We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

A Fine Structure Stimulation Strategy and Related Concepts

Clemens Zierhofer and Reinhold Schatzer
C. Doppler Laboratory for Active Implantable Systems
University of Innsbruck
Austria

1. Introduction

The auditory system provides a natural frequency-to-place mapping which is designated as *tonotopic organisation* of the cochlea. In the normal-hearing system, the acoustic signal causes a fluid pressure wave which propagates into the cochlea. At particular positions within the cochlea, most of the energy of the wave is absorbed causing mechanical oscillations of the basilar membrane. The oscillations are transduced into electrical signals (action potentials) in neurons by the action of the inner hair cells. Waves caused by low input frequencies travel further into the cochlea than those caused by high frequencies. Thus, each position of the basilar membrane can be associated with a particular frequency of the input signal (Greenwood, 1990). This natural form of frequency-to-place mapping, together with the fact that the positioning and fixation of an intrascalar electrode array is comparatively simple, is likely one of the most important factors for the success of cochlear implants as compared to other sensory neural prostheses.

1.1 The "Continuous Interleaved Sampling" stimulation strategy

In the late 1980s, Wilson and colleagues introduced a coding strategy for cochlear implants designated as "Continuous Interleaved Sampling" (CIS) strategy (Wilson et al., 1991). Supporting significantly better speech perception in comparison to all other coding strategies at the time, CIS became and still is the de-facto standard among CI coding strategy. CIS signal processing involves splitting up of the audio frequency range into spectral bands by means of a filter bank, envelope detection of each filter output signal, and instantaneous nonlinear compression of the envelope signals (map law).

According to the tonotopic principle of the cochlea, each stimulation electrode in the scala tympani is associated with a band pass filter of the external filter bank. High-frequency bands are associated with electrodes positioned more closely to the base, and low-frequency bands to electrodes positioned more deeply in the direction of the apex. For stimulation, charge-balanced current pulses - usually biphasic symmetrical pulses - are applied. The amplitudes of the stimulation pulses are directly derived from the compressed envelope signals. These signals are sampled sequentially, and, as the characteristic CIS paradigm, the stimulation pulses are applied in a strictly non-overlapping way in time. Typically, the pulse sampling rate per channel is within the range of 0.8-1.5 kpulses/sec.

Many investigators have aimed at finding optimum CIS parameters for best speech perception, e.g., for parameters such as number of channels, stimulation rate per channel, etc. (Loizou et al., 2000). It turns out that four channels seem to be the absolute minimum number of channels for reasonable speech intelligibility, and speech perception reaches an asymptotic level of performance at about ten channels. Stimulation rates higher than about 1.5 kpulses/sec per channel also do not substantially improve speech perception.

Other approaches to improve speech recognition are based on the idea that neurons should have some activity even in the absence of an acoustic input. In a normal-hearing system such an activity is present and is generally designated as "spontaneous activity", i.e., neurons produce action potentials and cause a type of noise floor of neural activity. Following the principle of stochastic resonance (Morse & Evans, 1999), such a noise floor could provide a more natural representation of the envelope signals in the spiking patterns of neurons. In the deaf ear, there is no or very little spontaneous activity of neurons. To combine the CIS strategy with principles of stochastic resonance, it has been suggested to introduce high-frequency pulse trains with constant amplitudes - so-called "conditioner pulses" - in addition to the CIS pulses (Rubinstein et al., 1999). However, such approaches have so far not found their way into broad clinical applications, partly because no substantial improvement in speech intelligibility has been found and partly for practical reasons, because the power consumption of the implants is considerably increased.

1.2 Temporal fine structure

According to the principles of the Fourier transform, each signal can be decomposed into a sum of sinusoids of different frequencies and amplitudes. Following Hilbert (Hilbert, 1912), an alternative way of signal decomposition is to factor a signal into the product of a slowly varying *envelope* and a rapidly varying *temporal fine structure*. Hilbert's decomposition is particularly useful in the case of band pass signals as used in a CIS filter bank. Considering the response of a band pass filter to a voiced speech segment, the envelope carries mainly pitch frequency information. Temporal fine structure information is, first of all, present in the position of the zero crossings of the signal and shows the exact spectral position of the center of gravity of the signal within its band pass region, including temporal transitions of such centers of gravity. For example, the temporal transitions of formant frequencies in vowel spectra are highly important cues for the perception of subsequent plosives of other unvoiced utterances. Furthermore, a close look at the details of a band pass filter output also reveals that the pitch frequency is clearly present in the temporal structure of the zero crossings.

It is known from literature that the neurons in the peripheral auditory system are able to track analogue electrical sinusoidal signals up to about 1 kHz (Hochmair-Desoyer et al., 1983; Johnson, 1980). However, CIS is entirely based on envelope information, and temporal fine structure information is largely discarded. For example, consider the response of an ideal CIS system (filter bank composed of ideal rectangular band pass filters) to a sinusoidal input signal with constant amplitude and a frequency which is swept over the whole input frequency range. Then one channel after the other will generate an output. The CIS response reflects the approximate spectral position of the input signal (i.e., the information, *which* band filter within the filter bank is responding), but within each responding filter, there is no further spectral resolution.

Smith and colleagues (Smith et al., 2002) described an experiment with normal-hearing subjects based on a vocoder-type signal processor similar to CIS. Audio input signals were split up by means of a filter bank, and both envelope and temporal fine structure information was extracted from each channel. So-called "auditory chimaeras" were then constructed by combining, across channels, the envelopes of one sound with the temporal fine structure of a different sound. By presenting these auditory chimaeras to normal-hearing listeners, the relative perceptual importance of envelope and fine structure was investigated for speech (American English), musical melodies, and sound localization. The main outcome was that for an intermediate number of 4 to 16 channels, the envelope is dominant for speech perception, and the temporal fine structure is most important for pitch perception (melody recognition) and sound localization. When envelope information is in conflict with fine structure information, the sound of speech is heard at a location determined by the fine structure, but the words are identified according to the envelope.

In the light of these results, standard CIS is a good choice for the encoding of speech in Western languages (e.g., for American English). However, regarding music perception and perception of tone languages (e.g., Mandarin Chinese, Cantonese, Vietnamese, etc.), CIS might be suboptimal due to the lack of temporal fine structure information.

1.3 Spatial channel interaction

One basic problem in cochlear implant applications is spatial channel interaction. Spatial channel interaction means that there is considerable geometric overlap of the electrical fields at the location of the excitable neurons, if different stimulation electrodes in the scala tympani are activated. Thus, the same neurons can potentially be activated if different electrodes are stimulated. Spatial channel interaction is mainly due to the conductive fluids and tissues surrounding the stimulation electrode array.

One approach to defuse spatial channel interaction is to employ particular electrode configurations which aim at concentrating electrical fields in particular regions. Bipolar or even multipolar configurations have been investigated. These configurations have in common that active *and* return electrodes are positioned within the scala tympani of the cochlea (Miyoshi et al., 1997; van den Honert & Kelsall, 2007; van den Honert & Stypulkowski, 1987). For example, a bipolar configuration uses single sink and source electrodes which are typically separated by 1-3 mm. The main disadvantage of these approaches is that comparatively high stimulation pulse amplitudes are necessary to achieve sufficient loudness. This is because the conductivity between the active electrodes represents a low-impedance shunt conductance which causes most of the current to flow within the scala tympani and only a small portion to reach the sites of excitable neurons.

Regarding power consumption, the most effective electrode configuration is the monopolar configuration. Monopolar means that only one *active* electrode is within the scala tympani, and a remote return electrode is positioned outside. E.g., a 12-channel CIS system employing monopolar stimulation uses a 12-channel electrode array within the scala tympani and a return electrode which may by positioned under the temporal muscle and is shared by all 12 electrodes. In CIS, only one electrode of the array is active at any given time (non-overlapping pulses).

Remarkably, although spatial channel interaction in bipolar or multipolar configurations is substantially less than in monopolar configurations, this advantage in general cannot be converted to better speech perception, as has been shown in a number of studies (Loizou et al., 2003; Stickney et al., 2006; Xu et al., 2005). Thus, monopolar configurations are widely used in clinical applications.

The channel separation in systems using monopolar electrode configurations may be improved by a closer positioning of the active electrodes to the excitable structures. Such special electrode arrays are designated as "modiolus-hugging" or "perimodiolar" arrays, and a variety of designs have been developed (Lenarz et al., 2000; Wackym et al., 2004). However, one basic requirement of modiolus-hugging electrode arrays is that after implantation the array should not exert any permanent mechanical pressure on the structures within the cochlea, because mechanical pressure can damage neural tissue or cause ossification if applied to bony structures. As a matter of fact, this requirement is difficult to meet in most approaches. Temporal bone studies have shown that modiolus-hugging electrodes increase the danger of basilar membrane perforations and spiral lamina fractures (Gstoettner et al., 2001; Richter et al., 2002; Roland et al., 2000). Besides, even if the array is closer to the modiolus, the surrounding conductive fluids will still cause broad potential distributions, and it is questionable whether a substantial improvement of channel separation is possible at all.

2. Supporting concepts

Fine structure stimulation strategies require a precise representation of temporal information in individual electrode channels. Maintaining the CIS paradigm of using non-overlapping pulses, the precision of representation is limited by the minimum phase duration of sequentially applied biphasic stimuli (Zierhofer, 2001). To relax such limitations and allow the implementation of fine structure stimulation with reasonable phase durations on multiple channels, two "supporting concepts" are investigated, i.e., simultaneous stimulation in combination with "Channel Interaction Compensation" (CIC) and the "Selected Groups" (SG) concept.

Both supporting concepts may generally be useful also for envelope-based stimulation without temporal fine structure. For example, the power consumption of a cochlear implant may be reduced without compromising the hearing performance.

As a first step, the supporting concepts have been evaluated with envelope-based coding strategies and compared to a standard CIS reference strategy.

2.1 Simultaneous stimulation

2.1.1 Channel interaction compensation

For the stimulation concept presented here, the *simultaneous* activation of two or more electrodes in the scala tympani against a remote return (ground) electrode is considered. As a basic requirement for simultaneous stimulation, pulses are 100% synchronous in time and have equal phase polarities. In the following, such pulses are designated as "sign-correlated".

The injection of a stimulation current into a single electrode will cause a particular voltage distribution within the scala tympani. If currents are simultaneously injected in more than one electrode, the individual voltage distributions are superimposed.

The basic idea of CIC is to replace a particular number of sequentially applied stimulation pulses (CIS paradigm) by the same number of simultaneously applied sign-correlated pulses. However, the amplitudes of the simultaneous pulses are reduced such that the electrical potentials at the position of the electrodes remain unchanged. Regarding the practical realization in a cochlear implant with limited space- and power resources, CIC based on general assumptions regarding voltage distributions is a computational challenge.

The computational cost for CIC can be reduced significantly, if a model of voltage distributions as shown in Fig. 1 is used (Zierhofer & Schatzer, 2008).
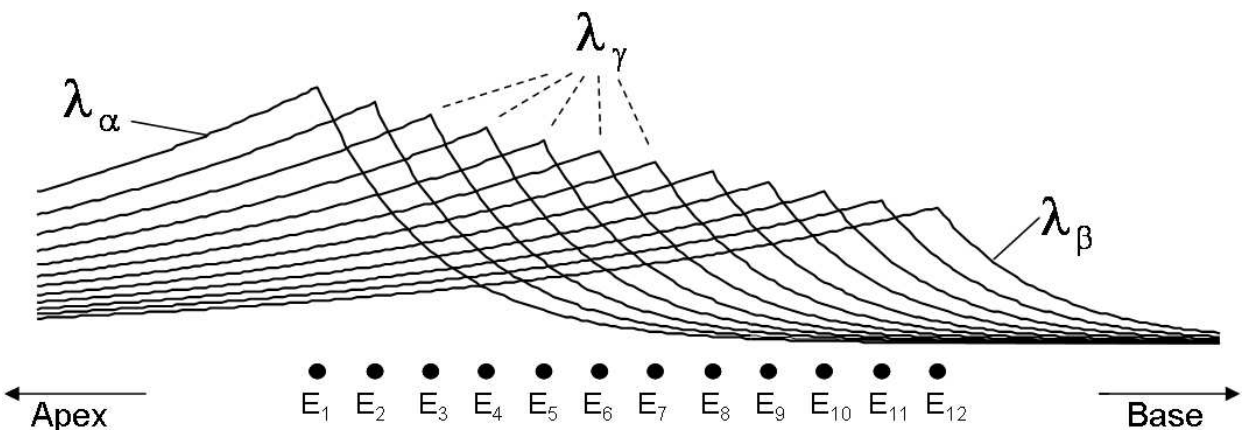

Fig. 1. Model of voltage distributions in the scala tympani as responses to unit currents in individual electrodes E_1, E_2, ..., E_{12} of a 12-channel array.

Each voltage distribution represents the response to a unit current in a single electrode and is composed of two branches with exponential decays (constants λ_α towards the apex, and λ_β towards the base). Due to the varying diameter of the scala tympani, the unit responses are place-dependent. The peaks of the distributions are also assumed to be located on an exponential characterized by constant λ_γ.

Assuming equally spaced electrodes (distance d), decay constants λ_α, λ_β, and λ_γ define CIC parameters α, β, and γ, i.e.,

$$\alpha = \exp\left(-\frac{d}{\lambda_\alpha}\right) \tag{1a}$$

$$\beta = \exp\left(-\frac{d}{\lambda_\beta}\right) \tag{1b}$$

$$\gamma = \exp\left(-\frac{d}{\lambda_\gamma}\right) \tag{1c}$$

Assuming M sequential amplitudes $I_{k,sequ}$ (k = 1, 2, ..., M) applied in M electrodes, the relation between the sequential amplitudes and the simultaneous amplitudes I_k (k = 1, 2, ..., M) is given by the set of linear equations

$$\begin{pmatrix} I_{sequ,1} \\ I_{sequ,2} \\ ... \\ I_{sequ,M} \end{pmatrix} = \mathbf{H} \begin{pmatrix} I_1 \\ I_2 \\ ... \\ I_M \end{pmatrix} \tag{2}$$

where **H** denotes the "interaction matrix"

$$\mathbf{H} = \begin{pmatrix} 1 & \gamma\alpha & (\gamma\alpha)^2 & ... & (\gamma\alpha)^{M-1} \\ \dfrac{\beta}{\gamma} & 1 & \gamma\alpha & ... & (\gamma\alpha)^{M-2} \\ \left(\dfrac{\beta}{\gamma}\right)^2 & \dfrac{\beta}{\gamma} & 1 & ... & (\gamma\alpha)^{M-3} \\ ... & ... & ... & ... & ... \\ \left(\dfrac{\beta}{\gamma}\right)^{M-1} & \left(\dfrac{\beta}{\gamma}\right)^{M-2} & \left(\dfrac{\beta}{\gamma}\right)^{M-3} & ... & 1 \end{pmatrix} \tag{3}$$

With (2), the simultaneous amplitudes are obtained with

$$\begin{pmatrix} I_1 \\ I_2 \\ ... \\ I_M \end{pmatrix} = \mathbf{H}^{-1} \begin{pmatrix} I_{sequ,1} \\ I_{sequ,2} \\ ... \\ I_{sequ,M} \end{pmatrix} \tag{4}$$

where $\mathbf{H}^{-1}$ represents the inverse matrix of **H**. It can be shown that, fortunately, matrix $\mathbf{H}^{-1}$ in general exists and is a tri-diagonal matrix. For a more detailed description the reader is referred to (Zierhofer & Schatzer, 2008).

2.1.2 Results

Sentence recognition in noise was assessed using the adaptive German Oldenburg sentence test (OLSA) in noise (Kollmeier & Wesselkamp, 1997). The test comprises 40 lists of 30 sentences each. The corpus is a closed set of nonsense sentences with a fixed name-verb-numeral-adjective-object word arrangement. As masker, unmodulated speech-shaped noise with the same spectral envelope as the long-term average spectrum of speech according to the standardized Comité Consultatif International Téléphonique et Télégraphique (CCITT) Rec. 227, was used (Fastl, 1993). The speech level was kept constant at comfortable loudness, while the level of the competing noise was varied adaptively. The speech reception threshold (SRT) was assessed at the 50% intelligibility point. The speech-noise mixture was fed into a MED-EL OPUS1 research processor via direct input (automatic gain control by-

passed). Galvanic separation between PC sound card and speech processor was provided via an audio isolation transformer. The results of two experiments are presented here.

Experiment 1 assessed the effect of varying the number of simultaneous electrodes. Variations included 2, 3, 4, half and all electrodes stimulated simultaneously with optimum subject-specific CIC parameters α, β, and γ. In this experiment, the distances between simultaneously stimulated electrodes were maximized. For example, for configuration P2 in a ten-channel setting, the electrode addresses of simultaneously stimulated electrodes are [1, 6], [2, 7], [3, 8], [4, 9], and [5, 10].

Group results for six MED-EL CI users are shown in Fig. 2. Because the data did not meet the equal-variance criterion for the parametric analysis of variance (ANOVA) test, they were analysed with a non-parametric Friedman repeated-measures (RM) ANOVA test on ranks. This test revealed a significant effect of the number of simultaneous electrodes on speech test performance ($p < 0.05$). Post-hoc multiple comparisons using Dunn's method indicated a significantly inferior test performance with the all-simultaneous setting *Pall* in comparison to CIS.
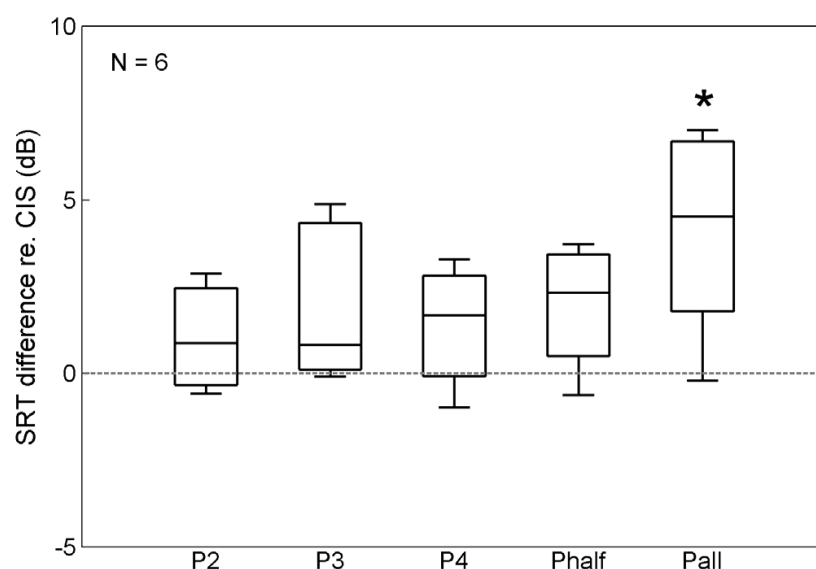

Fig. 2. SRT differences relative to CIS as a function of the number of simultaneous electrodes. Horizontal lines within the interquartile-range boxes indicate median SRT differences, whiskers encompass the ranges of SRT differences. The number of simultaneously activated electrodes was maximized. Only simultaneous stimulation of all electrodes significantly deteriorates test performance as compared to CIS (as marked by an asterisk).

Experiment 2 assessed the effect of varying the number of simultaneous channels for *adjacent* electrodes, in contrast to maximally separated electrodes as in the prior experiment. CIC parameters α, β, and γ were again individually adjusted.

Group means for five MED-EL CI users are shown in Fig. 3. A one-way RM ANOVA revealed no significant effect of electrode distances on mean SRTs ($F(2,12) = 0.016$, $p = 0.984$). Parametric test requirements for the ANOVA, i.e. normal distributions and equal variances, were verified with a Shapiro-Wilk ($p = 0.059$) and Levene test ($p = 0.963$), respectively.

However, because of the small sample size and the resulting low statistical power ($p = 0.049$) of this test, results have to be considered as preliminary.
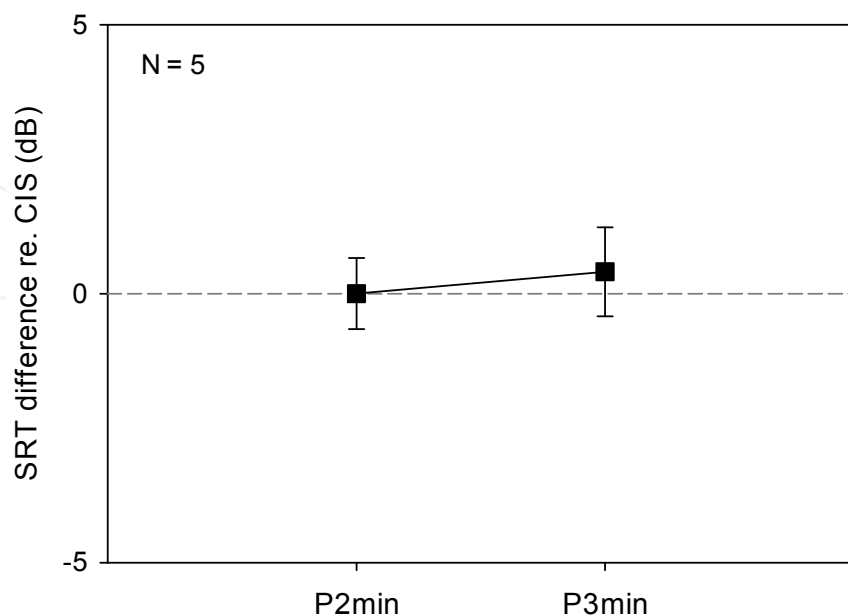

Fig. 3. SRT differences relative to CIS in a group of five subjects for two and three adjacent electrodes being stimulated simultaneously. Here, speech test performance is on a par to sequential stimulation with CIS, in contrast to the two- and three-electrodes settings P2 and P3 with maximum electrode separation in Fig. 2.

2.1.3 Discussion

Throughout all test conditions, including the CIS control condition, frame rates and phase durations were kept constant. In all configurations, the same amount of information was represented, which is why no advantage in performance of CIC as compared to CIS was expected.

Simultaneous stimulation on up to four maximally separated electrodes results in a slight, but not statistically significant, decrement in performance. When half of the electrodes are stimulated simultaneously, the decrement in performance is larger, but still not significant. The CIC configuration with all electrodes stimulated simultaneously leads to a statistically significant decrement in speech perception.

For CIC configurations with two or three neighboring electrodes the results are on a par with CIS and show no trend to a degradation of speech recognition. Thus, minimizing the distance between simultaneous electrodes is the recommended CIC configuration.

If the stimulation configuration is switched from a sequential to a simultaneous setting and the parameters are properly chosen, there should be no difference in loudness. Simultaneous stimulation *without* CIC leads to a percept that is too loud and requires an overall reduction of stimulation levels, e.g., by means of a volume control. However, this leads to an information loss in the stimulation pattern and speech perception is impaired (Zierhofer et al., 2009). On the other hand, *overcompensation* with CIC is also possible. If the spatial parameters α, β, and γ are set too high (e.g., too close to one), the perceived loudness

is too low and an overall increase of stimulation levels is necessary. Also in this case, speech perception is compromised. Thus, the subjective loudness is an indication for the quality of the choice of CIC parameters α, β, and γ.

Regarding fine structure stimulation strategies, simultaneous stimulation based on CIC represents a powerful tool. The overall number of stimulation pulses per second can be increased without reducing the phase durations. For example, in a CIC configuration with two simultaneous electrodes, the overall number of stimulation pulses per second can be doubled by filling up the temporal gaps between pairs of simultaneous pulses by additional pairs of simultaneous pulses.

2.2 The "Selected Groups" approach

2.2.1 Concept

It is well known that neurons show refractory behaviour. Immediately after an action potential has been elicited, a neuron cannot fire again during the *absolute refractory period* (typically about 0.8 ms in the neurons of the auditory system). After that, the excitation threshold decreases to the normal value within the *relative refractory period* (typically 7-10 ms).

Spatial channel interaction in combination with the refractory properties of the neurons lead to pronounced masking effects. For example, in response to a first stimulation pulse, most neurons will be activated in the immediate proximity of the stimulation electrode, and the number of activated neurons will decrease with increasing distance to the electrode. These neurons cannot be retriggered during the absolute refractory period. If a second stimulation pulse in another electrode is applied in this period, its efficacy will depend on the distance between the two electrodes. If the second pulse occurs in an adjacent electrode, it will elicit additional action potentials only when the amplitude is approximately equal to or higher than the amplitude of the first pulse. If it is lower, it will be almost entirely masked by the first pulse and largely unable to elicit additional action potentials. It is intuitively clear that the neural activation pattern will not differ substantially if, as opposed to both stimulation pulses, only the pulse with the larger amplitude is applied. However, as the distance between activated electrodes is increasing, the minimum amplitude required to elicit additional neural activity gradually decreases.

The basic idea of the SG approach is to detect and avoid pulses with high masking factors (Kals et al., 2010; Zierhofer, 2007). Neighbouring stimulation channels, i.e. channels with presumably high interaction, are arranged into groups. Within each stimulation frame, a particular number of active channels with the highest amplitudes in each group are selected for stimulation. The pulses with the smaller amplitudes are omitted. With SG, a uniformly distributed stimulation activity over all cochlear regions is ensured, and a clustering of pulses at particular electrode positions is avoided.

In a formal notation, groups of channels are represented within brackets. The index after the closing bracket denotes the number of channels within a group which are selected for stimulation. For example, $[1\ 2]_1\ [3\ 4]_1\ [5\ 6]_1\ [7\ 8]_1\ [9\ 10]_1\ [11\ 12]_1$ represents a twelve-channel configuration, where two adjacent channels each form a group, resulting in six "Selected Groups". Within each one of these six groups, the channel with the larger amplitude is dynamically selected for stimulation. In short, such a configuration is denoted as SG_1/2.

Note that CIS and NofM represent special cases of the SG concept (McKay et al., 1991; Wilson et al., 1988). E.g., a twelve-channel CIS strategy would be described by $[1]_1$ $[2]_1$ $[3]_1$ $[4]_1$ $[5]_1$ $[6]_1$ $[7]_1$ $[8]_1$ $[9]_1$ $[10]_1$ $[11]_1$ $[12]_1$, corresponding to the trivial case of SG with group size one and one active channel per group (in short, this would be denoted as SG_1/1). A 6of12 strategy is described by $[1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12]_6$, corresponding to SG with one single group comprising all twelve channels from which six are dynamically picked (in short, this is denoted as SG_6/12).

2.2.2 Results

Speech recognition with CIS and SG was assessed by measuring SRTs for German OLSA sentences in CCITT noise. Throughout all test conditions, including the CIS control condition, frame rates were kept constant. Two experiments were conducted.

Experiment 1 assessed the effect of group size variations, using the same phase durations as for CIS. Thus, according to the SG algorithm, stimulation pulses were simply discarded, and the resulting gaps were not filled with any additional pulses.

Individual and group means of the SRTs for nine tested ears (N = 9) are shown in Fig. 4. Mean SRTs for SG_1/2, SG_1/3, and SG_1/4 increased respectively by 0.3±0.5 dB, 0.6±0.9 dB, and 1.8±1.0 dB (mean±95% confidence interval) compared to the CIS control condition. A one-way RM ANOVA (Shapiro-Wilk normality test $p = 0.326$, Levene equal variance test $p = 0.758$) revealed a significant effect of the group size on mean SRTs ($p < 0.001$, $F(3,24) = 9.275$). Post-hoc Holm-Sidak multiple comparisons against CIS indicated a significant difference for SG_1/4 only ($p = 0.004$, $\alpha = 0.017$). Conditions SG_1/2 ($p = 0.163$, $\alpha = 0.050$) and SG_1/3 ($p = 0.189$, $\alpha = 0.025$) were not significantly different from CIS.
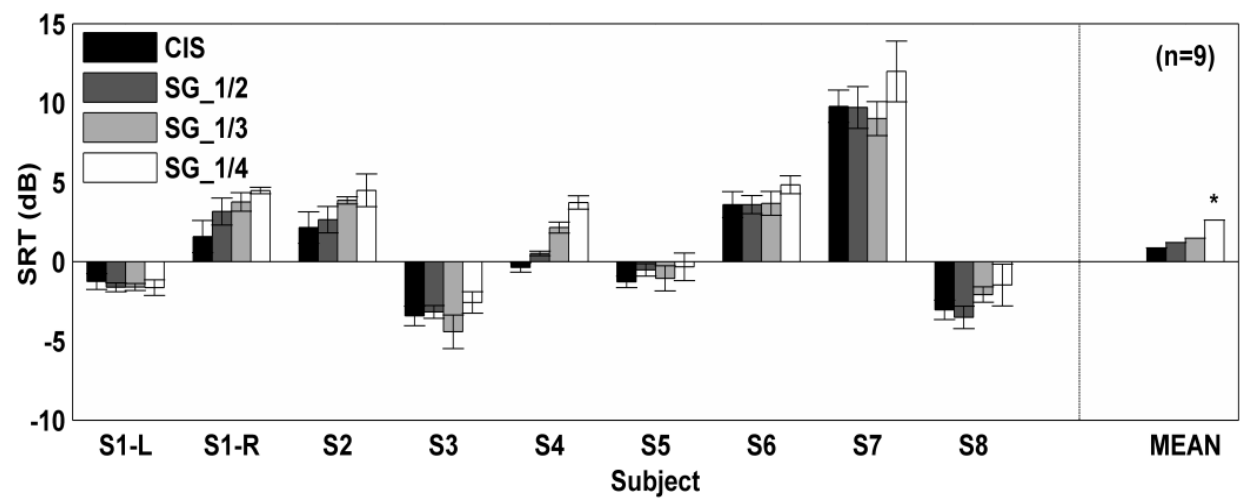

Fig. 4. SRTs for OLSA sentences in CCITT noise in eight subjects with CIS and three SG configurations including group sizes of two, three, and four. Error bars represent the standard deviations in the individual results. A significant difference among group means in comparison to CIS was found for setting SG_1/4. (From Kals et al., 2009. Used with permission of Elsevier B.V.)

Experiment 2 assessed the effect of doubling pulse phase durations with SG. Only one SG condition with a group size of two was used, and the phase durations of the stimulation

pulses were doubled compared to CIS. Configurations with doubled pulse phase duration are denoted with the suffix "2DUR". For double phase duration, new MCL and THR levels were determined.

Results are summarized in Fig. 5, showing individual and group means of SRTs for six tested ears (N = 6). The mean SRT for SG_1/2-2DUR was 0.5±0.9 dB lower than the mean SRT for CIS. However, this difference was not significant (p = 0.184, t(5) = 1.540; Shapiro-Wilk normality test p = 0.799). Due to the small sample size and low statistical power of the test (p = 0.155), however, results need to be regarded as preliminary. The doubling of phase durations resulted in a significant reduction of both MCL and THR currents. MCL and THR levels were lower by 5.0±2.2 dB (p = 0.002) and 7.1±4.3 dB (p = 0.008), respectively.
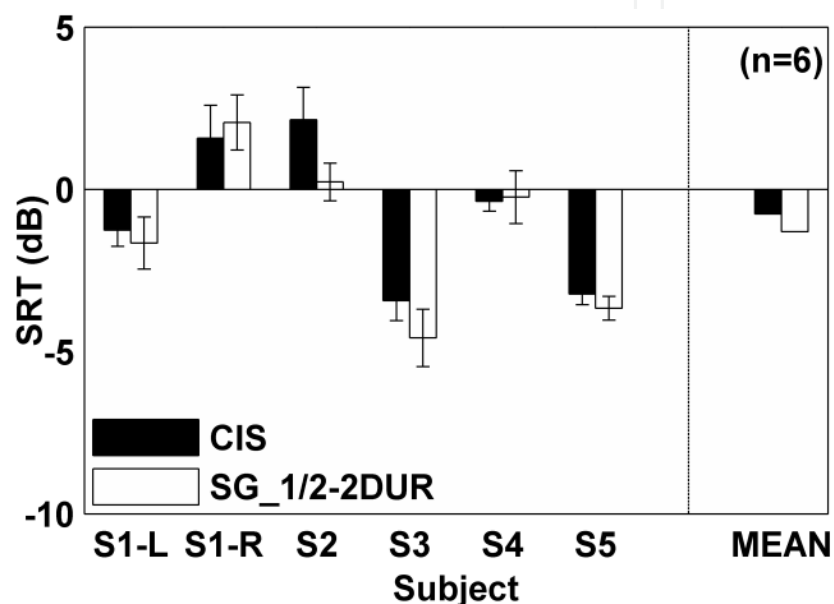

Fig. 5. SRTs for OLSA sentences in CCITT noise in five subjects with CIS and SG with doubled phase duration and a group size of two. Error bars represent the standard deviations in the individual results. No statistically significant differences between group means (α = 0.05) were found. (From Kals et al., 2009. Used with permission of Elsevier B.V.)

2.2.3 Discussion

For group sizes two and three, SG supports the same sentence recognition in noise as CIS. This suggests that a dynamic selection of up to one third of the stimulation pulses with SG still includes all relevant pulses, at least for this particular type of test materials. SG provides a simple and robust selection algorithm for increasing the efficacy of the applied stimulation pulses and does not require any subject-specific parameters. Hence, the application in current implant designs is straightforward. With a further pulse reduction (group sizes larger than three), SG starts to suppress relevant pulses, resulting in a significant decrement in performance.

The dynamic channel-picking with SG may be utilized to increase pulse phase durations and consequently reduce pulse amplitudes, yielding to less stringent implant supply voltage requirements. In our test setup, pulse amplitude reductions of about 40% with doubled phase durations and a group size of two supported sentence recognition scores which were

on a par with CIS. Such a reduction is likely to be beneficial for future designs of low-power or totally implantable cochlear implants. Additionally, longer phase durations resulted in larger electrode dynamic ranges, which may support better speech reception in CI users.

During testing, subjective sound quality was informally assessed with settings including both standard and longer phase durations. All test settings were presented multiple times in random order to the subjects, with live speech or speech test materials as input signals. Subjects did not know which setting was presented at any time. Most of them indicated SG_1/2-2DUR as their preferred setting.

3. Electric-acoustic pitch comparisons in unilateral CI subjects

In the peripheral auditory system, frequency information is thought to be encoded by a combination of two mechanisms. The first mechanism, or place principle, is based on the notion that the place of maximum excitation along the organ of Corti changes systematically as function of stimulus frequency, as previously mentioned (von Békésy, 1960; von Helmholtz, 1863). Low frequencies are "mapped" at the cochlear apex and high frequencies at the cochlear base. For humans and other mammals, cochlear frequency-place maps are well known (Greenwood, 1990, 1961). The second mechanism, or time principle, hypothesizes that stimulus frequency is encoded in the temporal structure of neural discharge patterns. Mediated by the phase locking properties of auditory nerve fibers (Johnson, 1980), neural discharges occur in synchrony to the individual harmonics of subharmonics of a stimulus waveform.

Cochlear implants implicitly assume that both place and time principles equally apply for electrical stimulation, as they do for normal acoustic stimulation. Frequency bands in implant speech processors are assigned to electrodes in tonotopic order. With CIS and related coding strategies, channel envelope modulations sampled by interleaved constant-rate pulse carriers represent a temporal code for periodic harmonic stimuli. However, due to the vastly different modality of neural excitation with electric and acoustic stimulation, the frequency-place map for electrical stimulation might differ from Greenwood's map for humans, and temporal cues to pitch mediated by phase locking may be encoded more robustly using other stimulus patterns than the CIS channel envelopes sampled at constant pulse rates. Although studies show that learning can compensate to a certain extent for a frequency-place mismatch (Svirsky et al., 2004), an optimal assignment of frequency bands based on a measured electrical frequency-place map may be beneficial in terms of CI performance and how fast asymptotic levels of performance are reached in newly implanted CI users.

In recent years, cochlear implants have become an effective treatment option for patients with single-sided deafness and intractable ipsilateral tinnitus (Van de Heyning et al., 2008). With near-normal hearing in one and a CI in the contralateral ear, direct comparisons of normal and electric hearing are possible in this unique subject population. At our laboratory, we were in the fortunate situation to study such a group of unilateral CI subjects implanted by Prof. Van de Heyning at the University Hospital Antwerp, Belgium. One aim of the study was to assess the frequency-place map for high-rate electrical stimulation, using a different methodology than similar studies in CI users with contralateral hearing (Boëx et al., 2006; Carlyon et al., 2010; Dorman et al., 2007; Vermeire et al., 2008). The second aim was to investigate the relative contributions of rate and place to pitch, in particular for low

frequencies, i. e. in the apical region. Preliminary results from this study have been presented by the authors and colleagues (Schatzer et al., 2009a; Vermeire et al., 2009).

In our study group, the frequency-place maps were between Greenwood and one octave below at basal and medial electrode positions, levelling off in the apical region to on average correspond to Greenwood. This levelling-off is likely to be a consequence of the absence of any temporal code in the constant-amplitude 1.5 kpulses/sec pulse trains that could be perceived by CI users (with electrical stimulation, there seems to be a canonical pitch saturation limit of 300-500 Hz, where a further increase of pulse rate does not result in a further increase of perceived pitch, as for lower rates), but may also be due to the cochlear anatomy in the apex. Whereas in the basal cochlear turn auditory nerve fibers innervating the organ of Corti take a radial course, in the apical turn they gradually take on a more tangential course. Additionally, the spiral ganglion holding the cell bodies of the bipolar auditory neurons does not reach all the way to the apical turn, forming a cluster of cell bodies at its apical end. As electrodes sit in close proximity to the organ of Corti, current spread on apical electrodes may target more innervating fibers or spiral ganglion cells as on basal electrodes, producing a broader and less distinct pitch percept.

The most interesting finding emerging from the second experiment was that for electrical stimulation place and rate, or in other words spatial and temporal cues, may have to match to produce a robust pitch percept. For instance, most subjects could only match electrical stimuli reliably to low-frequency pure tones ranging from 100 to 200 Hz, if those electrical stimuli had a correspondingly low pulse rate and were applied on electrodes inserted more than 360 degrees from the round window into the cochlea. Thus, in order to produce low pitch percepts in the F0 range, electrodes need to be placed beyond the first cochlear turn. This finding is of particular relevance to the design of coding strategies that represent within-channel temporal fine structure information as channel-specific rate codes on apical electrodes, as discussed in in the following section.

This observation is consistent with findings in normal-hearing listeners. Oxenham and colleagues (Oxenham et al., 2004) demonstrated that pitch discrimination thresholds for "transposed" acoustic stimuli, where the temporal information of low-frequency sinusoids is presented to locations in the cochlea tuned to high frequencies, are significantly deteriorated compared to discrimination thresholds for pure tones. Also, using harmonic transposed stimuli, they found that subjects were largely unable to extract the fundamental frequency from multiple low-frequency harmonics presented to high-frequency cochlear regions. These results also indicate the importance of a match between temporal and tonotopic cues for a robust pitch perception.

4. Temporal fine structure with Channel Specific Sampling Sequences

4.1 Concept

The basic idea of a stimulation strategy based on "Channel Specific Sampling Sequences" (CSSS) is to add fine time information to the envelope information already present in a speech processor, at least in the low-frequency filter bands (Zierhofer, 2003). For this, a particular stimulation pulse sequence composed of high-rate biphasic pulses (typically 5 kpulses/sec per channel) is associated to the filter channel. Such a sequence is designated

as "Channel Specific Sampling Sequence" and has a programmable length (number of pulses) and a programmable amplitude distribution. The length of a CSSS sequence can be anywhere from a single-pulse to a sequence of 16 pulses.

For stimulation, the sequences are triggered by the zero-crossings (e.g., negative-to-positive) of the band pass filter output signal. The individual sequences are weighted by the instantaneous envelopes of the output signal. Note that such a stimulation sequence contains both envelope- and temporal fine structure information. An example is shown in Fig. 6, where the sampling sequence is a pair of pulses. The sampling sequences are started at every other zero crossing of the band pass filter output (upper panel). The weights of the sequences are derived from the envelope of the band pass output signal.
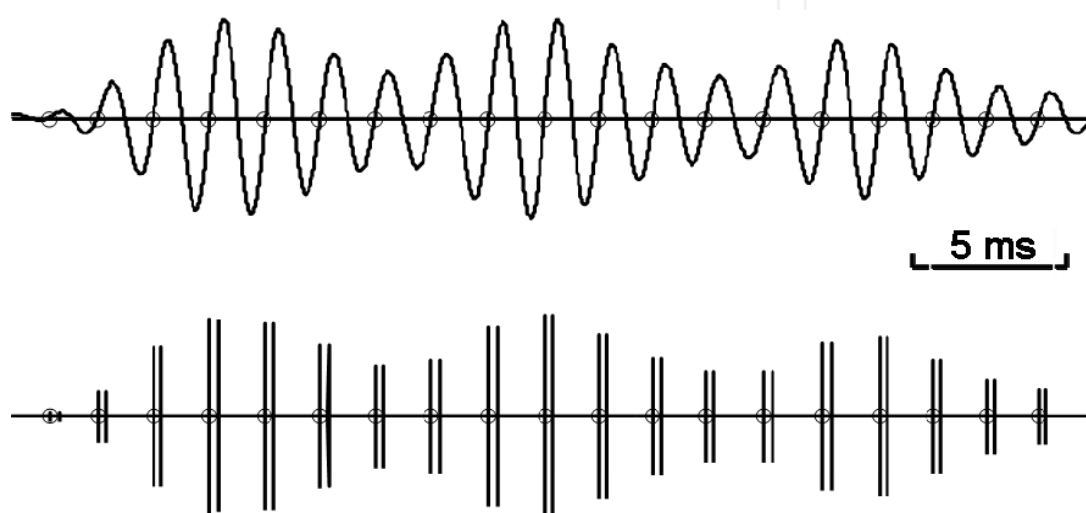
Fig. 6. Example for stimulation according to the CSSS concept. Upper panel: Output of a band pass filter from 450 to 603 Hz. Lower panel: CSSS stimulation sequence. Each vertical line represents a biphasic stimulation pulse. Here, the sampling sequences are double pulses, separated by 0.25 ms. The sequences are started at every other zero crossing of the band pass output.

Given the phase-locking limit of primary auditory neurons (Johnson, 1980) and the finding that in CI subjects the so-called "rate pitch" shows an upper boundary at about 300-500 pulses/sec in most subjects (Wilson et al., 1997; Zeng, 2002), but can also extend up to about 1 kpulses/sec (Hochmair-Desoyer et al., 1983; Kong & Carlyon, 2010), a stimulation based on CSSS seems appropriate only for stimulation channels below about 1 kHz. For higher-frequency channels, CIS stimulation based on envelope information only is applied. Thus, in practical implementations, a mixture between low-frequency CSSS channels and high-frequency CIS channels will be reasonable. For convenience, such a mixture between CSSS and CIS channels is designated as CSSS concept in the following.

4.2 Results

Effects of CSSS fine structure stimulation on speech recognition in competing-voice backgrounds and for the recognition of tonal languages have been investigated in recent studies, conducted at our laboratories and at Queen Mary Hospital in Hong Kong, China (Schatzer et al., 2010).

4.2.1 CSSS study in our laboratories

Subjects and strategy settings

In eight MED-EL implant subjects, SRTs were measured for OLSA sentences using a female-talker sentence as masker. Strategies were fitted and tested acutely with no listening experience for the subjects.

The overall frequency range for both CIS and CSSS settings was 80 to 8500 Hz. For CSSS, 4 CSSS channels from 80 to 800 Hz were arranged in two selected groups $[1\ 2]_1$ and $[3\ 4]_1$. Maximum pulse rates were 4545 and 1515 pulses/sec on CSSS and CIS channels, respectively. Sequences were double pulses on CSSS channels 1 and 2 and single pulses on CSSS channels 3 and 4.

Results and discussion

Individual and group results for the eight subjects are shown in Fig. 7. A paired *t*-test revealed a statistically significant difference among group means of 1.0 dB ($p = 0.008$, $t(12) = 3.192$, Shapiro-Wilk normality test $p = 0.675$), with better performance for the CSSS condition.
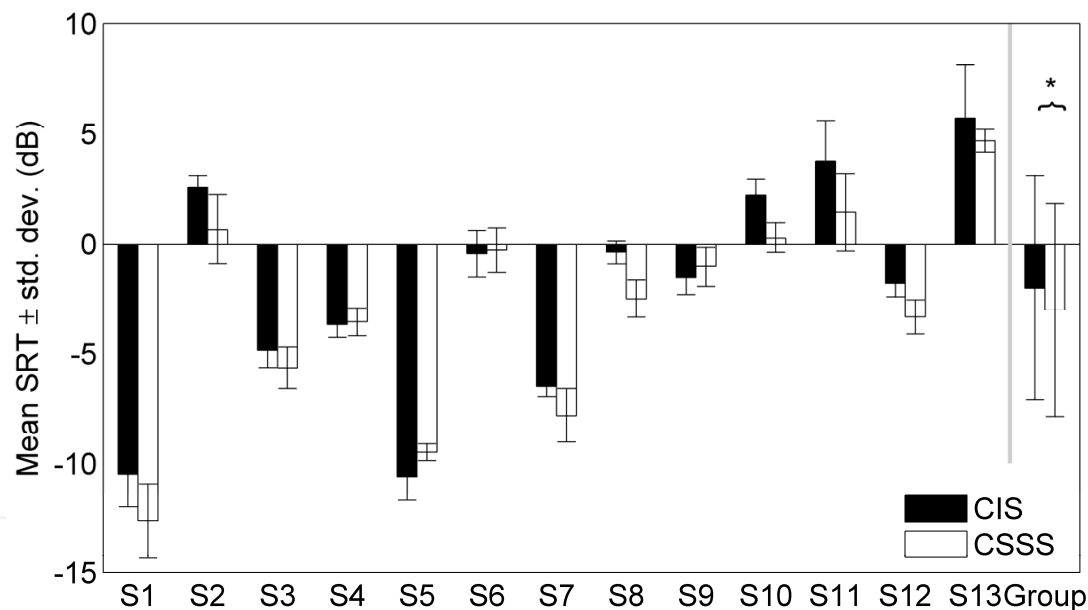
Fig. 7. Individual and group results for OLSA sentences presented with a competing female-talker masking sentence.

In this study a significant improvement in speech perception was found in acute tests. Results reported in literature often do not find an initial benefit for CSSS (Schatzer et al., 2009b), but an improvement over time (Lorens et al., 2010; Riss et al., 2011).

4.2.2 CSSS study in Hong Kong

Subjects and test materials

Twelve adult and experienced implant users participated in this study. All subjects are native speakers of Cantonese and implanted with MED-EL C40+ or PULSAR$_{CI}^{100}$ devices

with full electrode insertions. At the time of testing, all participants were users of the MED-EL TEMPO+ speech processor and had no prior exposure to temporal fine structure stimulation.

Test materials included the Cantonese lexical tone identification test and the Cantonese Hearing in Noise (CHINT) sentence test. The tone identification test was presented at a subject-specific constant SNR, i.e. CCITT noise was added on a per-subject basis to avoid ceiling effects, if necessary. Acutely compared test settings comprised CIS and CSSS.

The overall frequency range for both CIS and CSSS settings was 100 to 8500 Hz. For CSSS, 4 CSSS channels from 100 to 800 Hz were arranged in two selected groups $[1\ 2]_1$ and $[3\ 4]_1$. Maximum pulse rates were 4545 and 1515 pulses/sec on CSSS and CIS channels, respectively. Sequences on CSSS channels were quarter-sine shaped. For further details, the reader is referred to (Schatzer et al., 2010).

Results and discussion

Group results of CIS and CSSS performance for Cantonese tones and sentences are illustrated in Fig. 8. Mean identification scores of Cantonese lexical tones in 12 subjects were 59.2±15.2 % (mean±standard deviation) with CIS and 59.2±15.3 % with CSSS, indicating identical intelligibility with CIS and fine structure stimulation (paired t-test on percent-correct scores, $p = 0.972$, $t(11) = 0.036$; Shapiro-Wilk normality test $p = 0.873$). Mean identification scores for CHINT sentences were 54.2±27.7 % with CIS and 55.9±22.8 % with CSSS, indicating no significant difference (paired t-test, $p = 0.676$, $t(7) = 0.436$; Shapiro-Wilk normality test $p = 0.539$).
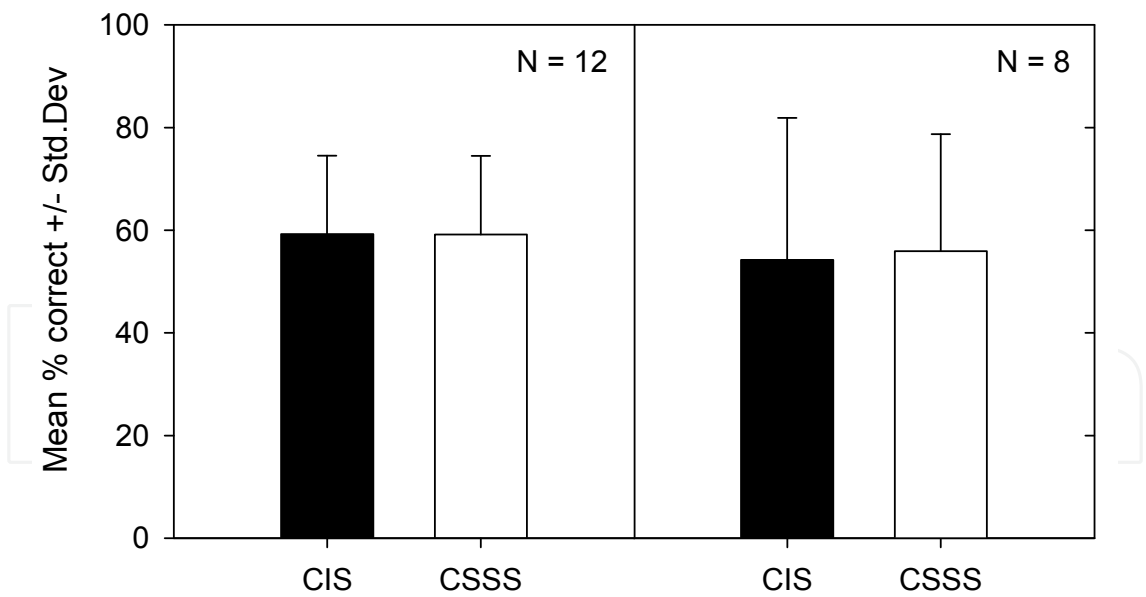

Fig. 8. Group results for Cantonese tone identification (left panel, 12 subjects) and CHINT sentences (right panel, 8 subjects) with CIS and CSSS, respectively.

These findings are consistent with results from studies in which experienced CI users were switched from CIS to fine structure stimulation, showing equal performance for the two coding strategies at baseline, and a statistically significant benefit with fine structure stimulation only after weeks or even months of fine structure use (Arnoldner et al., 2007).

4.2.3 Expanded pitch range with CSSS

When being switched from CIS to fine structure stimulation with CSSS, many CI recipients report an overall pitch shift towards lower frequencies. This effect was explored and quantified.

Subjects and methods

Seven experienced cochlear implant subjects participated in the study. The stimuli were 500 ms harmonic tone complexes with a spectral roll-off of 9 dB per octave, presented via the direct input of an OPUS1 research speech processor. All stimuli were carefully balanced for equal loudness before running the actual experiment. Additionally, in order to prevent the participants from basing their decisions on residual loudness differences, amplitudes were roved by 2 dB during the main experiment.

For pitch comparisons, a method of constant stimuli was utilized. Subjects were asked to compare the pitch of the harmonic tone complexes presented in two 500-ms intervals separated by 300 ms. In the first interval a harmonic tone was presented with the first strategy (either CSSS or CIS), in the second interval another harmonic tone was presented with the second strategy. Based on psychometric functions fitted to the data, points of subjective equivalence were derived. These indicated which pair of F0s elicited the same pitch when one of them was processed with the CSSS strategy and the other one was processed with the CIS strategy. For a more detailed description of the study the reader is referred to (Krenmayr et al., 2011).

Results and discussion

A summary of the results is illustrated in Fig. 9. For each one of the four reference frequencies, the frequency difference between CSSS and CIS required for identical pitch is plotted.
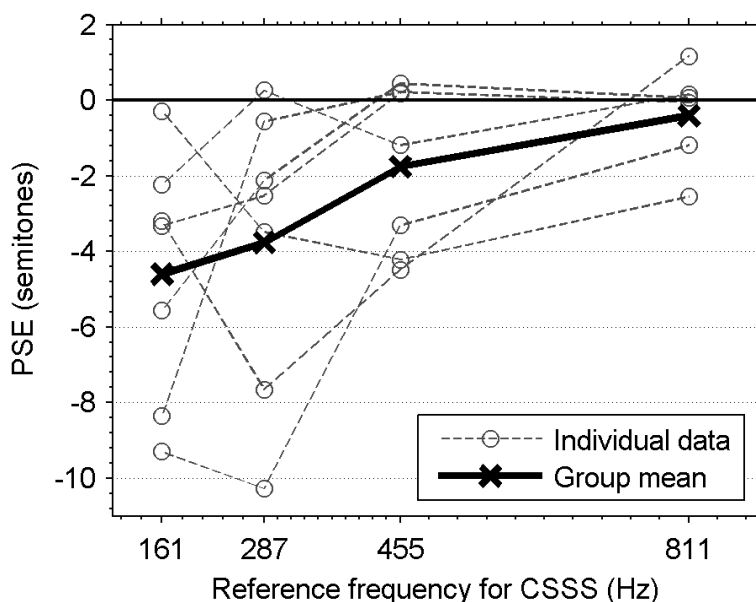

Fig. 9. Pitch differences between CSSS and CIS, expressed in semitones, as a function of F0. Dotted lines represent the individual data, the solid black line indicates the group mean. The pitch difference is most pronounced at low frequencies and vanishes towards higher frequencies, where the stimulation patterns of CSSS and CIS become more and more alike. (From Krenmayr et al., 2011. Used with permission of Maney Publishing.)

The results indicate that the pitch of a given harmonic tone complex decreases when presenting it with CSSS compared to when presenting it with CIS. On average, this pitch shift was 5 semitones or one perfect fourth at the lowest tested F0 (161 Hz). As expected, this pitch shift decreased at higher F0s. At 287 Hz it was 4 semitones or one major third, while at 455 Hz it was already reduced to 2 semitones or one major second. At the "blank trial" frequency of 811 Hz, the group mean was 0.5 semitones. Since these data were collected with one strategy as reference stimulus only, they can still contain possible subjective response bias. However, this response bias is expected to average out when calculating the group mean. Therefore the blank trial gives an estimate for the accuracy of the method.

The study quantitatively shows that explicitly stimulating the fine structure component of the lowest partials decreases the pitch of harmonic tone complexes with fundamental frequencies below 811 Hz. Since the pitches produced by both coding strategies converge at 811 Hz, it can be concluded that fine structure stimulation expands the range of perceivable pitches as compared to CIS.

As the present results are only based on a small number of subjects, further studies are needed to support the findings. However, the robustness of F0 representation with CSSS has also been demonstrated in another, more recent study (Bader et al., 2011). There, two pulse trains of *equal rates* have been applied in two neighbouring apical electrodes. In a ranking experiment, pitch percepts have been compared for dual-electrode pulse trains which were almost in phase (they were not completely in phase because of sequential stimulation) and dual-electrode pulse trains which were out of phase (phase shift π). Remarkably, the pitch percept in the latter case did not correspond to that of a stimulus with twice the rate of the individual pulse trains as could be expected in the case of spatial channel interaction, but just shifted slightly towards higher frequencies. Interestingly, Macherey and Carlyon found similar results of very low temporal channel interaction at low rates for phase-shifted pulse trains on electrode pairs which are not positioned in the apical, but the middle range of the cochlea (Macherey & Carlyon, 2010).

5. Conclusion

The introduction of the CIS strategy certainly represented a breakthrough in the field of cochlear implants. In its standard version, CIS provides an optimized level of speech perception for Western languages. CIS is first of all based on the tonotopic principle of hearing. The temporal information is limited to envelope signals.

Fine structure stimulation aims at representing temporal cues beyond envelopes and thus is based on the periodicity principle of hearing. Experiences with early cochlear implants (i.e., with single channel devices) and also more recent studies show that such temporal cues indeed can be used in a frequency range up to about 1 kHz. This essentially covers the ranges of the fundamental frequency F0 and the first formant frequency F1.

The need for a precise representation of both envelope and fine structure information lead to the development of two supporting concepts, i.e., simultaneous stimulation based on CIC and SG. Both concepts have been evaluated with envelope-based stimulation strategies

without temporal fine structure. Among CIC settings, configurations with neighbouring simultaneous channels resulted in speech test performance on a par with CIS. So far, only two and three neighbouring channels have been investigated, and studies with even more simultaneous channels are currently conducted. SG settings yielded no statistically significant difference to CIS for group sizes of two and three. However, also for these group sizes, a tendency to a reduction of speech perception scores was observed.

Implant subjects with unilateral deafness and contralateral normal hearing provide the possibility to directly compare electrical and acoustic hearing. An important finding from pitch matching studies in unilaterally deaf subjects was that the perception of low frequencies in the F0 range requires electrode insertion depths greater than 360 degrees, in combination with adequate low-frequency temporal cues. Thus, a sufficient insertion depth of the electrodes represents a *necessary condition*. If electrodes are not inserted deeply enough, low rate temporal cues alone do not produce low pitch percepts.

A strategy based on CSSS combines fine structure and envelope information. High-rate pulse packages are triggered at every other zero crossing of the band pass output signals, and the weights of the sequences are derived from the band pass envelopes. Practical configurations are mixtures of CSSS channels in the low frequency region (e.g., 100-800 Hz covering the F0 and F1 regions) and CIS channels for the higher frequencies (e.g., 800-8500 Hz). When being switched from a CIS to a CSSS stimulation setting, most patients perceive an immediate downward pitch shift. This effect has been quantified. On average, pitch shifts of five semitones at the low frequency end were measured. This is a clear indication that the additional temporal information represented in the low frequency CSSS channels can indeed be perceived despite the presence of spatial channel interaction. Besides, many patients are reporting an improvement in sound quality with fine structure stimulation. Typically, they report the sound being fuller and more natural as compared to their CIS settings. In acute tests, speech perception with CSSS is often on a par with CIS. However, there is increasing evidence that with fine structure stimulation, speech performance does improve over time, even in CI users with long-term implant experience.

6. Acknowledgements

7. References

Arnoldner, C.; Riss, D.; Brunner, M.; Durisin, M.; Baumgartner, W. D. & Hamzavi, J. S. (2007). Speech and music perception with the new fine structure speech coding strategy: preliminary results. *Acta Otolaryngol,* Vol. 127, No. 12, (Dec), pp. 1298-303, ISSN 0001-6489 (Print)

Bader, P.; Schatzer, R.; Vermeire, K.; Van de Heyning, P.; Visser, D.; Krenmayr, A.; Kals, M. & Zierhofer, C. (2011). Pitch of dual-electrode stimuli as a function of rate and electrode separation, *Conference on Implantable Auditory Prostheses*, Pacific Grove, CA, July 24-29, 2011

Boëx, C.; Baud, L.; Cosendai, G.; Sigrist, A. & Kós, M.-I. (2006). Acoustic to electric pitch comparisons in cochlear implant subjects with residual hearing. *J Assoc Res Otolaryngol,* Vol. 7, No. 2, (Jun), pp. 110-124, ISSN 1525-3961 (Print)

Carlyon, R. P.; Macherey, O.; Frijns, J. H.; Axon, P. R.; Kalkman, R. K.; Boyle, P.; Baguley, D. M.; Briggs, J.; Deeks, J. M.; Briaire, J. J.; Barreau, X. & Dauman, R. (2010). Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear. *J Assoc Res Otolaryngol,* Vol. 11, No. 4, (Dec), pp. 625-640, ISSN 1525-3961 (Print)

Dorman, M. F.; Spahr, T.; Gifford, R.; Loiselle, L.; McKarns, S.; Holden, T.; Skinner, M. & Finley, C. (2007). An electric frequency-to-place map for a cochlear implant patient with hearing in the nonimplanted ear. *J Assoc Res Otolaryngol,* Vol. 8, No. 2, (Jun), pp. 234-240, ISSN 1525-3961 (Print)

Fastl, H. (1993). A masking noise for speech intelligibility tests, *Proceedings of TC Hearing,* H-93–70, Acoust. Society of Japan

Greenwood, D. D. (1990). A cochlear frequency-position function for several species--29 years later. *J Acoust Soc Am,* Vol. 87, No. 6, (Jun), pp. 2592-2605, ISSN 0001-4966 (Print)

Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *J Acoust Soc Am,* Vol. 33, No. 10, pp. 1344-1356, ISSN 0001-4966 (Print)

Gstoettner, W. K.; Adunka, O.; Franz, P.; Hamzavi, J., Jr.; Plenk, H., Jr.; Susani, M.; Baumgartner, W. & Kiefer, J. (2001). Perimodiolar electrodes in cochlear implant surgery. *Acta Otolaryngol,* Vol. 121, No. 2, (Jan), pp. 216-219

Hilbert, D. (1912). *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, B. G. Teubner, Leipzig, Berlin

Hochmair-Desoyer, I. J.; Hochmair, E. S.; Burian, K. & Stiglbrunner, H. K. (1983). Percepts from the Vienna cochlear prosthesis. *Ann N Y Acad Sci,* Vol. 405, pp. 295-306, ISSN 0077-8923 (Print)

Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am,* Vol. 68, No. 4, pp. 1115-1122, ISSN 0001-4966 (Print)

Kals, M.; Schatzer, R.; Krenmayr, A.; Vermeire, K.; Visser, D.; Bader, P.; Neustetter, C.; Zangerl, M. & Zierhofer, C. (2010). Results with a cochlear implant channel-picking strategy based on "Selected Groups". *Hear Res,* Vol. 260, No. 1-2, (Feb), pp. 63-69, ISSN 1878-5891 (Electronic)

Kollmeier, B. & Wesselkamp, M. (1997). Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *J Acoust Soc Am,* Vol. 102, No. 4, pp. 2412-2421

Kong, Y. Y. & Carlyon, R. P. (2010). Temporal pitch perception at high rates in cochlear implants. *J Acoust Soc Am,* Vol. 127, No. 5, (May), pp. 3114-23, ISSN 1520-8524 (Electronic)

Krenmayr, A.; Visser, D.; Schatzer, R. & Zierhofer, C. (2011). The effects of fine structure stimulation on pitch perception with cochlear implants. *Cochlear Implants International,* Vol. 12, No. Suppl. 1, pp. S70-72, ISSN 1467-0100 (Print). Available online at www.maney.co.uk/journals/cim and www.ingentaconnect.com/content/maney/cii

Lenarz, T.; Battmer, R. D.; Goldring, J. E.; Neuburger, J.; Kuzma, J. & Reuter, G. (2000). New electrode concepts (modiolus-hugging electrodes), in *Updates in Cochlear Implantation.* vol. 57, C. S. Kim & S. O. Chang, Eds., ed Basel: Karger, 2000, pp. 347-353

Loizou, P. C.; Poroy, O. & Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *J Acoust Soc Am,* Vol. 108, No. 2, pp. 790-802

Loizou, P. C.; Stickney, G.; Mishra, L. & Assmann, P. (2003). Comparison of speech processing strategies used in the Clarion implant processor. *Ear Hear,* Vol. 24, No. 1, (Feb), pp. 12-19

Lorens, A.; Zgoda, M.; Obrycka, A. & Skarzynski, H. (2010). Fine Structure Processing improves speech perception as well as objective and subjective benefits in pediatric MED-EL COMBI 40+ users. *Int J Pediatr Otorhinolaryngol,* Vol. 74, No. 12, (Dec), pp. 1372-8, ISSN 1872-8464 (Electronic)

Macherey, O. & Carlyon, R. P. (2010). Temporal pitch percepts elicited by dual-channel stimulation of a cochlear implant. *J Acoust Soc Am,* Vol. 127, No. 1, (Jan), pp. 339-49, ISSN 1520-8524 (Electronic)

McKay, C.; McDermott, H.; Vandali, A. & Clark, G. (1991). Preliminary Results with a Six Spectral Maxima Sound Processor for the University of Melbourne/Nucleus Multiple-Electrode Cochlear Implant. *J Otolaryng Soc Austral,* Vol. 6, No. 5, pp. 254-359, ISSN 0030-6614

Miyoshi, S.; Sakajiri, M.; Ifukube, T. & Matsushima, J. (1997). Evaluation of the tripolar electrode stimulation method by numerical analysis and animal experiments for cochlear implants. *Acta Otolaryngol Suppl,* Vol. 532, pp. 123-125, ISSN 0365-5237 (Print)

Morse, R. P. & Evans, E. F. (1999). Additive noise can enhance temporal coding in a computational model of analogue cochlear implant stimulation. *Hear Res,* Vol. 133, No. 1-2, (Jul), pp. 107-119, ISSN 0378-5955 (Print)

Oxenham, A. J.; Bernstein, J. G. & Penagos, H. (2004). Correct tonotopic representation is necessary for complex pitch perception. *Proc Natl Acad Sci U S A,* Vol. 101, No. 5, (Feb 3), pp. 1421-1425, ISSN 0027-8424 (Print)

Richter, B.; Aschendorff, A.; Lohnstein, P.; Husstedt, H.; Nagursky, H. & Laszig, R. (2002). Clarion 1.2 standard electrode array with partial space-filling positioner: radiological and histological evaluation in human temporal bones. *J Laryngol Otol,* Vol. 116, No. 7, (Jul), pp. 507-513, ISSN 0022-2151 (Print)

Riss, D.; Hamzavi, J. S.; Katzinger, M.; Baumgartner, W. D.; Kaider, A.; Gstoettner, W. & Arnoldner, C. (2011). Effects of fine structure and extended low frequencies in pediatric cochlear implant recipients. *Int J Pediatr Otorhinolaryngol*, Vol. 75, No. 4, (Apr), pp. 573-8, ISSN 1872-8464 (Electronic)

Roland, J. T., Jr.; Fishman, A. J.; Alexiades, G. & Cohen, N. L. (2000). Electrode to modiolus proximity: a fluoroscopic and histologic analysis. *Am J Otol*, Vol. 21, No. 2, (Mar), pp. 218-225, ISSN 0192-9763 (Print)

Rubinstein, J. T.; Wilson, B. S.; Finley, C. C. & Abbas, P. J. (1999). Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res*, Vol. 127, No. 1-2, pp. 108-118

Schatzer, R.; Krenmayr, A.; Au, D. K.; Kals, M. & Zierhofer, C. (2010). Temporal fine structure in cochlear implants: preliminary speech perception results in Cantonese-speaking implant users. *Acta Otolaryngol*, Vol. 130, No. 9, (Feb 9), pp. 1031-9, ISSN 0001-6489 (Print)

Schatzer, R.; Vermeire, K.; Van de Heyning, P.; Voormolen, M.; Visser, D.; Krenmayr, A.; Kals, M. & Zierhofer, C. (2009a). A tonotopic map of the electrically stimulated cochlea from CI users with contralateral normal hearing, *Conference on Implantable Auditory Prostheses*, Lake Tahoe, CA, July 12-17, 2009

Schatzer, R.; Visser, D.; Krenmayr, A.; Kals, M.; Vermeire, K.; Neustetter, C.; Zangerl, M.; Bader, P. & Zierhofer, C. (2009b). Recognition of speech with a competing talker using fine structure in cochlear implants, *Conference on Implantable Auditory Prostheses*, Lake Tahoe, CA, July 12-17, 2009

Smith, Z. M.; Delgutte, B. & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, Vol. 416, No. 6876, (Mar 7), pp. 87-90, ISSN 0028-0836 (Print)

Stickney, G. S.; Loizou, P. C.; Mishra, L. N.; Assmann, P. F.; Shannon, R. V. & Opie, J. M. (2006). Effects of electrode design and configuration on channel interactions. *Hear Res*, Vol. 211, No. 1-2, (Jan), pp. 33-45, ISSN 0378-5955 (Print)

Svirsky, M. A.; Silveira, A.; Neuburger, H.; Teoh, S. W. & Suarez, H. (2004). Long-term auditory adaptation to a modified peripheral frequency map. *Acta Otolaryngol*, Vol. 124, No. 4, (May), pp. 381-386, ISSN 0001-6489 (Print)

Van de Heyning, P.; Vermeire, K.; Diebl, M.; Nopp, P.; Anderson, I. & De Ridder, D. (2008). Incapacitating unilateral tinnitus in single-sided deafness treated by cochlear implantation. *Ann Otol Rhinol Laryngol*, Vol. 117, No. 9, (Sep), pp. 645-652, ISSN 0003-4894 (Print)

van den Honert, C. & Kelsall, D. C. (2007). Focused intracochlear electric stimulation with phased array channels. *J Acoust Soc Am*, Vol. 121, No. 6, (Jun), pp. 3703-3716, ISSN 1520-8524 (Electronic)

van den Honert, C. & Stypulkowski, P. H. (1987). Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve. *Hear Res*, Vol. 29, pp. 195-206, ISSN 0378-5955 (Print)

Vermeire, K.; Nobbe, A.; Schleich, P.; Nopp, P.; Voormolen, M. H. & Van de Heyning, P. H. (2008). Neural tonotopy in cochlear implants: an evaluation in unilateral cochlear

implant patients with unilateral deafness and tinnitus. *Hear Res,* Vol. 245, No. 1-2, (Nov), pp. 98-106, ISSN 0378-5955 (Print)

Vermeire, K.; Schatzer, R.; Visser, D.; Krenmayr, A.; Kals, M.; Neustetter, C.; Bader, P.; Zangerl, M.; Van de Heyning, P. & Zierhofer, C. (2009). Contributions of temporal and place cues to pitch in the apical region, *Conference on Implantable Auditory Prostheses*, Lake Tahoe, CA, July 12-17, 2009

von Békésy, G. (1960). *Experiments in Hearing*, McGraw-Hill Book Company, Inc., ISBN 978-0070043244, New York

von Helmholtz, H. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Friedrich Vieweg und Sohn, Braunschweig, Germany,

Wackym, P. A.; Firszt, J. B.; Gaggl, W.; Runge-Samuelson, C. L.; Reeder, R. M. & Raulie, J. C. (2004). Electrophysiologic effects of placing cochlear implant electrodes in a perimodiolar position in young children. *Laryngoscope,* Vol. 114, No. 1, (Jan), pp. 71-76, ISSN 0023-852X (Print)

Wilson, B. S.; Finley, C. C.; Farmer, J. C., Jr.; Lawson, D. T.; Weber, B. A.; Wolford, R. D.; Kenan, P. D.; White, M. W.; Merzenich, M. M. & Schindler, R. A. (1988). Comparative studies of speech processing strategies for cochlear implants. *Laryngoscope,* Vol. 98, No. 10, (Oct), pp. 1069-1077, ISSN 0023-852X (Print)

Wilson, B. S.; Finley, C. C.; Lawson, D. T.; Wolford, R. D.; Eddington, D. K. & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature,* Vol. 352, No. 6332, (Jul 18), pp. 236-238, ISSN 0028-0836 (Print)

Wilson, B. S.; Zerbi, M.; Finley, C. C.; Lawson, D. T. & van den Honert, C. (1997). *Speech processors for auditory prostheses: Relationships between temporal patterns of nerve activity and pitch judgments for cochlear implant patients*, Eighth Quarterly Progress Report, NIH project N01-DC-5-2103, Neural Prosthesis Program, National Institutes of Health, Bethesda, MD

Xu, L.; Zwolan, T. A.; Thompson, C. S. & Pfingst, B. E. (2005). Efficacy of a cochlear implant simultaneous analog stimulation strategy coupled with a monopolar electrode configuration. *Ann Otol Rhinol Laryngol,* Vol. 114, No. 11, (Nov), pp. 886-893, ISSN 0003-4894 (Print)

Zeng, F. G. (2002). Temporal pitch in electric hearing. *Hear Res,* Vol. 174, No. 1-2, (Dec), pp. 101-6, ISSN 0378-5955 (Print)

Zierhofer, C. (2001). Analysis of a Linear Model for Electrical Stimulation of Axons - Critical Remarks on the "Activating Function Concept". *IEEE Trans Biomed Eng,* Vol. 48, No. 2, pp. 173-184, ISSN 0018-9294 (Print)

Zierhofer, C.; Kals, M.; Krenmayr, A.; Vermeire, K.; Zangerl, M.; Visser, D.; Bader, P.; Neustetter, C. & Schatzer, R. (2009). Simultaneous Pulsatile Stimulation with "Channel Interaction Compensation", *9th European Symposium on Paediatric Cochlear Implantation*, Warsaw, Poland, May 14-17, 2009

Zierhofer, C. M. (2003). Electrical nerve stimulation based on channel specific sampling sequences, U.S. Patent 6,594,525

Zierhofer, C. M. (2007). Electrical stimulation of the acoustic nerve based on selected groups, U.S. Patent 7,283,876

Zierhofer, C. M. & Schatzer, R. (2008). Simultaneous intracochlear stimulation based on channel interaction compensation: analysis and first results. *IEEE Trans Biomed Eng,* Vol. 55, No. 7, (Jul), pp. 1907-16, ISSN 1558-2531 (Electronic)
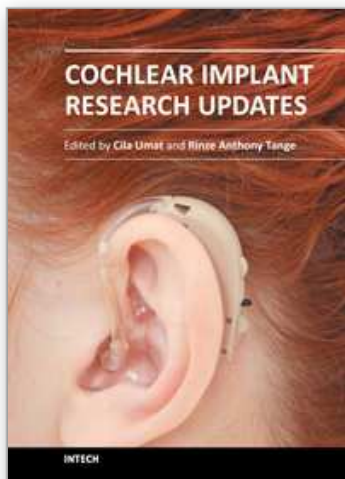
Cochlear Implant Research Updates

Edited by Dr. Cila Umat

For many years or decades, cochlear implants have been an exciting research area covering multiple disciplines which include surgery, engineering, audiology, speech language pathology, education and psychology, among others. Through these research studies, we have started to learn or have better understanding on various aspects of cochlear implant surgery and what follows after the surgery, the implant technology and other related aspects of cochlear implantation. Some are much better than the others but nevertheless, many are yet to be learnt. This book is intended to fill up some gaps in cochlear implant research studies. The compilation of the studies cover a fairly wide range of topics including surgical issues, some basic auditory research, and work to improve the speech or sound processing strategies, some ethical issues in language development and cochlear implantation in cases with auditory neuropathy spectrum disorder. The book is meant for postgraduate students, researchers and clinicians in the field to get some updates in their respective areas.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds