REFERENCES

Bisgaard, S. (**2010**). **"**Coping with emergent hearing loss, Expectations and Experiences of new, adult hearing aid users. An anthropological study in Denmark" Johann Wolfgang Goethe-Universität Frankfurt am Main, Sprach- und Kulturwissenschaften.

Ceylan, D. and Hudemann, W. (**2011**). "A questionnaire: Satisfaction by using hearing Aids - according to the psychosocial aspects expectations, motivation, social life, getting use to and instruction" Master Thesis, University of Southern Denmark, Faculty of Humanities.

Acoustic simulation of cochlear implant hearing

ANJA CHILIAN[1,2], ELISABETH BRAUN[1,2], TAMAS HARCZOS[1,3]

[1] *Fraunhofer Institute for Digital Media Technology IDMT, D-98693 Ilmenau, Germany*

[2] *Institute of Biomedical Engineering and Informatics, Faculty of Computer Science and Automation, Ilmenau University of Technology, D-98693 Ilmenau, Germany*

[3] *Institute for Media Technology, Faculty of Electrical Engineering and Information Technology, Ilmenau University of Technology, D-98693 Ilmenau, Germany*

One aim in current cochlear implant (CI) research is to improve and optimize speech processing strategies. During the development of new strategies acoustic simulations of CI hearing have widely been used for evaluation. These models usually take audio signals as input and mimic the effects of CI signal processing. In the present paper a new algorithm of acoustic simulation is presented, which transforms stimulation patterns of any cochlear implant directly into an audio signal. Therefore it is independent of the CI strategy used for generating the stimulation pattern. Technical aspects like current spread and physiological aspects including loudness perception and phase locking capabilities of the simulated CI listener can be configured. The presented algorithm was used to evaluate and compare two different CI speech processing strategies in terms of speech intelligibility and pitch discrimination. The results show that acoustic simulation can help estimate the amount of useful information in a CI stimulation pattern and hence be a help in evaluating CI strategies.

INTRODUCTION

A cochlear implant (CI) is an electronic device to restore partial hearing in patients with severe to profound hearing loss. It bypasses the damaged part of the auditory system by direct electrical stimulation of the auditory nerve. Advances in the field of CI research over the last decades have resulted in good speech perception abilities of most CI users in quiet environments. However, speech recognition in noise and music perception still remain challenging.

One factor determining performance in CI users is the speech processing strategy (also called CI strategy), which translates sounds into electrical stimuli. Therefore several approaches aim to improve and optimize these CI strategies. Their evaluation is often performed with CI users in clinical studies, which can be very time-consuming and expensive. In addition, intra- and interindividual variability has to be taken into account. Consequently, a simpler way to evaluate and compare performances of CI strategies has to be found. One possible solution is the use of acoustic simulations of cochlear implant hearing. These simulations generate an acoustic sig-

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.

nal, which is similarly degraded as the signal presented through a CI. Hence, normal-hearing listeners are able to gain insight into the sound perception of a CI user.

Typical approaches of acoustic simulations of cochlear implant hearing were described by Shannon *et al.* (1995) and Dorman *et al.* (1997). In their algorithms the input audio signal is bandpass filtered into m frequency bands. The temporal envelope is extracted from each band and is subsequently used to modulate m carrier signals. The modulated carrier signals are finally recombined to generate an acoustic waveform. Shannon *et al.* and Dorman *et al.* used bandpass filtered noise and sine waves as carrier signals. Acoustic simulations based on this principle are thus commonly referred to as noise-band and tone vocoders.

In numerous studies the vocoder approach has been modified to simulate additional processing parameters and perceptual phenomena. Modified acoustic simulations were used in many experimental investigations with normal-hearing listeners. One possible application is the comparison of the performance of new speech processing strategies. However, this is often impractical, because current acoustic simulations mimic only one specific CI strategy. To simulate possible outcomes of new speech processing strategies a modification of the algorithm is necessary.

In this paper we present a more general algorithm of acoustic simulation. Instead of processing a sound signal, the novel algorithm uses a CI stimulation pattern as input. Therefore, the algorithm itself is independent of the CI strategy used for generating the stimulation pattern. Furthermore, different technical and physiological aspects were modeled including current spread, loudness perception and frequency perception. The developed algorithm was used to evaluate and compare two different CI speech processing strategies in terms of speech intelligibility and pitch discrimination.

ACOUSTIC SIMULATION ALGORITHM

Overview

The acoustic simulation algorithm transforms a CI stimulation pattern directly into an audio signal. Different steps of signal processing were carried out to mimic technical and physiological phenomena influencing speech perception in CI users. A block diagram of the main components of the algorithm is shown in Fig. 1. The stimulation pattern forms the input of the algorithm and can be generated using any speech processing strategy. The electrical stimuli are first used to calculate current spread in the cochlea mimicking the electrode-tissue-interface. Afterwards, simplified models of perception are applied. Firstly, the current amplitudes are converted to simulate loudness perception in CI users. Secondly, transformed values are used for amplitude modulation of bandpass filtered carrier signals to simulate frequency perception and synthesize an audio signal. In addition, several simulation parameters have been implemented (see left-hand side of Fig. 1), which can be adjusted to simulate different situations regarding speech perception of a CI user.

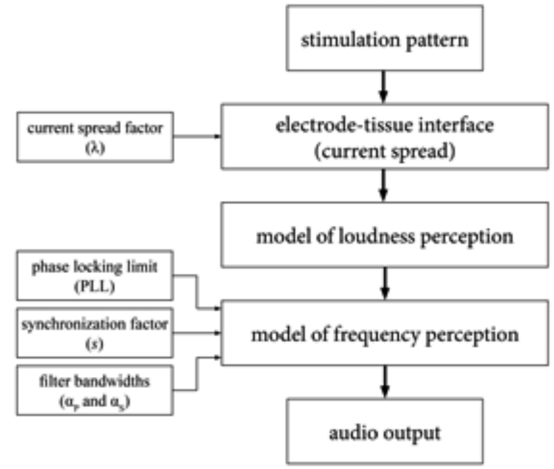

Fig. 1: Block diagram of the acoustic simulation algorithm. On the left-hand side simulation parameters are illustrated that can be configured to simulate different capabilities of a CI user.

Electrode-tissue interface

First, based on the stimulation pattern, effects at the interface between the electrodes and the cochlear tissue are modeled. Stimulation current of the active electrode is spread widely along the cochlea due to the good conductivity of cochlear fluids. This current spread limits the frequency resolution and therefore can degrade speech recognition abilities of CI users.

As the stimulus current level is known from the stimulation pattern, a direct calculation of current spread is possible. We used an approach similar to de la Torre Vega *et al.* (2004). Each CI electrode is assigned to a section along the cochlea. Current spread is calculated for each cochlear section using a simplified one-dimensional model. In several studies current spread has been measured (e. g. Black *et al.*, 1983, Kral *et al.*, 1998), which identified an exponential decay of current with increasing distance to the active electrode. Hence, the following equation is used to calculate current spread (Bingabr *et al.*, 2008):

$$I_x = I_E \cdot \exp\left(-\frac{x}{\lambda}\right), \quad \text{(Eq. 1)}$$

where I_x denotes the current at the distance x to the active electrode and I_E is the current expressed by the active electrode. The current spread factor λ is one of the simulation parameters. Bingabr *et al.* (2008) reported appropriate values of $\lambda = 2..4$ mm and $\lambda = 8..11$ mm for bipolar and monopolar stimulation, respectively. If current spread factor is set to $\lambda = 0$ mm no current spread will be modeled and only the stimulation current of the active electrode will be used for further calculations.

Model of loudness perception

Loudness perception in electrical stimulation depends on the stimulation current and the pulse width of the stimuli. We assume a constant pulse width of all electrical stimuli so that loudness perception is modeled to depend only on the current level. Calculated values of current spread for each cochlear section are therefore transformed using a function based on physiological data of loudness perception.

Fu and Shannon (1998) related the loudness function in acoustic hearing to the loudness function in electrical hearing. They assumed that the relation between loudness L and sound pressure P in acoustic hearing can be described as:

$$L = k_1 \cdot P^{p_1}, \quad \text{(Eq. 2)}$$

with the exponent $p_1 = 0.6$. For electrical hearing they obtained following power function describing the relation between loudness L and current level I:

$$L = k_2 \cdot I^{p_2}, \quad \text{(Eq. 3)}$$

with the exponent $p_2 = 2.72$ and the proportionality factors k_1 and k_2. Combining Eq. 2 and Eq. 3 yields:

$$P = k_3 \cdot I^{\frac{p_2}{p_1}}, \quad \text{(Eq. 4)}$$

where k_3 is the proportionality factor. In our model the power function in Eq. 4 is used to transform the calculated values of current spread into a value proportional to sound pressure.

Model of frequency perception

After transformation of the given stimulation current according to the current spread and to the loudness perception, an audio signal is synthesized simulating the frequency perception. Frequency information in electrical stimulation can be coded by both place and rate of stimulation. To simulate these two mechanisms, we extend the signal synthesis of the general vocoder approach by combining two different carrier signals. This approach is analogous to that of de la Torre Vega *et al.* (2004).

The signal synthesis in the model of frequency perception is illustrated in Fig. 2. In common with the vocoder approach, modulation and bandpass filtering of carrier signals is applied. To simulate place pitch perception in the present algorithm white noise is used as carrier (see bottom of Fig. 2). It is first bandpass filtered and then amplitude modulated using calculated values from previous stages of the simulation. Pulse trains are used as carrier to additionally simulate temporal pitch perception (see top of Fig. 2). For this carrier amplitude modulation is applied before bandpass filtering.

Audio signals generated using the two carrier signals are then combined according to synchronization factor s, which describes the supposed capability of the CI user for synchronization with the stimulation rate of each electrode (de la Torre Vega *et al.*, 2004). The synchronization factor is a number between 0 and 1, where $s = 1$ indicates very good capability for synchronization with the rate of stimulation. Therefore, both place pitch perception and temporal pitch perception are simulated. For $s = 0$ no synchronization capability is simulated. Hence, only place pitch perception is modeled.

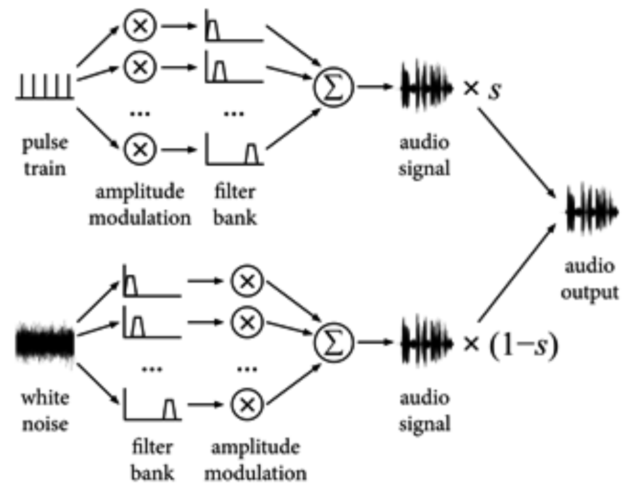

Fig. 2: Block diagram illustrating signal synthesis in the model of frequency perception. Two different carrier signals are combined according to the synchronization factor s.

The filter bank, which is used for audio signal synthesis, consists of Butterworth bandpass filters. The center frequencies of the filters are consistent with the center frequencies of the CI electrodes. To determine edge frequencies of the filters the center frequency *f* is first transformed into a position *x* along the cochlea using following equation (Greenwood, 1990):

$$x = \frac{1}{a} \cdot \log\left(\frac{f}{A} + k\right), \quad \text{(Eq. 5)}$$

with $a = 0{,}06$, $A = 165{,}4$ and $k = 0{,}88$. Then, two simulation parameters α_P and α_S are implemented to determine filter bandwidths. These parameters describe defined sections along the cochlea as shown in Fig. 3. While α_P determines the distance between positions along the cochlea associated with upper and lower passband frequencies (x_{pass1} and x_{pass2}) from the position of the center frequency x_c, α_S defines the distance between positions along the cochlea associated with upper and lower stopband frequencies (x_{stop1} and x_{stop2}) from the position of the center frequency x_c. The calculated positions of the edge frequencies are subsequently transformed into frequency values using the inversion of Eq. 5. For sufficient overlapping of the filter bands the parameter values should be $\alpha_P = 0.75$ mm and $\alpha_S = 4.5$ mm, which results in 4th-order Butterworth filters.

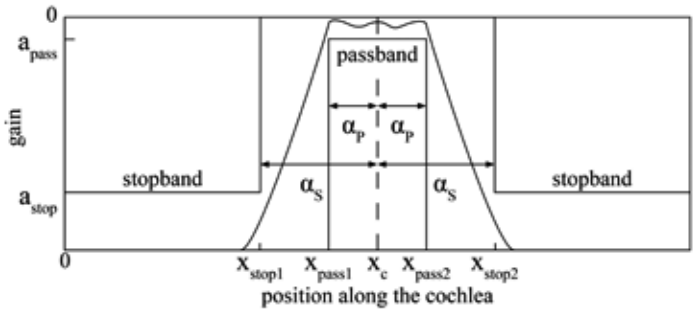

Fig. 3: Bandpass filter specification illustrating parameters to determine the filter bandwidth.

Temporal pitch perception with cochlear implants is additionally limited. Most CI users cannot discriminate changes above the so-called phase locking limit. Therefore, we used an additional simulation parameter, which determines the phase locking limit (PLL). Up to this frequency temporal pitch perception, i. e., synchronization with the rate of stimulation, is possible. For electrodes with center frequencies above PLL the sound signal is generated using only white noise as carrier signal. Below PLL two different sound signals are generated using white noise and pulse carrier, which are eventually combined according to the synchronization factor s.

HEARING TESTS

The presented algorithm of acoustic simulation was used to compare two different CI speech processing strategies. These strategies produced the stimulation patterns, which were processed by the acoustic simulation algorithm to generate audio data. The synthesized sound signals were presented to normal-hearing listeners to measure speech intelligibility and pitch discrimination.

Subjects and methods

Thirty volunteer normal-hearing listeners (18 men, 12 women) participated in this study. Their ages ranged from 19 to 30 years. All subjects were native German speakers. Tests of speech intelligibility were performed using words consisting of one syllable (Freiburg monosyllables). Each subject was presented with 15 monosyllables and the proportion of correctly repeated words determined word recognition scores. Pitch discrimination was measured using the sung vowel /a/ produced by a male singer. An adaptive three-alternative forced-choice test was employed using the 1-up-2-down rule. The subjects were asked to indicate the vowel which was different from the other two presented and whether it was lower or higher pitched. The base tone was D#3 (155.6 Hz) and the starting difference was 6 semitones.

The test material was processed using ACE (Advanced Combination Encoder) and SAM (Stimulation based on Auditory Modeling) strategy. While ACE is a common n-of-m strategy, SAM is a novel CI strategy based on neurophysiological models of the auditory system. In SAM, therefore, several psychoacoustic phenomena are ac-

counted for inherently. Additionally, the coder is not restricted to a pre-defined channel stimulation rate. The generated stimulation patterns using these two strategies were processed by the acoustic simulation algorithm. Simulation parameters were set as follows: current spread factor $\lambda = 0.5\ \text{mm}$, synchronization factor $s = 0.9$, phase locking limit $\text{PLL} = 3000\ \text{Hz}$, parameters determining filter bandwidths $\alpha_\text{P} = 0.75\ \text{mm}$ and $\alpha_\text{S} = 4.5\ \text{mm}$.

Results

Box whisker plots of measured average word recognition scores and pitch discrimination using ACE and SAM strategy can be seen in Fig. 4. The central marks of the boxes indicate the median. The 25th and 75th percentiles are represented by the bottom and the top of the boxes, respectively. The whiskers extend to the most extreme data points, which are not considered outliers. Outlying points lay at least 1.5 inter quartile ranges from either end of the box.

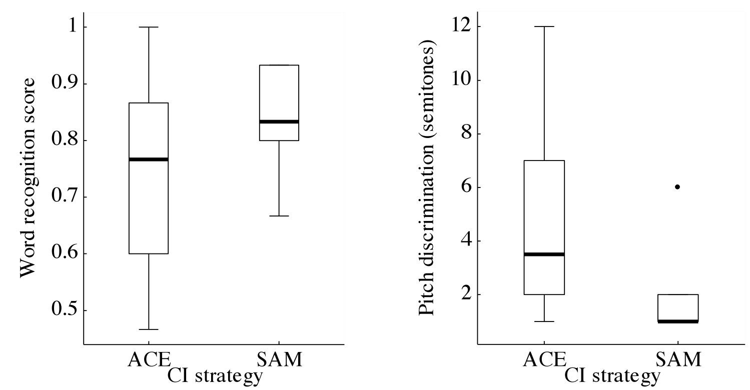

Fig. 4: Box whisker plots of measured word recognition scores (left) and pitch discrimination (right) using ACE and SAM strategy.

Speech intelligibility tests showed an increase of the average word recognition scores using SAM when compared to ACE. Further analysis indicated that this increase was statistically significant.

Measured pitch discrimination of sung vowels using SAM was significantly better than the pitch discrimination using ACE. These results may indicate that the SAM strategy is more efficient in preserving pitch information.

CONCLUSION

In this paper we presented a new algorithm of acoustic simulation of CI hearing. Physiological phenomena like current spread, loudness perception and frequency perception were included in the model. In contrast to the vocoder approach of Shannon *et al.* (1995) and Dorman *et al.* (1997) we developed a more general algorithm using CI stimulation patterns as input. Therefore, the new acoustic simulation can be used to compare different CI strategies without modifying the algorithm or its parameterization. At the same time, the simulation can be configured to mimic individual capabilities of CI users. Consequently, investigating specific influencing factors of speech intelligibility like current spread or phase locking ability is possible.

The results of this study indicate that the acoustic simulation algorithm can be used to estimate the amount of useful information in a CI stimulation pattern. Hence, it might help evaluating speech processing strategies. However, the acoustic simulation is only intended to measure trends in speech recognition performance and pitch discrimination. Exact predictions of performance regarding speech perception of a CI user are currently not possible. Further work should compare results of normal-hearing listeners using the acoustic simulation with actual CI user performance to asses the validity of the algorithm.

ACKNOWLEDGMENT

The underlying research reported in the paper was funded by the grant B514-09020 of the Thuringian Ministry of Education, Science and Culture.

REFERENCES

Bingabr, M., Espinoza-Varas, B., and Loizou, P. C. (**2008**). “Simulating the effect of spread of excitation in cochlear implants“ Hearing Research, **241**, 73-79.

Black, R. C., Clark, G. M., Tong, Y. C., and Patrick, J. F. (**1983**). “Current distributions in cochlear stimulation” Ann. N.Y. Acad. Sci., **405**, 137-145.

Dorman, M.F., Loizou, P.C., and Rainey, D. (**1997**) “Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs“ J. Acoust. Soc. Am., **102**, 2403-2411.

Fu, Q. J. and Shannon, R. V. (**1998**). “Simulating effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners“ J. Acoust. Soc. Am., **104**, 2570-2577.

Greenwood, D. D. (**1990**). “A cochlear frequency-position function for several species - 29 years later“ J. Acoust. Soc. Am., **87**, 2592-2605.

Kral, A., Hartmann, R., Mortazavi, D., and Klinke, R. (**1998**). “Spatial resolution of cochlear implants : The electrical field and excitation of auditory afferents” Hearing Research, **121**, 11-28.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). “Speech recognition with primarily temporal cues” Science, **270**, 303-304.

de la Torre Vega, Á., Martí, M. B., de la Torre Vega, R., and Quevedo, M. S. (**2004**). “Cochlear implant simulation version 2.0 : Description and usage of the program“ University of Granada, Spain.

Benefits of common vocabulary in hearing aid fitting

CARSTEN DAUGAARD[1], SØREN L. JØRGENSEN[1] AND LONE ELMELUND[2]

[1] *DELTA, Technical-Audiological Laboratory, DK-5000 Odense C, Denmark*

[2] *Institute of Biology, University of Southern Denmark, Campusvej 55 DK 5230 Odense M*

BACKGROUND

Modern hearing aids are very sophisticated devices and through the fitting process they can be adjusted to fit the hearing loss of a large variety of people. However in order to set the fitting parameters right, the communication between the hearing aid professional and the user has to be successful. The challenge here is to understand and map the experience of the user in order to transfer it to the fitting software.

Today, hearing aid manufacturers has taken up the challenge by designing software handles whose function is less technical and more related to commonly experienced hearing aid problems. They have also added expert assistants to the software, mapping common user complaints into the traditional technical software handles. When it comes to perceived sound quality, however, the challenge lies first and foremost in understanding the user’s perception, to decode the sound experience of the user so to speak. For this challenge the hearing aid professional must be experienced enough to understand the user’s language of sound perception. Hearing aid professionals know that this can prove to be a complicated problem. As with many other perceptual experiences we are not used to express sound experiences in many more words than soft, loud, annoying or pleasant.

A common vocabulary between the user and hearing aid professional would probably make the task easier so rather than relying on the hearing aid professional’s skills to understand the user’s desire, the user’s vocabulary of sound perception could be trained. Inspiration for this alternative approach can be found in the sensory evaluation discipline, where selected panels train their ability to express differences in selected sound attributes (Bech and Zacharov, 2006).

Attributes in sensory evaluation

Sensory evaluation is a systematic approach to assess the sensory impression of a given object, i.e. food products, perfumes, sound. The goal of sensory evaluation is to have a panel of trained assessors, known as a listening panel, which is able to consistently and repeatable evaluate objects in a range of attributes, describing the object. In other words, to establish a “sense-o meter” to evaluate how humans experience the object to be tested.

A central part of the descriptive analysis process of sensory evaluation is to establish specific traits of the object that can be explained and evaluated on a scale. Every

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.