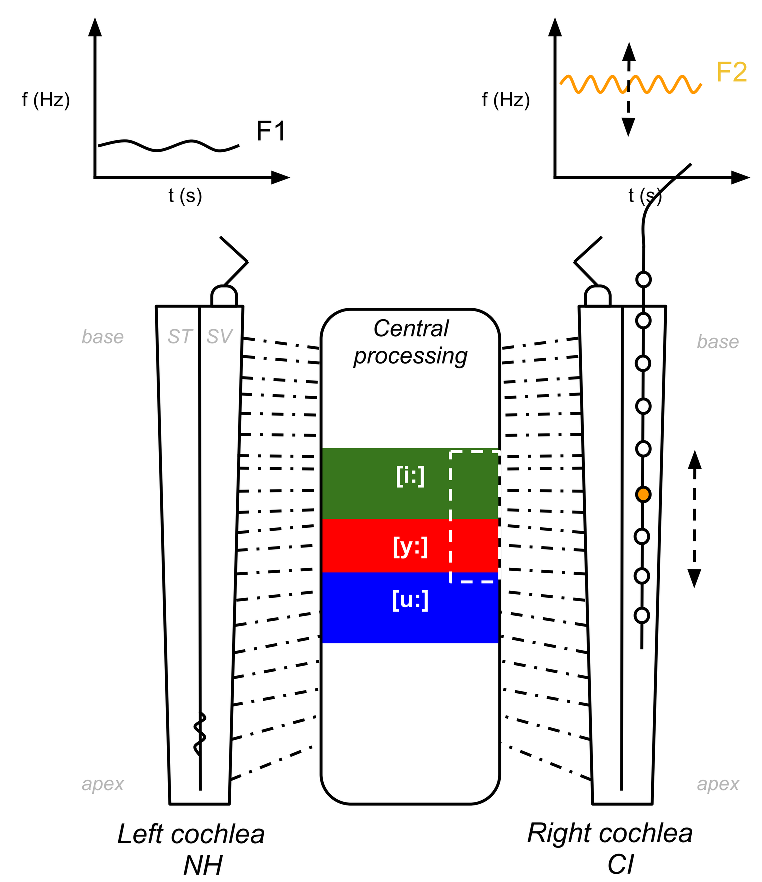

Fig.1 If the implant is perfectly inserted, the three vowels of this example ([u:], [y:], [i:]) should be perceived identically when presenting the second formant (F2, *yellow*) in either the normal hearing (NH) or CI side. If there is a shift towards the base, the perceived vowel map obtained by varying F2 (*dashed white rectangle*) should also show a shift.

Methods

Subjects
8 NH subjects, 5 bimodal (BM) and 6 single-sided-deaf (SSD) CI users participated, all German native speakers.

Stimuli
- Two-formant vowels produced using a Matlab-based Klatt synthesizer (Klatt, 1980) and mixed with consonants to form a /t/-/vowel/-/k/ stimulus. F1=[250,400] Hz, F2=[600,800,..,2200] Hz.

- For normal-hearing (NH) listeners, a noise vocoder (Litvak et al., 2007) was used in the right channel to simulate a perfect insertion ('Voc1') and two different mismatches ('Voc2' and 'Voc3', specified in figure 2). Vocoder training was achieved with an audiobook.

Procedure
- Subjects had to categorize (forced choice) the perceived stimuli into different vowel propositions using a Matlab GUI.

Results: NH listeners

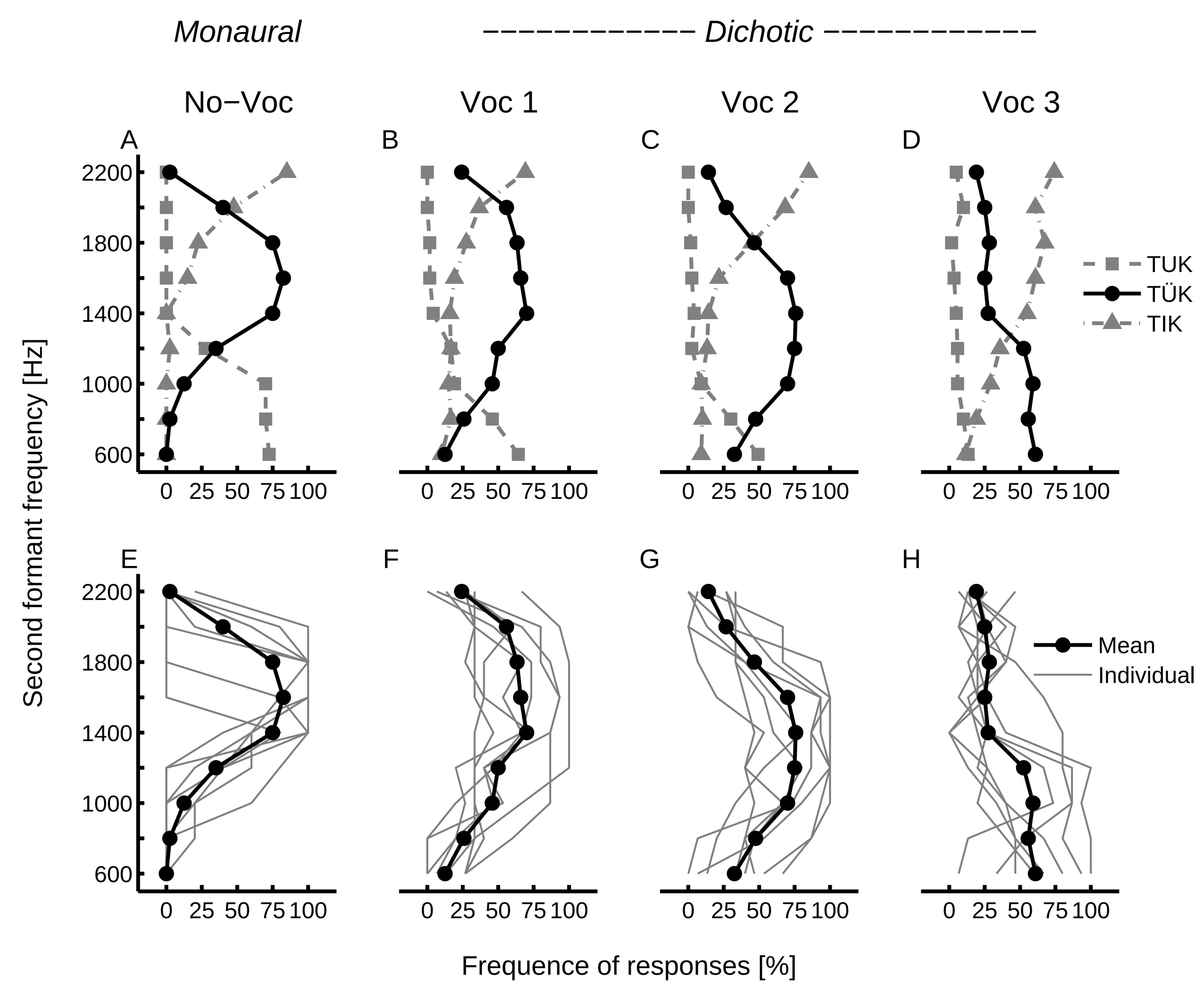

Fig.2 *Top panel:* Mean results (N=8) of the categorization test for the NH listeners. Only the results for F1=250 Hz are shown here. *Bottom:* Individual and mean results for the mid-F2 vowel ([y:] for F1=250 Hz).

- *Monaural*, No-Voc: F1 and F2 => Left Channel
- *Dichotic*, Voc1: F1 => Left, F2 => Right, vocoded with no shift
- *Dichotic*, Voc2: F1 => Left, F2 => Right, vocoded with ≈ 0.45 oct shift
- *Dichotic*, Voc3: F1 => Left, F2 => Right, vocoded with ≈ 0.85 oct shift

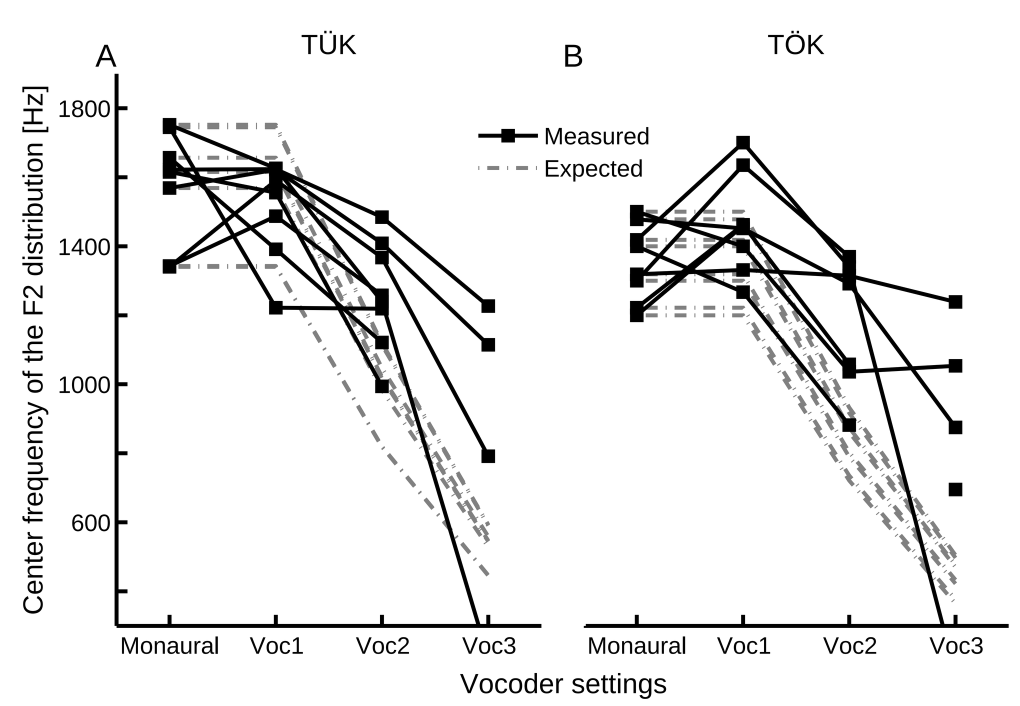

Fig.3 Fitted center frequencies for individual NH listeners (N=8) based on a Gaussian fit applied to the mid-F2 vowel distributions. (*A*) Category "TÜK" (F1=250 Hz). (*B*) Category "TÖK" (F1=400 Hz). Dashed gray lines show expected centers for each individual calculated from the "Monaural" condition data and the vocoder settings.

Main results:
- NH listeners were able to fuse formants, even when one is vocoded.

- Simulating a shift with the vocoder has an effect: the low-F2 vowel progressively disappears, and the two other vowel distributions move downwards, using this representation.

- Individual variability increases when simulating a large mismatch (Fig.2H and Fig.3).

- Small mismatches can be estimated using a Gaussian fit of the mid-F2 vowel distribution (Fig.3).

Results: CI listeners

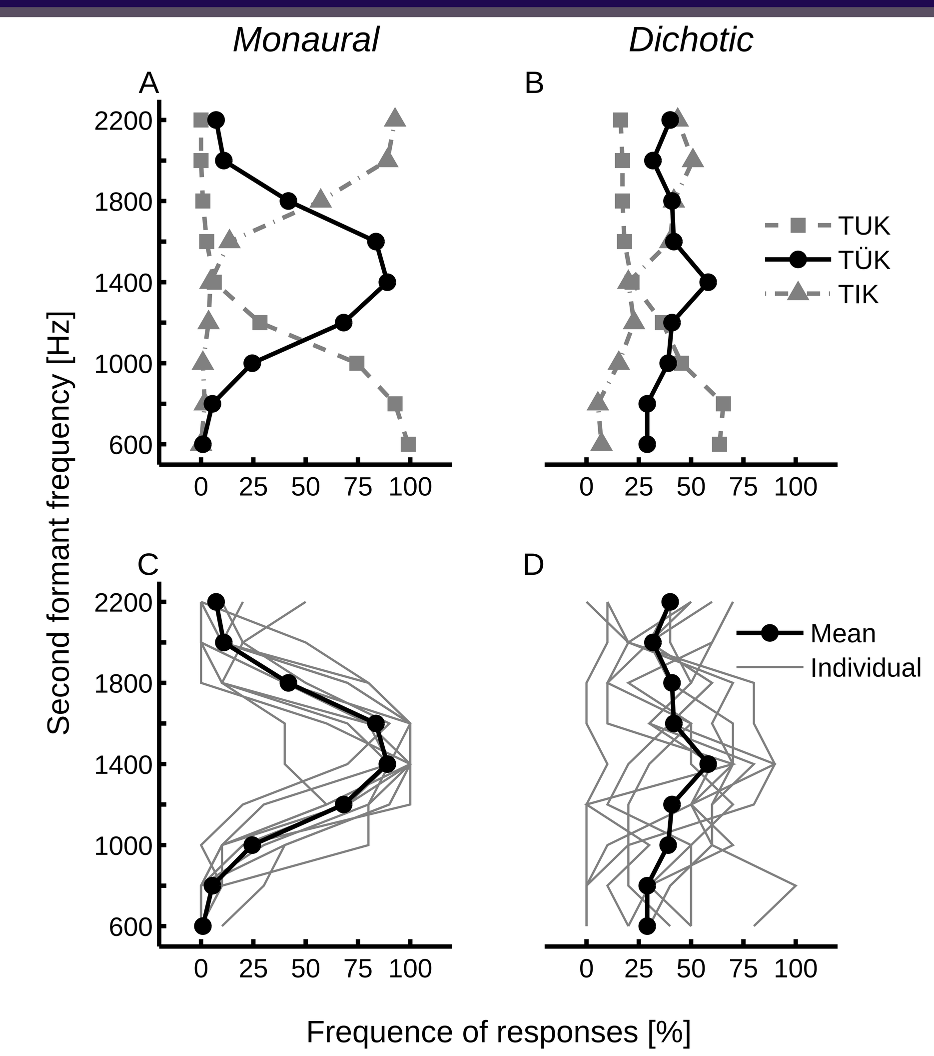

Fig.4 *Top:* Mean results (N=11) of the categorization test for the CI listeners. For the monaural condition, F1 and F2 were presented to the non-CI ear, with BM subjects wearing the hearing aid. *Bottom:* Mean and individual results for the mid-F2 vowel.

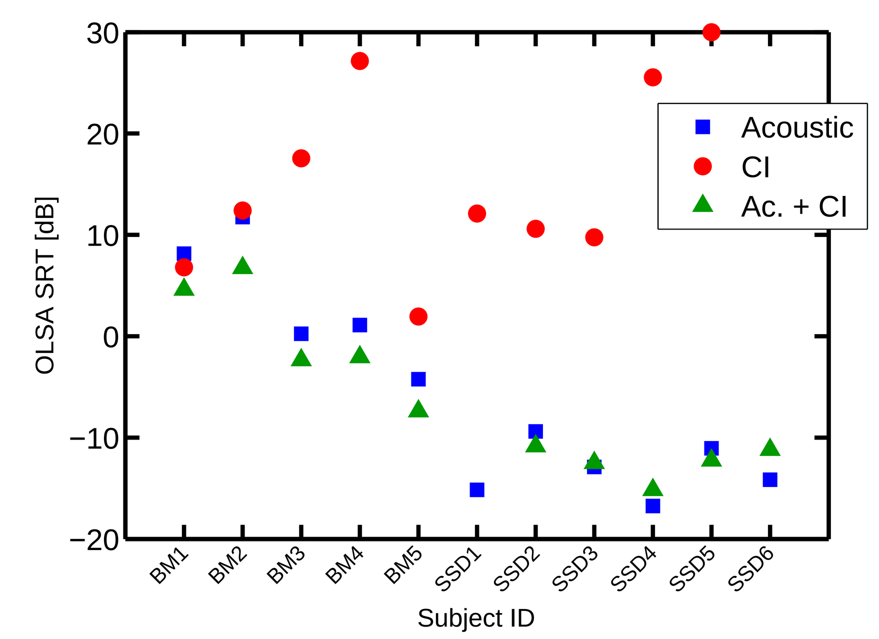

Fig.6 Speech reception thresholds of the CI listeners, using the Oldenburg sentence test (OLSA) and the International Speech Test Signal (ISTS) as interferer.

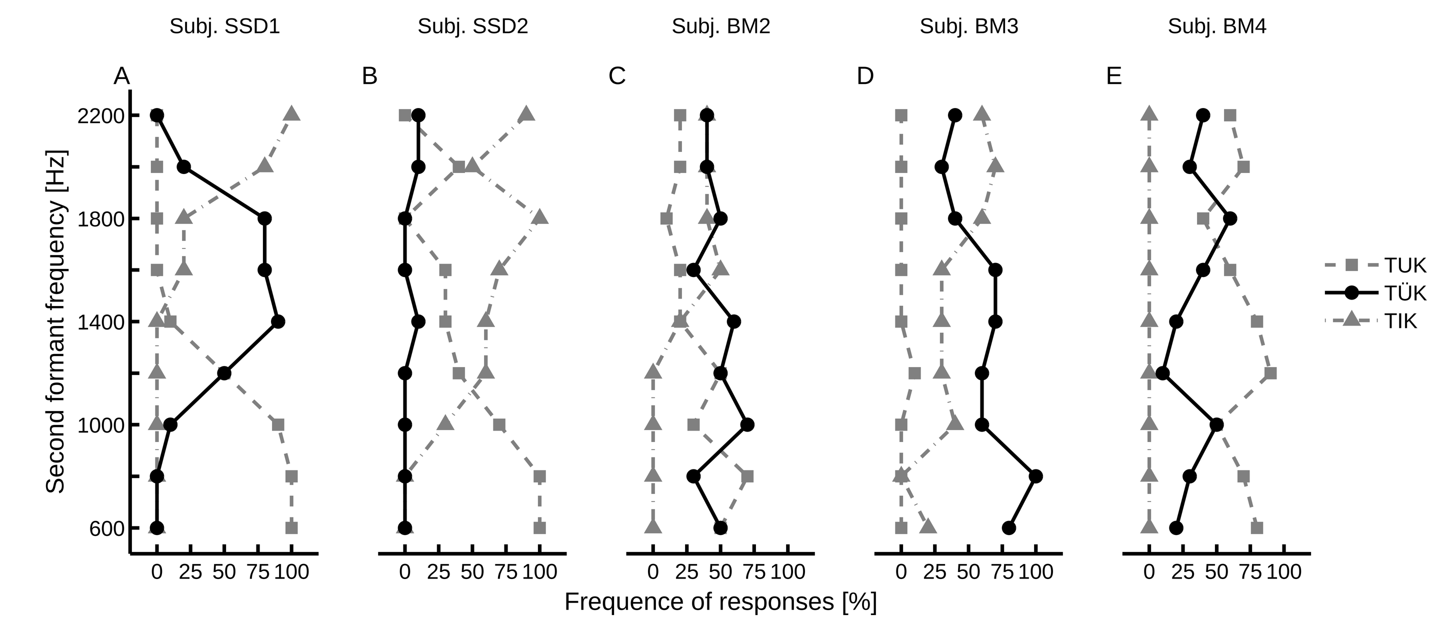

Fig.5 Individual results of the categorization test for a subset of CI listeners for the *Dichotic* condition (F1 => non-implanted ear, F2 => CI).

Main results:
- Similar performance to NH listeners for the monaural condition (fig.4A), even for BM subjects using hearing aids.

- Trend to see an acclimatization to the frequency map in the mean results, especially for the low and high F2 vowels ([u:] and [i:], Fig.4B).

- High variability for the dichotic condition, because of the difficulty to fuse formants having different percepts (Fig.4D and 5).

Insertion depth:
With the exception of one subject (BM2), all insertions depths, determined from CT scans, ranged from 332 to 375 degrees (Table 1).

=> The obtained vowel distributions are not directly related to the CI insertion depth (Fig.5 and Table 1).

Speech reception thresholds:
Speech reception thresholds were measured with the OLSA test in ISTS background. The acoustic ear was significantly better than the electric ear for this population ($p<0.001$, Wilcoxon rank sum test, Fig.6).

For the combined acoustic and CI condition, speech level in the CI was adjusted by adding the SNR between the "Acoustic" and "CI" conditions, to encourage subjects to use cues from both sides. With this method, there was a small average benefit of adding electric information for BM listeners, but no benefit for SSD listeners.

=> This population of CI listeners relies mainly on the information from the acoustic ear.

CI subjects' data

Subject ID	Age at surgery [yrs]	Duration of deafness at surgery [yrs]	Duration of implant use [mths]	Electrode; strategy	Side	Insertion depth angle [deg]	PTA of the non-implanted ear [dB HL]
BM1	74	3	15	Helix; HiRes Fidelity 120	right	350	38
BM2	73	20	21	1j; HiRes S	right	452	50
BM3	65	11	15	1j; HiRes Fidelity 120	left	355	20
BM4	60	37	19	1j; HiRes S	left	355	28
BM5	37	5	20	Helix; CIS	right	332	25
SSD1	59	5	16	Helix; HiRes Fidelity 120	right	375	12
SSD2	62	2	19	Helix; HiRes Fidelity 120	right	353	18
SSD3	53	6	20	Helix; HiRes Fidelity 120	right	355	18
SSD4	31	2	19	1j; HiRes Fidelity 120	left	375	13
SSD5	36	2	19	Helix; HiRes Fidelity 120	right	360	18
SSD6	35	20	15	1j; HiRes S	right	337	10

Table 1: Data from the CI subjects. Pure Tone Average threshold (PTA) is the mean of the thresholds at 500 Hz, 1 kHz and 2 kHz. Insertion depth was determined from post-operative CT scans.

Discussion, conclusions

- NH listeners' results suggest that this new procedure could be a tool to indicate the existence of a mismatch. However, quantitative estimation of the mismatch remains challenging.

- CI listeners could perform the task within a reasonable amount of time (30 minutes), and reliably for the acoustic *Monaural* condition.

- Large individual differences were observed for the dichotic acoustic-electric stimulation (poor vowel discrimination). Difficulty to fuse electric and acoustic percepts might be an explanation, as suggested by the speech perception results.

Overall, these results suggest that place mismatches can be derived from such vowel spaces, but the method remains limited by the individual variability and the difficulty to achieve spectral fusion.

References:
- Carlyon et al. (2010). Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear. *JARO*, 11(4):625-640
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *JASA*, 67(3):971-995
- Litvak et al. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *JASA*, 122(2):982-991

pants with the resynthesized voices in a phonemic restoration paradigm.

Results
Data show, as expected, an overall decrease of intelligibility as the spectral resolution is reduced. However the effect of pitch presence or absence does not follow exactly our expectations. The addition of pitch showed improvement of restoration only at 8- and 6-bands spectral resolution, which is fully due to improvement of intelligibility when the interruptions are filled with noise.

Conclusion
With the addition of pitch, NH listeners may be better able to discriminate the speech from the noise, and may therefore avoid interpreting the latter as (spurious) speech cues, thus yielding fewer errors. Moreover, the addition of pitch to the spectrally degraded speech also provides more bottom-up cues that seem to trigger the top-down repair of the missing segments (especially at 8- and 6-bands). When the speech is intelligible enough, adding new speech features does not enhance restoration any further. When the speech is too degraded, adding pitch is not sufficient to improve intelligibility. The present results suggest that phonemic restoration also depends on the amount of speech features available in the speech segments. Adding pitch information to CIs could lead to the improvement of top-down repairs in noisy listening situation.

PS - 492

Interaural Place-Mismatch Estimation With Two-Formant Vowels in Unilateral Cochlear-Implant Users

François Guérit[1]; **Sébastien Santurette[1]**; Josef Chalupper[2]; Iris Arweiler[2]; Torsten Dau[1]
[1]Technical University of Denmark; [2]Advanced Bionics European Research Center GmbH

Background
For patients with one cochlear implant (CI) and residual hearing in the opposite ear, a default frequency-to-electrode map is typically used despite large individual differences in electrode-array insertion depth. This non-individualized fitting rationale might partly explain the variability in long-term speech-reception benefit among CI users. Knowledge about the electrode-array location is thus crucial for adequate fitting. Although electrode location can theoretically be determined from CT scans, these are often unavailable in audiological practice. Moreover, existing behavioral procedures such as interaural pitch-matching are rather tedious and time-consuming. Here, an alternative method using two-formant vowels was developed and tested.

Methods
Eight normal-hearing (NH) listeners were presented synthesized two-formant vowels embedded between consonants /t/ and /k/, with first-formant frequencies (F1) at 250 and 400 Hz and second-formant frequencies (F2) between 600 and 2200 Hz. F1 was presented unaltered to the left ear, while F2 was presented to the right ear via a vocoder system simulating 3 different CI insertion depths. In each condition, the listeners indicated in a forced-choice task which of 6 vowels they perceived for different [F1, F2] combinations. Ten CI users (5 bimodal and 5 single-sided deaf) performed the same task for F1 presented acoustically to the non-CI ear and F2 presented either acoustically to the same ear or electrically to the CI ear.

Results
After some training, all NH listeners were able to fuse the unaltered F1 and vocoded F2 into a single vowel percept, and vowel distributions could be reliably derived in 7 listeners. Vocoder simulations of reduced CI insertion depth led to clear vowel-distribution shifts in these listeners. However, these shifts were overall smaller than their theoretical value, with high across-subject variability. Vowel distributions could be derived for all CI users in the monaural acoustic condition, indicating an ability to perform the task reliably. Despite this, large individual differences were observed for dichotic bimodal stimulation, with listeners showing either basal or apical shifts, or generally-poor vowel discrimination.

Conclusion
The two-formant-vowel method is a fast and clinic-friendly candidate to derive interaural place mismatches from a simple vowel-recognition task. However, it remains unclear whether the measured “vowel spaces” in CI users are directly related to insertion depth, and whether they are influenced by the ability to fuse acoustic and electric stimuli or habituation to the CI. The comparison of the present results to CT-scan and speech-intelligibility data in the same listeners will shed light on the validity of the proposed method.

PS - 493

The Effect of Spectral Resolution on Temporal Processing Abilities in Simulated Electric Hearing

Fawen Zhang[1]; **Chelsea Blankenship[1]**; Jing Xiang[2]; Qiang-jie Fu[3]
[1]University of Cincinnati; [2]Cincinnati Childrens Hospital Medical Center; [3]University of California

Background
Gap detection threshold (GDT, the shortest intervals a person can perceive) is a commonly used measure of temporal processing resolution. Normal GDT is critical for speech encoding. It has been reported that the large variability in cochlear implant users’ speech perception may be related to the temporal processing deficits in some cochlear implant (CI) users (Muchnik et al., 1994; Fu et al., 2002). Unlike the GDT measured using direct electric stimulation through the electrode, the GDT measured via clinical processors may also reflect additional limitations imposed by limited spectral resolution due to CI processing. The purpose of this study is to determine how spectral resolution may affect GDTs in normal hearing (NH) subjects listening to acoustic simulation of CI processing. Additionally, the neural correlates of gap detection were examined using the late auditory evoked potential (LAEP). If a correlation between behavioral and LAEP measures exists, a clinically useful outcome would be to use the