COMPARING THE PERFORMANCE OF THE SPEAK STRATEGY (SPECTRA 22) AND THE CIS STRATEGY (MED-EL) IN QUIET AND IN NOISE

Philip Loizou[1], Sharon Graham[2], John Dickins[2], Michael Dorman[3] and Oguz Poroy[1]

[1]University of Arkansas at Little Rock, Little Rock, AR 72204-1099
[2]Arkansas Otolaryngology Center, Little Rock, AR 72205
[3]Arizona State University, Tempe, AZ 85287

[1] http://giles.ualr.edu/asd/cimplants/

UALR

Abstract

- The SPEAK and CIS strategies have been proven to be successful sound processing strategies for cochlear implants. These two strategies are based on two different principles. The SPEAK strategy estimates the envelope amplitudes of 20 bandpass filters, and selects the ones with the largest amplitude for low pulse rate (180-300 pps) stimulation. The SPEAK strategy is therefore based on the idea that the largest amplitudes are the ones that are perceptually most important. In contrast, the CIS strategy estimates the envelope amplitudes of six (or eight) bandpass filters and uses all the amplitudes for high pulse rate (> 800 pps) stimulation. At issue is whether either one of the two strategies performs better on speech recognition tasks or offers any advantages (e.g., more robust in noise).

Abstract

- The aim of the present study is to evaluate the strengths and weaknesses of the SPEAK and CIS strategies in both quiet and noisy environments. Several questions will be addressed in this study regarding the differences between the two strategies. For instance, it is of question whether the CIS strategy, that provides a good representation of rapid envelope variations through the use of high pulse rate stimulation, performs better on consonant identification than the SPEAK strategy. It is also of question whether the SPEAK strategy, that selects (on the average) six spectral maxima provides a better spectral representation and therefore performs better on vowel recognition than the CIS strategy. Finally, it is of question whether the SPEAK strategy or the CIS strategy performs better in noise.

Methods

- **Subjects**

  A total of 18 postlingually deafened patients were used in this study

  – 11 Nucleus patients fit with the Spectra 22 processor using the SPEAK processing strategy. All patients were from the Arkansas population.

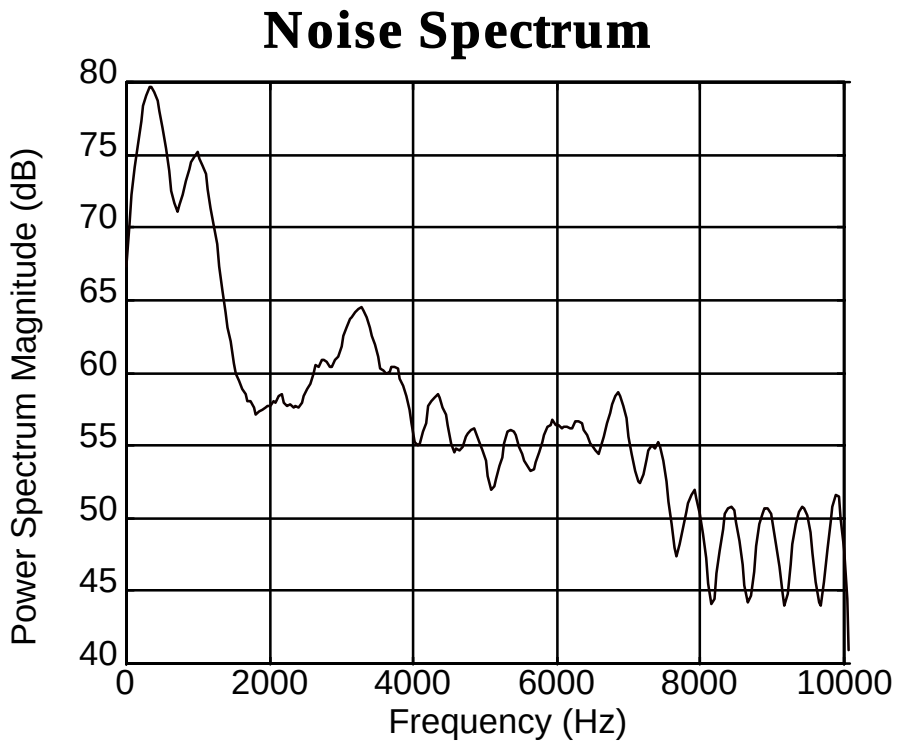
  – 7 Ineraid patients fit with the Med-El processor using the CIS strategy

Speech Materials

- Patients were tested on vowel, consonant and sentence recognition in a sound only paradigm
- Consonant and vowel stimuli were taken from the Iowa laser videodisc. These stimuli were digitized at a 20 kHz sampling rate.
  - 16 medial consonants /p,t,k,f,s,sh,b,d,g,v,th,z,j,l,m,n/ spoken by a male speaker in a/C/a context.
  - 8 vowels in h/V/d context spoken by a male speaker:
    - » heed, hod, head, hid, hood, hud, had, who'd
- Sentences were taken from the H.I.N.T. database (Nilsson et al., 1994). A different sentence list was used for each test condition.
- The same speech materials were used for Nucleus and Ineraid patients.

Noise stimuli

- Speech-shaped noise was added at 5 dB, 10 dB and 15 dB S/N



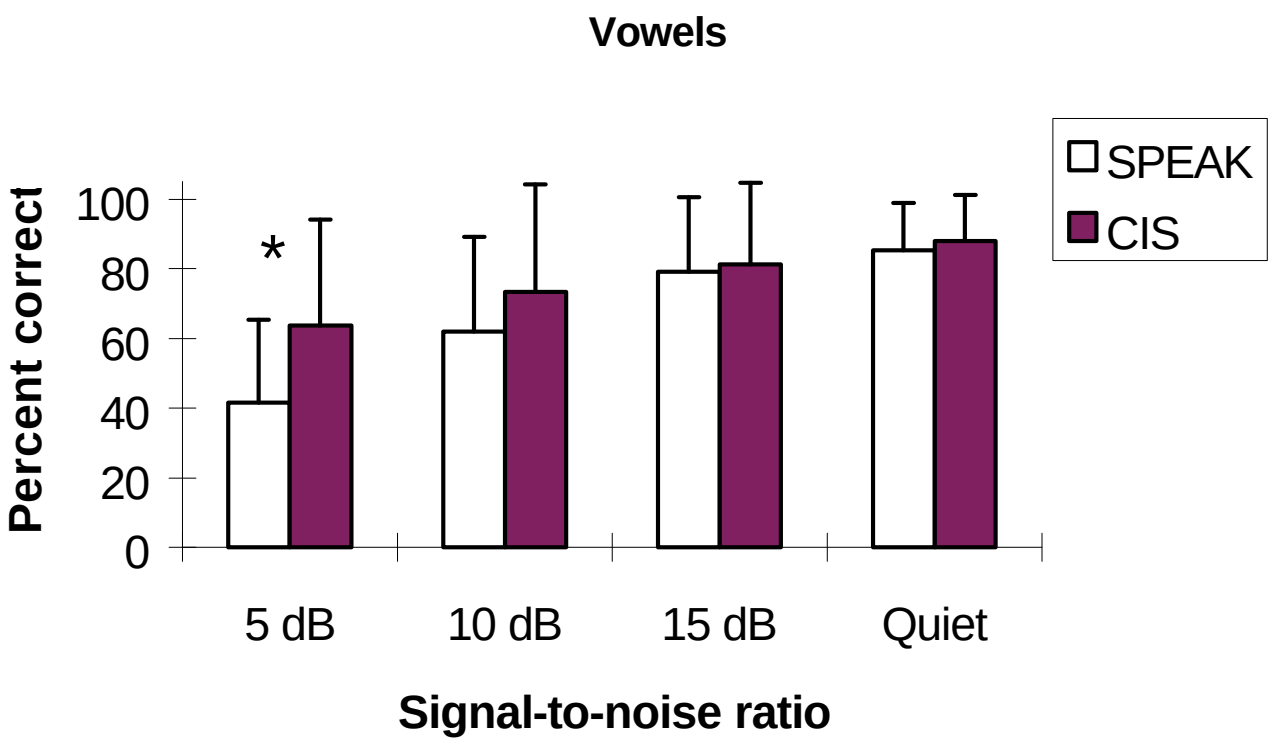
Procedure

- All the stimuli in noise (at 5 dB, 10 dB and 15 dB S/N) were completely randomized with the stimuli in quiet.
- Practice was given before each test. All practice stimuli were presented in quiet only. Practice for the sentence material consisted of the presentation of 10 sentences with concurrent display of the sentences. Practice for the vowel and consonant tests consisted of two runs through the stimulus lists with visual indication of item identity.
- The Nucleus patients were tested in sound-field and were sitted one meter from the loudspeaker. The Ineraid patients were tested using the processor's auxiliary input jack.
- Responses were collected with custom software using a computer display of response alternatives and a mouse as a response key.

Results

- A two-way ANOVA with repetition was performed for each test material and each strategy.
  - For all test materials, there was a significant effect of noise condition and subjects, but no interaction between subjects and noise condition.
- For the intersubject comparisons, the t test for independent samples was used for comparing CIS versus SPEAK strategies in different noise conditions
- The results for each set of test materials (vowels, consonants and sentences) are discussed separately.

Vowels

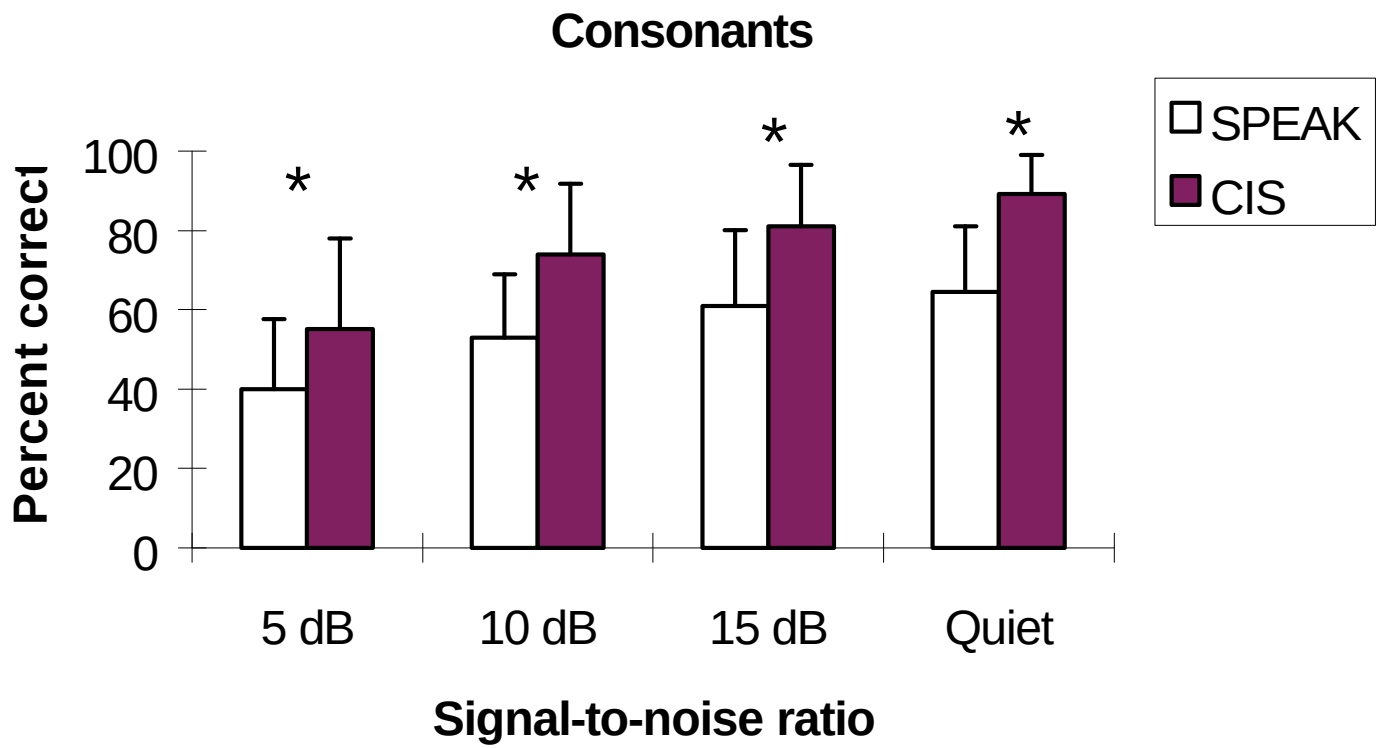


* $p < 0.005$

Vowels

- Difference between the CIS and SPEAK strategies was statistically significant ($p < 0.005$) only for the 5 dB S/N condition
- At extremely low S/N conditions (5 dB) vowel recognition seemed to be less affected with the CIS strategy than with the SPEAK strategy
  - Mean score of the SPEAK strategy dropped from 85% correct in quiet to 42% correct in 5 dB S/N
  - Mean score of the CIS strategy dropped from 88% correct in quiet to 64% correct in 5 dB S/N

Consonants

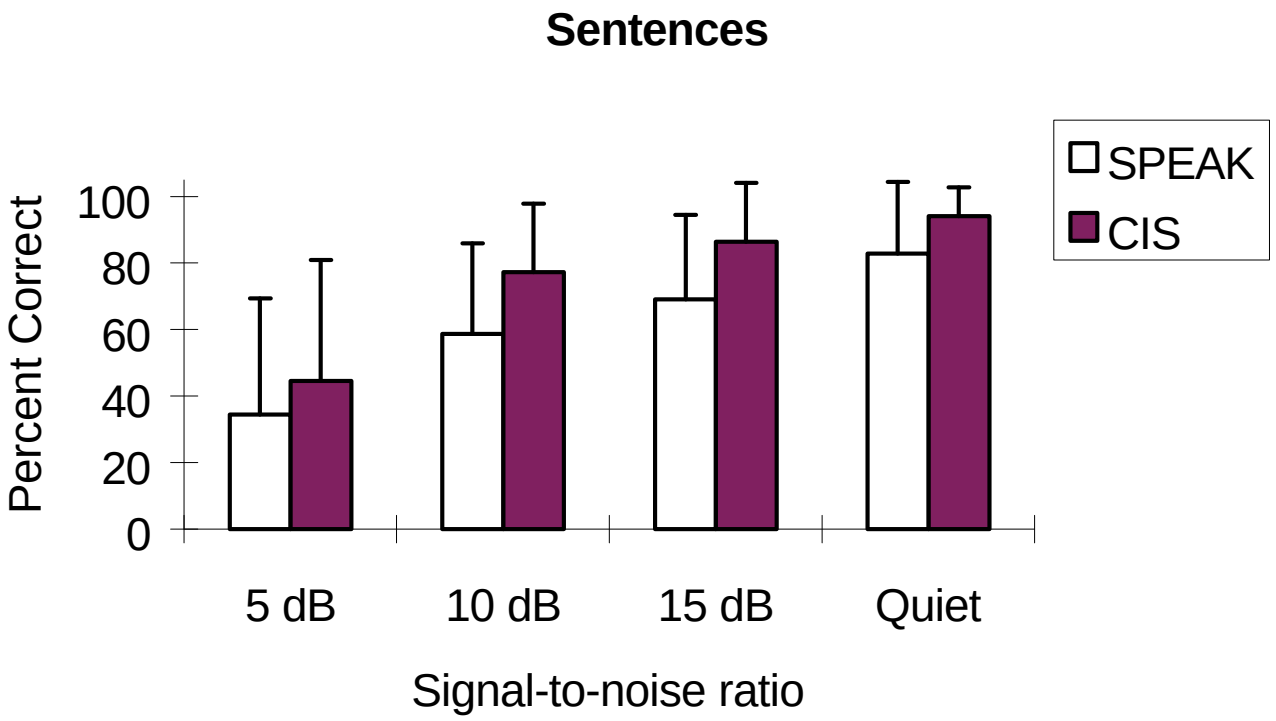


* $p < 0.005$

Consonants

- The results with the CIS strategy were superior to the SPEAK strategy in all testing conditions
- The difference between the CIS and SPEAK strategies was statistically significant ($p<0.005$) in all conditions

Sentences


Sentences
Percent Correct
100
80
60
40
20
0
5 dB
10 dB
15 dB
Quiet
Signal-to-noise ratio
SPEAK
CIS

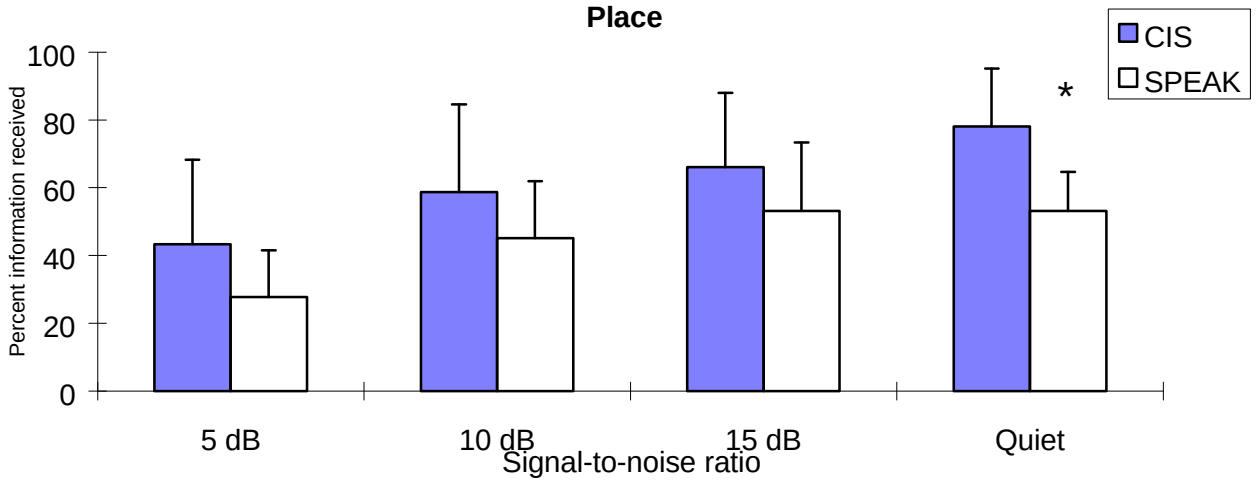
Sentences

- The mean scores of the CIS strategy were higher than the SPEAK strategy in all test conditions
- The difference, however, between the the CIS and SPEAK strategies was not statistically significant in any condition
- This is consistent with the findings of Kiefer *et al.* (1996) who found no significant difference between the CIS and SPEAK strategies with the easy Innsbruck sentences
- The lack of statistical significance may be partially due to a ceiling effect

Feature Analysis

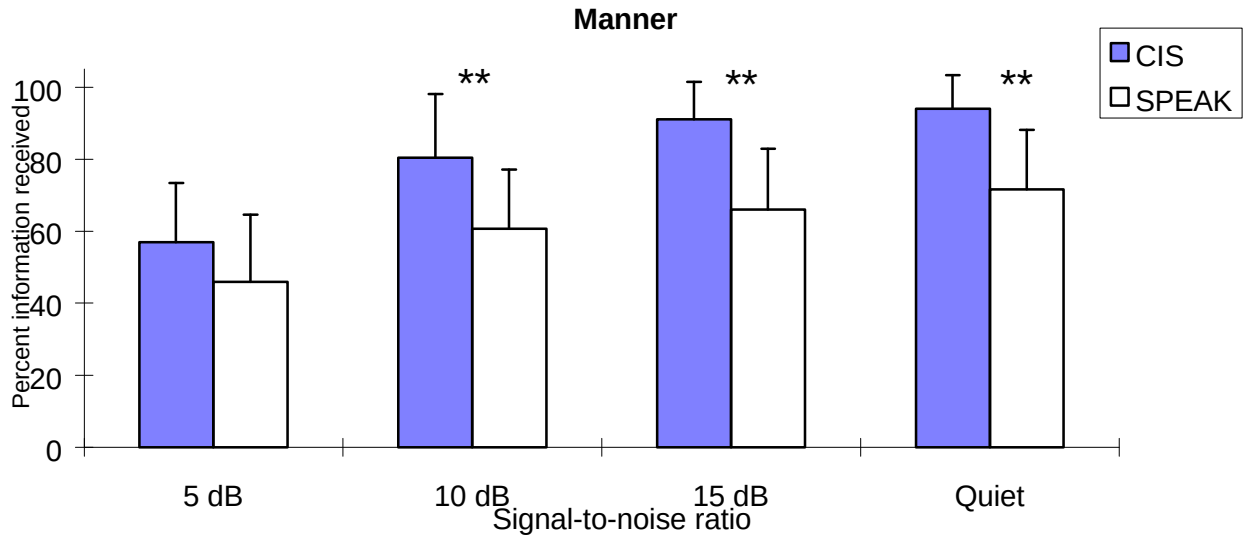
- The consonant confusion matrices were analyzed for information transmitted on the features of place, manner and voicing (Miller and Nicely, 1955)
- **Place**
  - The mean percent information received in place was higher with the CIS strategy than the SPEAK strategy
  - The difference between the CIS and SPEAK strategies was statistically significant ($p < 0.005$) only in quiet

Place


Place
CIS
SPEAK
*
Percent information received
100
80
60
40
20
0
5 dB
10 dB
15 dB
Quiet
Signal-to-noise ratio

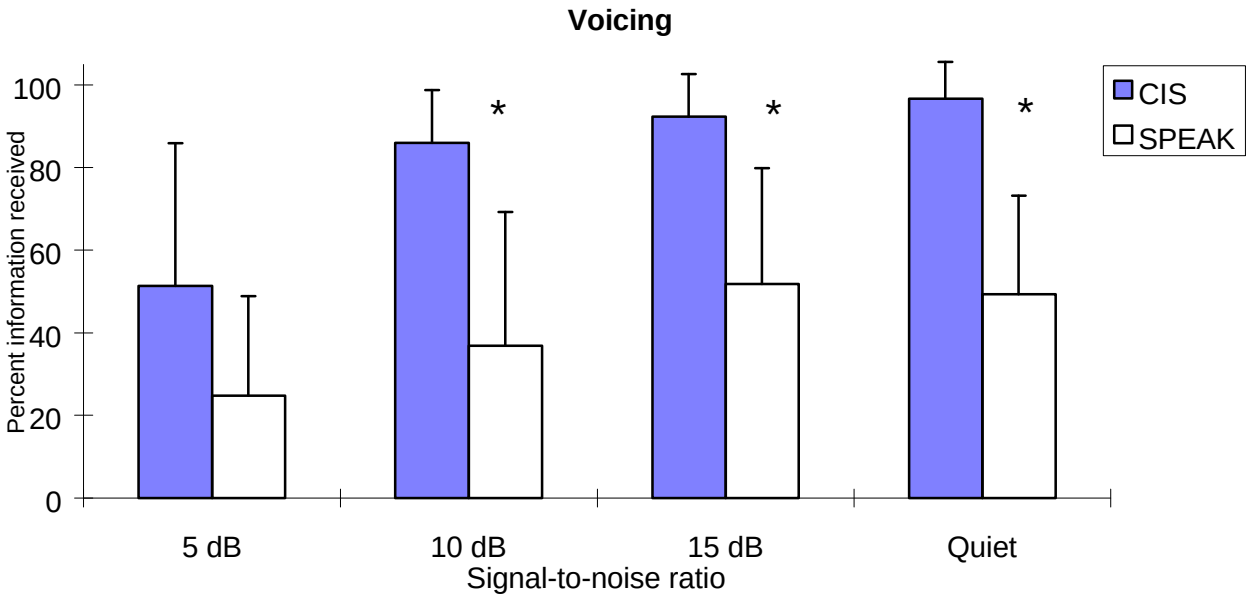
* $p < 0.005$

Manner


Manner
CIS
SPEAK
**
**
**
100
80
60
40
20
0
Percent information received
5 dB
10 dB
15 dB
Quiet
Signal-to-noise ratio

* $p < 0.05$

Voicing



* $p < 0.005$

Feature Analysis

- **Manner**
  - The mean percent information received in manner was higher with the CIS strategy than the SPEAK strategy
  - The difference between the CIS and SPEAK strategies was statistically significant ($p < 0.05$) at 10 dB S / N, 15 dB S / N and in quiet.
- **Voicing**
  - The mean percent information received in voicing was higher with the CIS strategy than the SPEAK strategy
  - The difference between the CIS and SPEAK strategies was statistically significant ($p < 0.005$) at 10 dB S / N, 15 dB S / N and in quiet.

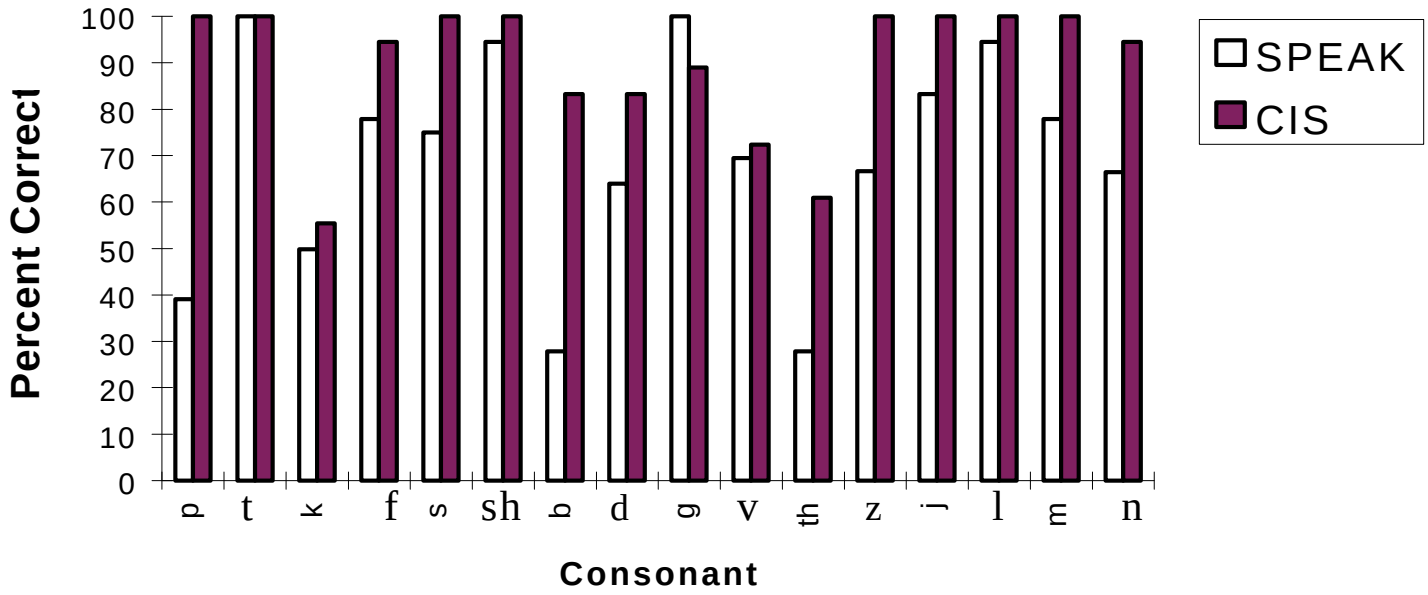
Feature Analysis: Summary

- Overall, the information transmitted in place, manner and voicing features was higher with the CIS strategy than the SPEAK strategy
- The difference between CIS and SPEAK strategies was especially large in the voicing feature
- High performance on place, manner and voicing cues was maintained with the CIS strategy even at a 10 dB S/N ratio.
- The greatest reduction in percent information transmitted in place, manner and voicing features occurred at 5 dB S/N with both strategies

Analysis of Consonant Scores

- The Figure below shows the mean percent correct scores of the 16 consonants in quiet by six “good” Nucleus patients (SPEAK) and by six “good” Ineraid patients (CIS)

Consonant recognition in quiet



Analysis of Consonant Scores

- The lowest consonant scores obtained with the SPEAK strategy occurred with: /b/, /th/ (as in *<u>th</u>en*), /p/ and /k/
  - /b/ was the most difficult to identify. Six of the eight patients identified /b/ with 0% accuracy
- The lowest consonant score obtained with the CIS strategy occurred with /k/. This is consistent with previous findings by Dorman and Loizou (1996). The most common error response for /k/ was /t/.

Discussion

- One possible explanation why the CIS strategy performs better than the SPEAK strategy on consonant recognition is the fact that the CIS strategy uses high pulse rate (>800 pps) stimulation.
- High pulse rate stimulation better represents rapidly changing temporal variations which are known to be important in consonant recognition
- The SPEAK strategy uses a pulse rate that varies between 180 and 300 pps depending on the number of maxima selected. The CIS strategy typically uses a pulse rate of 800 pps or higher.

Discussion

- One would expect that consonant recognition with the SPEAK strategy would improve by increasing the pulse rate.
- Indeed, Lawson *et al.* (1996) found out that consonant recognition improved significantly when the pulse rate of the n-of-m processor increased from 250 pulses / sec to 833 pulses / sec. The slower (250 pps) n-of-m processor produced a significant decrease in consonant recognition performance in six of the nine conditions tested with five Nucleus patients.

Discussion

- Is it possible that the Nucleus patients used in this study were all poor-performing patients? If that was true, then it would explain the low performance on consonant recognition.
- To answer this question, we compared the vowel and consonant scores (in quiet) with those reported by other researchers in other studies (see Table).
- As shown in the Table below, the performance of the Nucleus patients used in this study fell very close within the range of performance of other Nucleus patients.
- This suggests that our patients are representative of the Nucleus patients' population

Comparative results on vowel and consonant recognition in quiet using the SPEAK strategy

Study	Vowels (% correct)	Consonants (% correct)
Lawson *et al.* (1996) (n=5)	--	68 % (s.d.=8)
Fishman *et al.* (1996) (n=11)	80 % (s.d.=10)	68 % (s.d.=11)
Present study (n=11)	85 % (s.d.=13)	65 % (s.d.=16)

Conclusions

- The performance of the SPEAK and CIS strategies on vowel, consonant and sentence recognition was compared in quiet and in noise
- The mean performance of the CIS strategy was higher than the SPEAK strategy in both quiet and in noise for all test materials
- **In quiet:**
  - The difference between CIS and SPEAK strategies was only significant ($p < 0.005$) in consonant recognition.
  - No significant difference was found in vowel and sentence recognition. The lack of significance may be partially due to ceiling effects. Kiefer *et al.* (1996) have also found no significant difference between the CIS and SPEAK strategies in recognition of easy sentences in both quiet and noisy conditions.

Conclusions

- **In noise:**
  - No statistical difference was found between the two strategies in sentence recognition.
  - Performance in vowel recognition was significantly ($p<0.005$) higher with the CIS than the SPEAK strategy at 5 dB S/N ratio.
  - Performance in consonant recognition was significantly ($p<0.005$) higher with the CIS than the SPEAK strategy in all noise conditions.
- Performance with the CIS strategy on consonant recognition was superior to the SPEAK strategy. This may be due to the fact that higher pulse rate (>800 pps) stimulation is used in the CIS strategy.

References

- Dorman, M. and Loizou, P. (1996). "Improving consonant intelligibility for Ineraid patients fit with CIS processors by enhancing contrast among channel outputs," *Ear and Hearing*, **17**, 308-313.
- Fishman, K., Shannon, R. and Slattery, R. (1996). "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," *J. Speech Hear. Res.*, in press.
- Kiefer, J., Muller, J. et al. (1996). "Speech understanding in quiet and in noise with the CIS speech coding strategy (MED-EL Combi-40) compared to the Multipeak and Spectral Peak strategies (Nucleus)," *ORL*, **58,** 127-135.
- Lawson, D., Wilson, B., Zerbi, M. and Finley, C. (1996). Speech processors for auditory prosthesis, Third Quarterly Progress Report, *NIH Contract N01-DC-5-2103*.
- Nilsson, M., Soli, S. and Sullivan, J. (1994). "Development of the Hearing in Noise Test for the measurement of speech perception thresholds in quiet and in noise," *J. Acoust. Soc. Am.*, **95**, 1085-1099.