ISSN 0005–1144
ATKAAF 47(1–2), 69–74 (2006)

Branko Somek, Siniša Fajt, Ana Dembitz, Mladen Ivković, Jasmina Ostojić

Coding Strategies for Cochlear Implants

UDK 004.934:616.281-007.843
616.281-007.843:004.934
IFAC 5.9.2

Professional paper

The application of cochlear implants in the correction of hearing and speech impediments in deaf persons is known more than 25 years. This paper compares 3 different speech-coding strategies. These are the multipeak (MPEAK) coding strategy, the new spectral-peak (SPEAK) coding strategy and the high-rate continuous interleaved sampling (CIS) coding strategy. Results of several investigations show that the SPEAK and CIS coding strategies, based on spectral signal analysis, allow for better speech understanding in quiet as well as in noise, than MPEAK coding strategy, which relays on speech feature extraction.

Key words: CIS speech-coding strategy, MPEAK speech-coding strategy, SPEAK speech-coding strategy, cochlear implant

1 INTRODUCTION

The procedure for application of the cochlear implant (called the artificial cochlea in Croatia) can in principle be divided into three basic phases: 1. Diagnostic, 2. Operation, 3. Rehabilitation. The most complex, or rather the most difficult and the longest phase is the rehabilitation of the auditory capabilities of the person who has the cochlear implant. This process takes place gradually (»step by step«) and lasts several years, usually two. Cochlear implants have long proved to be successful in the treatment of deaf or severely hearing-impaired patients. Listening conditions in everyday life vary greatly from the ideal soundproof chamber. Competing environmental noise frequently impairs speech understanding. Several investigators have therefore considered it necessary to test sentence understanding under conditions of competing noise, and have been made to develop new speech coding strategies aimed at improving understanding in the presence of noise. These results are still preliminary, due to the relatively small number of patients and the great inherent intrasubject variability of results.

2 THE IMPLANT SYSTEM

The cochlear implant system consists of the implant components (receiver/stimulator and a series of electrodes) and the external components (microphone and speech processor). The system receives acoustic signals trough the microphone in the headset, and the acoustic parameters are processed in the speech processor. The acoustic parameters determine electrical stimulation parameters that are coded and emitted by a radio frequency to the receiver/stimulator located under the skin trough a coil above the skin. The receiver/stimulator decodes the information and transforms it into electrical stimulation impulses through a series of electrodes inside the cochlea. During rehabilitation it is necessary to determine the amplitude of the stimulating impulse for each individual patient, i.e. for each individual channel of the patient because each patient perceives volume and tone range differently. Psychophysiological and acoustic parameters are stored in the form of database in the speech processor. The administration of psychophysiological tests, the creation, storage and operation of data is conducted trough the DPS – Diagnostic Programming System.

3 PATIENT SELECTION CRITERIA

The conditions for receiving a cochlear implant are:

1. Deafness in both ears (the hearing threshold in both ears must be above 90 dB)
2. Inability to discern speech using any type of hearing aid
3. No inflammation in outer or middle ear
4. Openness of the cochlea, unhindered access to the fenestra rotunda and adequate pneumatiza-

tion of the mastoid antrum shown by computerized tomography or magnetic resonance imaging
5. The normal functioning of the vestibulocochlear nerve as verified by a promontory-test or an electroaudiogram
6. Positive mind about speech and communication ability (by speech pathologist)
7. Regular psychiatric and intelligence test result (child psychologist's report)
8. A positive attitude to this method of rehabilitation on the part of the user of cochlear implant and his/her family
9. The standard medical conditions for operating under general anesthetics must be satisfied.

In addition to the above, the following guidelines are also recommended:

- that the deafness developed after speech development (post-lingual deafness)
- that the person has not been deaf for more than 8 to 13 years
- that the person receiving the cochlear implant is not younger than 2 or older than 50–60 years.

Post-lingual deafness, duration of deafness and age are not mandatory conditions, but if they fulfilled, the rehabilitation period is shorter and the results in terms of hearing are much better.

4 CODING STRATEGIES FOR ELECTRICAL STIMULATION OF THE AUDITORY NERV

Initial concepts in speech coding for cochlear implants were based on the presentation of the analog waveform to a single electrode. To cope with the limited dynamic range for electrical stimulation of the auditory nerve, the incoming signal had to be compressed.

In spite of the fact that this analog signal itself preserves most of the original temporospectral information, the signal transfers to the auditory nerve is handicapped by the limited maximal firing frequency of the auditory nerve in response to the electrical stimulation. High synchronization of nerve fibers and the neural refractory period only allow for frequency transmission up to 1 kHz via temporal coding alone. For frequencies above 1 kHz, the spectral information cannot be sufficiently transferred by temporal coding alone. Multichannel implants have been developed to make use of the tonotopic organization in the cochlea and thus transmit more spectral information to the auditory nerve. The compressed analog strategy compresses the incoming microphone signal to the narrow dynamic range evoked hearing. The compressed signal is then filtered into several frequency bands for simultaneous analog stimulation of separate intracochlear electrodes. The presented signal still contains the exact temporal information and can make use of the tonotopic organization by stimulating spatially separated electrodes in the cochlea. The simultaneous stimulation on different electrodes is complicated by uncontrolled channel interactions and current summation. To avoid uncontrolled channel interactions, pulsatile coding strategies have been developed, which deliver nonsimultaneous pulses on multiple-electrode arrays. A pulsatile speech coding strategy was introduced based on the explicit extraction of features important for the recognition of speech. This speech coding strategy (F_0F_2) extracts the fundamental frequency of the voice (F_0) and the energy in the frequency region of the second formant (F_2). The fundamental frequency determines the frequency of the pulse ordered sequence, whereas the value of F_2 is used to select the electrode to be stimulated. Later strategies added information on first formant ($F_0F_1F_2$ strategy) and additional information 3 higher-frequency bands that partially overlapped with F_2 – MPEAK.

The SPEAK coding strategy no longer relies on the explicit extraction of speech features but is based on the spectral analysis of the signal. The signal is passed trough a digital filter bank of 20 frequency-specific filters. The filter output is continuously scanned and up to 6 channels with the highest energy is chosen for stimulation at varying rates. With an average of 250 Hz/s and channel. The average length of each analysis cycle is 4 ms. The increase of channels stimulated per cycle from 4 with the MPEAK to an average of 6 with the SPEAK strategy, as well as more flexible evaluation of the maxima and the stimulation of the corresponding electrodes, have been proven advantageous for the majority of patients.

The CIS strategy was designed to combine the advantages of nonsimultaneous pulses on multiple electrodes with the preservation of the fine temporal structure of the original analog signal by using brief pulses at high rate. With the CIS strategy, the signal is passed trough a digital filter bank (frequency-specific filters). The amplitude of each pulse is derived from the envelopes of the filter output signals, which are formed by rectification and low-pass filtering. With the high rate-rate CIS strategy, the channels are stimulated at 1500 pulses/s and channel. Irrespective of the theoretical advantages and disadvantages of each coding strategy, it's merits and drawbacks are finally determined by the empirical performance of the patients.

The Nucleus 22-channel (and recent 24) cochlear implant is currently the most commonly used multi-channel system. The Nucleus implantable electrode

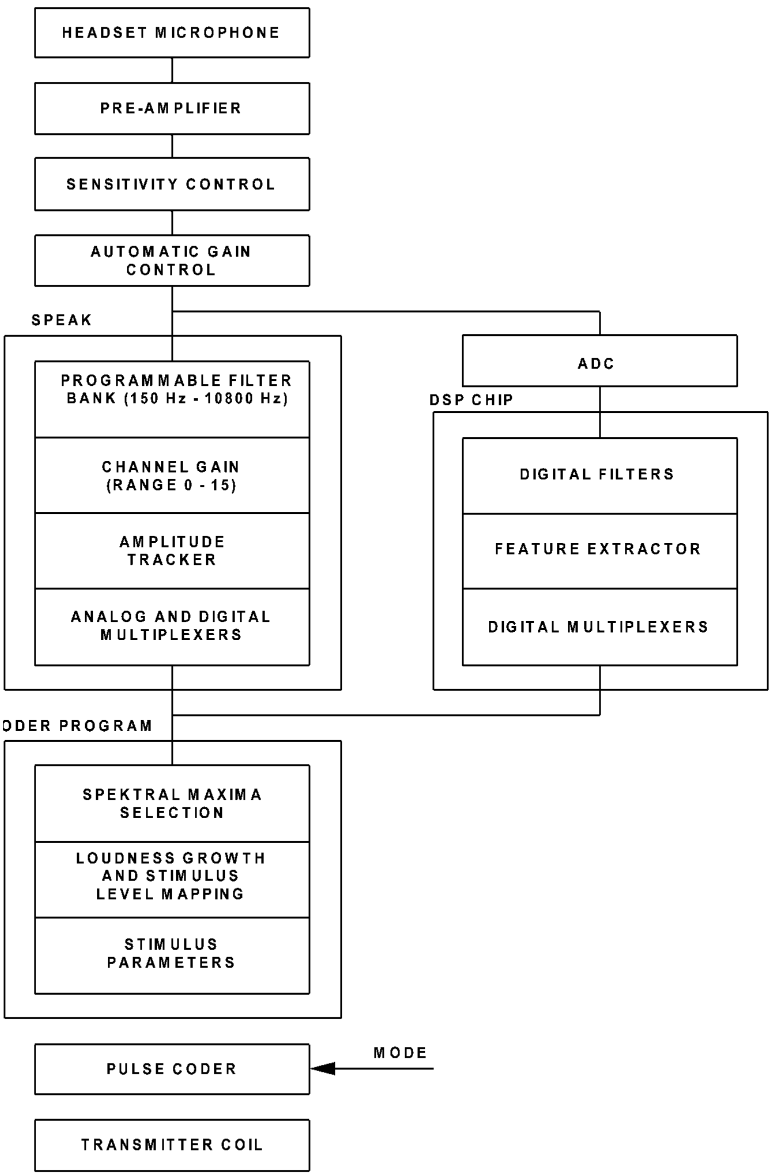

Fig. 1 The multipeak – MPEAK coding strategy

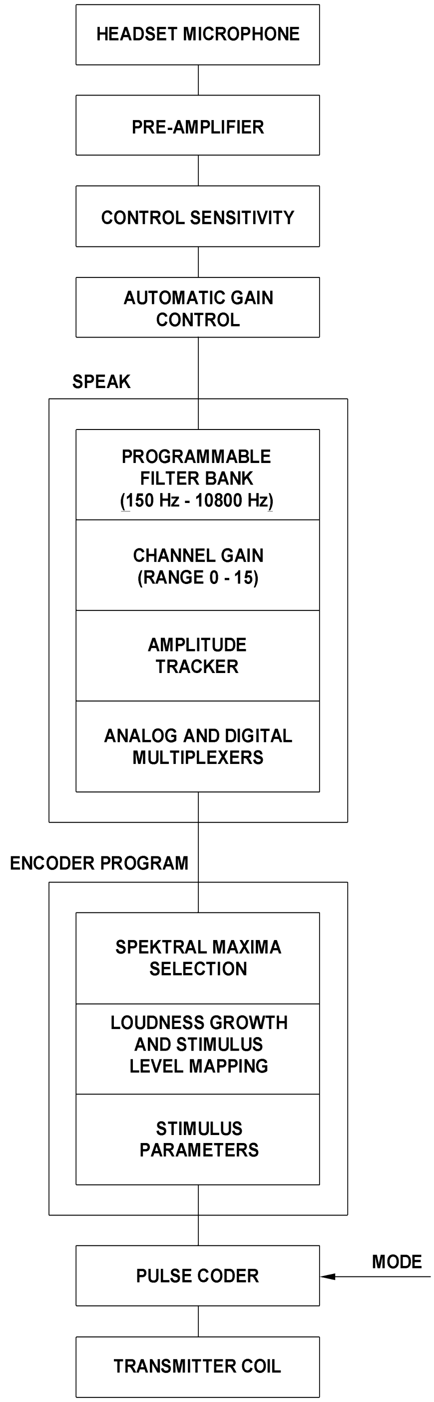

Fig. 2 The spectral-peak – SPEAK coding strategy

array consists of platinum-iridium band electrodes placed in a silastic carrier (Clark, 1987). Several generations of speech processors have been employed with the Nucleus multi-channel cochlear implant. The initial Nucleus speech processors used a feature-extraction scheme in which selected key features of speech were presented through the implanted electrode array. An early speech processing strategy, the F_0-F_1-F_2 strategy, primarily conveyed vowel information, including the first and the second formant frequencies and their amplitudes, as well as voice pitch. A later coding scheme, the MULTIPEAK strategy, presented these acoustic features along with additional information from three high-frequency spectral bands to aid in consonant perception. The current Nucleus speech processing strategy is the Spectral Peak (SPEAK) strategy. This strategy uses a vocoder in which a filterbank consisting of 20 filters covering the center frequencies from 200–10 000 Hz is employed. Each filter is al-

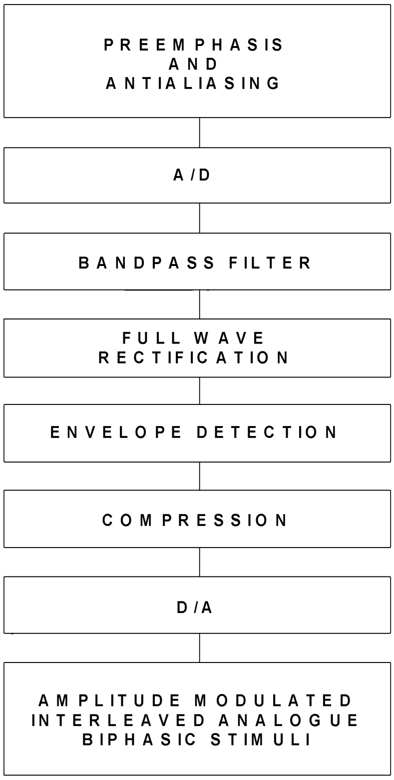

Fig. 3 The high-rate continuos interleaved sampling – CIS coding strategy

located to an active electrode in the array. The filter outputs are scanned and the electrodes that are stimulated represent filters that contain speech components with the highest amplitude. Depending on the acoustic input, the number of spectral maxima detected, and thus the number of electrodes stimu-

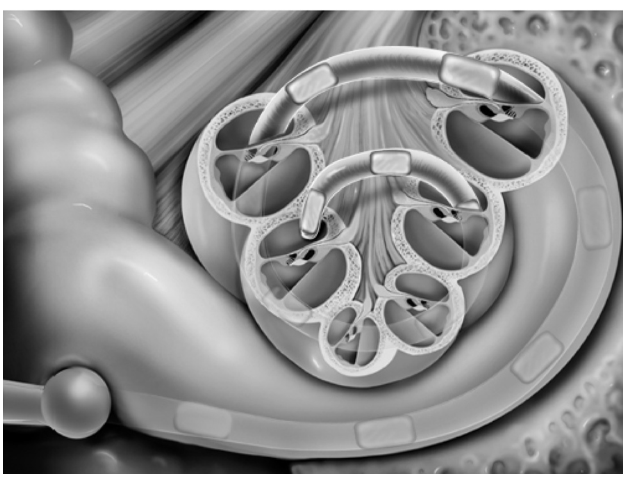

Fig. 4 Position of electrodes inside the cochlea near basilar membrane

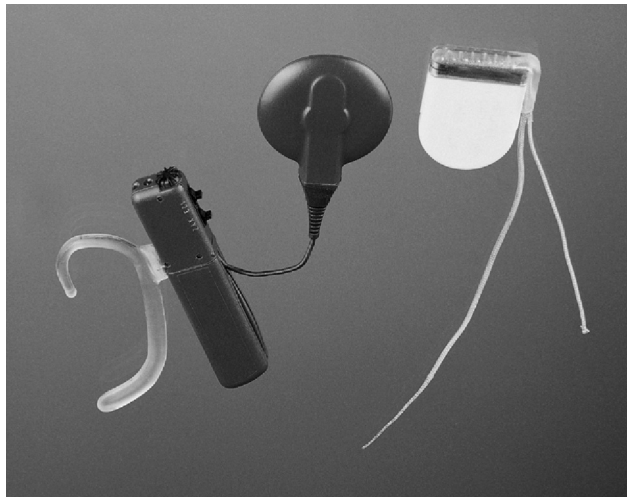

Fig. 5 The cochlear implant system consists of the implant components (right: receiver/stimulator and a series of electrodes) and the external components (left: microphone and speech processor)

lated, on each scan cycle can vary from one to 10, with an average of six per cycle. The rate at which the electrodes are stimulated varies adaptively between 180–300 pulses per second. The Clarion multichannel cochlear implant has an eight-channel electrode array that utilizes a radial bipolar configuration through electrode pairs positioned adjacent to the osseous spiral lamina in a 90-degree orientation (Schindler et al., 1986). The Clarion multi-channel cochlear implant offers two types of speech processing strategies: Simultaneous Analog Stimulation (SAS) and Continuous Interleaved Sampling (CIS). Both strategies represent the waveform or envelope of the speech signal (Wilson et al., 1991). The Clarion SAS strategy first compresses the analog signal into the restricted range for electrically-evoked hearing, and then filters the signal into a maximum of eight channels for presentation to the corresponding electrodes. Speech information is conveyed via the relative amplitudes and the temporal details contained in each channel. The CIS strategy filters the incoming speech into eight bands, obtains the speech envelope and compresses the signal for each channel. Stimulation consists of interleaved digital pulses that sweep rapidly through the channels at a rate of 833 pulses per second when using all eight channels for a maximum pulse rate of 6 664 pulses per second ($8 \times 833 = 6\,664$). With the CIS strategy, rapid changes in the speech signal are tracked by rapid variations in pulse amplitude. The pulses are delivered to consecutive channels in sequence to avoid channel interaction. The MED-EL COMBI 40-Cochlear Implant system utilizes the CIS (continuous interleaved sampling)-strategy that provides both spectral and temporal resolution. Up to eight active electrodes can be utilized. The electrode ar-

Fig. 6 The cochlear implant system consists: 1. speech processor, 2. microphone, 3. housing of battery, 4. battery, 5. emitter with magnet, 6. receiver/stimulator, 7. series of electrodes, 8 reference electrode

ray used has the capability of deep insertion into the apical regions of the cochlea (Gstöttner, Baumgartner, Franz & Hamzavi, 1997). The MED-EL has the capacity to provide the most rapid stimulation rate of any of the currently available implants (maximum of 12 000 biphasic pulses per second) (Hochmair, 1996).

5 THE COMPARISON CODING STRATEGIES

The results of studies which compares sentence understanding in quiet and noise show that the spectral-peak (SPEAK) and continuous-interleaved-sampling (CIS) strategies, based on spectral signal analysis, allow for better speech understanding in quiet as well as in noise, than the multipeak (MPEAK) coding strategy, which relays on speech feature extraction. In the intrasubject comparisons, the mean results in noise with the CIS strategy were better than both the MPEAK and the SPEAK strategies. Understanding in quiet was not significantly different between the CIS and the SPEAK strategies; both strategies were significantly better than the PEAK strategy in quiet. These results are still preliminary, due to the relatively small number of patients and the great inherent intersubject variability of results. Comparing different studies is difficult due the differences in study design. Besides differences for example in test material, testing conditions, follow-up and training of patients, the selection criteria may influence the test results to a large extent. Selection may be based on explicitly stated criteria or on various factors such as the availability of subjects and their cooperation. These factors may also influence selections without being explicitly stated. If these restricting factors are taken into accounts, careful comparisons can still be useful. This finding has to be interpreted with care because of the great inherent variability in performance that can only partially be predicted on the basis of the preoperative evaluation. In small numbers of subjects, results may still be influenced by casual factors. Some today and future implants should be flexible enough to incorporate various speech coding strategies, including high-rate stimulation strategies. This would allow for intrasubject comparisons and the evaluation and optimization of different speech coding strategies on a general as well as on individual basis.

6 CONCLUSIONS

Comparison of speech coding strategies is very difficult and complex procedure. It is very important and necessary to choose appropriate speech coding strategies for exact applications and concrete situation and every patient separately. The main problem was that a speech coding strategies was chosen before operations and implementation and couldn't be changed after it. Some new recent and future cochlear implant systems enable all coding strategies and patient can compare and chose coding strategy that is the most adequate and appropriate. This possibility is necessary for further investigations because enables research in same conditions. Speech coding strategies are closely connected with particular language and it is necessary to adapt rehabilitation for Croatian language.

REFERENCES

[1] K. Alberger, **Cochlear Implants bei Kinder**. Wiener Medizinische Wochenschrift, Heftn 1/2 1994.

[2] A. L. Boiter, J. A. Brimacombo, **Cohlear Implants, Rehabilitative Adiology: Children and Adults.** 2nd ed. alpiner JG, Veterans Administration Center Denver, Colorado, 1992. pp. 418–440.

[3] C. Boëx, M. Pelizzone, P. Montandon, **Improvements in Speech Recognition with the CIS Strategy for the Ineraid Multichannel Intracohlear Implant.** Advances in Cochlear Implants, Proceedings of the 3rd International Cochlear Implant Conference, Innsbruck, Austria, April 1993.

[4] M. Ivković, S. Fajt, A. Dembitz, **Signal Dynamics of Applied Cochlear Implant Simulation Methods.** ICECOM '97., Proceedings ICECOM'97 14th International Conference on Applied Electromagnetics and Communications, Dubrovnik, Croatia, October 15–17 1997, pp. 189–192.

Strategije kodiranja kod umjetne pužnice. Primjena umjetne pužnice kao pomagala pri slušanju i govoru u gluhih osoba poznata je više od 25 godina. Ovaj članak uspoređuje 3 različite strategije kodiranja govora. To su MPEAK strategija kodiranja više vršnih vrijednosti, nova SPEAK strategija kodiranja spektralnih vršnih vrijednosti i CIS strategija kodiranja kontinuirano-proširenim uzorkovanje velike brzine. Rezultati raznih istraživanja ukazuju da SPEAK i CIS strategije kodiranja, zasnovane na analizi spektra signala, omogućavaju bolju razumljivost govora u uvjetima tišine kao i buke, nego MPEAK strategija kodiranja koja se zasniva na izdvajanju karakterističnih frekvencijskih područja za govor.

Ključne riječi: CIS strategija kodiranja govora, MPEAK strategija kodiranja govora, SPEAK strategija kodiranja govora, umjetna pužnica

AUTHORS' ADDRESSES

Branko Somek
Siniša Fajt
Faculty of Electrical Engineering and Computing
Unska 3, HR-10000 Zagreb, Croatia

Ana Dembitz
Mladen Ivković
Zagreb University Clinical Hospital Centre
Šalata 4, HR-10000 Zagreb, Croatia

Jasmina Ostojić
Ministry of Labor and Social Welfare
Prisavlje 14, HR-10000 Zagreb, Croatia

Received: 2006-03-30