A Novel Stimulation of Cochlear Implant in Music Perception: Fine Structure Model by Using Nicstream

Fatih Öğüt, M.D., Erkan Zeki Engin, M.D., M. Akif Kılıç, M.D., Zafer Dicle, M.D., Mehmet Engin, M.D.

OBJECTIVES: We modified the software “Nicstream” to increase the quality of the music perception of the Nucleus-24 implant system, and thus, to improve the quality of the music perception in cochlear implant (CI) users by adjusting the cut-off frequency of the filters.

MATERIALS AND METHODS: In this study, Continuous Interleaved Sampling (CIS) and Advanced Combinational Encoder (ACE) digital signal processing strategies for the Nucleus-24 implant system were used. We modified the current software to design a novel stimulation cochlear implant in music perception at the fine structure model with different division of the frequency scale that was based on Helmholtz. We examined a musical melody and eight notes of a musical instrument. First, the music quality of the output files was determined subjectively by five musicians, who listened the output music signals and gave a score for melody, harmony, rhythm, and pitch discrimination of the musical melody and the notes of the musical instrument. A score between 0 (poor result) and 10 (good result) was accorded. Then the sounds were objectively evaluated by signal processing using a software program (Hypersignal) to obtain digitized sonograms.

RESULTS: The scores of melody, harmony and pitch discrimination of the CIS using Helmholtz filter frequencies were considerably higher. Spectrographic analyses and evaluation of the sounds showed that CIS strategy is more appropriate than ACE strategy for CI listeners.

CONCLUSION: To improve the quality of the music perception in CI users, the software of the Nucleus-24 implant system must be modified by adjusting the cut-off frequency of the filters using frequency scales such as Helmholtz. The optimum parameters can be determined for individual musical instruments and different kinds of music.
The provision of important useful information about speech and music through the electrical stimulation of the cochlea is a complex task. These technics depend on several factors including the function of a normal auditory system, the residual capacity of the impaired ear, the informational basis of speech and music mechanisms, and both the software and the hardware of the processing technology.

All hearing prostheses developed so far comprise an external part (the sound processor) and a separate internal part (the implanted electrodes, and electronic stimulator, if required). An important research problem is to optimize the way sounds are processed externally so that the resulting electric stimulation provides as much auditory information as possible.

The resolution of complex spectral patterns is one of the important aspects in perceiving speech. The speech processor typically resolves the frequency components in the signal using band-pass filtering, and the spectral shape information in the signal is then represented in the pattern of stimulation across the electrode array. The spectral resolution of cochlear implant (CI) listeners depend first the number of channels (the ability of providing spectral detail), and also on the way in which the acoustic signal is processed and represented in the electrical stimulation; second, the ability of the individual CI listener to perceive this electrical representation of the spectral information.

The number of stimulating channels in current systems is generally limited to 22, depending on the device and speech processing strategy. Thus, speech processors do not preserve the fine spectral detail. Many studies have been made to understand the effect of the number of channels on speech performance of CI listeners. These studies have shown that, in quiet listening conditions, high levels of speech understanding can be obtained at 4 to 8 spectral bands and increase in the number of channels beyond 8 result in little improvement in performance. However, when listening in background noise, the increase in the number of channels improves intelligibility depending on the level of noise.

While many CI users enjoy success regarding speech understanding, most are still frustrated by their inability to accurately hear music that a typical music listener experiences; implant recipients enjoy music less post-implantation than they did prior to hearing loss. Moreover, the ability for an accurate differentiation of pitch information is crucial for the understanding and production of a tone language (e.g. Cantonese or Mandarin), in which changes in the fundamental frequency pattern within a phonemic segment determine the lexical meaning. Despite the limited amount of sensory information provided by a CI, postlingually deafened CI users process musical information through the same neural mechanisms as normal hearing controls, and brain functions processing musical irregularities are still active.

It seems very likely that a cognitive representation of the major-minor tonal system is established before the patients lose their hearing, and this representation is still a basic reference for musical analysis. Thus, an improved accuracy of music perception in CI users by modifying their processors can be provided by an exact representation of musical regularities.

In this paper, we modified the software “Nicstream” to increase the quality of the music perception of the Nucleus-24 implant system by adjusting the cut-off frequency of the filters using different division of the frequency scale that was based on Helmholtz. Helmholtz proposed that the ear contains an array of sympathetic resonators, like piano strings, which serve to give the ear its fine frequency discrimination.

**MATERIALS AND METHODS**

In this study, Continuous Interleaved Sampling (CIS) and Advanced Combinational Encoder (ACE) digital signal processing strategies for the Nucleus-24 implant system were used. These approaches involve band-pass filtering the input signal into N bands and modulating N sinusoids, the center frequencies of the filters, by the envelope of each band. The sum of the modulated sinusoids is then presented to a loudspeaker.

The filters in these strategies implemented for the Nucleus-24 implant system use a combination of linear and logarithmic division of the frequency scale and the channel lower and upper cut-off frequencies (default values) for signal processors with different number of channels are given in. In this study, we
used a different division of the frequency scale that was based on Helmholtz and the optimum channel number for this frequency scale was investigated. We tried to determine the frequency scale similar to that of musical tones.

The dynamic range of the envelope signal is compressed by a Loudness Growth Function (LGF). It is a logarithmically shaped function that maps the envelope amplitude to a magnitude in the range 0 to 1. The default and ±20 changed values of the LGF were tested for both of the frequency scales.

All of these states were examined for a musical melody and eight notes of a musical instrument, and the optimum parameters for musical perception at these sounds were determined.

The program was written in MATLAB; it takes WAV audio files as input and generates WAV files as output.

First, the music quality of the output files was determined subjectively by a jury of five different musicians. Each musician listened the output music signals recorded on a CD in privacy in order not to be affected by the others. The listeners gave a score for: 1) melody; 2) harmony; 3) rhythm; and 4) pitch discrimination (PD) of the musical melody and the notes of the musical instrument. A score between 0 (poor result) and 10 (good result) was accorded.

We examined the output WAV files with the use of the Hypersignal software. The spectrographic analyses of the music and the eight musical notes of a musical instrument was made.

**RESULTS**

The music test scores of the musicians are presented in Table 1.

For the spectrographic analyses in the first two figures, Hamming window with a frame size of 256 points and 99% overlap was used with 16 kHz sampling frequency.

**Fig. 1.** The spectrographic analysis of musical melody: **A.** Original. **B.** After processing with ACE strategy (default filter frequencies) for 12 channels. **C.** After processing with CIS strategy (default filter frequencies) for 12 channels. **D.** After processing with CIS strategy (Helmholtz filter frequencies and increasing +20 of LGF value) for 12 channels.
The spectrographic analysis of the musical melody is shown in Fig. 1a; the signal showed no frequency component upon the frequency of 2.3 kHz.

The spectrographic analyses of the signals after processing the musical melody with ACE strategy (default filter frequencies), CIS strategy (default filter frequencies), and CIS strategy (Helmholtz filter frequencies and increasing +20 of LGF value) for 12 channels are shown in Fig. 1b-d.

In Fig. 1b and 1c, the energy at high frequencies is more powerful than that in Fig. 1d (Helmholtz design).

The spectrographic analysis of the music instrument at eight notes is shown in Fig. 2a; the signal showed no frequency component upon the frequency of 2.5 kHz.

The spectrographic analyses of the signals after processing the music instrument at eight notes with ACE strategy (default filter frequencies), CIS strategy (default filter frequencies), and CIS strategy (Helmholtz filter frequencies and increasing +20 of LGF value) for 12 channels are shown in Fig. 2b-d. In Fig. 2b and c, the energy at high frequencies is more powerful than that in Fig. 2d.

Table 1. The music test outcomes of the musicians for the original and processed sounds with different strategies

<table>
<thead>
<tr>
<th></th>
<th>Melody</th>
<th>Harmony</th>
<th>Rhythm</th>
<th>Pitch discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>9.6</td>
<td>9.2</td>
<td>9.6</td>
<td>9.4</td>
</tr>
<tr>
<td>ACE</td>
<td>2.0</td>
<td>2.8</td>
<td>5.0</td>
<td>3.6</td>
</tr>
<tr>
<td>CIS</td>
<td>3.0</td>
<td>3.4</td>
<td>5.2</td>
<td>5.4</td>
</tr>
<tr>
<td>Proposed</td>
<td>5.0</td>
<td>5.2</td>
<td>5.6</td>
<td>6.4</td>
</tr>
</tbody>
</table>

Fig. 2. The spectrographic analysis of music instrument at 8 notes: A. Original. B. After processing with ACE strategy (default filter frequencies) for 12 channels. C. After processing with CIS strategy (default filter frequencies) for 12 channels. D. After processing with CIS strategy (Helmholtz filter frequencies and increasing +20 of LGF value) for 12 channels.
For the spectra analyses in Fig. 3 and 4, Hamming window with a frame size of 1024 points and 99\% overlap was used with 16 kHz sampling frequency.

The spectra analyses of note ‘fa’ for CIS strategy (default filter frequencies) and CIS strategy (Helmholtz filter frequencies) are shown in Fig. 3a and 3b, respectively.

**DISCUSSION**

The spectral resolution of CI listeners depends on the number of channels and also on the way in which the acoustic signal is processed and represented in the electrical stimulation. The ability of the individual CI listener to perceive this electrical representation of the spectral information is also important. Speech processors do not preserve the fine spectral detail. With the current software and processing techniques, many CI users enjoy successful speech understanding.

However, listening to music often becomes unpleasant after cochlear implantation. Twenty-five CI recipients (86\%) presented lower scores of listening habits after implantation, and 11 (38\%) stated that they did not like listening to music.\[8\]

Alternative approaches to signal processing may be necessary in order to improve song recognition abilities of CI recipients to the levels of normal hearing adults. New perspectives for improving music perception in CI subjects have been explored, involving changing the parameters of the code strategy. New signal processing strategies, designed to improve pitch and music perception, must be researched.

**Fig. 3.** The spectra analyses of notes “fa” with CIS strategy for 12 channels: **A.** Default filter frequencies. **B.** Helmholtz filter frequencies and increasing +20 of LGF value.

**Fig. 4.** The spectra analyses of note “fa” for reducing the minimum bandwidth.
In our program, a fine structure model was designed by adjusting the cut-off frequency of the filters using different division of the frequency scale that was based on Helmholtz. Helmholtz proposed that the ear contains an array of sympathetic resonators, like piano strings, which serve to give the ear its fine frequency discrimination.

The output music signals, which were processed by the default and our fine structure model parameters, were assessed perceptually by five musicians. The scores of melody, harmony and pitch discrimination of the CIS using Helmholtz filter frequencies were significantly higher. There was not a significant change in the evaluation of the rhythm.

The results of spectrographic analyses and evaluation of the sounds show that CIS strategy is more appropriate than ACE strategy for CI listeners and 12 channels give good performance for music perception, while 4 to 7 channels are still enough for speech perception.

As shown in Fig. 1c, d and Fig. 2c, d, with default values, the energy at high frequency is more powerful than our design, whereas the original sounds have nearly no component at high frequencies. Thus, these frequencies can be accepted as noise, which means that SNR (Signal-to-Noise Ratio) is better in our design.

At the default values, only one or two harmonics can be effective (Fig. 3a), whereas in our design it was found that more harmonics could be effective (Fig. 3b). This situation is very important for music perception because the melody of the music can be determined by the harmonics. Thus, a more accurate temporal fine structure can be supplied by our design.

The bandwidth of the band-pass filters are the integer times of 125 Hz in all these applications. The spectra analyses of note ‘fa’ in our design is shown in Fig. 4, when this frequency is reduced to 62.5 Hz. In this situation, the music perception improves because of a more accurate temporal fine structure. However, it is not possible because the instrument stage does not support this requirement.

 Loudness Growth Function is an important parameter for perception. Increasing this parameter improves the intelligibility of the sound. However, if it is too much increased, some high frequency musical notes will be lost (Fig. 2d). Therefore, selection of this parameter is critical. It is advisable that these analyses be made for all the musical instruments, after which global optimum parameters can be determined.

In this study, processing of the sound was conducted on the computer. In the future, these strategies will be applied to CI listeners, and in the light of their mapping, these strategies and parameters will be better adjusted.

To improve the quality of the music perception in CI users we only modified the software of the Nucleus-24 implant system by adjusting the cut-off frequency of the filters. These adjustments should also be applied to CI listeners with different mappings. The optimum parameters must be determined for different musical instruments and different kind of music. To obtain a more accurate temporal fine structure, the hardware of the speech processor must be developed; for eliciting more accurate frequencies of the musical tones and their harmonics the rate of the signal processing must be increased.

REFERENCES
