

# Development of a loudness normalisation strategy for combined cochlear implant and acoustic stimulation

Tom Francart<sup>a,b,1,\*</sup>, Hugh McDermott<sup>a</sup>

<sup>a</sup>*Bionics Institute, Melbourne, Australia*

<sup>b</sup>*KULeuven, Leuven, Belgium*

---

## Abstract

Users of a cochlear implant together with a hearing aid in the non-implanted ear currently use devices that were developed separately and are often fitted separately. This results in very different growth of loudness with level in the two ears, potentially leading to decreased wearing comfort and suboptimal perception of interaural level differences. A loudness equalisation strategy, named 'SCORE bimodal', is proposed. It equalises loudness growth for the two modalities using existing models of loudness for acoustic and electric stimulation, and is suitable for implementation in wearable devices. Loudness balancing experiments were performed with six bimodal listeners to validate the strategy. In a first set of experiments, the function of the loudness models used was validated by balancing the loudness of four harmonic complexes of different bandwidths, ranging from 200 Hz to 1000 Hz, separately for each ear. Both the electric and acoustic loudness models predicted the data well. In a second set of experiments, binaural balancing was done for the same stimuli. It was found that SCORE significantly improved binaural balance.

---

---

\*Corresponding author

*Email addresses:* tom.francart@med.kuleuven.be (Tom Francart),  
hmcdermott@bionicsinstitute.org (Hugh McDermott)

## 1. Introduction

Many recently implanted users of a cochlear implant (CI) have residual hearing in the non-implanted ear, which is often stimulated using a hearing aid (HA). This configuration is called bimodal stimulation. As there are presently no commercial devices specifically designed for bimodal stimulation, standard CIs and HAs are used together, which leads to several problems. For example, in many cases the two devices are fitted at separate places (e.g., the CI clinic and a HA dispenser). There are currently no generally accepted fitting procedures for bimodal devices. In what follows we discuss the growth of loudness with level for current clinically fitted bimodal devices, and propose the SCORE bimodal strategy, which equalises loudness growth at the two ears using models of loudness perception.

In normal and impaired hearing, an increase in sound pressure level generally leads to an increase in loudness. Similarly for CI stimulation, an increase in current leads to an increase in loudness. The relationship between level and loudness is called the loudness growth function (LGF). LGFs vary with the spectrum of the signal, but also with characteristics of the hearing impairment, and, if CIs or HAs are used, with the signal processing employed in the device. For normal hearing (NH) listeners, LGFs are monotonic and the same for the two ears, leading to consistent changes of perceived interaural level difference with angle of incidence. NH listeners can use this direct relationship to localise sound sources (Moore, 2012) and it can aid in segregation of spatially separated sound sources, which in turn leads to improved perception of speech in noise (Akeroyd, 2006). For hearing impaired (HI) listeners with symmetric hearing losses, LGFs can be similar at the two ears when using the same HAs with similar compression parameters bilaterally. For HI listeners with severely asymmetric hearing losses, of which most bimodal listeners are an extreme example, LGFs cannot easily be equalised across the ears with current commercial HAs and CI sound processors. This is reflected in the performance of bimodal listeners on localisation tasks. Bimodal listeners are sensitive to interaural level difference cues (Francart et al., 2008a), but with commercial devices localisation performance is usually very poor (Ching et al., 2007).

McDermott and Varsavsky (2009) presented an analysis of differences in the LGF between acoustic stimulation with a hearing impairment and electric stimulation, using models of loudness perception. They found that while large parts of the LGFs were similar for a low-frequency noise-band stimulus,

there were large differences between acoustic and electric stimulation for a 1-kHz tone. These differences are related to three main factors: 1) differences in the signal processing in the HA and CI sound processors, 2) suboptimal fitting of the two devices, and 3) the limited frequency range of the residual hearing. We will discuss these three factors in the following paragraphs.

CIs and HAs were designed separately without their combined use in mind. Both device types were designed primarily for optimal speech perception. The HA may contain automatic gain control and fast compression, which change the LGF, but the HA is not usually fitted to normalise loudness, i.e., the loudness evoked by a certain sound is not the same as for a NH listener, or even to equalise LGFs across ears. The CI sound processor also contains automatic gain control, implemented differently and usually with parameter settings very different from those in the HA. It also contains non-linear processing such as maxima selection and instantaneous compression designed for optimal speech perception. This and the very different nature of electrical stimulation leads to a LGF very different from normal and from the one in the acoustically stimulated ear.

While several of the aforementioned issues could be partially resolved by appropriately fitting the two devices to work better together, this is currently not feasible because 1) the intrinsic operation of the compression systems in the two devices is often different, which makes it hard to achieve identical function by tuning parameters, 2) devices are often fitted by different clinicians or even at different places, which makes it hard to coordinate the fitting, 3) there are currently no well-established and verified fitting methods for bimodal stimulation, and 4) equalising loudness growth across ears would require extensive loudness balancing across frequencies and levels, which would be very time-consuming, and impossible to conduct for a large part of the clinical population. If any specifically bimodal fitting method is used in the clinic, it is usually the method proposed by Ching et al. (2004), which involves adjusting the level of the acoustic stimulus in order to loudness-balance running speech stimuli across the ears, either at one intensity (for linear hearing aids) or two (for compression hearing aids). Unfortunately this does not solve the inherent problems with current commercial devices.

Another aspect of the problem with differences in loudness growth is due to the nature of the residual hearing: often there is only residual hearing at low frequencies (e.g., up to 1 or 2 kHz), and a dead region (Moore, 2004) at higher frequencies. This means that broadband sounds will not be balanced if their high-frequency part falls in the dead region and can therefore not

be perceived acoustically. Consider, for instance, the influence of bandwidth on loudness. For NH listeners, for a signal with fixed overall level, loudness increases with increasing bandwidth (Fastl and Zwicker, 2007). However, once the highest frequencies in the signal exceed the edge frequency of a dead region in impaired ears, loudness will decrease. Additionally, signals with only frequency content beyond the edge frequency of the dead region will not elicit any loudness at all.

In the current article we describe the development of a sound processing strategy designed to normalise loudness perception for electric and acoustic stimulation. The strategy is called SCORE bimodal (Stimulus Control to Optimise Recipient Experience). We build on previous theoretical work (McDermott et al., 2003) and on a loudness-normalisation strategy designed exclusively for electric stimulation (Varsavsky and McDermott, 2012). The proposed strategy consists of add-on processing to current systems: it can be inserted into the processing of current commercial HAs and CI sound processors with only minor modifications to other components. SCORE bimodal uses loudness models at three different stages to normalise loudness: it estimates (1) the loudness of the signal received by the microphones of the two devices for a NH listener, using a loudness model for NH (Moore and Glasberg, 1996), (2) the loudness of the signal processed by the HA using a loudness model for HI (Moore and Glasberg, 1997), and (3) the loudness of the electrical stimulation pattern generated by the CI sound processor (McKay et al., 2003). The overall output level of the signals from the two devices is adjusted to match the predicted CI and HI loudness to the predicted NH loudness. By normalising loudness, the LGFs at the two sides become similar, the HA is inherently fitted, and broadband perceived interaural level cues are made more reliable. Note that if normal loudness is not considered the ideal or desired target, it can easily be transformed to any desired loudness function.

We evaluated the functionality of the SCORE bimodal strategy by having six bimodal listeners adjust the level of stimuli of different bandwidths so they sounded equally loud. In experiment 1, we validated the acoustic and electric loudness models separately by performing monaural loudness balancing and comparing the results to model predictions. In experiment 2, we assessed if SCORE improved binaural loudness balance by performing binaural loudness balancing experiments, both for the standard clinically used processing (ACE and a linear HA fitted according to the NAL-RP rule), and with application of the SCORE bimodal processing. We started from the existing clinical fitting

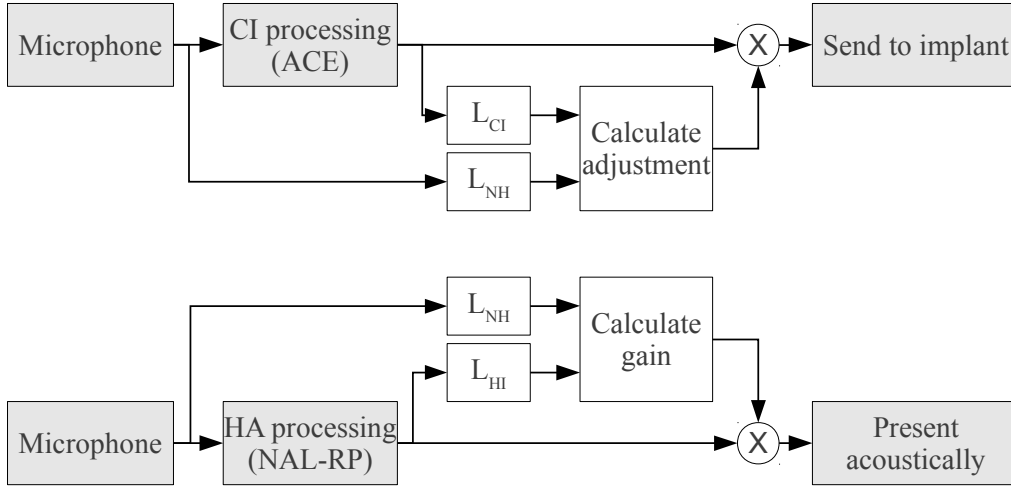


Figure 1: Block diagram of SCORE bimodal processing. Grey blocks indicate the existing clinical processing, and white blocks indicate the add-on SCORE processing.  $L_{NH}$ ,  $L_{HI}$ , and  $L_{CI}$  indicate respectively loudness models for normal hearing, impaired hearing and cochlear implant (electrical) stimulation.

for all subjects and minimised the time spent fitting the loudness models, as to reflect results that would be obtained in a typical clinical setting.

## 2. The SCORE bimodal strategy: signal processing

### 2.1. Overview

We used two signal processing schemes. The first one, termed ACE bimodal, consisted of the standard clinical Advanced Combination Encoder (ACE) processing combined with linear HA processing, fitted according to the NAL-RP rule (Byrne et al., 1990). The second scheme, termed SCORE bimodal, consisted of ACE bimodal plus extra loudness model-based processing. In the block diagram of figure 1, the standard ACE bimodal processing is indicated by the grey blocks. In the next sections, the ACE and SCORE strategies are described in more detail. For brevity, we refer to the strategies as ACE and SCORE, dropping the “bimodal” suffix.

All processing was performed on frames of samples with an overall duration of 6.9 ms, at a sample rate of 16 kHz. Each frame was processed by each of the signal processing blocks in the diagram. While we had the entire signal available in our Matlab-based implementation, no “future” blocks were used, to enable implementation of the strategy on a real-time system.

## 2.2. ACE

For current purposes, the sound processing strategy can be considered as a way to convert an acoustic signal into a pattern of electrical pulses. The operation of SCORE does not depend on the way in which this is done. The ACE strategy (Vandali et al., 2000; McDermott et al., 1992) was selected because it is the one the present subjects were familiar with, and it is used in Cochlear’s commercial sound processor for the majority of CI recipients. Other sound processing strategies could be used. The implementation of the ACE strategy was provided by Cochlear in the form of the Nucleus Matlab Toolbox version 4.31. When a CI using the ACE strategy is fitted to a recipient, threshold (T) and comfortable (C) levels are determined for each electrode using pulse trains at the channel stimulation rate (typically 900 pps). These and other parameters relevant to the fitting are stored in a MAP. The current level of a pulse is expressed in current units (CU). The current level of each pulse always lies between the T and C levels, and is determined based on channel magnitudes according to  $CL = M * (C - T) + T$ , with CL the number of current units (CU), and M the magnitude between 0 and 1<sup>1</sup>. The difference between C and T levels will be called the dynamic range, and can be different for each individual electrode. The microphone response of the Freedom sound processor was measured using an artificial head and applied to our stimuli before the ACE processing.

## 2.3. HA

A linear HA was simulated using a 512-coefficient finite impulse response filter, designed for a sampling rate of 16 kHz. The desired aided thresholds were calculated from the unaided audiogram according to the NAL-RP rule (Byrne et al., 1990). Hearing thresholds were measured using the experimental computer and insert earphones, and the finite impulse response filter was designed to produce the required aided thresholds (Dillon, 2001, chapter 4). The result was verified by measuring aided thresholds using the simulated HA and the insert earphones.

---

<sup>1</sup>The magnitude in each channel is obtained by transforming the level in each channel of the filter bank after envelope detection by an instantaneous compression function. Swanson (2008) provides a good description of the whole process.

#### 2.4. SCORE

The loudness models used in the current study are described in detail by McDermott and Varsavsky (2009). They are briefly summarised below.

The loudness as perceived by a NH listener was calculated using the loudness model described by Moore and Glasberg (1997), but we replaced the middle-ear transfer function by the one proposed by Moore and Glasberg (2004). For each time frame the model calculates the short-term spectrum, filters it by the middle-ear transfer function, calculates the excitation pattern, transforms it to specific loudness, and sums the specific loudness across frequencies to obtain the total loudness in sones. While updated versions of this loudness model exist (Moore and Glasberg, 2004), we opted to use the older version because of its lower computational complexity, which makes it feasible to implement in current CI sound processors and HAs, and the lower number of parameters to be considered. Moreover, differences between the older and newer version of this model are largest at levels near threshold, which is less important for the current application.

Similarly, the loudness as perceived through the subject’s hearing-impaired ear was estimated using the loudness model described by Moore and Glasberg (1997). While this model uses a large number of parameters, it can be fitted based solely on the unaided audiogram, using default values for all other parameters. While potentially better performance could be achieved by individually measuring all other parameters in psychophysical experiments, this would not be feasible in a clinical context because of the time requirements.

In the “Calculate gain” block in figure 1, the target (NH) loudness was available ( $L_{NH}$ ), as well as the loudness perceived by the HI listener of the signal processed through the linear HA ( $L_{HI}$ ). A gain value was calculated for each frame such that the loudness of the amplified/attenuated acoustic signal would be equal to the target loudness. A formula similar to those given by Launer and Moore (2003) was used, with the main difference that the total loudness was modified instead of the specific loudness. The detailed algorithm is described in appendix A.

To estimate the loudness of the electrical stimulation pattern generated by ACE, a simplified version of the model developed by McKay et al. (2003) was used. In this model, the current level of each pulse is converted to a loudness contribution using an electrode-specific LGF. Then all loudness contributions within a time frame are added. In the published model, LGFs are assumed to consist of a combination of a linear and exponential part of the form  $\log(L) = a * CL + 0.03 * b * e^{(CL-C_0)/b} + k$ , with  $L$  the loudness

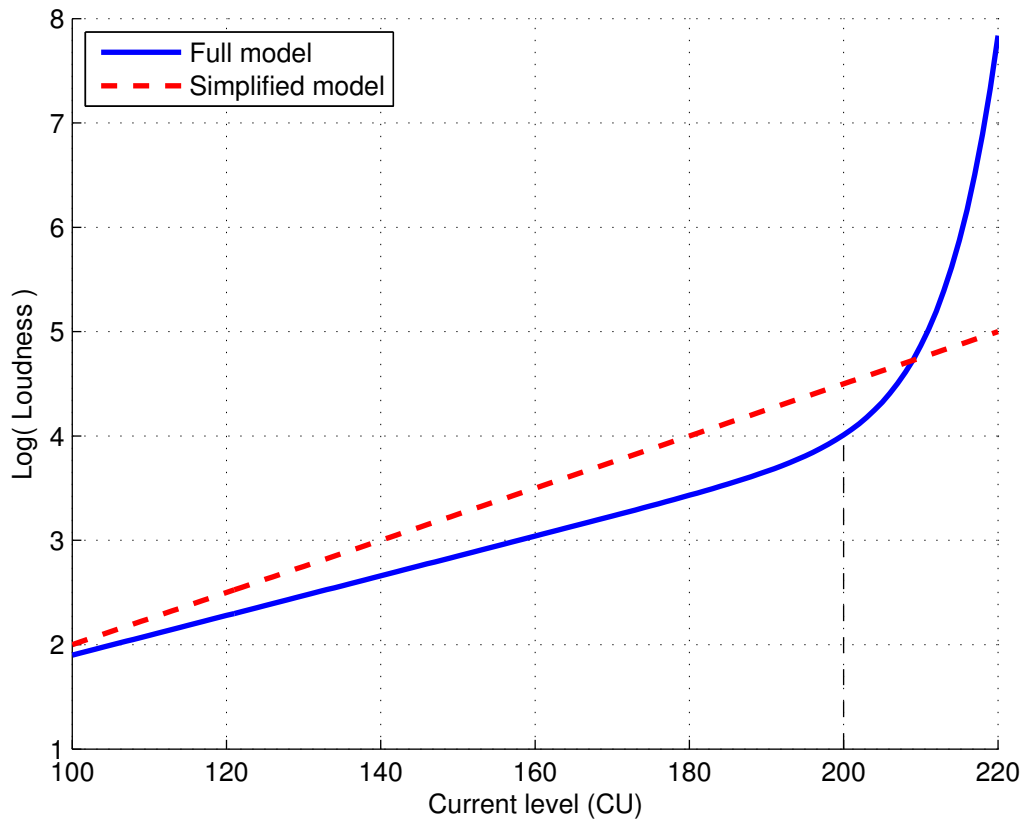


Figure 2: Example LGF for a single pulse. The full line was plotted according to McKay et al. (2003), with  $a = 0.019$ ,  $c_0 = 200$ ,  $b = 0.21$ , and  $k = 0$ . The dashed line was plotted according to the simplified model with  $a = 0.025$ .



contribution of a pulse, CL the current level,  $a$  the slope of the linear part,  $b$  determining the shape of the exponential part,  $c_0$  the knee point where the linear part changes into the exponential part, and  $k$  a scaling factor. An example LGF is shown in figure 2, for  $a = 0.019$ ,  $c_0 = 200$ ,  $b = 0.21$ , and  $k = 0$ . The slope  $a$  is usually about 0.019, and  $c_0$  generally corresponds to the C level of a 500 pps pulse train (McKay et al., 2003). Determining the full LGF for one electrode is not feasible in a clinical setting, let alone for each electrode. We assumed that LGFs had the same shape on all electrodes and that operation was mainly in the linear region, and thus simplified the LGF to  $\log(L) = a * CL + k$ . As the current required for a certain loudness percept decreases with increasing stimulation rate, and given that clinical strategies usually operate at higher rates than 500 pps (usually at 900 pps), we expect this assumption to hold for most subjects. The slope  $a$  was determined in a simple loudness balancing task using broadband signals (see below), which is feasible in a clinical setting, and the scaling factor  $k$  was used to obtain loudness values that were similar to the units (sones) provided by the acoustical model. More details are provided in section 5.2.

In the “Calculate adjustment” block, the target (NH) loudness is available ( $L_{NH}$ ), as well as the loudness of the signal processed through ACE ( $L_{CI}$ ). The current-level adjustment  $A$  is calculated such that the loudness of the adjusted electric signal is equal to the target loudness:  $A = \frac{\log(L_{NH}/L_{EL})}{a}$ .  $A$  is expressed in terms of percentage of the dynamic range (%DR) for each electrode. For instance, if the dynamic ranges for electrodes 5 and 10 were respectively 20 CU and 30 CU, and an adjustment of  $A = 10$  %DR was required, the current level of a pulse destined for these electrodes would be increased by 2 and 3 CU. Modified levels were always limited to the C level for each electrode.

The “Calculate gain” and “Calculate adjustment” blocks yield gains and adjustments to be applied to a specific frame. We assume that by normalising the instantaneous loudness in this way, we would also normalise long-term loudness (Glasberg and Moore, 2002). To avoid audible artefacts in the signal, the gains (in dB) and adjustments (in %DR) were smoothed according to a set of heuristics and automatic-gain-control-like processing with attack and release times of 5 ms and 50 ms respectively. The heuristics included a rule to not apply any adjustment if the loudness estimated by any of the models (NH, HI, and CI) was below a threshold value, a rule to limit the maximal change in adjustment between subsequent frames, and a rule to limit the maximal adjustment. By using the same type of smoothing for the

Subject	Age (y)	CI use (months)	CI side	Aetiology	$a$	$k$
S26	75	17	R	Unknown	0.025	-4.93
S27	78	25	L	Unknown	0.010	-2.37
S30	60	28	R	Unknown	0.021	-4.25
S31	62	17	R	Unknown	0.034	-6.67
S32	54	84	R	Menieres	0.017	-3.57
S34	53	28	R	Unknown	0.030	-6.16

Table 1: Subject information: “Age” is in years at the time of testing. “CI use” is the number of months of implant use at the time of testing. “CI side” is left (L) or right (R); the HA was on the other side.  $a$  and  $k$  are the parameter values used for the electrical loudness model.

two devices, we intended to keep the LGFs largely the same.

### 3. General methods

#### 3.1. Subjects

Six subjects were recruited who used a CI in one ear and a HA in the other ear on a daily basis. The unaided pure-tone audiograms of their non-implanted ears are shown in figure 3. Relevant subject information is listed in table 1. Permission to conduct the studies was obtained from the local hospital ethical committee, and each participant provided informed consent.

#### 3.2. Apparatus

The subjects’ own HAs or sound processors were not used. All stimuli were presented under direct computer control using the APEX 3 program developed at ExpORL, KU Leuven (Francart et al., 2008b). APEX 3 controlled an RME Fireface 400 sound card (Audio AG, Haimhausen, Germany) for acoustic stimulation, and the Cochlear NICv2 interface connected to an L34 experimental processor for electric stimulation. The L34 was set up to start stimulating when a trigger pulse was received from the sound card, and the clocks of the sound card and L34 were synchronised. In this way stimulation for the HA and CI was synchronised with uncertainty smaller than 200 ns. The electric stimulus was delayed by 1.5 ms relative to the acoustic stimulus to compensate for the estimated acoustic travelling-wave delay (Francart et al., 2009). For acoustic stimulation, an Etymotic ER-3A insert earphone connected to the sound card was used. The insert phone was

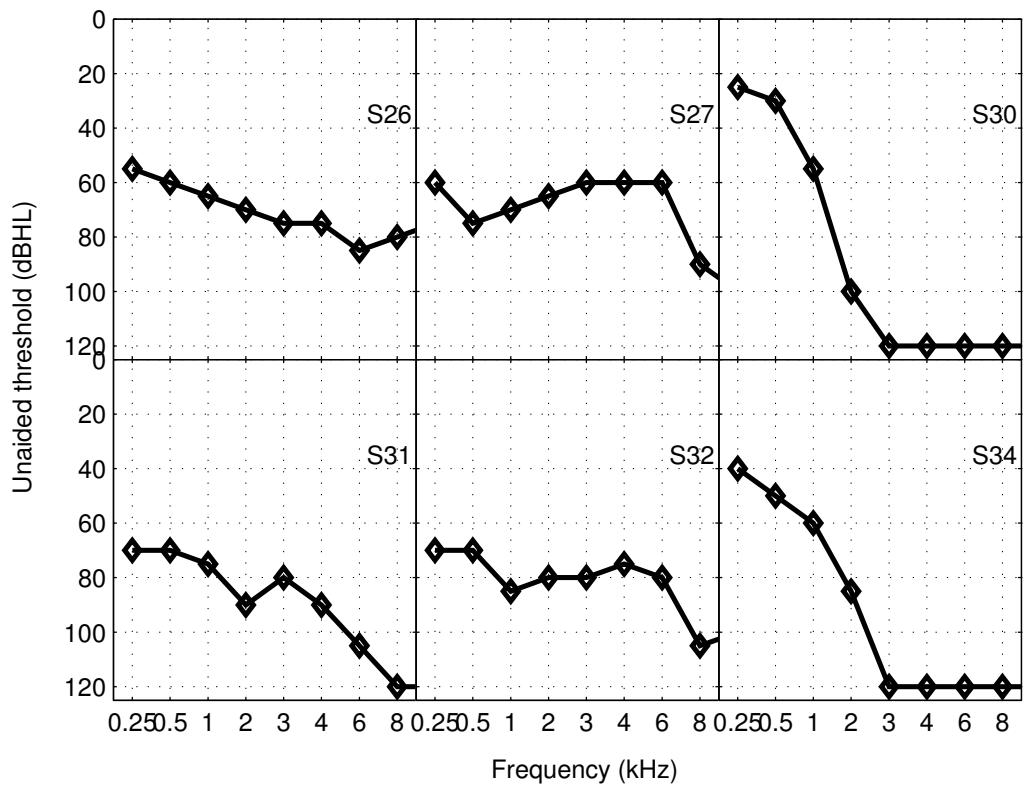


Figure 3: Pure tone unaided audiograms of the non-implanted ear for each subject.

calibrated using a HA-1 2cc coupler. Its frequency response was measured, and inverse filtering at the end of the acoustic processing chain was applied in order to compensate for its non-flat response. All signal processing and stimulus preparation was done using Matlab (The Mathworks, Natick, MA, USA).

### 3.3. Stimuli

The stimuli were harmonic complexes with a fundamental frequency of 200 Hz and a varying number of equal amplitude harmonics. In the following, they are labelled in terms of the frequency of their highest harmonic: S400, S600, S800, and S1200 consisted respectively of 2, 3, 4, and 6 harmonics. The levels of the four stimuli were set to produce equal loudness for a NH listener, using the acoustic loudness model described above (Moore and Glasberg, 1997). To investigate overall level effects experimentally, the loudness of each stimulus was set respectively to 8 and 16 sones. All loudnesses are expressed in terms of binaural stimulation, assuming that switching from monaural to binaural stimulation results in a doubling in loudness. This means that 8 sones corresponded to a fairly soft level, and 16 sones to a comfortable level, similar to that of conversational speech. The resulting spectra are shown in figure 4. All stimuli had a duration of 1 s and had 50-ms sinusoidal ramps to avoid audible clicks at the onset and offset.

These eight stimuli were processed through ACE. This resulted in activation of a different number of electrodes for each stimulus at a certain NH equivalent loudness. Figure 5 illustrates the distribution of current across electrodes. In one of the last stages of ACE processing, magnitudes between 0 and 1 are converted to current levels between T and C level. A magnitude of 0 is mapped to the T level and a magnitude of 1 is mapped to the C level. The cumulative magnitude shown in figure 5 is the sum of all magnitudes on an electrode during one second. This type of plot can be compared to the spectrum of an acoustic signal. It is clear that with increasing maximal stimulus frequency, the number of activated electrodes increases, and that the distribution of current across electrodes is not flat. Also, the cumulative magnitude is higher for the 16-sones stimuli than for the 8-sones ones.

### 3.4. Fitting

For the CI, the subjects' clinical fitting was used, and the volume was set to a level that was judged as comfortable, yielding a loudness similar to that obtained with their own sound processor for speech at 60 dB A. For

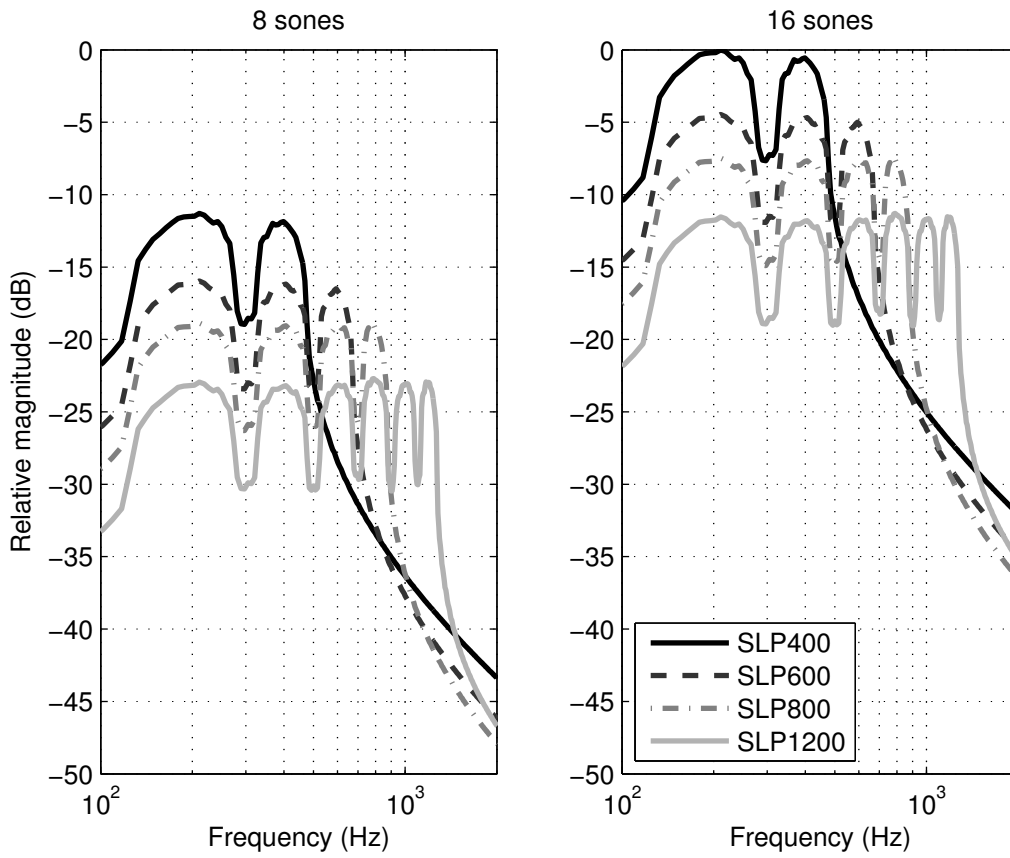


Figure 4: Spectra of the acoustic stimuli at the input of the CI and HA processing. The acoustic stimuli were set at equal NH loudness.

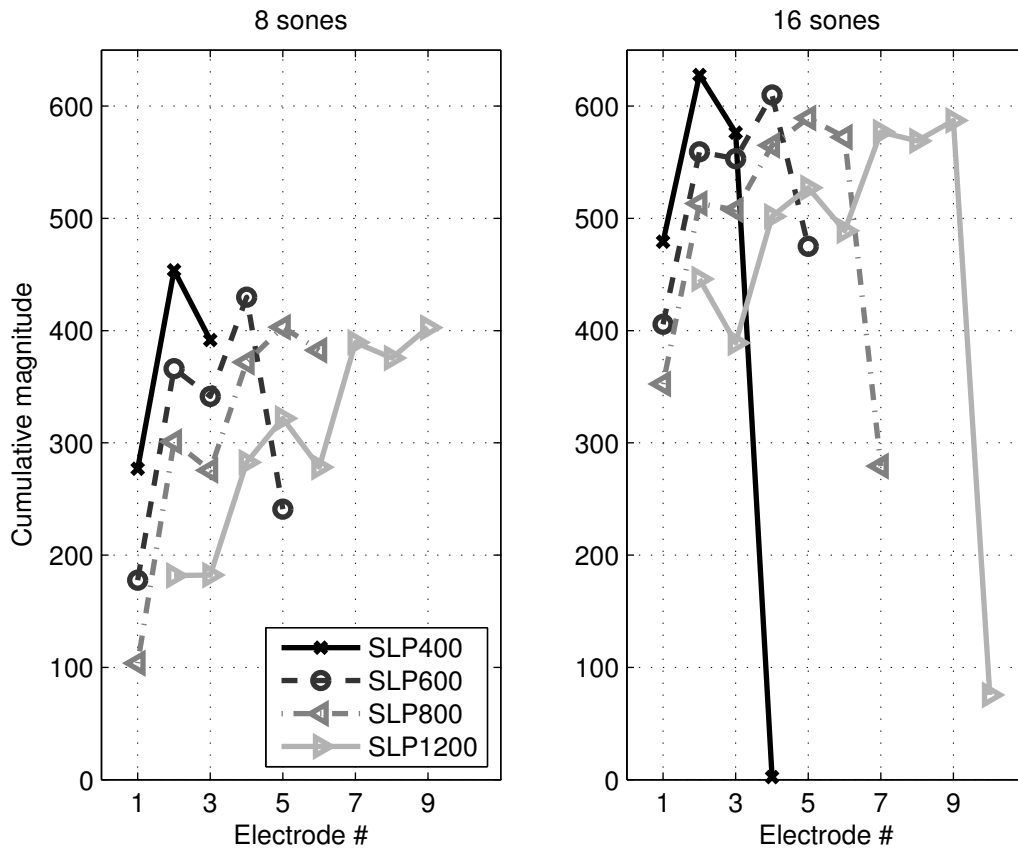


Figure 5: Cumulative magnitudes of the four different stimuli for a loudness of 8 or 16 sones. Electrodes are numbered from apex to base. The cumulative magnitude is the sum of all magnitudes on a single electrode during one second.

all subjects except S26, a channel stimulation rate of 900 Hz was used, with monopolar (MP1+2) stimulation, an inter-phase gap of 8  $\mu\text{s}$ , a phase width of 25  $\mu\text{s}$ , and 8 maxima. For S26 the phase width was 75  $\mu\text{s}$ , and therefore only 6 maxima were selected, as a total pulse width of 158  $\mu\text{s}$  did not allow for 8 maxima at a channel stimulation rate of 900 Hz.

The overall gain of the HA was adjusted such that a noise shaped like the long-term average speech spectrum according to ANSI-3.5 and a sentence uttered by a male and female speaker, all at 60 dB A, yielded a percept that was balanced in loudness with the corresponding electric stimulus. Usually the same overall HA gain was found for all three stimuli. If this was not the case, the average gain across these three stimuli was used. This was done separately for ACE and SCORE.

## 4. Experiment 1 - Monaural balancing

### 4.1. Methods

To validate the acoustic and electric loudness models, we performed monaural loudness balancing experiments, separately for the acoustically and electrically stimulated ear, and compared the results to model predictions. The procedures used were identical for the two cases. Two stimuli were balanced against each other by adjusting the level of one of them until they sounded equally loud. In the acoustic case this was done by applying an adjustable overall gain to the comparison stimulus. In the electric case this was done by adding a proportion of the dynamic range to each pulse.

The S1200 stimulus was balanced in loudness against the other three stimuli using an adaptive procedure. In each trial the subject heard two intervals (the reference and the comparison stimulus in random order) and had to indicate which one was louder, the first or second. The level of the comparison stimulus was adjusted in a 2-interval, 2-alternative forced-choice adaptive 1 up/1 down procedure. Initially, the step size was 6 dB / 20 %DR, which was reduced to 3 dB / 10 %DR after 2 reversals. The procedure was ended after 10 reversals. The result of an adaptive run was calculated as the mean of the parameter value at the last 6 reversals.

First the adaptive procedure was run twice with S1200 as the reference and one of the other three stimuli as the comparison stimulus, with starting levels alternately above and below the expected outcome. Then the reference and comparison stimuli were swapped and another two adaptive runs were completed, with the level of the former comparison stimulus set to the average

level obtained from the previous two balancing runs. The values reported below are averages of the four runs conducted for each stimulus pair.

Here is an example of acoustic balancing between S1200 and S400:

- Reference S1200, comparison S400. Start level +6 dB adjustment of the comparison stimulus. Example result of adaptive run (average of last 6 reversals): +2 dB adjustment.
- Reference S1200, comparison S400. Start level -4 dB adjustment. Result: +1.5 dB.
- Reference S400 (adjusted by  $(2 + 1.5)/2 = 1.75$  dB), comparison S1200. Start level +6 dB adjustment. Result: -1 dB.
- Reference S400 (adjusted by 1.75 dB), comparison S1200. Start level -7 dB adjustment. Result: -3 dB.
- The final result is calculated as  $(2 + 1.5 + (1.75 - (-1)) + (1.75 - (-3)))/4 = 2.75$  dB adjustment.

#### 4.2. Results

The monaural balancing results and model predictions are shown in figure 6 for the electric model and in figure 7 for the acoustic model. The black bars indicate the adjustment needed to make the stimulus indicated on the horizontal axis equal in loudness to stimulus S1200. The white bars indicate predictions made using the loudness model. The closer the white bar is to the corresponding black one, the better the model prediction. As the input stimuli were set at equal loudness for a NH listener (according to the acoustic loudness model for NH), the psychophysical results can be interpreted as loudness errors due to a combination of signal processing in the CI and hearing impairment of the subject; for a NH listener, we would expect adjustments close to 0 dB for all stimuli. If subtracting the model prediction from the measurement result brings it closer to 0, we would expect the proposed SCORE processing to bring loudness relations between stimuli closer to normal. Note, however, that if the model prediction was inaccurate for S1200, this would result in a vertical shift of all adjustment predictions. Assuming that this did not occur, then, for the electric model, application of SCORE would improve loudness relations for all subjects and conditions except for S26 (all stimuli at 8 and 16 sones) and S32 (all stimuli at 8 sones).



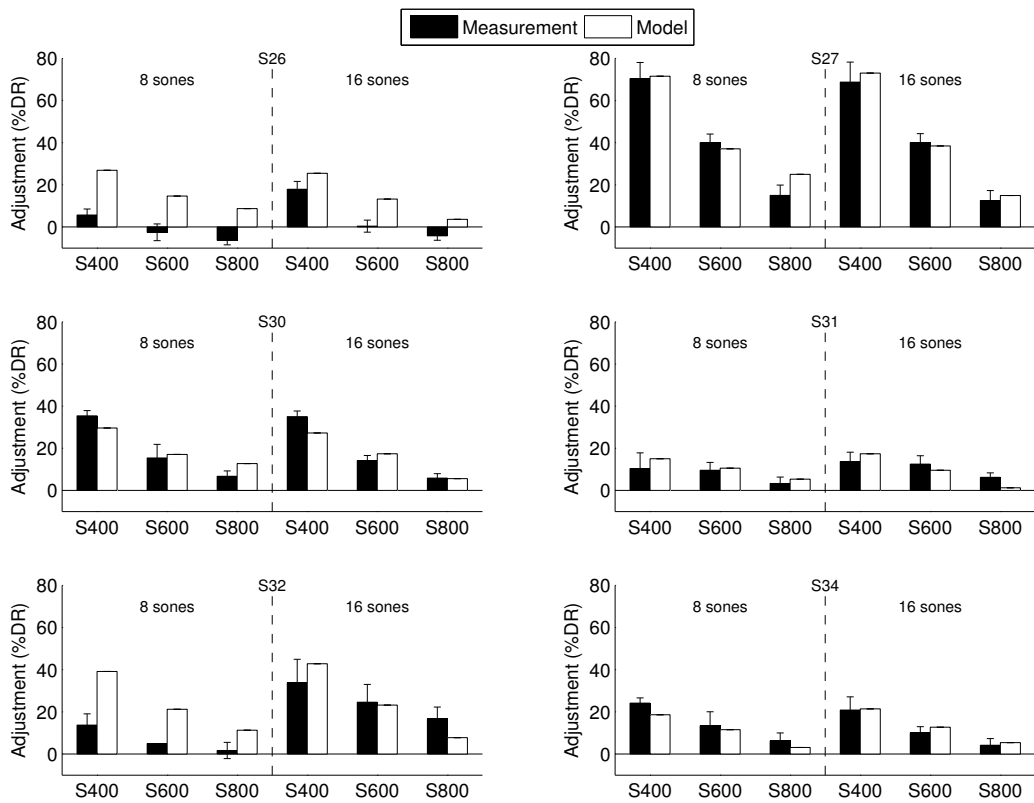


Figure 6: Monaural balancing results for electric stimulation, for each subject. The black bars indicate the adjustment needed to make the stimulus indicated on the horizontal axis equally loud as stimulus S1200. The white bars indicate predictions made by the loudness model. The error bars indicate one standard deviation.

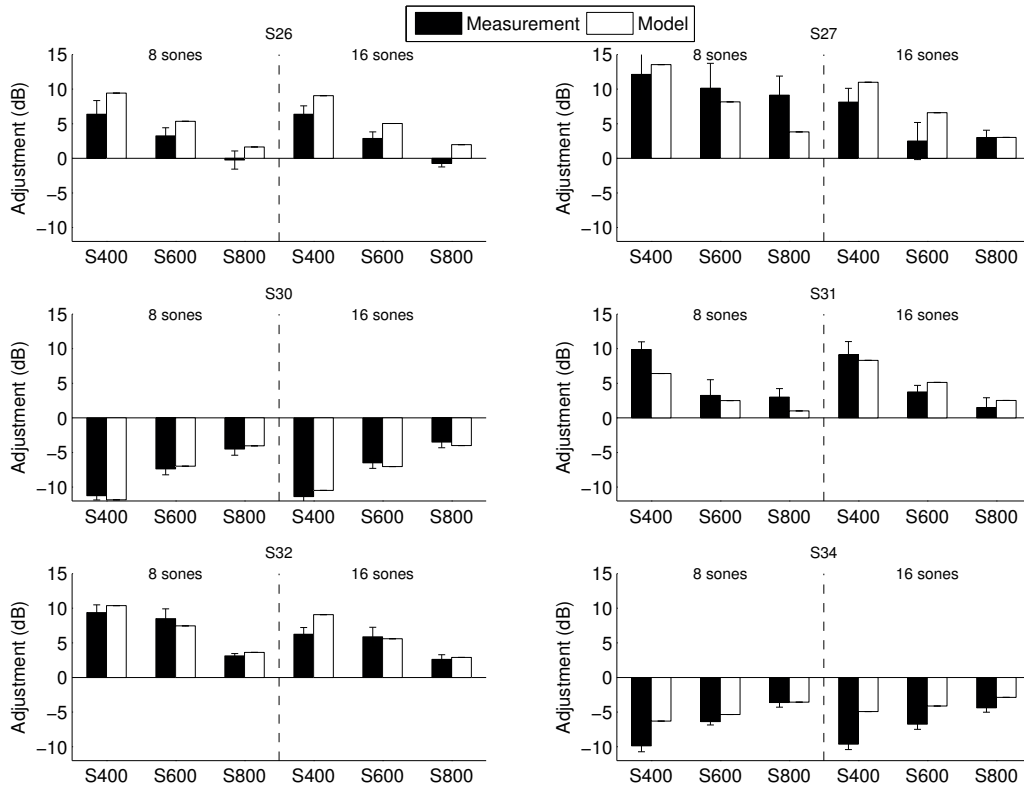


Figure 7: Monaural balancing results for acoustic stimulation. The black bars indicate the adjustment needed to make the stimulus indicated on the horizontal axis equally loud as stimulus S1200. The white bars indicate predictions made by the loudness model. The error bars indicate one standard deviation.

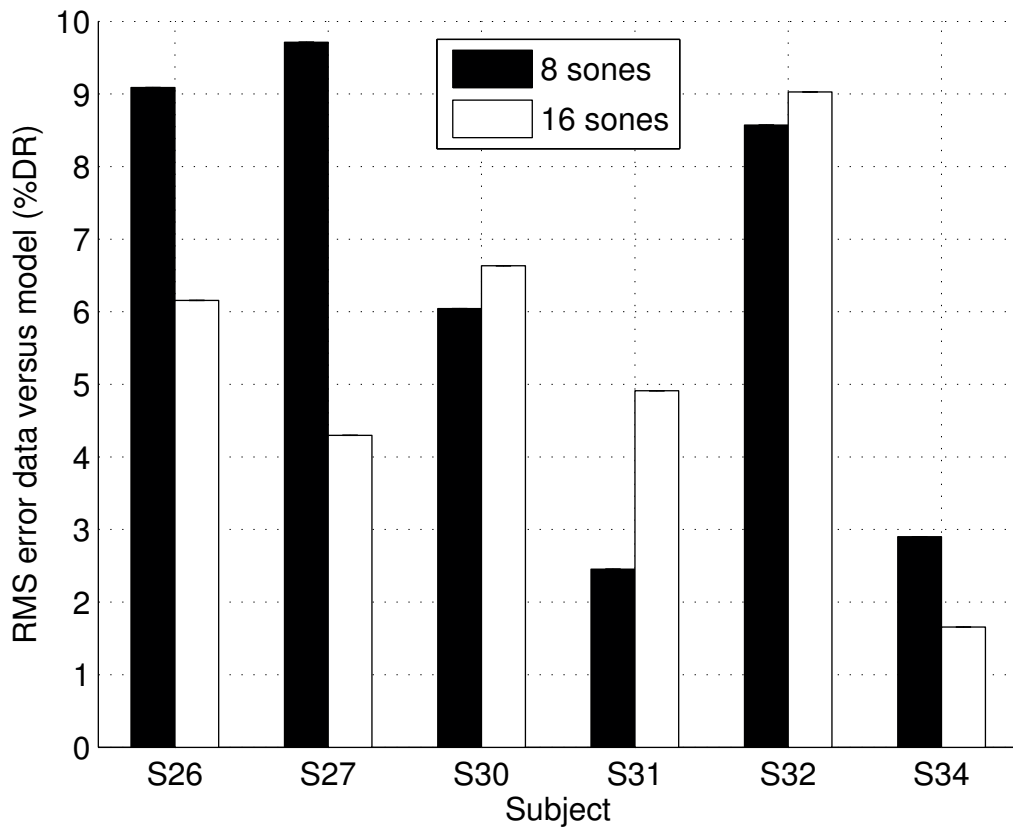


Figure 8: RMS error of the difference in adjustment between adjacent stimuli and the corresponding model predictions, for monaural electric stimulation. Smaller values indicate better performance of the model.

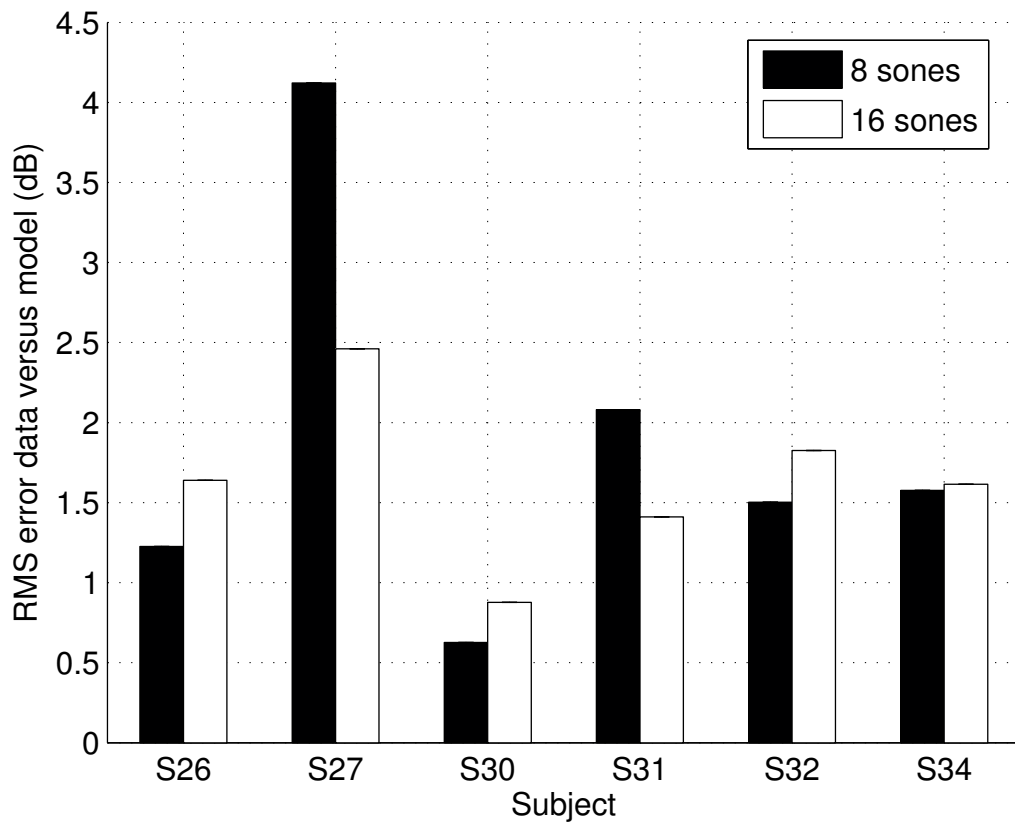


Figure 9: RMS error of the difference in adjustment between adjacent stimuli and the corresponding model predictions, for monaural acoustic stimulation. Smaller values indicate better performance of the model.

An ANOVA of adjustment with factors subject, stimulus type, model or measurement (MM), and target loudness indicated a significant effect of subject ( $F(5, 57) = 28.7, p < 0.01$ ) and of stimulus ( $F(2, 57) = 17.5, p < 0.01$ ), but no effect of MM ( $p = 0.25$ ) or of any other factors or interactions. A similar analysis for the acoustic model indicates that SCORE would improve loudness relations in all conditions, except stimulus S800 at 16 sones for S26, and stimulus S600 at 16 sones for S27. A similar ANOVA indicated an effect of subject ( $F(5, 57) = 48.8, p < 0.01$ ), but no effect of any other factors or interactions.

In this task, no absolute loudness estimates were obtained from the subjects, so an incorrect estimate of the loudness of S1200 would result in a vertical shift of the predicted adjustments for all three stimulus pairs. We therefore considered the difference in adjustment between stimulus pairs, to evaluate whether the model correctly predicted changes in loudness with changes in bandwidth. We calculated the RMS error of (1) the difference between balancing adjustments for adjacent members of our set of stimulus bandwidths, and (2) the difference between the corresponding model predictions. The RMS error rather than the absolute error was used in order to emphasise large errors. RMS errors for each subject for the two target loudnesses for the electric model are shown in figure 8. Consider, for example, the results for S26 at 8 sones. The balancing adjustments for [S400 S600 S800 S1200] were [5.7 -2.5 -6.3 0] %DR (note that the value of 0%DR for S1200 was not measured, it is clear that an adjustment of 0%DR is required to make a stimulus equally loud to itself), and the corresponding differences [-8.2 -3.8 6.3] %DR. The adjustments predicted by the model were [27 15 8.7] %DR, with corresponding differences of [-12 -5.9 -8.7] %DR. The RMS difference between [-8.2 -3.8 6.3] and [-12 -5.9 -8.7] is 9.1 %DR. The RMS errors for 8 and 16 sones were compared across all subjects using a Wilcoxon signed rank test, which indicated no statistically significant difference ( $p = 0.69$ ). This indicates that the model prediction errors were not significantly different between the two target loudnesses. RMS errors ranged from 2 to 10 %DR with a mean of 5 %DR.

Figure 9 shows the same analysis for the acoustic model. A Wilcoxon signed rank test indicated no significant difference in RMS error between the two target loudnesses ( $p = 1.00$ ). RMS errors ranged from 0.6 to 4.1 dB with a mean of 1.7 dB across subjects and loudnesses.

## 5. Experiment 2 - Binaural balancing

### 5.1. Methods

To assess if SCORE improved binaural balance we conducted binaural balancing experiments for the standard ACE processing and for SCORE, and compared the results. In contrast to the monaural experiments, the electric and acoustic stimuli were presented simultaneously. The stimulus prior to processing was the same for the two ears, and therefore should ideally yield a balanced percept, like for a NH listener. We measured the acoustic gain adjustment needed for a balanced percept and compared these values between the ACE and SCORE conditions.

In a preliminary experiment, it was determined whether the binaural stimulus yielded a fused sound image, which was described to the subjects as “a single sound image, as opposed to two different sound images”. If the image was fused, the subject was instructed to indicate whether the stimulus sounded as if it were located to the left, right of the middle of the head. If not fused, the subject was instructed to indicate whether the stimulus was louder on the left or right side. There was complete fusion for S27 and S32, fusion for most stimuli for S26, and no fusion for S30, S31, and S34. The last three subjects were the ones with the worst unaided audiograms at 1000 Hz and beyond.

For the main binaural balancing experiment, an interleaved single-interval, 2-alternative forced-choice 1up/2down and 2up/1down adaptive procedure (Jesteadt, 1980) was used to determine the acoustic gain for which the binaural stimulus was balanced. In each trial the subject heard a single stimulus and had to answer “left” or “right”. The adaptive procedure adjusted the gain of the acoustic stimulus. Initially the step size was 6 dB, which was decreased to 3 dB after two reversals. The procedure was stopped after 6 reversals of each of the two interleaved tracks. The gain for which the stimulus was balanced was calculated by averaging the mean of the last four reversals from each of the two interleaved tracks.

The interleaved adaptive procedure required two starting levels. These were obtained by first running a short 1up/1down adaptive procedure which ended after four reversals, and calculating the starting levels for the interleaved tracks by adding and subtracting 6 dB to the mean of its last 2 reversals.

Binaural balancing was done for each of the four stimuli, at two loudnesses (8 and 16 sones) and in two processing conditions (ACE and SCORE). In

the SCORE condition, two runs of the interleaved adaptive procedure were completed for each stimulus and the results were averaged. To reduce testing time, usually only one run of the interleaved adaptive procedure was done in the ACE condition for each stimulus, as the main outcome measure of the current study was the balancing results with application of SCORE.

### 5.2. Model fitting

In the binaural balancing experiment, the SCORE processing was applied, which required setting some model parameter values. While the acoustic loudness model for HI listeners only required the subject’s audiogram, for the electric model two parameters needed to be set: the slope of the loudness growth function  $a$  and  $k$ .  $a$  and  $k$  were determined numerically, assuming that stimulus S1200 had a loudness of 16 sones and minimising the RMS error between the psychophysical monaural balancing results at 16 sones and the model predictions. Note that 16 sones corresponds to a comfortable loudness, which was the target of the initial volume setting of the sound processor (see section 3.4).

$k$  mainly serves to set the units of the electric loudness model to sones, thus making the results comparable to those of the acoustic loudness model.  $a$  is a scaling factor of the adjustment. If the clinical fitting were ideal, i.e., T and C levels corresponded to exactly the same loudness in all subjects, and all current levels fell within the linear region of the LGF,  $a$  would be the same for all subjects. The  $k$  and  $a$  values that were used for each subject are shown in table 1.

### 5.3. Results

The binaural balancing results are shown in figure 10. The bars indicate the adjustments needed for the sound to be perceived either in the middle of the head or equally loud in the two ears, with and without application of SCORE. The average adjustment across all stimuli and levels was subtracted from the results as it corresponds to a change of the “volume” control of the HA. As the stimulus at the input of the processing was identical at the left and right sides, the adjustment would ideally be 0 dB. Deviations from 0 dB indicate distorted interaural loudness relations. For ACE there were large errors (adjustments), up to  $\pm 10$  dB, which is a large proportion of most subjects’ acoustic dynamic ranges. For SCORE the remaining errors were usually much smaller. We performed a repeated-measures ANOVA of the absolute value of the adjustment with factors stimulus type, processing

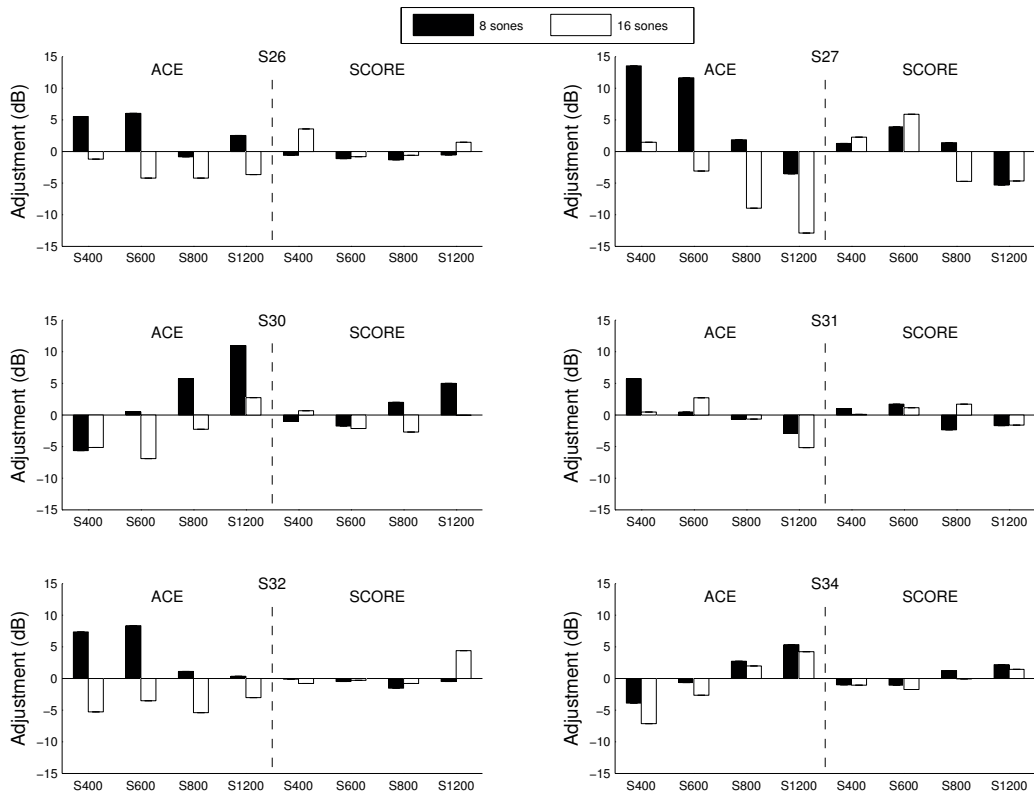


Figure 10: Binaural balancing adjustments for each subject, at 8 and 16 sones, for processing conditions ACE and SCORE. The bars indicate the adjustment of the acoustic side to binaurally balance the stimulus indicated on the horizontal axis.



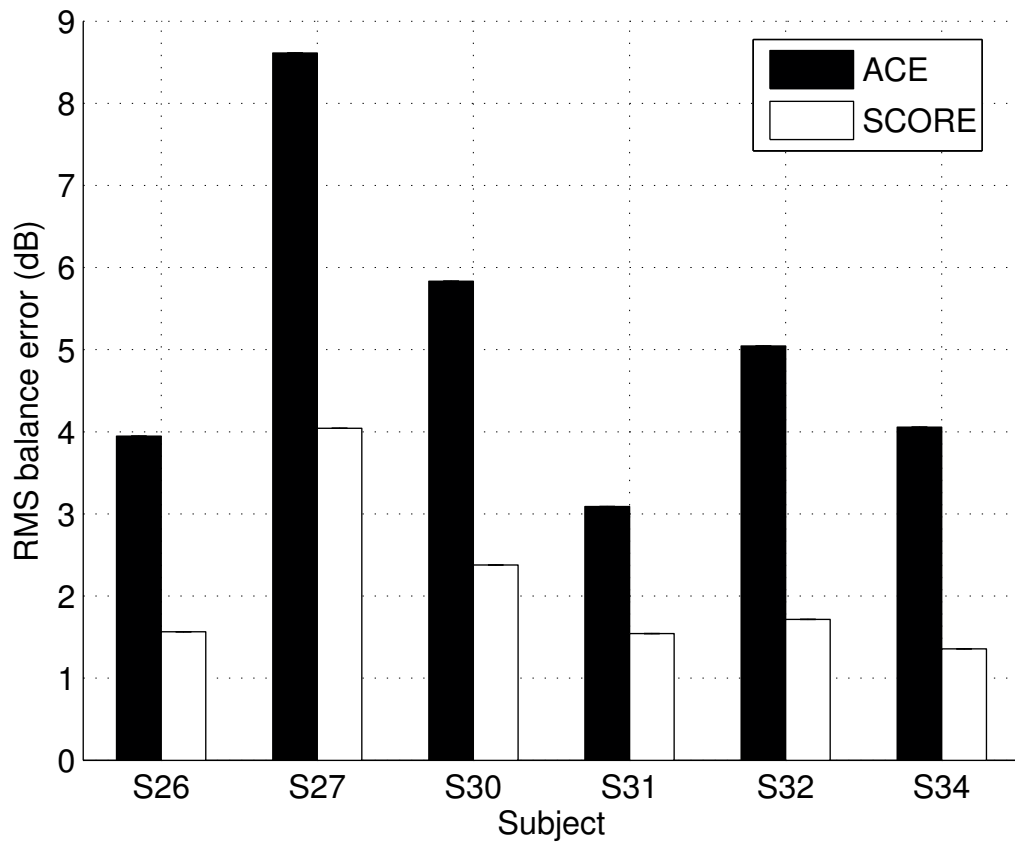


Figure 11: RMS error of the binaural balancing adjustments for each subject in the two processing conditions.

(ACE or SCORE), and loudness (8 or 16 sones). There was a significant effect of processing ( $F(1, 7) = 29.8, p < 0.01$ ), and no effect of stimulus type or loudness, or their interaction.

RMS balance errors for ACE and SCORE are shown in figure 11. The average RMS errors for ACE and SCORE respectively were 5.1 dB and 2.1 dB. In spite of some remaining model prediction errors in the monaural balancing experiments, SCORE generally improved loudness balance for all subjects. This benefit was significant in a Wilcoxon signed rank test ( $p = 0.03$ ).

## 6. Discussion

A loudness normalisation strategy for combined electric and acoustic stimulation was proposed and its function was evaluated using monaural and binaural loudness balancing experiments. The monaural loudness balancing results were compared with predictions made using loudness models, and RMS differences between measurements and model predictions were calculated. Binaural balancing was performed in signal processing conditions ACE and SCORE and a significant improvement in binaural balance was found with application of SCORE.

While the acoustic loudness model was validated during its development, the current results additionally indicate that it is well-suited for application to subjects with a severe to profound hearing impairment, at both soft and comfortable levels. While there were prediction errors ranging from 0.6 to 4.1 dB, its application in the SCORE strategy should bring loudness closer to normal.

Considering the electric loudness model, as stimuli with increasing bandwidth produced an increasing number of CI pulses per time window, the electric loudness balancing results for S26 at 8 sones seem to contradict the premise of the simplified electric loudness model. This model predicts that greater current expressed as a fraction of the dynamic range of the clinical map integrated over a certain time window should result in greater loudness. As the subject found that S600 and S800 at 8 sones and S800 at 16 sones before adjustment were louder than S1200, this results in poor model predictions. This kind of result could be caused by differences between the LGFs of individual electrodes, or differences in loudness between stimuli at individual T (or C) levels on different electrodes. It should also be noted that this subject's map was different from that of the others. Because of facial nerve stimulation at high current levels, the phase width of the pulses

was increased to 75  $\mu$ s to achieve sufficient loudness, which only allowed for 6 maxima to be selected instead of the usual 8. Even with these wide pulses, the overall loudness percept through his CI was soft. This could mean that we were operating in a different part of the LGF for this subject, which might invalidate the simplified electric loudness model. For S32, a value of  $a$  fitted to the whole dataset (8 and 16 sones instead of only 16 sones) would have yielded a much better fit at 8 sones. This could be due to operation in different regions of the LGF, or due to measurement errors in the psychophysical procedures, such as bias due to differences in pitch or sharpness between stimuli of different bandwidths.

The electric loudness model was simplified and its single remaining parameter was fitted based on the monaural loudness balancing data at 16 sones. While this might have decreased the reported RMS errors at 16 sones, it should be noted that 1) it concerns only one parameter value, while there were 3 different balancing results (degrees of freedom) at 16 sones, 2) the data at 8 sones were not used for the determination of the parameter value, and 3) the binaural balancing data were not used for the determination of the parameter value. The results of the electrical monaural balancing experiment indicate that the electrical model performed well for all but one subject (who had an unusual MAP), and for a second subject at soft levels. Even for these subjects, improvements in binaural balance were found with application of SCORE.

The method used for determining the remaining parameter of the electric loudness model consisted of a simple task (“which sound was louder, the first or second?”). Therefore it should be practical to incorporate this task into a clinical fitting protocol. Apart from the unaided audiogram, this is the only subject-dependent parameter that needs to be set. Note that while balancing between electric and acoustic stimulation is still required during fitting, it only needs to be done at one intensity, which is much less involved than attempting to equalise loudness growth across ears by setting the compression parameters of the devices.

The range of values for the  $a$  parameter of the electric model (see table 1) is comparable to that found in other studies (McKay et al., 2003). In the current simplified model, the  $a$  parameter depends on the full, non-simplified LGF of each electrode and the clinical fitting. If (contrary to our assumptions) the clinical T and C levels encompass a region of the LGF that is non-linear, the obtained value for  $a$  will be an average across that range. Consider, for instance, the LGF in figure 2: if the clinical C level was higher

than 200 CU, the full LGF would have been linearly approximated with a value of  $a$  larger than the value for the full model, as indicated by the dashed line in the figure. Additionally, the procedures used in the clinic to obtain T levels and especially C levels may provide quite different values depending on the way the listener interprets the instructions and whether they like to listen to loud sounds. This usually means that the clinically measured dynamic range needs to be multiplied by a scaling factor to obtain the actual dynamic range. The level adjustments calculated by SCORE are expressed in percentage dynamic range, and the formula for the adjustment contains  $a$  in the denominator (see section 2.4). This means that the scaling factor to obtain the actual dynamic range is incorporated in the measured value of the  $a$  parameter.

SCORE adjusts the total loudness of the acoustic and electric signals to the total loudness for a NH listener. While it could be argued that instead the specific loudness over a range of centre frequencies should be normalised, as has been proposed for HAs (Launer and Moore, 2003), this would not be practical and probably not desirable for bimodal listeners. There are no practical loudness models for electric stimulation that have been validated on a frequency-specific basis, so currently we cannot adjust the specific loudness for electric stimulation. Additionally, for bimodal stimulation the same acoustic stimulus often leads to activation of different regions in the two cochleas, such that normalising specific loudness at each ear would not necessarily normalise interaural balance. Next to these practical issues, normalising total loudness can have other advantages. Current bimodal listeners are implanted because their speech perception with only acoustic stimulation is poor. As a consequence, they use their acoustic hearing mainly to provide cues complementary to the electric stimulation, and to perceive binaural cues. To be able to use the complementary cues, the acoustic signal needs to be audible and there needs to be binaural fusion. Given that similarity is one of the auditory grouping cues (Bregman, 1990), equal loudness may improve binaural fusion. For listeners with a broad high-frequency dead region, binaural balance for a broadband sound is not achieved by normalising specific loudness. It should be noted though that normalising the loudness is not beneficial for signals with only high-frequency content, that fall entirely in a high-frequency dead region of the acoustically stimulated ear. In this case there is no energy at frequencies that can be made audible. SCORE copes with this by not adjusting the level if the loudness estimated by the model for HI is below a threshold value.

The results from the binaural balancing experiments indicate that the application of SCORE improved binaural balance, which will likely lead to improved listening comfort and better perception of interaural level differences, leading to improved sound-source localisation ability. It may also result in improved perception of interaural time differences, as this requires binaural balance (Francart et al., 2009, 2011). While in the current study SCORE was evaluated for bimodal stimulation (HA and CI in opposite ears), with some modifications it may also be suitable for hybrid stimulation (HA and CI in the same ear). In this case, the specific loudness integrated across the respective frequency ranges of acoustic and electric hearing could be normalised, instead of the total loudness, as is done for SCORE bimodal. To our knowledge, the SCORE strategy is the first strategy and fitting method specifically designed for combined electric and acoustic hearing.

## **Acknowledgements**

We are grateful to our subjects, who inexhaustibly and cheerfully participated in numerous test sessions. We thank Ruth English for her help with the psychophysical testing, and Cochlear Ltd for financial support. Hamish Innes-Brown provided valuable feedback on an earlier version of the manuscript. We thank Brian Moore and one anonymous reviewer for their constructive remarks to improve the manuscript. Tom Francart was sponsored by a Post Doctoral Fellowship of the Fund for Scientific Research of the Flemish Government and a Marie Curie International Outgoing Fellowship of the European Commission, grant agreement number PIOF-GA-2009-252730. The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program.

## **References**

- Akeroyd, M.A., 2006. The psychoacoustics of binaural hearing. *Int J Audiol* 45 Suppl 1, 25–33.
- ANSI-3.5, 1997. ANSI S3.5, methods for calculation of the speech intelligibility index. American National Standards Institute, New York .
- Bregman, A.S., 1990. Auditory scene analysis. MIT Press, Cambridge, Massachusetts.

- Byrne, D., Parkinson, A., Newall, P., 1990. Hearing aid gain and frequency response requirements for the severely/profoundly hearing impaired. *Ear Hear* 11, 40–49.
- Ching, T.Y., Incerti, P., Hill, M., 2004. Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears. *Ear Hear* 25, 9–21.
- Ching, T.Y., van Wanrooy, E., Dillon, H., 2007. Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review. *Trends Amplif* 11, 161–192.
- Dillon, H., 2001. *Hearing aids*. Thieme Medical Pub, Stuttgart.
- Fastl, H., Zwicker, E., 2007. *Psychoacoustics: Facts and Models*. Springer, New York.
- Francart, T., Brokx, J., Wouters, J., 2008a. Sensitivity to interaural level difference and loudness growth with bilateral bimodal stimulation. *Audiol Neurootol*. 13, 309–319.
- Francart, T., Brokx, J., Wouters, J., 2009. Sensitivity to interaural time differences with combined cochlear implant and acoustic stimulation. *J Assoc Res Otolaryngol* 10, 131–141.
- Francart, T., Lenssen, A., Wouters, J., 2011. Sensitivity of bimodal listeners to interaural time differences with modulated single- and multiple-channel stimuli. *Audiol Neurootol* 16, 82–92.
- Francart, T., van Wieringen, A., Wouters, J., 2008b. APEX 3: a multi-purpose test platform for auditory psychophysical experiments. *J Neurosci Methods* 172, 283–293.
- Glasberg, B.R., Moore, B.C.J., 2002. A model of loudness applicable to time-varying sounds. *J Audio Eng Soc* 50, 331–342.
- Jesteadt, W., 1980. An adaptive procedure for subjective judgments. *Percept Psychophys* 28, 85–88.
- Launer, S., Moore, B.C.J., 2003. Use of a loudness model for hearing aid fitting. v. on-line gain control in a digital hearing aid. *Int J Audiol* 42, 262–273.

- McDermott, H., McKay, C.M., Richardson, L.M., Henshall, K.R., 2003. Application of loudness models to sound processing for cochlear implants. *J Acoust Soc Am* 114, 2190–2197.
- McDermott, H., McKay, C.M., Vandali, A.E., 1992. A new portable sound processor for the University of Melbourne/Nucleus limited multielectrode cochlear implant. *J Acoust Soc Am* 91, 3367–3371.
- McDermott, H., Varsavsky, A., 2009. Better fitting of cochlear implants: modeling loudness for acoustic and electric stimuli. *J Neural Eng* 6, 65007.
- McKay, C.M., Henshall, K.R., Farrell, R.J., McDermott, H., 2003. A practical method of predicting the loudness of complex electrical stimuli. *J Acoust Soc Am* 113, 2054–2063.
- Moore, B.C.J., 2004. Dead regions in the cochlea: conceptual foundations, diagnosis, and clinical applications. *Ear Hear* 25, 98–116.
- Moore, B.C.J., 2012. *An Introduction to the Psychology of Hearing*. Emerald, Bingley, UK. 6th edition.
- Moore, B.C.J., Glasberg, B.R., 1996. A revision of Zwicker’s loudness model. *Acta Acust United Ac* 82, 335–345.
- Moore, B.C.J., Glasberg, B.R., 1997. A model of loudness perception applied to cochlear hearing loss. *Audit Neurosci* 3, 289–311.
- Moore, B.C.J., Glasberg, B.R., 2004. A revised model of loudness perception applied to cochlear hearing loss. *Hear Res* 188, 70–88.
- Swanson, B., 2008. *Pitch Perception with Cochlear Implants*. Ph.D. thesis. The University of Melbourne.
- Vandali, A.E., Whitford, L.A., Plant, K.L., Clark, G.M., 2000. Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear Hear* 21, 608–624.
- Varsavsky, A., McDermott, H., 2012. Application of real-time loudness models can improve speech recognition for cochlear implant users. *IEEE Trans Neural Syst Rehabil Eng* in press.

## A. Calculation of acoustic adjustments

The acoustic adjustment needed to equalise the total loudness of a frame to a target loudness was calculated as follows. The symbols used correspond to those used by Launer and Moore (2003) and Moore and Glasberg (1997).

Most values are calculated for a set of frequency bins  $f$ . Let

$L_T$  the target loudness, in sones,

$E_{HI}(f)$  the HI excitation level, for each frequency bin  $f$ , in linear units,

$A_{IHC}(f)$  the attenuation factor due to inner hair cell loss,

$E_{THRQ}(f)$  the HI hearing threshold, in linear units,

$\alpha$  and  $C$ : constants of the loudness model,

$G$  the current gain relative to NAL-RP, in linear units, initially set to 1

$G$  is calculated iteratively as follows:

Let  $AT$  be the set of values of  $f$  for which  $G \frac{E_{HI}(f)}{A_{IHC}(f)} > E_{THRQ}(f)$

Calculate  $G$  as

$$G = \left( \frac{L_T/C + \sum_f^{AT} (E_{THRQ}(f))}{\sum_f^{AT} \left( \frac{E_{HI}(f)}{A_{IHC}(f)} \right)^\alpha} \right)^{1/\alpha} \quad (1)$$

Calculate the resulting HI total loudness after application of  $G$

If total new total loudness is more than 5% off target, repeat steps above

Note that to save computation time the HI excitation pattern is not calculated again for every iteration. This is an approximation, but should not have a large effect for  $G$  values in a reasonable range.