Issues in the implementation of the CIS strategy

Philip C. Loizou and Oguz Poroy

Department of Electrical Engineering
University of Texas at Dallas

www.utdallas.edu/~loizou/cimplants

Research supported by R01 DC03421 from NIDCD.

Abstract

- The Continuous Interleaved Sampling (CIS) strategy is currently supported by most cochlear implant (CI) devices. Results reported with the CIS strategy, however, were not consistent across the different devices. We suspect that the difference in performance obtained with the CIS strategy in different devices can be attributed in part by the way the CIS strategy is implemented.

  To investigate the effect of the CIS implementation on performance, we collected some data on the implementation of the logarithmic amplitude mapping function. There are at least three different ways of implementing the log function: (1) using look-up tables, (2) using a Taylor series approximation to the log function, and (3) using the power function with different exponents. Using look-up tables is by far the simplest approach, and different CI manufacturers may use different table sizes. In this study, we systematically varied the size of the table containing the mapped values from 128 to 32768 and examined vowel and consonant recognition.

Introduction

- Recent studies showed contradictory results on the effect of higher stimulation rates on speech recognition. Vandali *et al.* (2000) found no benefit when using higher rates of stimulation, while others (Loizou *et al.*, 2000; Brill *et al.*, 1997; Kiefer *et al.*, 2000; Wilson *et al.*, 2000) found a significant (positive) benefit.
- We suspect that the difference in performance obtained with the CIS strategy in different devices can be attributed in part by the way the CIS strategy is implemented.
- The CIS implementation could differ in many ways: analog preprocessing (hardware), wordlength (16 bit vs. 24 bit) of the DSP, envelope detection (lowpass filter vs. Hilbert transform) and mapping function (table look-up, Taylor series approximation)
- In this study, we investigated the effect of the mapping function implementation on vowel and consonant recognition.

Mapping function

- In a typical CIS implementation, the acoustic envelope amplitude x is mapped to the electrical amplitude y according to:

$$y = A\,x^{p} + B$$

In a fixed-point implementation, x takes values in the range $0 <= x < 1$ or equivalently $0 < x < 32768$ in integer format assuming a 16-bit A/D.

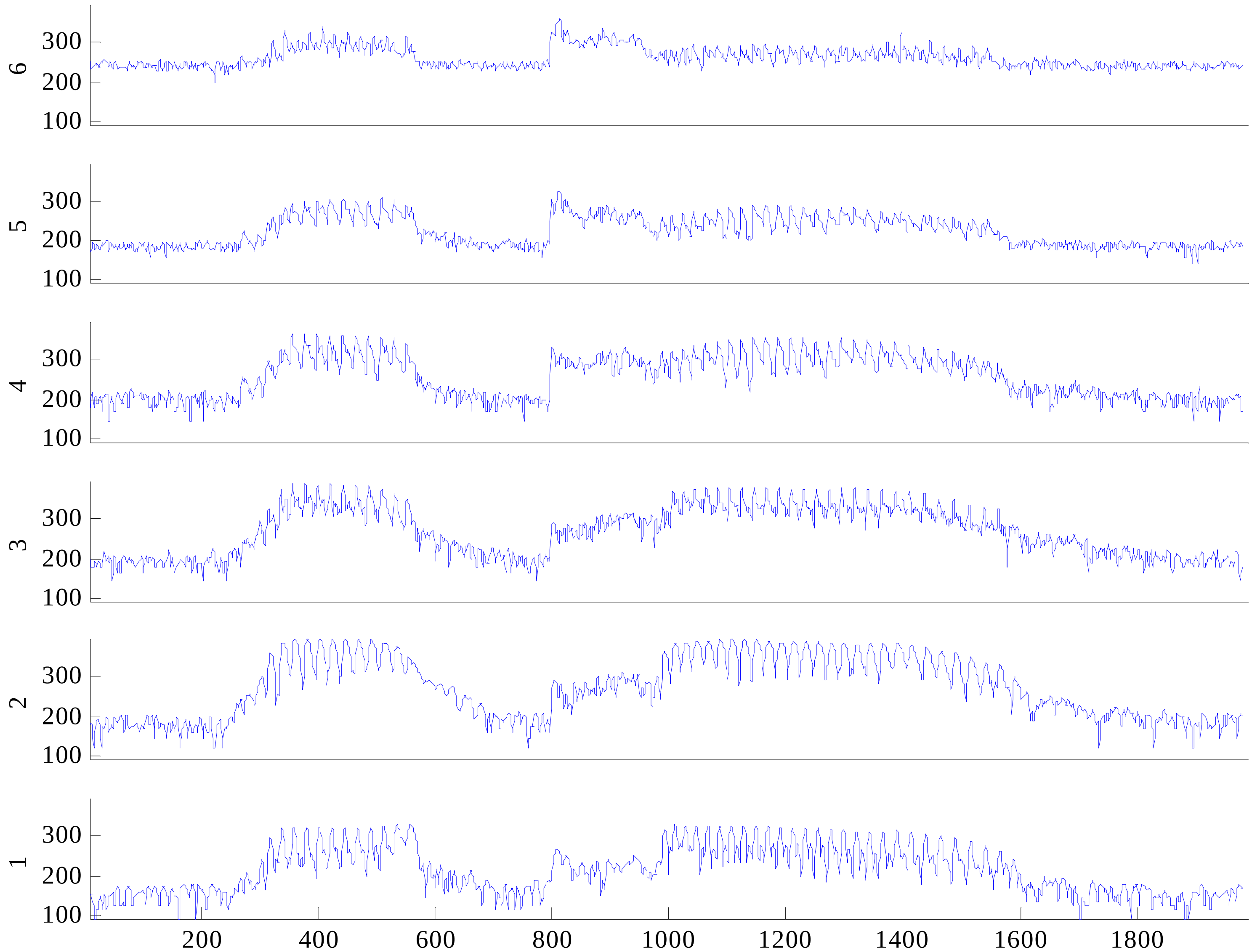
This implementation of the mapping function would require a table of 32768 values for y.

UTD

Limiting the size of the mapping table

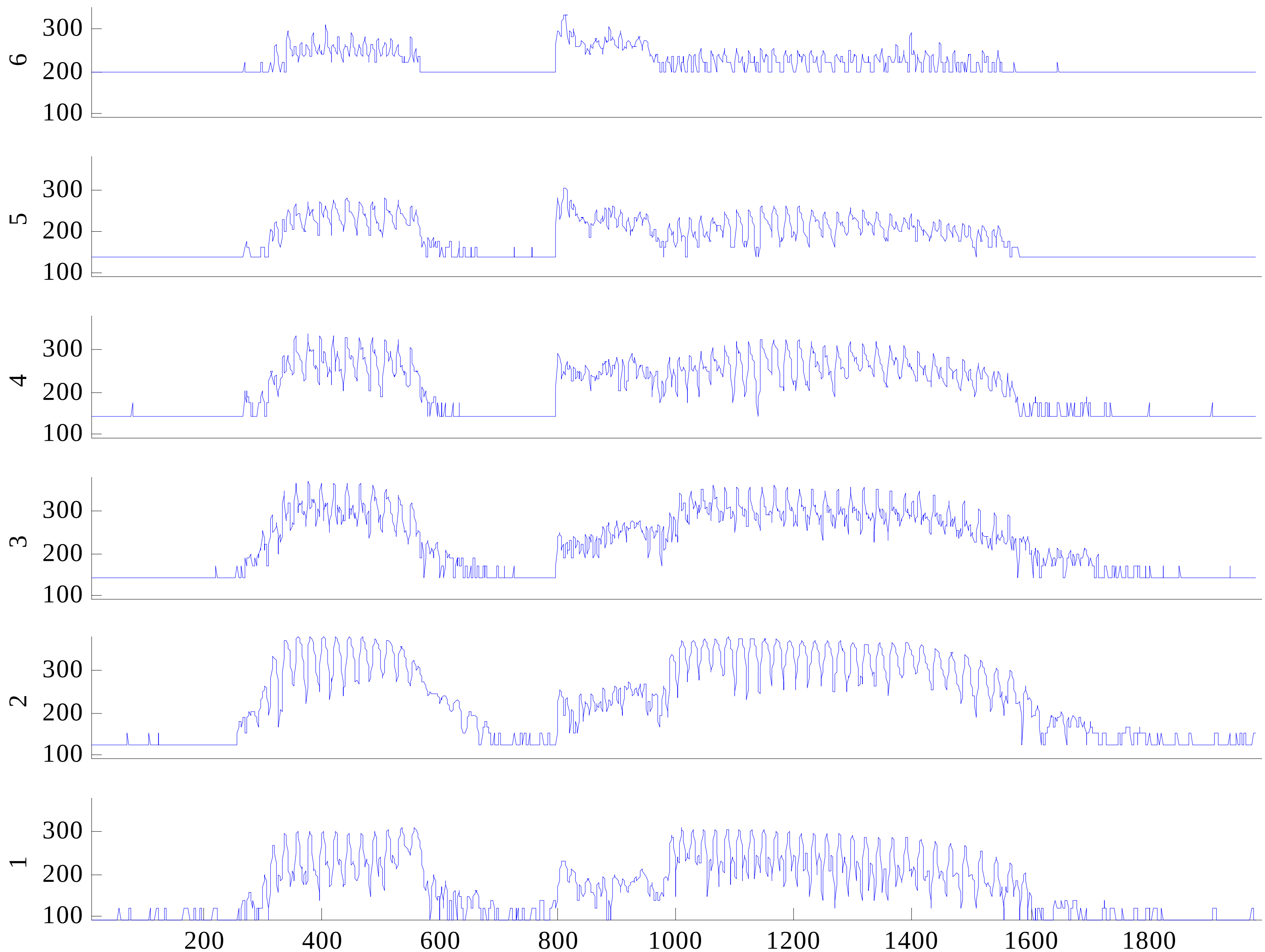
- To limit the integer *x* to lie in the range 0<= *x* < Xmax, we can pre-multiply *x* by Xmax / 32768.
- Then, a table with only Xmax (Xmax << 32768) values would be required for storing the *y* values.
- Xmax defines the size of the mapping table.
- Note that the above pre-multiplication of *x* essentially quantizes the input signal to Xmax levels.
- To investigate the effect of the size of the mapping table on phoneme recognition, we varied Xmax from 128 to 32768 levels.

Envelope amplitudes of /ata/ using Xmax=32768

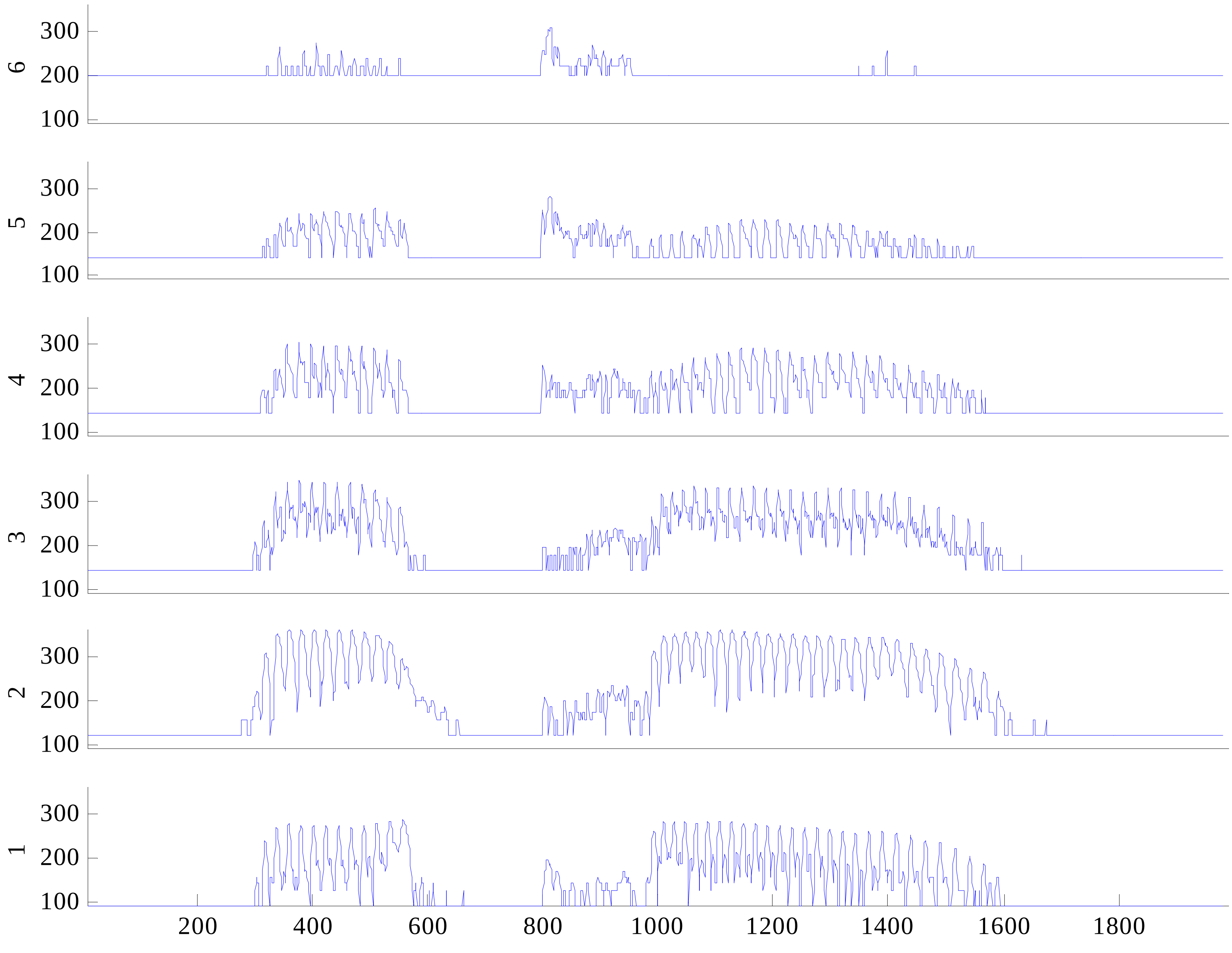


As

Envelope amplitudes of /ata/ using Xmax=4096



Envelope amplitudes of /ata/ using Xmax=1024



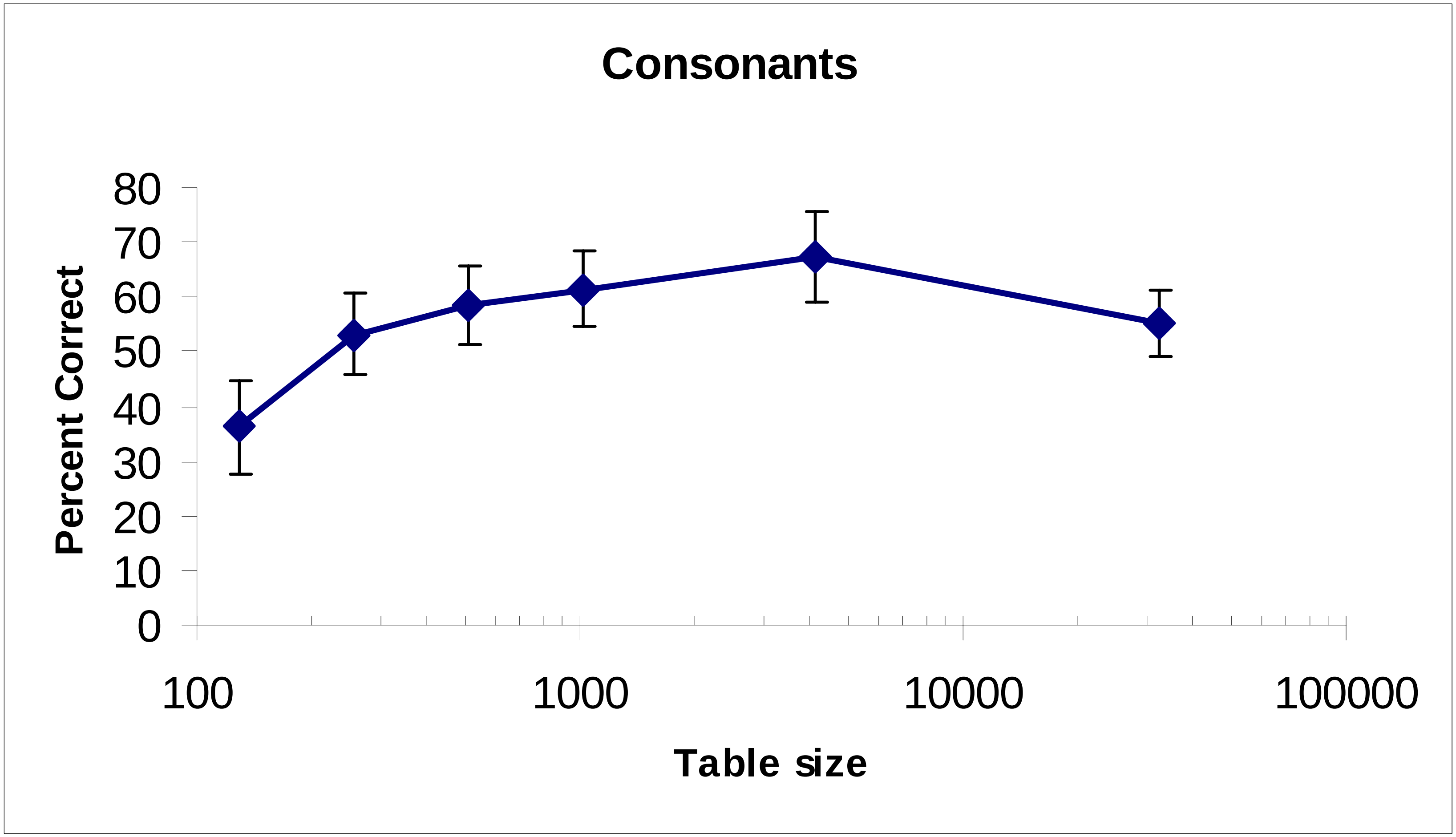
Methods

- **Subjects**
  Five postlingually deafened adults wearing the Med-El/CIS-link device. All subjects have been wearing their device for more than a year.
- **Speech Material**
  - 11 vowels in /hVd/ context from the Hillenbrand *et al.* (1995) database produced by 3 male and 3 female speakers.
  - 20 consonants in /vCv/ context (v=/i, a, u/) produced by a female speaker (recorded at the House Ear Institute)

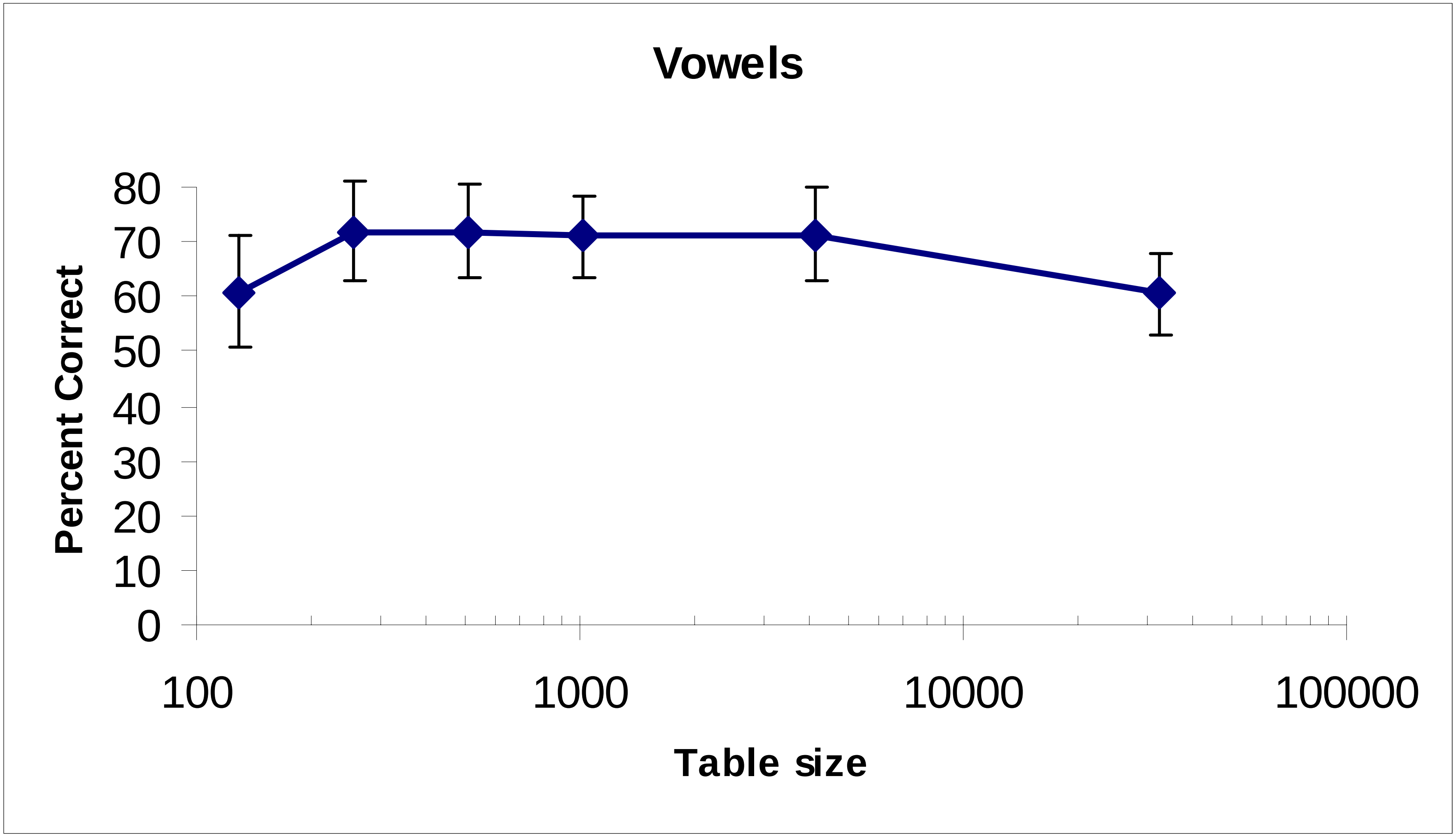
Procedure

- The vowel and consonant stimuli were first processed off-line through the CIS strategy. The envelope amplitudes were mapped according to the power function using tables of size Xmax=128, 256, 512, 1024, 32768.
- The exponent *p* was set to -0.0001 for log mapping. The CIS channel amplitudes were saved in a file, and then presented to the CI listeners via our laboratory processor.
- The order of Xmax conditions was counterbalanced between subjects to avoid order effects.

Consonant recognition as a function of mapping table size



Vowel recognition as a function of mapping table size



Discussion

- The effect of the table size in the implementation of the mapping function was larger for consonants than for vowels.
- Consonants
  - Use of a larger size table did not yield higher performance. The score obtained with Xmax=4096 was significantly higher ($p < 0.05$) than the score obtained with Xmax=32768.
- Vowels
  - The table size needs to be at least 256 for high performance. The score obtained with Xmax=256 was not significantly different than the score obtained with Xmax=32768.

Discussion

- Consonant recognition suffered the most, because the size of the compression table affects the coding of temporal-envelope cues. When the table size is small, the low envelope amplitudes are lost, as they are mapped to threshold. This can be seen in the example waveforms of /ata/.
- The size of the table also affects the effective spectral contrast as shown in the Figure below. We suspect that this is because the operating point in the compression function changes when the size of the table changes.

Histograms of spectral dynamic range for /ata/

Table size = 1024

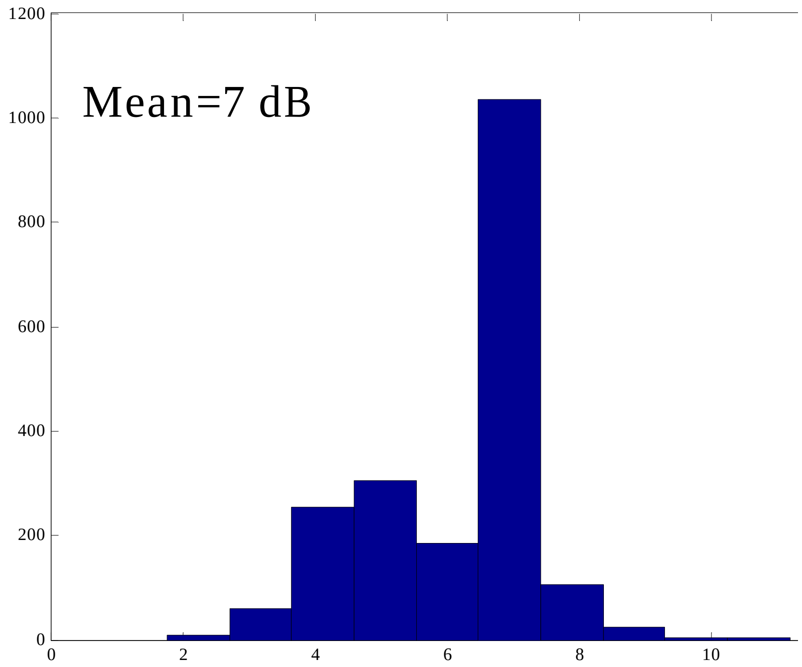
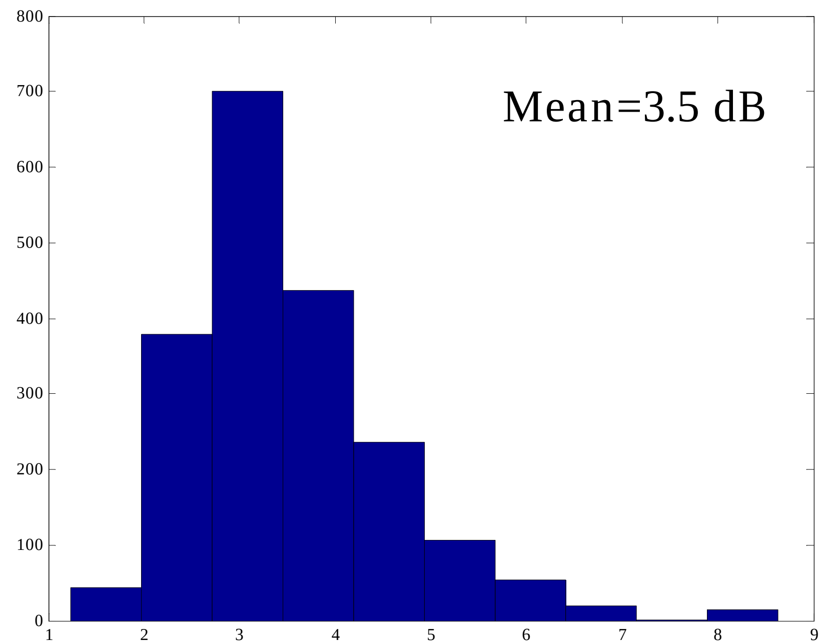


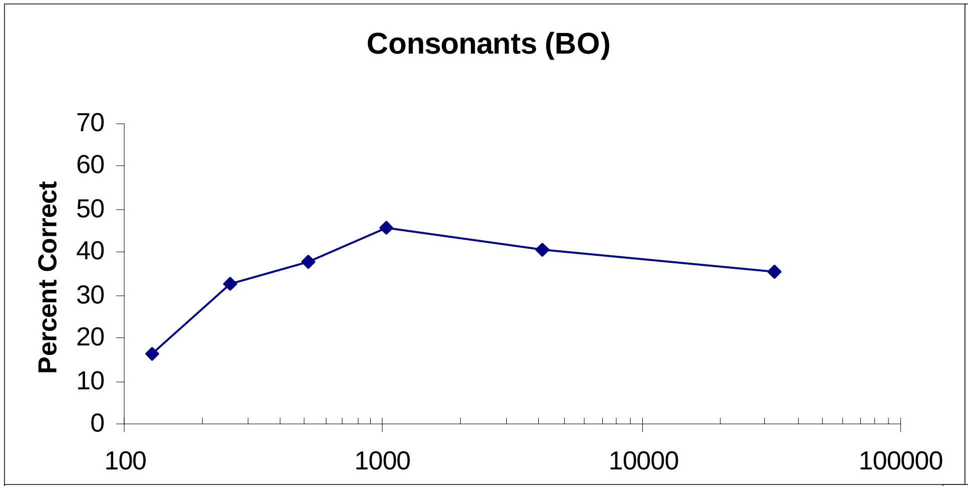
Table size = 32768

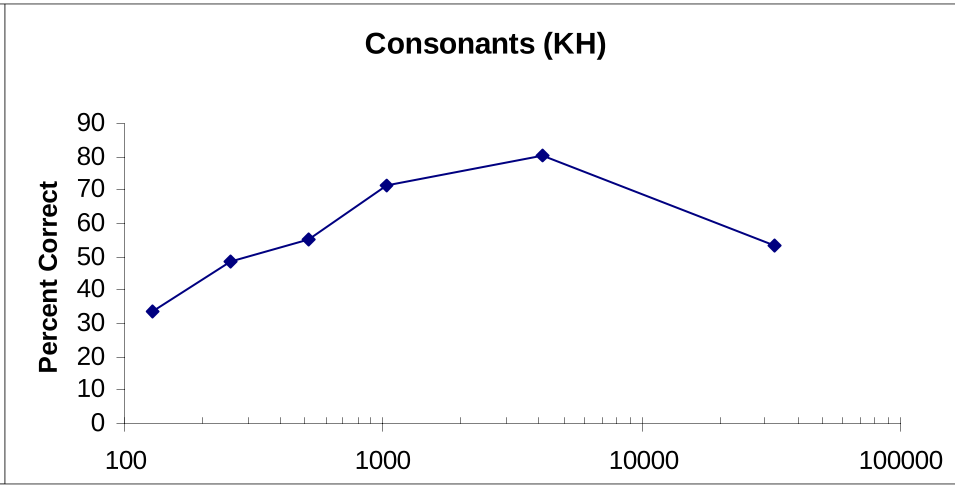


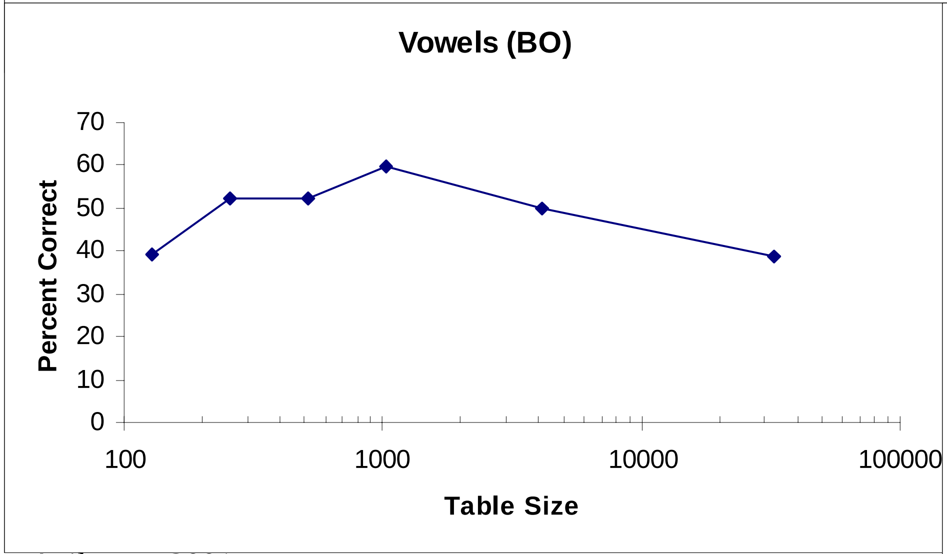
Spectral dynamic range (dB)

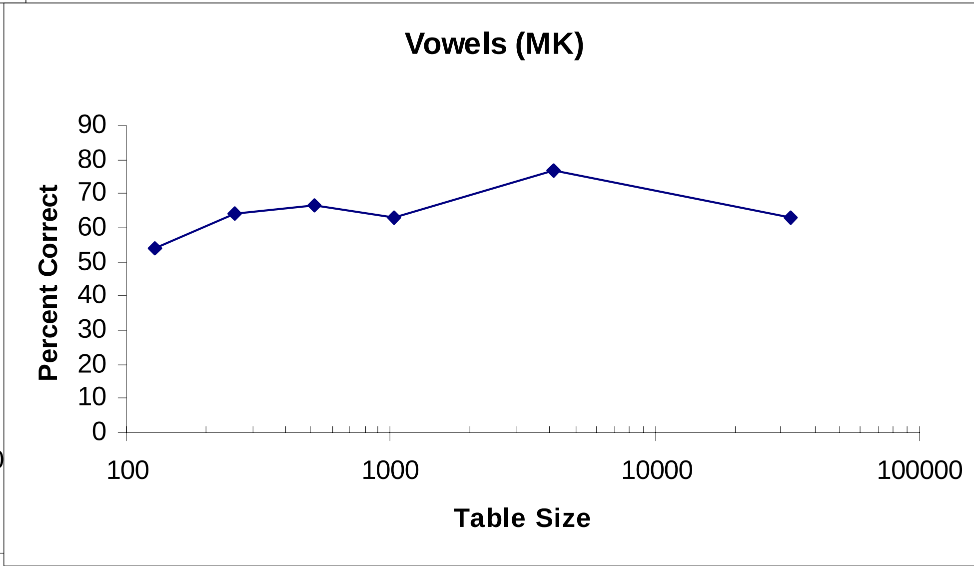
UTD

Individual data









Conclusions

- There was a significant effect of the mapping table size on vowel and consonant recognition.
- The size of the mapping table affects the coding of temporal-envelope cues (hence, consonant recognition suffered the most) and also the spectral contrast.
- Our results suggest that the mapping function does not need to be defined with high resolution. In fact, for consonants, it is best if it is not. A table size of 4096 yielded the maximum performance on consonant recognition for most subjects. We suspect that the maximum performance was obtained due to increased spectral contrast.
- There was subject variability in the effect of table size on vowel and consonant recognition, suggesting that the table size can be optimized for individual subjects.

References

- Brill SM, Gstottner W, Helms J, von Ilberg C, Baumgartner W, Muller J, Kiefer J. (1997). Optimization of channel number and stimulation rate for the fast continuous interleaved sampling strategy in the COMBI 40+, Am J Otol, 18 (6 Suppl): S104-106.
- Kiefer, J von Ilberg, C. Rupprecht, V. Hubner-Egner, J. Knecht, R. (2000). Optimized speech understanding with the continuous interleaved sampling speech coding strategy in patients with cochlear implants: effect of variations in stimulation rate and number of channels. Annals of Otology, Rhinology and Laryngology, 109(11), 1009-1020.
- Loizou, P., Poroy, O. and M. Dorman (2000)."The effect of parametric variations of cochlear implant processors on speech understanding," Journal of Acoustical Society of America , 108(2), 790-802.
- A. Vandali, L. Whitford, K. Plant and G. Clark (2000). “Speech perception as a function of electrical stimulation rate using the Nucleus 24 cochlear implant system,” Ear Hear, 21, 608-624.
- B. Wilson, R. Wolford and D. Lawson (2000), Speech Processors for Auditory Prostheses, Sixth Quarterly Report.