SPECTRO-TEMPORAL DISCRIMINATION IN COCHLEAR IMPLANT USERS

Elisabet Molin, Arne Leijon*

KTH, Royal Institute of Technology
Dept. of Signals, Sensors and Systems
Osquldasväg 10, SE-100 44 Stockholm, Sweden

Helene Wallsten

KI, Karolinska Institute
Huddinge University Hospital
SE-171 77 Stockholm, Sweden

ABSTRACT

A novel test method was designed to measure the spectro-temporal discrimination ability of cochlear implant (CI) users. The test signals are bandpass filtered, speech weighted noise, with the long term spectrum of speech. The goal of the test is to measure the amplitude difference between spectral bands in two presented signals which is required for the listener to just discriminate between the two sounds. Twenty CI users were tested with the spectro-temporal discrimination test and a conventional speech recognition test.

For test stimuli differing in only two spectral bands CI and normal hearing users show nearly equally good results on the spectro-temporal test. For spectra with four and more bands the spread in the CI users' results is great.

1. INTRODUCTION

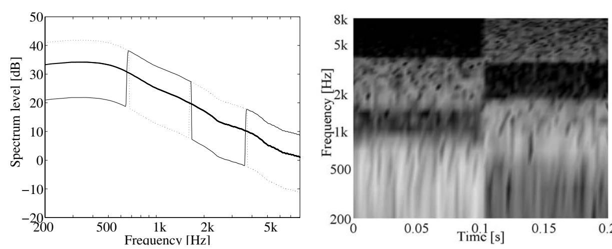
This article introduces a spectro-temporal discrimination test for cochlear implant (CI) users which might help explain the great speech recognition variability found among CI users, [1],[2].

Cochlear implant users may have a reduced spectro-temporal discrimination ability, due to a decreased resolution in the auditory system caused by the absence of well-functioning hair cells in combination with other factors, such as the the distance between the electrodes and neurons in the cochlea etc. A spectro-temporal test noise consists of two halves, that are the spectral complements of each other (fig.1). The noise is divided into a certain number of spectral bands. An alternative noise signal is created by swapping places of the two halves and the just noticeable amplitude difference between these two noises is the output of the test for each number of spectral bands (1,2,4,8,16 or 32). Many spectral bands will put higher requirements on the resolution of the cochlea.

The spectro-temporal discrimination test is language independent and designed to mainly test the function of the peripheral hearing. A benefit of the method is that its discrimination index can be calculated using a theoretical model of CI signal transmission [3]. Thus, the test can be used to evaluate the theoretical model. Results from a study on cochlear implant (CI) users comparing their results on this test and that on an ordinary speech test are shown. A normal hearing (NH) reference group is tested on the spectro-temporal discrimination test. The goal of the study is to answer the questions: How well do CI users perform compared to NH persons on the spectro-temporal test? Can the spectro-temporal discrimination test explain the differences in speech-recognition ability of CI users?

*This study was supported by the Swedish Research Council

2. SPECTRO-TEMPORAL STIMULI



(a) Filter spectra. (b) Spectrogram of a test signal.

Fig. 1. (a) Spectra for the "Odd" (dotted line) and "Even" (thin line) parts of the spectro-temporal test signals with four spectral bands at PVR 20 dB. The total loudness of both signals is the same, regardless of the PVR. The thick black curve shows spectrum for PVR 0 dB. (b) Spectrogram of an "OddEven" spectro-temporal test signal with four spectral bands at 20 dB PVR. Bright colours indicate high spectral contents. Each half of the signal is 100 ms. The odd half correspond to the dotted line in (a) and the even half corresponds to the thin line in (a).

The spectro-temporal discrimination test signals are created using Gaussian white noise shaped as the long-term spectrum of speech [4]. Two noise parts, one "Odd" and one "Even", are created using two filters with alternating pass bands and stop bands. Fig.1 shows the two filter frequency responses used to create noise parts with four spectral bands. Indicated in the figure is also the peak-to-valley ratio (PVR), which is a measure in dB of the spectral change between the two noise parts.

The final spectro-temporal test sounds are created by taking one of each band filtered part described above (one odd and one even), but in different time order (creating an "OddEven" and an "EvenOdd" signal). A spectrogram of the final OddEven signal for four bands and PVR 20 dB is found in fig.1.

Test signals were created to test the just noticeable difference for 1, 2, 4, 8, 16 and 32 spectral bands. Ten versions were created for each of the 23 PVR values 0, 0.33, 0.66, 1.0, 2.0, 3.0,..., 19.0 and 20.0 dB. For 1 spectral band the spectro-temporal test reduces to an intensity discrimination test.

3. SPECTRO-TEMPORAL DISCRIMINATION

An adaptive, modified up-down $3I3AFC_6$ procedure, that converges to 70.7% correct responses and $d' = 2.2$ [3], is used to find the just noticeable PVR for each type of spectro-temporal noise (4, 8, 16, 32, 1 and 2 spectral bands). With an adaptive method the test runs until the estimated standard deviation of the result is less than one dB and at least six test reversals have been made. The PVR-value is decreased in steps of 1 dB in the interval 1 to 20 dB. Below 1 dB the step size is $\frac{1}{3}$ dB.

Three consecutive test signals are presented at a fixed level (70 dB SPL) from a loudspeaker placed in front of the listener. Two of these signals are of the same type.

The listener is asked to indicate on a graphical user interface [5] which signal deviates from the other two. Each of the three test signals has a duration of 200 ms, and they are separated by a 300 ms pause. The test person begins each test with a practice run on four frequency bands, starting with 20 dB PVR. Then the real test starts, and the just noticeable difference is found (always starting from 15 dB) for 4, 8, 16, 32, 1 and 2 spectral bands. The CI users were tested twice on separate occasions. CI users were tested monaurally; any aid on the non-CI ear was turned off during the test. NH listeners used both ears and only performed the spectro-temporal test once.

4. SPEECH TEST

The CI users are tested using speech material referred to as Hagerman's sentences [6]. This test consists of a closed set of 50 proper Swedish words. Using these words, ten grammatically correct 5-word sentences are constructed. All sentences are semantically meaningless. The sentences are presented at a fixed level (70 dB SPL) in competing noise. An adaptive method [7] is used to set a new noise level for each sentence, converging to the signal-to-noise ratio (SNR) when the user has 40% correctly recognised words in a sentence. For users with poor speech recognition ability no competing noise is necessary to get a level of 40% correct answers or lower. In these cases the noise is turned off and the percentage of correctly repeated words are noted.

5. COCHLEAR IMPLANT USERS

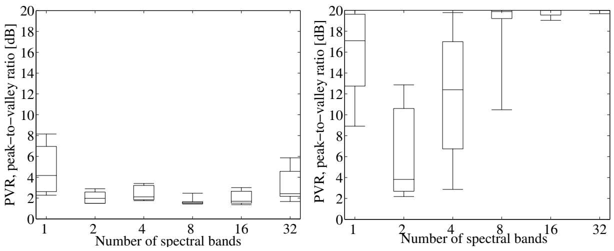
The CI users were selected from those available at Karolinska University Hospital, Huddinge. All users had implants that had been working for at least one year, and they had no neurological diseases. The participants were between 15 and 80 years old. The users were found by doing two searches in the clinic's database: One for CI users with open-set phonetically balanced word recognition results worse than, or equal to, 50% correct responses, which yielded 15 persons. An identical search for CI users with word recognition results better than 50% gave 20 persons. In the study 2 CI users with word recognition results less than or equal to 50% participated. Eighteen users with results better than 50% took part in the tests. The average age of the users was 59 years and the extremes were 35 and 75 years. A total of 11 women and 9 men participated in the study. Seventeen CI users completed the spectro-temporal discrimination test and retest as well as Hagerman's speech test (10 women and 7 men). Eight CI users used a Med-El implant with 12 electrodes, 10 used Nucleus with 22 electrodes and one used an Advanced Bionics implant with 16 electrodes. One implant brand is at the moment unknown.

6. NORMAL HEARING LISTENERS

A reference group of ten NH persons (25-35 years old) have done the spectro-temporal discrimination test.

7. RESULTS

The results of the 10 NH listeners on the spectro-temporal discrimination test (fig.2(a)) show that results average about 3 dB PVR for 2, 4, 8, 16 and 32 spectral bands. For one band (when the test reduces to an intensity test), the result is a little worse. Fig.2(b) shows the 20 CI user's results. For 2 spectral bands the results are good, not much worse than for NH listeners. However, for 4 spectral bands a much greater PVR is required for CI users than for NH users. At 8 bands only a few CI users succeed in discriminating the sounds. For 16 and 32 bands none of the CI users managed to discriminate the sounds. For 1 band, i.e. the intensity test, the results are also worse than for NH listeners. Test-retest PVR differences were generally (75-percentile) less than 2 dB.



(a) Normal hearing. (b) Cochlear implant.

Fig. 2. Peak-to-valley ratio in dB at the just noticeable spectral difference ($d' = 2.2$) plotted as a function of the number of spectral bands. Indicated in the figures are the median values, quartiles and 10% and 90% percentiles for 20 CI users and 10 NH listeners. For 17 CI users the average of test and retest results are plotted. For 3 users only one test result was available, this result was plotted.

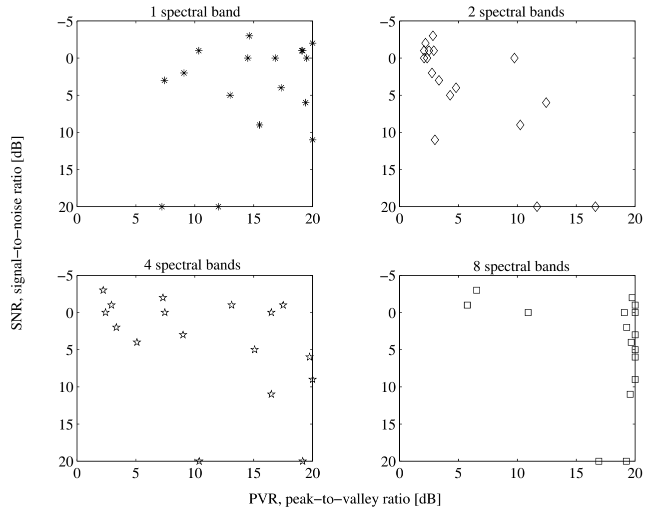


Fig. 3. Speech recognition ability, expressed as the signal-to-noise ratio (SNR) required for 40% speech recognition with Hagerman's sentences at 70 dB SPL, plotted versus the peak-to-valley ratio required for discrimination. Results for 1, 2, 4 and 8 bands are shown. Each marker indicates one of the 17 users. All SNR greater than 20 dB are set to 20 dB in the plots. Low SNR indicates better speech recognition performance.

Fig.3 shows the speech recognition ability plotted versus the just noticeable PVR for all 17 CI users in separate graphs for 1, 2, 4 and 8 spectral bands. The results for 2 spectral bands show a weak trend that the better speech performance a user had, the lower was also the just noticeable PVR difference.

8. DISCUSSION

Fig.2 shows that CI users do not perform very well on more than 2 spectral bands, despite the fact that they all use more than two electrodes in their implants. Results for 1 spectral band show a big variance for CI users. An explanation to this could be the automatic volume control of the CI processor.

Fig.3 shows that almost all CI users participating in this study have good results on the speech test. To hear the words in Hagerman's sentences is not a problem as long as the noise does not exceed the signal level too much (SNR ≈ 0 dB is a common result). Still, fig.2 shows that the spectral resolution of CI users is very limited. Apparently low spectral resolution (2 spectral bands) is enough to succeed very well in this speech test. [8] and [9] also showed that good speech recognition can be achieved with very limited spectral resolution.

Two users have poor speech recognition results (PVR$<$ 50%), and for two spectral bands (fig.3) it can be seen that they also require a larger PVR than users with better speech recognition scores. To be able to explore this trend further the study needs to be extended to include more users with poor speech recognition results.

A possible problem with the spectro-temporal test was detected when testing some NH listeners outside this study. It was discovered that some did not understand the test procedure. With some more training these persons would perhaps be able to complete the test. None of the tested CI users had any problems with the test method. On the contrary, many commented on the fact that the spectro-temporal test very clearly showed their limitations in frequency resolution.

9. THE ELECTRICAL CI MODEL

An electrical model of the peripheral hearing of a CI user (see fig.4) has been developed [10]. Input sounds (e.g. OddEven or EvenOdd spectro-temporal test signals) are passed through the model and produce different patterns in the nerve bundles. The input sound is processed through a filter bank that determines the corresponding current pulse amplitudes in the electrodes. Using the assumption that the cochlea has uniform resistivity the current spread from the electrodes and the excitation patterns in the "neurons" are calculated. To reduce the complexity of the calculations the 30 000 neurons of an ordinary ear are grouped into bundles containing several neurons. The impulse pattern in the neuron bundles is the output of the model.

The modelled neural patterns of two sounds are used to calculate a difference measure with the same meaning as the discrimination index, d'. For the spectro-temporal discrimination test this value is 2.2. The settings of the CI model can be modified so that $d' = 2.2$ when the inputs are spectro-temporal test signals with the PVR set to the just noticeable difference for a particular user. Examples of model changes that can be tried in order to reproduce a user's result are altering the distance between the electrodes and neurons in the model and varying the assumed number of functioning neurons in the bundles. Speech test material can also be processed through the model to give an estimate of the amount of transmitted phonetic information.

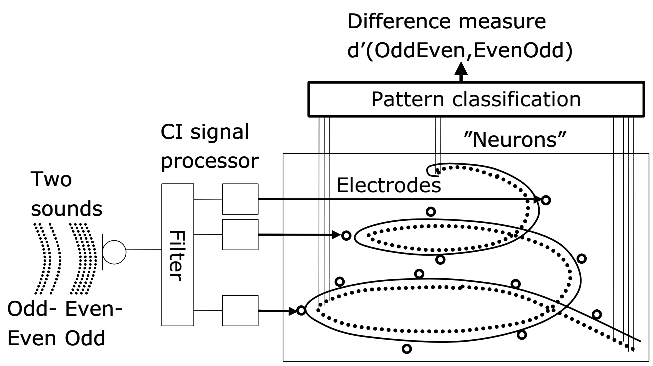


Fig. 4. Schematic picture of an electrical model of the peripheral hearing of a CI user [10]. An input sound is converted into the appropriate current pulse amplitude to each implant electrode. Each open circle in the picture represents an electrode. The current density was calculated at each nerve bundle. A nerve bundle contains a number of neurons and are indicated by the dots in the figure. The output of the model is an impulse pattern in the nerve bundles. After running two spectro-temporal test signals through the model their impulse patterns are used to calculate a difference measure corresponding to the discrimination index in the psycho-acoustic test.

10. CONCLUSION

For 2 spectral bands CI and NH users show nearly equally good results on the spectro-temporal test (fig.2). For four and more bands the spread in the CI users' results is great.

The spectro-temporal discrimination test can not explain the variations in speech recognition ability shown among CI users. However, there is a weak trend for 2 bands in fig.3 which indicates that good results on the spectro-temporal test is associated with a good speech recognition. To explore this trend further the study needs to be extended to include more users with poorer speech recognition results.

11. REFERENCES

[1] Anne Geers, Chris Brenner, and Lisa Davidson, "Factors associated with development of speech perception skills in children implanted by age five," *Ear and Hearing*, vol. 24, pp. 24–35, 2003.

[2] Jill B Firszt, Laura K Holden, Margaret W Skinner, Emily A Tobey, Ann Peterson, Wolfgang Gaggl, Christina L Runge-Samuelson, and P Ashley Wackym, "Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems," *Ear and Hearing*, vol. 25, no. 4, pp. 375–387, 2004.

[3] Niek J. Versfeld, Huanping Dai, and David M. Green, "The optimum decision rules for the oddity task," *Perception and Psychophysics*, vol. 58, no. 1, pp. 10–21, 1996.

[4] ANSI-S3.5, "Methods for calculation of the speech intelligibility index," Tech. Rep., American National Standards Institute, New York, 1997.

[5] Claes Björnemo, "Clinical test methods for spectro-temporal discrimination in cochlear implant users," M.S. thesis, KTH, Royal Institute of Technology, May 2003.

[6] Björn Hagerman, "Sentences for testing speech intelligibility," *Scandinavian Audiology*, vol. 11, pp. 79–87, 1982.

[7] Björn Hagerman and Catharina Kinnefors, "Efficient adaptive methods for measuring speech reception threshold in quiet and in noise," *Scandinavian Audiology*, vol. 24, pp. 71–77, 1995.

[8] Robert V Shannon, Fan-Gang Zeng, Vivek Kamath, John Wygonski, and Mikael Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, pp. 303–304, 1995.

[9] Qian-Jie Fu and Robert V Shannon, "Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners," *J. Acoust. Soc. Amer.*, vol. 105, no. 5, pp. 2570–2577, 1998.

[10] Torbjörn Nilsson, "A model of sound coding for cochlear implants," M.S. thesis, KTH, Royal Institute of Technology, 2004.