Spectral contrast enhancement improves speech intelligibility in noise for cochlear implants

Waldo Nogueira[a)]
Department of Otolaryngology, Medical University Hannover, Cluster of Excellence Hearing4all, Hannover, Germany

Thilo Rode
HörSys GmbH, Hannover, Germany

Andreas Büchner
Department of Otolaryngology, Medical University Hannover, Cluster of Excellence Hearing4all, Hannover, Germany

(Received 13 April 2015; revised 17 December 2015; accepted 4 January 2016; published online 11 February 2016)

Spectral smearing causes, at least partially, that cochlear implant (CI) users require a higher signal-to-noise ratio to obtain the same speech intelligibility as normal hearing listeners. A spectral contrast enhancement (SCE) algorithm has been designed and evaluated as an additional feature for a standard CI strategy. The algorithm keeps the most prominent peaks within a speech signal constant while attenuating valleys in the spectrum. The goal is to partly compensate for the spectral smearing produced by the limited number of stimulation electrodes and the overlap of electrical fields produced in CIs. Twelve CI users were tested for their speech reception threshold (SRT) using the standard CI coding strategy with and without SCE. No significant differences in SRT were observed between conditions. However, an analysis of the electrical stimulation patterns shows a reduction in stimulation current when using SCE. In a second evaluation, 12 CI users were tested in a similar configuration of the SCE strategy with the stimulation being balanced between the SCE and the non-SCE variants such that the loudness perception delivered by the strategies was the same. Results show a significant improvement in SRT of 0.57 dB ($p < 0.0005$) for the SCE algorithm.
 [http://dx.doi.org/10.1121/1.4939896]

[MAS] Pages: 728–739

I. INTRODUCTION

Cochlear implant (CI) users achieve reasonable hearing performance in quiet, however, perception in noisy environments still remains a challenge. In realistic listening environments, speech signals are degraded by noise and CI users are much more sensitive to noise than normal hearing (NH) listeners (Dorman *et al.*, 1998; Stickney *et al.*, 2004).

One possibility to improve speech perception in noisy environments is to emphasize relevant features of speech. Several studies have examined the emphasis of spectral peaks as a means of enhancing speech for hearing-impaired listeners (Baer *et al.*, 1993; Moore, 2003). In these studies, the difference in amplitude between spectral peaks and valleys (spectral contrast) had been increased with the goal to highlight certain regions of the frequency spectrum around important spectral features, such as formant frequencies. It has been shown that such enhancement of contrastive changes in the speech spectrum can improve speech intelligibility for hearing impaired people (Baer *et al.*, 1993; Kluender *et al.*, 2003; Alexander *et al.*, 2011).

Spectral contrast enhancement (SCE) has been proposed as a method to improve speech intelligibility for CI users (Loizou and Poroy, 2001; Bhattacharya *et al.*, 2011). There might be at least two potential reasons to expect an improvement in speech recognition for CI listeners when enhancing the spectral contrast of speech. First, CIs provide limited spectral resolution. One measure of spectral resolution is spectral modulation detection threshold, defined as the smallest detectable spectral contrast in the spectral ripple stimulus, which is spectrally modulated noise. CI speech processors use relatively broad analysis filters, which combined with the lack of specificity of electrical stimulation, will produce very broad electrical excitation patterns. This may increase the spectral modulation detection threshold and, thus, predict reduced speech intelligibility (Litvak *et al.*, 2007). SCE can be used to compensate for the low resolution of the analysis filters and, therefore, reduce the width of the electrical excitation patterns delivered to the electrodes (Loizou and Poroy, 2001). Second, in many realistic sound scenarios, increasing spectral contrast of speech can improve the signal-to-noise ratio (SNR) to some extent.

CI speech recognition in noise is adversely affected by a reduced internal representation of spectral contrast compared to NH listeners (Bhattacharya *et al.*, 2011). The auditory system uses nonlinear processing to provide the necessary spectral and temporal resolution. It has been suggested (Alexander *et al.*, 2011) that neural mechanisms, such as suppression and adaptation, or higher level processes with similar properties can account for enhanced spectral contrast

[a)]Also at: Department of Otolaryngology, Medical University Hannover, Cluster of Excellence Hearing4all, Hannover, Germany. Electronic mail: nogueiravazquez.waldo@mh-hannover.de

in NH. For example, outer hair cells (OHCs) in the cochlea are responsible for active mechanisms observed in the peripheral auditory system (Meddis *et al.*, 2001). In CIs, the inner hair cells (IHCs) and OHCs are bypassed and, therefore, the active mechanisms of the cochlea need to be modeled in the speech processor. In current CI devices, the only common active functionality is implemented in the automatic gain control (AGC). Most AGCs implemented in current commercial CIs only apply temporal processing and are not frequency specific (Boyle *et al.*, 2009; Khing *et al.*, 2013), providing a very simple processing in comparison to the complex processing applied by the NH system (Nogueira *et al.*, 2007; Harczos *et al.*, 2013). Only the adaptive dynamic range optimization (ADRO) algorithm used by some Nucleus® (Sydney, Australia) CI users (James *et al.*, 2002) and the AGC used in Oticon-Medical (Vallauris, France) CIs (Bozorg-Grayeli *et al.*, 2015) are frequency specific. Furthermore, in CIs, spectral resolution is degraded by the interface between the electrodes and the auditory nerve. The broad spread of neural excitation created by electrical stimulation in the cochlea (Chatterjee and Shannon, 1998; Cohen *et al.*, 2003; Cohen, 2009), suboptimal electrode placement (Finley *et al.*, 2008), and poor neural survival (Long *et al.*, 2014) may reduce the number of independent spectral channels that transmit information (Friesen *et al.*, 2001). In addition, the narrow electrical dynamic range and the low sensitivity to perceive intensity changes contribute to degrade the perception of spectral contrast even further (Nelson *et al.*, 1996; Loizou *et al.*, 2000). It is hypothesized that applying SCE in the sound coding strategy may partially compensate for the wide electrical field, and will also provide a more natural modelling of some nonlinear mechanisms observed in NH.

When listening to speech in noise, relevant features of speech are masked by the noise. It has been shown that CI users typically require more contrast in the input spectra than NH listeners to understand speech in noise (Loizou and Poroy, 2001). This might be explained by the fact that the strong compression applied by CI devices reduces dynamic range (Loizou *et al.*, 2000), which combined with the channel interaction caused between electrodes produces spectral smearing. These negative effects might be partially compensated by enhancing spectral contrast prior to the channel interaction induced along with CI stimulation.

Many CI stimulation strategies are based on the principle of a so-called N-of-M strategy. These strategies stimulate fewer channels (N) per cycle than the total number of active electrodes (M) ($N$of$M$; $N < M$). In contrast, the continuous interleaved sampling (CIS) strategy stimulates all electrodes in each stimulation cycle. In the simplest form, $N$of$M$ strategies, such as the advanced combinational encoder (ACE), only the N channels with the highest amplitudes are stimulated. For this reason, the ACE strategy can already be considered as a SCE algorithm. One characteristic of the ACE strategy is that it often selects many consecutive adjacent spectral bands for stimulation (Nogueira *et al.*, 2005). This might cause pronounced spectral smearing through channel interaction. Instead, if the bands selected were more separated from each other, channel interaction would be reduced and, additionally, spectral smearing should become lower. It is hypothesized that applying SCE before the $N$of$M$ band selection will increase the likelihood of selecting bands that are more spectrally separated.

In previous studies, SCE implementations have already been applied to improve speech intelligibility in noise for hearing impaired people using acoustic stimulation. Alexander *et al.* (2011) proposed a real- time SCE algorithm that has been shown to provide significant improvements in vowel and consonant identification scores in background noise conditions. They evaluated their SCE algorithm in 166 NH subjects. Turicchia and Sarpeshkar (2005) showed that enhanced spectral contrast is an emergent property of a companding strategy and had speculated that their strategy had the potential to improve speech performance in noise. In a study investigating a very similar companding algorithm, Oxenham *et al.* (2007) tested NH listeners using sentences processed through a noise-excited envelope vocoder. They showed improved performance by 6 percentage points in a sentence test based on the IEEE/TIMIT corpus (IEEE, 1969) averaged across subjects and different SNRs. Bhattacharya and Zeng (2007) implemented the same algorithm as a front-end processing stage before applying the CIS or the ACE strategies. They found that a companding algorithm causing spectral and temporal enhancement significantly improved the recognition of both phonemes and sentences in noise using parameters similar to those used in the study of Oxenham *et al.* (2007). In a follow-up study, Bhattacharya *et al.* (2011) investigated the effectiveness of the companding strategy when enhancing spectral and temporal contrast in CI users. In that study, they confirmed that CI subjects indeed obtained improvements in speech intelligibility in noise.

In the studies of Bhattacharya and Zeng (2007) and Bhattacharya *et al.* (2011), SCE was implemented as a front-end algorithm. That means that the audio signal was spectrally enhanced before being processed and analyzed by the CI sound coding strategy. For this reason, they did not have accurate control over the real spectral contrast applied to the electrical signal as the CI processor incorporates AGC and other signal processing algorithms that can modify the action of the front-end spectral contrast enhancer. Moreover, the evaluation of the SCE strategy was performed in five Nucleus® and two Clarion CI users (one CI user's deafness was pre-lingual in origin and therefore excluded from the statistical analysis). The variety of speech processors and implants implies that probably different amounts of spectral enhancement for the same presentation level were delivered to each CI user. For this reason, the results are difficult to quantify and generalize.

The present work evaluates a novel SCE algorithm implemented and integrated into the ACE sound coding strategy in Nucleus® CI users. Section II describes the design and implementation as well as an objective analysis of the SCE algorithm. Section III presents the evaluation results in CI users. Section IV discusses the results in CI users and its relation to the objective analysis of the algorithm. Conclusions are made in Sec. V.

II. METHODS

A. The reference signal processing method

The reference algorithm is the ACE® strategy used in Cochlear Nucleus® devices. The general principle of this strategy is depicted in the block diagram presented in Fig. 1. A more detailed description of the strategy can be found in Nogueira *et al*. (2005).

The signal from the microphone is digitized with a sampling frequency (FS) of 15.659 kHz and sent through an AGC. The AGC implemented in the ACE strategy includes a slow-acting automatic sensitivity control (ASC) followed by a fast acting compression limiter. The goal of the AGC is to compress the incoming acoustic dynamic range such that it can be conveyed into the small electrical dynamic range of a CI recipient (Khing *et al*., 2013). Next, a filter bank implemented as a fast Fourier transform (FFT) is applied to the compressed signal. The FFT uses a buffer size of 128 samples weighted by a Hanning window. A hop size of 16 samples is applied, resulting in an analysis rate of 979 Hz that is converted into the patient specific channel stimulation rate. The first two FFT bins are neglected, which results in spectral range from 184 Hz to FS/2. An estimation of the envelope is calculated for each spectral band of the audio signal. The envelopes E_k ($k = 1,\ldots, M$) are obtained by computing the magnitude of the complex FFT bins. Each band is allocated to one electrode and represents one channel. For each frame of the audio signal, out of the M channels, N channels with the highest amplitudes are selected.

Next, the loudness growth function (LGF) is applied to the N selected channels. This function maps the envelope amplitudes E_k ($k = 1,\ldots, M$) to the corresponding electrodes, compressing the acoustic amplitudes into the subject's dynamic range between the measured threshold level (THL) and the maximum comfortable loudness (MCL) level for electrical stimulation. These levels are given in current levels (CL). For the CI users that participated in the evaluation, the relation between current levels and micro-ampere values (I[μA]) is given by

$$I = \alpha \cdot 100^{\mathrm{CL}/255}, \qquad (1)$$

with $\alpha = 17.5\ \mu$A. Finally, for each frame, the electrode contacts are stimulated sequentially from basal to apical electrodes. The frame rate determines the rate of stimulation on a single channel, also known as channel stimulation rate.

CI users have the possibility to adjust the volume delivered by the speech processor. In the implementation of the ACE strategy, a volume control l is used to adjust the stimulation level SL based on the dynamic range DR between the MCL and the THL (i.e., DR = MCL − THL). For each electrode, the stimulation level SL is obtained as

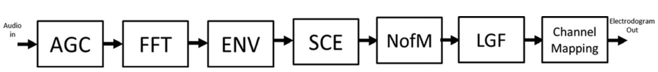

FIG. 1. Block diagram for the SCE strategy.

$$\mathrm{SL} = \mathrm{THL} + \mathrm{DR}\left(1 - l_{\mathrm{range}} + \frac{l_{\mathrm{range}} l}{l_{\max}}\right), \qquad (2)$$

where $l_{\max}$ denotes the highest volume setting and l_{range} denotes the volume range defined as the difference between $l_{\max}$ and $l_{\min}$ using the volume setting "0." Typically, $l_{\max} = 10$ and $l_{\min} = 0$.

B. The SCE strategy

The SCE strategy (Fig. 1) incorporates a new processing stage just before the *NofM* band selection.

The goal of this stage is to enhance peaks with respect to valleys in the envelope domain. The algorithm has been designed so that no other spectral features, such as spectral tilt, are modified with respect to the ACE strategy. The algorithm works as follows:

(1) Locate the spectral local maxima/peaks for each spectral audio frame. The frequency band of the peak is denoted by p and the envelope magnitude of the peak in this specific frame is denoted by E_p, where $p = 1,\ldots, P$. P was set to 3 because the peaks will coincide with formants. It has been shown that 2–3 formants are enough to disambiguate some speech sounds, such as vowels and some consonants.

(2) Define band intervals of FFT bins between the envelope peaks E_p, and between the peaks and the boundaries (frequency bands $k = 1$ and $k = 22$). Band intervals are denoted using the index i.

(3) Locate the spectral valleys for each audio frame. The frequency band of the valley is denoted by ν and the envelope of the valley is denoted by E_ν, where $\nu = 1,\ldots, V$. V is equal to P, P + 1, or P − 1, depending on whether the local peak corresponds or not to the frequency boundaries.

(4) Estimate the original contrast for each interval i as

$$\mathrm{SC}_i = E_{p_i} - E_{\nu_i}, \qquad (3)$$

with E_{p_i} and E_{ν_i} representing magnitudes on a dB scale.

(5) Compute the enhanced spectral contrast SR_i as

$$\mathrm{SR}_i = \mathrm{SC}_i(1 + \mathrm{SCE}_{\mathrm{factor}}), \qquad (4)$$

where $\mathrm{SCE}_{\mathrm{factor}}$ is in percentage units.

(6) Apply the spectral contrast gain SR_i to the envelopes E_k contained in each interval i using a modified version of the algorithm presented by (Loizou and Poroy, 2001)

$$E_{\mathrm{SCE}_k} = \left(\frac{E_k - E_{\nu_i}}{E_{p_i} - E_{\nu_i}}\right) SR_i + E_{p_i} - SR_i, \qquad (5)$$

where k denotes the frequency band ($k = 1, \ldots, M$). For example, a $\mathrm{SCE}_{\mathrm{factor}} = 0$ would not change the given contrast while a $\mathrm{SCE}_{\mathrm{factor}} = 1$ would double it on a dB scale. Using this algorithm, audio frames with little spectral contrast have their contrast less enhanced when compared to frames with high contrast.

Figure 2 shows as an example the original and the enhanced spectral envelopes using a $SCE_{factor} = 1$ for a single frame of a vowel "a" before applying the *N*of*M* block used in the ACE strategy.

C. Software/hardware implementation

The ACE and the SCE strategies were implemented in Simulink® to be run in real-time on a Speedgoat® xPC® target machine (Goorevich and Batty, 2005). The calibrated stimuli were played back from a standard personal computer (PC) with a Fireface 800® (RME, Haimhausen, Germany) soundcard through an audiometric loudspeaker on 0° azimuth at 1 m distance from a Cochlear® System5 microphone array. The playback setup was positioned in a sound proof anechoic chamber. The signal being picked up by the speech processor microphone was transmitted to the SpeedGoat® xPC® system for further real-time processing. This was performed by a Simulink-validated model of the ACE strategy, which subsequently computed the electrical stimuli delivered to the CI. For the experimental evaluation of SCE, three different configurations of the SCE strategy, which differed in the amount of SCE, were used. The three strategies are denoted by SCE0, SCE05, and SCE1 and correspond to $SCE_{factors}$ 0, 0.5, and 1, respectively. SCE0 means no spectral enhancement and is equivalent to the clinical ACE strategy.

All study participants used their own clinical map, i.e., their clinical stimulation rate, comfort and THLs, number of maxima, and frequency-place allocation table.

Figure 3(a) presents the envelopes obtained for one frame of the phoneme /a/ and Fig. 3(b) presents the selected bands using SCE0, SCE05, and SCE1. The envelopes are given in decibel (dB) units normalizing the maximum amplitude value to 96 dB.

D. Objective evaluation

There are different reasons why SCE could potentially improve speech intelligibility in CIs. First, SCE will cause bands surrounding local maxima in the spectrum to be

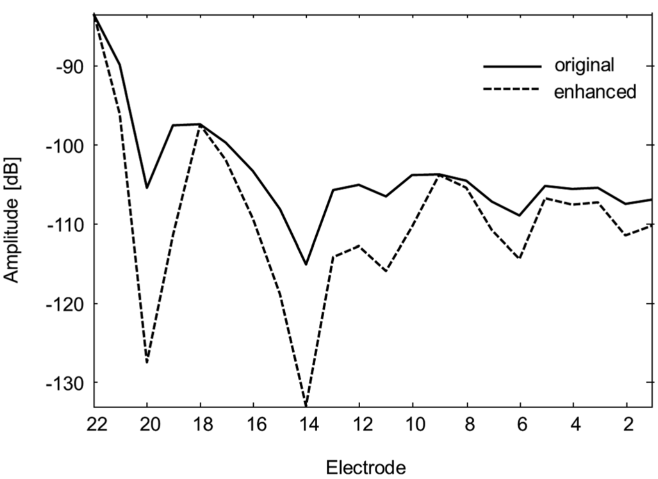

FIG. 2. Example of SCE for a single spectral frame of an audio signal. Three local maxima are kept with respect to the original signal and the valleys are reduced. The original contrast is enhanced based on the original contrast found in each interval using a $SCE_{factor} = 1$.

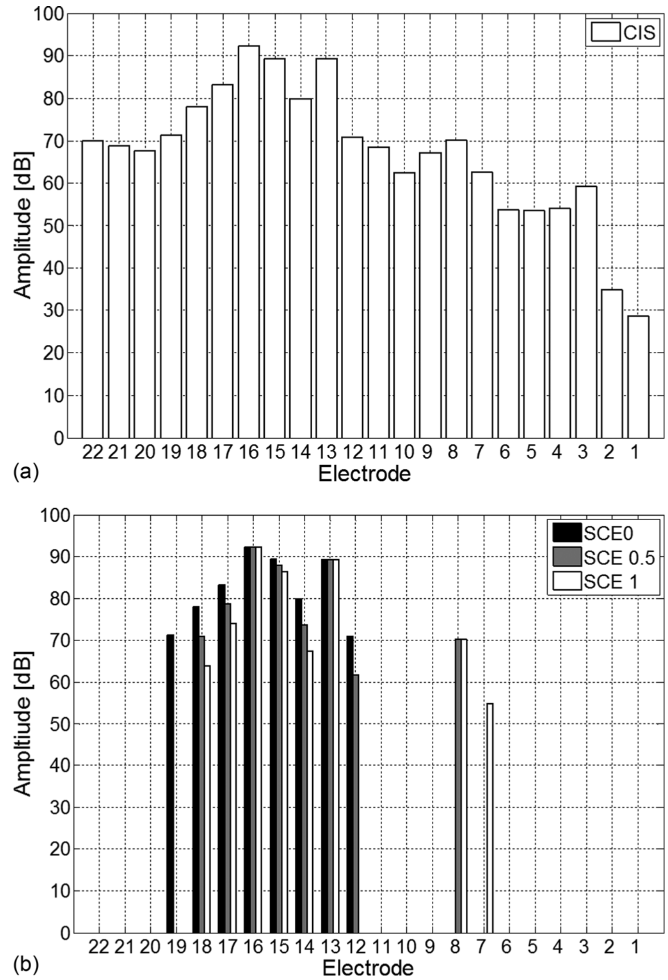

FIG. 3. (a) Frequency band decomposition for one frame of the phoneme /a/ using CIS. (b) Selected frequency bands using ACE/SCE0, SCE05, and SCE1. The bands are given in decibel units normalizing the maximum amplitude value to 96 dB. Bands selected by the ACE/SCE0, SCE05, and SCE show that different SCE_{factor} values cause a different selection of bands on the one hand, and different amplitudes of the selected bands depending on the SCE factor on the other hand.

attenuated, so subsequently the amount of SCE (i.e., the SCE factor) will impact which bands are selected by the *N*of*M* strategy for stimulation. It is possible that a particular SCE enhancement causes a band selection that contains more perceptually relevant information than a standard band selection based on a simple peak picking algorithm (Nogueira *et al.*, 2005). Second, in scenarios with a relatively flat spectrum background noise, enhancing the spectral contrast should enhance the local SNR, at least if the target speech has higher energy than the noise.

Additionally, SCE attenuates some components of the sound that might result in lower stimulation currents yielding potential power savings for the device. Sections II D1–II D3 present how SCE influences band selection, current savings, and SNR of speech signals.

1. Band selection

The SCE algorithm is applied before the *N*of*M* selection and this will affect the selection of bands depending on the SCE factor. The ACE strategy tends to select groups of consecutive frequency bands or clusters of bands, preserving the likelihood of channel interaction between adjacent electrodes inside the cochlea. Using SCE before the band selection should decrease the number of times that clusters of bands are selected (Nogueira *et al.*, 2005). An experiment that involves counting the number of clusters of different lengths that are selected by SCE0, SCE05, and SCE1 was conducted. Twenty sentences of the Oldenburg sentence test (OLSA; Kollmeier and Wesselkamp, 1997) in quiet and in noise (SNR = 5 dB) were processed using a channel stimulation rate of 1000 Hz and selecting eight bands in each frame for all strategies. The maximum cluster length is therefore eight, when all selected bands are sequenced consecutively. The minimum possible cluster length is one, which occurs when one selected band is separated from other selected bands by at least one band.

The results presented in Figs. 4(a) and 4(b) show that increasing the $SCE_{factors}$ results in fewer broad clusters, therefore, reducing the channel interaction between adjacent bands.

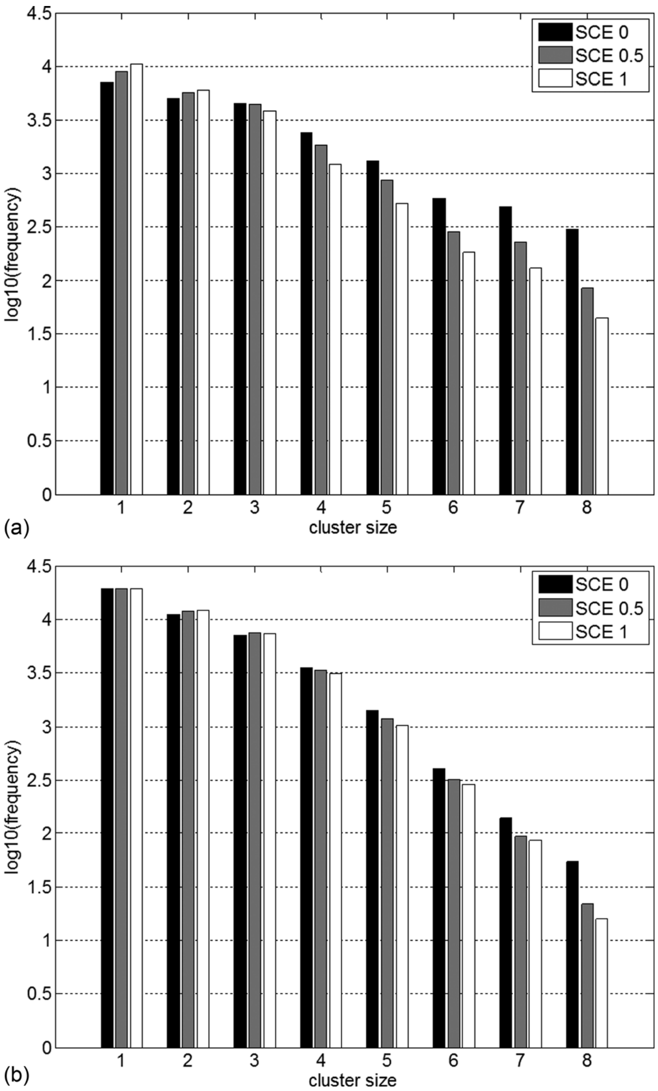

FIG. 4. Histogram of cluster length or number of consecutive bands for various SCE factors (0,0.5,1) (a) without noise and (b) with noise at an SNR = 5 dB.

2. Analysis of average stimulation current

The power consumption is related to the amount of stimulation current of the CI device. Electrodograms or patterns of stimuli over the electrodes and time for ten sentences of the OLSA sentence test were analyzed. The sentences were processed with the SCE strategy for different $SCE_{factors}$ (0, 0.5, and 1). The stimulation current delivered by each SCE implementation was estimated by averaging all currents applied to all electrodes over the whole duration of the sentences. Next, the percentage decrease in current for each SCE implementation with respect to the ACE strategy was computed using the same map for all strategies, i.e., THL and MCL levels. The map was set to THL = 150 and MCL = 200 current levels. Figure 5 presents the current savings in percentage for the different SCE factors.

Figure 5 shows that the SCE strategy reduces the amount of current delivered to the electrodes compared to the ACE strategy. This reduction in current can have two consequences: (1) SCE can produce power savings, (2) SCE might sound softer in level than the ACE strategy and this could negatively impact speech intelligibility performance (Firszt *et al.*, 2004). Note that this will be investigated in experiment 2.

3. SNR analysis

The effects of the SCE algorithm on SNR were analyzed by mixing speech signals with noise at different levels. The mixed signals were processed with the SCE algorithm using different $SCE_{factors}$ ranging from 0 to 4. Both the speech and noise signals were extracted from the OLSA sentence test. Next, the SNR was estimated after applying the *NofM* band selection using the different $SCE_{factors}$ (see Fig. 1). The SNR was estimated as follows: First, a noisy OLSA sentence with a known SNR was processed and the SCE gains (SR_i, with $i = 1, \ldots, M$) for each frame of the audio signal were stored. Second, the original clean speech signal and the original OLSA noise signal were processed separately by the SCE algorithm using the stored gains SR_i. This technique, however, is only applicable if all processing operations up to the *NofM* block (included) are linear (see in Fig. 1). This does not hold for the envelope detection block. As a workaround,

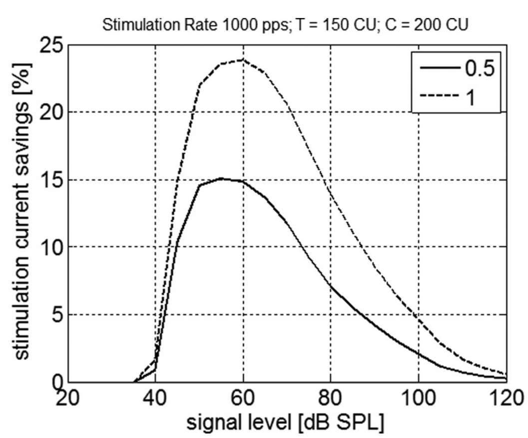

FIG. 5. Percentage of stimulation current savings using the SCE algorithm with SCE factors 0.5 and 1 in comparison to the ACE/SCE0 strategy.

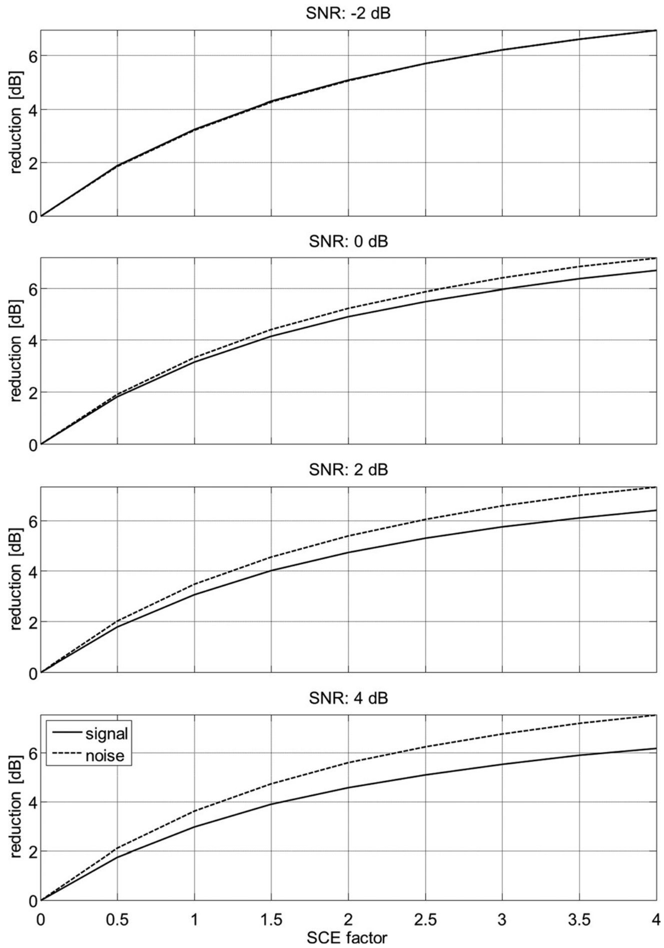

FIG. 6. SNR estimation for different $SCE_{factors}$. The SNR of the input signals was set to −2, 0, 2, and 4 dB from top to bottom, respectively. Signal and noise were processed individually with the same SCE gains to estimate the resulting dependence of SNR on the SCE factor.

the stored SCE gains SR_i from the envelope domain were converted into the linear FFT domain. This was achieved using Table 5 provided in Nogueira *et al.* (2005). The same SR_i gain value was used for all the FFT bands corresponding to the same envelope band. Next, the SR_i gains were applied to the individual FFT frequency bins of the speech and noise signals individually. The root-mean-square (RMS) power of each frame of the individual signals was estimated and averaged across all frames for speech and noise. Next, the effect of speech level reduction was defined as the difference between the averaged RMS power for the speech signal processed with $SCE_{factors} = 0$ vs all other $SCE_{factors}$ values (0.5, 1, 2, and 4). The effect of noise reduction was obtained in the same way as the effect of speech reduction but using the noise signal instead of the speech signal. The SNR improvement of using a SCE_{factor} different than 0 can be estimated by subtracting the noise level reduction from the estimate of speech level reduction. Figure 6 presents the averaged speech level reduction and noise level reduction estimation for ten sentences of the OLSA test for different $SCE_{factors}$ and for different input SNRs of the mixed signal (−2, 0, 2, and 4 dB), which are within the range of SRTs typically observed in CI users using the OLSA sentence test.

Figure 6 shows that the SCE algorithm causes a larger reduction in RMS for the background noise than for the target speech signal for input SNRs larger than or equal to 0 dB. This analysis also shows that the target speech RMS level is reduced by the SCE algorithm and this might cause a softer sound perception using SCE.

III. RESULTS

A. Experiment 1: Speech recognition in noise with ACE and SCE

1. Subjects

Twelve adult CI subjects participated in the evaluation of the SCE strategy. The subjects were native German speakers, post-lingually deaf, and had at least one year of experience with their CI. All subjects had either a CI24RE (Freedom), RE-24 CA, CI422, or CI512 series implant. The details for the subjects are presented in Table I. The study protocol was approved by the institutional medical ethics committee. All subjects gave their informed consent to participate in the study.

2. Procedure

Speech intelligibility in noise was measured by means of the OLSA test. Lists of 20 OLSA sentences were used to determine the speech reception threshold (SRT) of each patient with each strategy. Prior to testing, the volume control of the real-time xPC system was adjusted to satisfy the

TABLE I. Details for subjects participating in experiment 1.

Patient identification	Age	Duration of deafness	Cause of deafness	Implant experience in years	Electrode type	Stimulation rate
P1	34	1.25	Trauma	2.6	CI512	900
P2	64	1.5	Sudden hearing loss	2.5	CI512	900
P3	47	30.67	Otitis media	3.3	CI512	900
P4	82	37.17	Trauma	10.8	RE-24 CA	1200
P5	85	0	Unkwown	0.7	CI422	900
P6	63	0	Toxic	1.5	CI422	1200
P7	77	0	Unknown	3.5	RE-24 CA	900
P8	77	0	Unknown	2.9	CI422	900
P9	56	0	Unknown	1.1	CI422	900
P10	57	0.75	Sudden hearing loss	7.4	RE-24 CA	900
P11	68	9.42	Sudden hearing loss	4.3	RE-24 CA	900
P12	68	16.84	Morbus Meniere	4.1	CI512	900

loudness requirements of the CI users using the ACE/SCE0 strategy. The same volume setting was used to evaluate the SCE05 and SCE1 strategies. Each condition (SCE0, SCE05, and SCE1) was repeated twice and tested in an ABC-CBA pattern with a randomized order of the conditions. The test was blinded for the research audiologist performing the tests and for the study participants. One list of 20 sentences without background noise was used to train the CI users in the task, 2 additional lists of 20 sentences in speech shaped noise (OLSA noise) were used for additional training. The noise level was set constant to 60 dB sound pressure level (SPL) and the level of the speech was varied until 50% speech recognition was achieved. The resulting SNR was used as the SRT score.

3. Results

Figure 7 presents the results for each strategy and CI user. The resulting data were compared using one-way repeated measures of analysis of variance (ANOVA) to test for the effect of processing scheme. If a statistically significant effect was identified by the ANOVA, pairwise comparisons were made using two-tailed *t*-tests with Bonferroni-corrected *p*-values. Following each ANOVA, three pairs of data were compared. Therefore, a correction factor of 3 was applied to all *p*-values.

The results of the ANOVA indicated no significant effect for the processing scheme (SCE0, SCE05, or SCE1), Wilk's lambda $= 0.795$, $F(2,10) = 1.28$, $p < 0.318$, $\eta^2 = 0.205$. It has to be remarked that the SCE05 and SCE1 strategies used less stimulation current with no significant difference in speech intelligibility with respect to the SCE0 strategy.

B. Experiment 2: Speech recognition in noise with loudness-compensated SCE

In experiment 2, the volume on the SCE05 strategy was modified to match the loudness perception caused by the ACE/SCE0 strategy. Because the SCE strategy uses less current than the ACE strategy, it is hypothesized that it might sound softer and, therefore, speech intelligibility might be inferior when using the SCE processing as described in experiment 1.

1. Subjects

Twelve additional CI subjects participated in the evaluation of a modified version of the SCE strategy. All subjects were German native speakers, post-lingually deaf and had more than one year of experience with their CI. Details of the subjects participating in experiment 2 are presented in Table II.

2. Procedure

A loudness balancing procedure was used to match the loudness perception induced by the SCE configuration. A single sentence from the OLSA sentence test played in a loop was used for the balancing procedure. The new loudness matched strategy has been termed loudness-compensated spectral contrast enhancement (LCSCE) strategy. Loudness matching was achieved by first setting the most comfortable

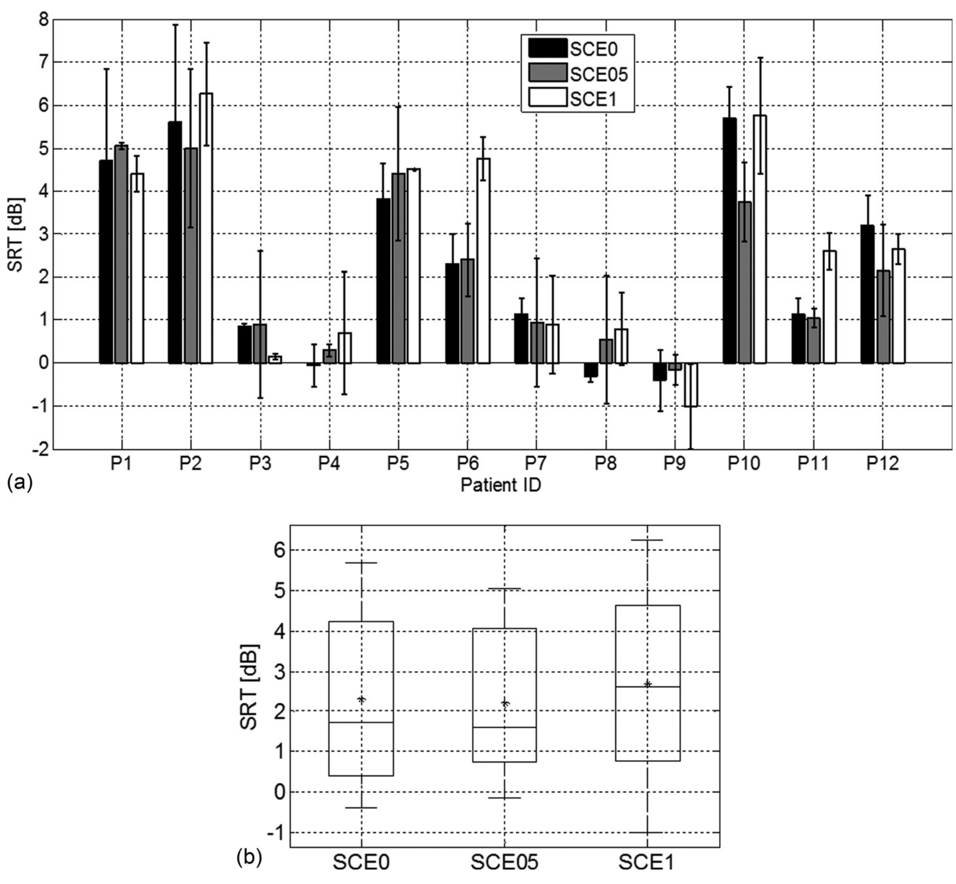

FIG. 7. (a) SRT by subject (average and standard deviation) for the OLSA sentence test. (b) Averaged SRT for all subjects, the horizontal line in the box plot indicates the median value and the asterisk indicates the mean value.

TABLE II. Details for subjects participating in experiment 2.

Patient identification	Age	Duration of deafness	Cause of deafness	Implant Experience in years	Electrode type	Stimulation rate
P13	55	4.92	Unknown	2.0	CI422	900
P14	52	19	Unknown	1.9	CI422	900
P15	20	0	Genetic	6.6	RE-24CA	1200
P16	78	8.92	Sudden hearing loss	3.5	CI512	900
P17	61	1.08	Otosklerose	6.8	RE-24 CA	1200
P18	76	0l	Unknown	1.6	CI422	900
P19	64	13.08	Cholesteatom	1.9	CI512	900
P20	24	1	Unknown	1.2	CI422	900
P21	65	0	Sudden hearing loss	1.8	CI24RE	900
P22	75	9.17	Sudden hearing loss	5.8	RE-24CA	900
P23	67	7.76	Sudden hearing loss	7.5	RE-24 CA	500
P24	75	6.42	Sudden hearing loss	2.2	RE-24CA	900

volume (MCV) for the ACE/SCE0 strategy. Only the volume control (l) of the xPC system was modified [see Eq. (2)]. Next, the SCE05 was matched in loudness by the subject with respect to the ACE/SCE0. The loudness matching was repeated twice, starting from 30% below and above the volume setting of the ACE/SCE0 strategy, and the outcome was averaged. The CI user compared the loudness delivered by both strategies and matched the volume setting of the SCE05 strategy using a slider in a graphical user interface. Table III shows the loudness balancing results for each individual CI user.

From Table III it can be concluded that the SCE05 generally sounds softer than the ACE/SCE0 strategy.

3. Results

Exactly the same procedure as described in Sec. III A2 to evaluate speech performance with each strategy was used. The three strategies being compared were the ACE/SCE0, SCE05, and LCSCE05. Because of testing time available, only one SCE condition together with its loudness balanced version could be tested. The protocol used for the evaluation is based on an ABC-CBA pattern with additional training lists. Clinical experience with the OLSA test shows that adding more conditions to the protocol causes patients to become tired, decreasing the accuracy of the results. From experiment 1, the SCE05 condition was arbitrarily selected. The same statistical analysis as used in experiment 2 was applied. The speech intelligibility results for each CI subject are presented in Fig. 8(a) and the averaged results using a box plot are presented in Fig. 8(b).

The results of the ANOVA indicated a significant effect of the processing scheme, Wilk's lambda $= 0.216$, $F(2,10) = 18.13$, $p < 0.0005$, $\eta^2 = 0.784$. Follow-up comparisons indicated that the pairwise comparisons (LCSCE05 vs SCE0/ACE) and (LCSCE05 vs SCE05) were significant. The LCSCE05 obtained a significant improvement of 0.57 dB in SRT with respect to the SCE0 [$p < 0.0005$; $t(11) = 6.247$] and of 0.5 dB with respect to the SCE05 [$p = 0.008$; $t(11) = 3.265$]. No significant difference could be observed between SCE05 and ACE [$p = 0.657$; $t(11) = 0.657$]. These results suggest that SCE improves speech intelligibility if the loudness is balanced with respect to the ACE and the SCE05 strategies.

TABLE III. Results of the MCV obtained from the loudness balancing procedure for each CI user.

Patient identification	Volume SCE0 (MCV)	Volume SCE05		
		Start value: 30% below MCV	Start value: 30% above MCV	Mean value
P1	6	7	7.615	7.3075
P2	7	8	8.48	8.24
P3	5.4	6.4	7.3	6.85
P4	6	6.36	6.72	6.54
P5	4	7.24	7.96	7.6
P6	6	6.76	6.84	6.8
P7	6	6.36	6.48	6.42
P8	6	6.28	6.8	6.54
P9	5	6.4	6.8	6.6
P10	5	5.7	6.3	6
P11	6	6.99	6.84	7.14
P12	6	6.6	8.16	7.38

IV. DISCUSSION

This study investigated whether SCE can improve speech intelligibility in *NofM* strategies for CI users. In a first evaluation with OLSA sentences in noise, 12 CI users were tested using a standard coding strategy (ACE/SCE0) and two spectrally enhanced versions (SCE05 and SCE1). No significant differences in SRT were observed between the three conditions. However, an objective analysis of the stimulation patterns shows a clear reduction in electrical stimulation current when using SCE05 and SCE1, which probably causes a softer loudness percept when listening to speech with these strategies. In a second evaluation, 12 CI users were evaluated using a loudness-compensated version of the SCE strategy (LCSCE05). Results show a significant average improvement of 0.57 dB SRT for the LCSCE05 with respect to the ACE/SCE0 strategy. Eleven out of 12 CI users obtained an improvement with the

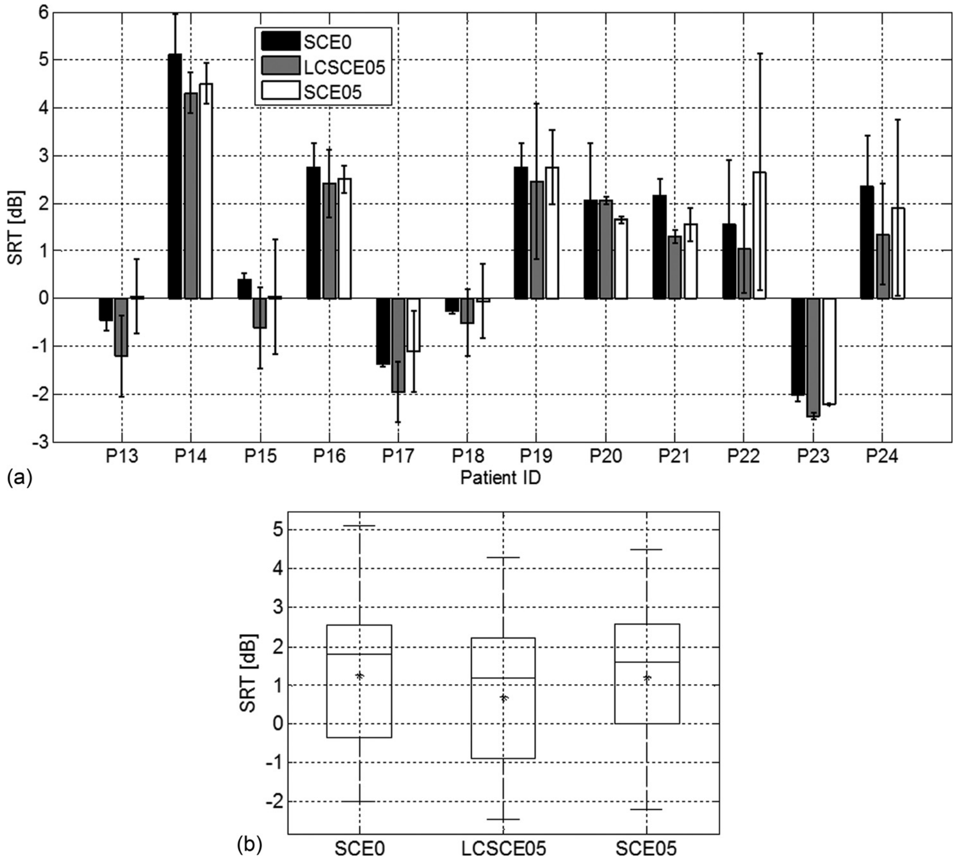

FIG. 8. (a) SRT for the OLSA sentence test by subject represented as average and standard deviation of two lists of the OLSA for each condition (SCE0, SCE05, and LCSCE05). (b) Averaged SRT for all subjects, the horizontal line in the box indicates the median value.

LCSCE05 strategy. Again, no significant differences in performance were observed between the ACE/SCE0 and the SCE05 strategies.

There are at least three possible explanations for the results obtained: (1) The SCE algorithm provides an objective SNR improvement. (2) *N*of*M* strategies select groups of consecutive bands or clusters and this might smear the spectral information transmitted to the auditory nerve. Applying spectral enhancement before band selection reduces the amount of times that larger clusters are selected for stimulation. (3) The SCE algorithm enhances certain features of the sound that are important for speech intelligibility.

A. Effect on SNR improvement

It has been shown that SCE reduces the energy of the noise signal with respect to the speech signal when both are mixed together leading to an improvement in SNR. Single channel noise reduction (SCNR) algorithms do not cause an improvement in speech intelligibility in hearing aid listeners at relatively low SNRs (Sarampalis, 2009), but at positive SNRs seem to be beneficial for CI users (Hochberg *et al.*, 1992; Dawson *et al.*, 2011; Mauger *et al.*, 2012; Buechner *et al.*, 2014; Nogueira *et al.*, 2015). For example, CI users obtained significant improvements in speech perception of 24 percentage points in babble noise (Hu *et al.*, 2007), 2.1 dB SRT (Dawson *et al.*, 2011), and 19 percentage points (Mauger *et al.*, 2012) in speech weighted noise. It is important to mention that SCNR algorithms achieve objective improvements in SNR and speech intelligibility index (SII). However, this does not automatically lead to an improvement in intelligibility in human subjects. This has been shown in errors in speech intelligibility prediction produced by the SII, depending whether the noise is fluctuating or stationary (Rhebergen *et al.*, 2006), and using models of the effects of modulations on intelligibility (Jorgensen and Dau, 2011). The SCE strategy presented in this paper shows smaller benefits in speech intelligibility than those reported in the literature using SCNR algorithms. In the present study, it can be observed that the SRT improvement obtained from SCE cannot be directly explained by the obtained objective SNR improvement. Therefore, this SNR increased may not be the only effect contributing to the increased speech intelligibility in noise. Both algorithms, SCNR and SCE, may be compatible and it remains a question to investigate whether both algorithms can provide combined additive benefits when working together.

B. Effect on band selection

It has been shown that the SCE strategy causes not only an attenuation of some of the spectrum's energy, but also a different band selection compared to using the ACE/SCE0 strategy. In another strategy called PACE (psychoacoustic advanced combination encoder)[1] (Nogueira *et al.*, 2005) using a psychoacoustic model of simultaneous masking instead of a simple peak-picking algorithm to select bands, it was shown that the selection of consecutive bands or clusters within each stimulation cycle occurred less often than with the ACE strategy. It was hypothesized that selecting bands for stimulation that are more separated from each other reduces channel interaction, leading to a more efficient transfer of information from the electrodes to the auditory nerve. An objective experiment has shown that SCE, in the same manner as the PACE strategy, selects narrower clusters than the ACE strategy. However, in the studies of Nogueira *et al.* (2005) and Büchner *et al.* (2008), only minor improvements in speech intelligibility could be observed when using the PACE strategy with respect to the ACE strategy. Therefore, other reasons may contribute to the improvements in speech intelligibility observed when using the SCE strategy.

C. Effect on enhancement of certain features of sound

Formants are important features to recognize vowels and some consonants. For example, the relative amplitudes or spectral contrast of formant energies affect the percepts arising from $F1$ and $F2$ in close proximity (Chistovich and Lublinskaya, 1979; Sagi *et al.*, 2010). The electrode–nerve interface of each CI user will modify the percepts of $F1$ and $F2$ in a different manner which can lead to potential confusions between vowels and between some consonants. Enhancing spectral contrast will emphasize $F1$ and $F2$ peak amplitudes by attenuating spectral valleys in between, and can therefore improve the discrimination of vowels and some consonants.

Moreover, it has been shown that higher stimulation levels improve modulation detection thresholds for single channel stimulation (Green *et al.*, 2012) and these are strongly correlated with speech recognition in CIs (Fu, 2002). Therefore, it could be that the higher stimulation levels used in the LCSCE05 condition induce a better coding of the speech modulations.

D. Comparison with previous studies

Previous studies have shown that SCE produces improvement in speech intelligibility in noise for hearing impaired people (Turicchia and Sarpeshkar, 2005; Alexander *et al.*, 2011) and NH listeners using a vocoder (Oxenham *et al.*, 2007). Bhattacharya and Zeng (2007) and Bhattacharya *et al.* (2011) evaluated SCE in CI users as a front end signal processor without having full control on the real spectral contrast achieved in the stimulation pattern delivered to the electrodes. The present study, however, implemented the SCE algorithm integrated into the sound coding strategy. Furthermore, the evaluation was conducted in CI users of the same implant device and strategy, allowing a more controlled evaluation of the algorithm. As in previous studies, an improvement in speech performance when using SCE was observed, but only when the loudness between the SCE05 and the ACE/SCE0 strategies was balanced. Without loudness compensation SCE does not improve speech intelligibility, probably because this processing also reduces the RMS of the target speech signal. Bhattacharya and Zeng (2007) showed an average improvement in SRT of around 1–2 dB for a speech presentation level of 70 dB SPL for all conditions. These results are slightly better than the results presented in this paper, however, the algorithm they used enhanced both temporal and spectral contrast of the audio signal. Speech signals vary with time and, therefore, the SCE algorithm presented in this manuscript will not only enhance spectral contrast but also temporal fluctuations as a by-product. For this reason, the temporal enhancement cannot be well controlled using the SCE algorithm.

E. Variability of the SRT in the OLSA sentence test

Clinical experience with the OLSA sentence test in noise has shown that the test–retest variability of the OLSA increases with poorer SRT performance. It has been shown that the reproducibility of the OLSA becomes unreliable for CI users with SRTs above 5 dB (Hey *et al.*, 2014). In the present study, CI users reached performance levels ranging from -2 dB to $+6$ dB SRT. Given the relatively small improvement obtained by the SCE strategy and the relatively large variability of the OLSA, it is remarkable that 11 out of 12 CI users obtained an improvement with the SCE strategy, which in turn led to an overall statistical significant improvement for the LCSCE05 strategy.

F. Real-time implementation

An important aspect when designing new signal processing strategies for CIs is the complexity of the algorithms. The SCE algorithm here presented is computationally much less expensive than the algorithm used by Bhattacharya and Zeng (2007), which requires around 50 analysis and synthesis filters, and additional compressing and expanding stages within each filter band. The SCE was implemented in a real-time Speedgoat xPC system. All the strategies were implemented using Simulink. The whole system has been validated to produce the same stimulation patterns as the commercial Freedom/System5 processor. Given the low complexity and latency introduced by the SCE algorithm, the implementation on actual behind-the-ear processors seems feasible.

G. Future directions

The SCE strategy has been evaluated in a speech intelligibility task under stationary background noise. In future evaluations, it should be investigated whether the same improvements are observed in other types of background noises and using different loudspeaker configurations simulating other sound environments. The SCE strategy has been implemented within the ACE strategy framework, which—by means of the peak picking algorithm—already applies SCE to some extent. Implementing the SCE algorithm within a CIS strategy, where no bands are discarded for stimulation, could produce larger improvements in speech intelligibility than those observed within the ACE strategy.

In the present study, it can be observed that SCE reduces loudness perception and this causes a reduction in speech intelligibility. For a moderate $\mathrm{SCE}_{\mathrm{factor}}$ a significant improvement in speech intelligibility was observed when loudness was compensated. It is possible that applying loudness compensation for larger $\mathrm{SCE}_{\mathrm{factors}}$ produces further improvements in speech intelligibility. Additionally, it might be interesting to investigate whether the amount of SCE can be optimized to each CI user. One idea could be to configure the strength of the SCE based on the spectral resolution of each CI user.

Spectral resolution could be measured based on a spectral modulation threshold task (Litvak *et al.*, 2007) or based on backward telemetry measures from the CI, such as spread of excitation (Cohen *et al.*, 2003; Cohen, 2009). Litvak *et al.* (2007) showed that spectral modulation detection correlates with vowel identification performance. Therefore, one can think that enhancing spectral modulations, for example, by means of SCE, can improve vowel identification. However, a CI user who has very poor spectral resolution capabilities, i.e., cannot discriminate between several electrodes, will probably not benefit from SCE.

Recently, it has been shown that the enhancement of spectral contrast in the current time interval in combination with the enhancement of spectral changes over time is beneficial for hearing impaired people (Chen *et al.*, 2012, 2013). The implementation of such an enhancement in a CI coding strategy can possibly lead to further improvements in speech intelligibility.

V. CONCLUSIONS

SCE is achieved by attenuating the valleys of the spectrum and thus increasing the spectral peak-to-valley ratio in the signal. The consequence of this processing is that, generally, the amount of current delivered by the implant is decreased. This paper shows that SCE (using $SCE_{factors}$ 0.5 and 1) without any loudness compensation does not change speech intelligibility, but does reduce current consumption. Additionally, SCE provides a significant improvement in speech intelligibility in noise compared to the ACE strategy when the loudness of the SCE-processed signal is balanced to the stimuli generated by the standard ACE strategy. Given the small improvements produced by SCE in objective SNR, two potential reasons might explain why speech intelligibility improves. First, the smearing induced by the electrode–nerve interface of each CI user might be slightly compensated by providing SCE. Second, SCE causes a different band selection for stimulation, which might be more perceptually relevant than without SCE.

ACKNOWLEDGMENTS

The authors would like to thank the subjects who have participated in the experiments and the two anonymous reviewers for their comments on different versions of this manuscript. This work was supported by the DFG Cluster of Excellence EXC 1077/1 "Hearing4all" and Cochlear Corporation.

[1]PACE has been commercialized in the Nucleus® system as MP3000.

Alexander, J. M., Jenison, R. L., and Kluender, K. R. (**2011**). "Real-time contrast enhancement to improve speech recognition," PloS One **6**(9), e24630.

Baer, T., Moore, B. C., and Gatehouse, S. (**1993**). "Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times," J. Rehabil. Res. Dev. **30**(1), 49–72.

Bhattacharya, A., Vandali, A., and Zeng, F.-G. (**2011**). "Combined spectral and temporal enhancement to improve cochlear-implant speech perception," J. Acoust. Soc. Am. **130**(5), 2951–2960.

Bhattacharya, A., and Zeng, F.-G. (**2007**). "Companding to improve cochlear-implant speech recognition in speech-shaped noise," J. Acoust. Soc. Am. **122**(2), 1079–1089.

Boyle, P. J., Büchner, A., Stone, M. A., Lenarz, T., and Moore, B. C. J. (**2009**). "Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants," Int. J. Audiol. **48**(4), 211–221.

Bozorg-Grayeli, A., Guevara, N., Bebear, J.-P., Ardoint, M., Saaï, S., Hoen, M., Gnansia, D., Romanet, P., and Lavieille, J.-P. (**2015**). "Clinical evaluation of the xDP output compression strategy for cochlear implants," Eur. Arch. Otorhinolaryngol. (in press).

Büchner, A., Nogueira, W., Edler, B., Battmer, R.-D., and Lenarz, T. (**2008**). "Results from a psychoacoustic model-based strategy for the Nucleus-24 and Freedom cochlear implants," Otol. Neurotol. **29**(2), 189–192.

Buechner, A., Dyballa, K.-H., Hehrmann, P., Fredelake, S., and Lenarz, T. (**2014**). "Advanced beamformers for cochlear implant users: Acute measurement of speech perception in challenging listening conditions," PloS One **9**(4), e95542.

Chatterjee, M., and Shannon, R. V. (**1998**). "Forward masked excitation patterns in multielectrode electrical stimulation," J. Acoust. Soc. Am., **103**(5), 2565–2572.

Chen, J., Baer, T., and Moore, B. (**2013**). "Effect of spectral change enhancement for the hearing impaired using parameter values selected with a genetic algorithm," J. Acoust. Soc. Am. **133**(5), 2910–2920.

Chen, J., Baer, T., and Moore, B. C. J. (**2012**). "Effect of enhancement of spectral changes on speech intelligibility and clarity preferences for the hearing impaired," J. Acoust. Soc. Am. **131**(4), 2987–2998.

Chistovich, L. A., and Lublinskaya, V. V. (**1979**). "The 'center of gravity' effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli," Hear. Res. **1**(3), 185–195.

Cohen, L. T. (**2009**). "Practical model description of peripheral neural excitation in cochlear implant recipients: 4. Model development at low pulse rates: General model and application to individuals," Hear. Res. **248**(1–2), 15–30.

Cohen, L. T., Richardson, L. M., Saunders, E., and Cowan, R. S. C. (**2003**). "Spatial spread of neural excitation in cochlear implant recipients: Comparison of improved ECAP method and psychophysical forward masking," Hear. Res. **179**, 72–87.

Dawson, P. W., Mauger, S. J., and Hersbach, A. A. (**2011**). "Clinical evaluation of signal-to-noise ratio-based noise reduction in Nucleus® cochlear implant recipients," Ear Hear. **32**, 382–390.

Dorman, M. F., Loizou, P. C., and Fitzke, F. (**1998**). "The identification of speech in noise by cochlear implant patients and normal-hearing listeners using 6-channel signal processors," Ear Hear. **19**(6), 481–484.

Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., Hullar, T. E., and Skinner, M. W. (**2008**). "Role of electrode placement as a contributor to variability in cochlear implant outcomes," Otol. Neurotol. **29**(7), 920–928.

Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gaggl, W., Runge-Samuelson C. L., and Wackym, P. A. (**2004**). "Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems," Ear Hear. **25**, 375–387.

Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Am. **110**(2), 1150–1163.

Fu, Q. J. (**2002**). "Temporal processing and speech recognition in cochlear implant users," Neuroreport. **13**, 1635–1640.

Goorevich, M., and Batty, M. (**2005**). "A new real-time research platform for the Nucleus® and Nucleus® Freedom™ cochlear implants," in *CIAP 2005*, poster presentation.

Green, T., Faulkner, A., and Rosen, S. (**2012**). "Variations in carrier pulse rate and the perception of amplitude modulation in cochlear implant users," Ear Hear. **33**(2), 221–230.

Harczos, T., Chilian, A., and Husar, P. (**2013**). "Making use of auditory models for better mimicking of normal hearing processes with cochlear implants: The SAM coding strategy," IEEE Trans. Biomed. Circuits Syst. **7**(4), 414–425.

Hey, M., Hocke, T., Hedderich, J., and Müller-Deile, J. (**2014**). "Investigation of a matrix sentence test in noise: Reproducibility and discrimination function in cochlear implant patients," Int. J. Audiol. **53**(12), 895–902.

Hochberg, I., Boothroyd, A., Weiss, M., and Hellman, S. (**1992**). "Effects of noise and noise suppression on speech perception by cochlear implant users," Ear Hear. **13**(4), 263–71.

Hu, Y., Loizou, P. C., Li, N., and Kasturi, K. (**2007**). "Use of a sigmoidal-shaped function for noise attenuation in cochlear implants," J. Acoust. Soc. Am. **122**(4), EL128–EL134.

IEEE Subcommittee on Subjective Measurements. (**1969**). "IEEE recommended practices for speech quality measurements," IEEE Trans. Audio Electroacoust. **17**, 227–246.

James, C. J., Blamey, P. J., Martin, L., Swanson, B., Just, Y., and Macfarlane, D. (**2002**). "Adaptive dynamic range optimization for cochlear implants: A preliminary study," Ear Hear. **23**(1), 49S–58S.

Jorgensen, S., and Dau, T. (**2011**). "Predicting speech intelligibility based on the signal-to-noise envelope power ratio after modulation-frequency selective processing," J. Acoust. Soc. Am. **130**(3), 1475–1487.

Khing, P. P., Swanson, B. A., and Ambikairajah, E. (**2013**). "The effect of automatic gain control structure and release time on cochlear implant speech intelligibility," PLoS One **8**(11), e82263.

Kluender, K. R., Coady, J. A., and Kiefte, M. (**2003**). "Sensitivity to change in perception of speech," Speech Commun. **41**(1), 59–69.

Kollmeier, B., and Wesselkamp, M. (**1997**). "Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment," J. Acoust. Soc. Am. **102**(4), 2412–2421.

Litvak, L. M., Spahr, A. J., Saoji, A. A., and Fridman, G. Y. (**2007**). "Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners," J. Acoust. Soc. Am. **122**(2), 982–991.

Loizou, P. C., Dorman, M. F., Tu, Z., and Fitzke, J. (**2000**). "Recognition of sentences in noise by normal-hearing listeners using simulations of speak-type cochlear implant signal processors," Ann. Otol. Rhinol. Laryngol. Suppl. **185**, 67–68.

Loizou, P. C., and Poroy, O. (**2001**). "Minimum spectral contrast needed for vowel identification by normal hearing and cochlear implant listeners," J. Acoust. Soc. Am. **110**(3), 1619–1627.

Long, C. J., Holden, T. A., McClelland, G. H., Parkinson, W. S., Shelton, C., Kelsall, D. C., and Smith, Z. M. (**2014**). "Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding," J. Assoc. Res. Otolaryngol. **15**(2), 293–304.

Mauger, S. J., Dawson, P. W., and Hersbach, A. A. (**2012**). "Perceptually optimized gain function for cochlear implant signal-to-noise ratio based noise reduction," J. Acoust. Soc. Am. **131**, 327–336.

Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (**2001**). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**(6), 2852–2861.

Moore, B. C. J. (**2003**). "Speech processing for the hearing-impaired: Successes, failures, and implications for speech mechanisms," Speech Commun. **41**, 81–91.

Nelson, D. A., Schmitz, J. L., Donaldson, G. S., Viemeister, N. F., and Javel, E. (**1996**). "Intensity discrimination as a function of stimulus level with electric stimulation," J. Acoust. Soc. Am. **100**(4), 2393–2414.

Nogueira, W., Büchner, A., Lenarz, T., and Edler, B. (**2005**). "A psychoacoustic 'NofM'-type speech coding strategy for cochlear implants," EURASIP J. Appl. Signal Process. **2005**(18), 3044–3059.

Nogueira, W., Kátai, A., Harczos, T., Klefenz, F., Buechner, A., and Edler, B. (**2007**). "An auditory model based strategy for cochlear implants," Conf. Proc. IEEE Eng. Med. Biol. Soc. **2007**, 4127–4130.

Nogueira, W., Lopez, M., Rode, T., Doclo, S., and Buechner, A. (**2015**). "Individualizing a monaural beamformer for cochlear implant users," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 19–24, pp. 5738–5742.

Oxenham, A. J., Simonson, A. M., Turicchia, L., and Sarpeshkar, R. (**2007**). "Evaluation of companding-based spectral enhancement using simulated cochlear-implant processing," J. Acoust. Soc. Am. **121**(3), 1709–1716.

Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (**2006**). "Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise," J. Acoust. Soc. Am. **120**(6), 3988–3997.

Sagi, E., Meyer, T. A., Kaiser, A. R., Teoh, S. W., and Svirsky, M. A. (**2010**). "A mathematical model of vowel identification by users of cochlear implants," J. Acoust. Soc. Am. **127**(2), 1069–1083.

Sarampalis, A., Kalluri, S., Edwards, B., and Hafter, E. (**2009**). "Objective measures of listening effort: Effects of background noise and noise reduction," J. Speech, Lang. Hear. Res. **52**(5), 1230–1240.

Stickney, G. S., Zeng, F. G., Litovsky, R., and Assmann, P. (**2004**). "Cochlear implant speech recognition with speech maskers," J. Acoust. Soc. Am. **116**(2), 1081–1091.

Turicchia, L., and Sarpeshkar, R. (**2005**). "A bio-inspired companding strategy for spectral enhancement," IEEE Trans. Speech Audio Process. **13**(2), 243–253.