Making use of auditory models for better mimicking of normal hearing processes with cochlear implants: first results with the SAM coding strategy

TAMÁS HARCZOS[1,2,*], ANJA CHILIAN[1,3], ANDRAS KATAI[1], FRANK KLEFENZ[1], IZET BALJIĆ[4], PETER VOIGT[5], AND PETER HUSAR[1,3]

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

[2] *Electronic Media Technology Lab, Faculty of Electrical Engineering and Information Technology, Ilmenau University of Technology, Ilmenau, Germany*

[3] *Institute of Biomedical Engineering and Informatics, Faculty of Computer Science and Automation, Ilmenau University of Technology, Ilmenau, Germany*

[4] *HELIOS Hospital Erfurt, Department of Otolaryngology, Erfurt, Germany*

[5] *Cochlear-Implant Rehabilitationszentrum Thüringen, Erfurt, Germany*

Stimulation based on auditory modeling, or SAM, is a new speech-processing strategy for cochlear implants that we developed recently at Fraunhofer IDMT. SAM incorporates active cochlear filtering along with the mechanoelectrical transduction of the inner hair cells, so that several psychoacoustic phenomena are accounted for inherently. SAM was tested with a group of five CI users: We investigated speech perception in quiet and in the presence of noise or reverberation, pitch discrimination abilities (for pure tones and sung vowels), and consonant discrimination. We also asked for subjective quality rating for speech and music snippets. Tests were repeated with the everyday strategy of the implantees and results were compared. This paper presents the test results in detail and compares outcomes with those of the previously published simulation studies. Results are encouraging, although more tests would be needed to increase statistical significance.

INTRODUCTION

Increased processing speeds make applications using auditory models that mimic some properties of the human ear viable. The idea of using models of the human auditory system in cochlear implants (CIs) is not new (see Wilson *et al*., 2010), but still fairly uncharted. In Harczos *et al.* (2013) we presented a novel sound-processing strategy, SAM (Stimulation based on Auditory Modeling), which was based on hydromechanical and neurophysiological models of the human ear and could be employed in auditory prostheses.

SAM incorporates active cochlear filtering (basilar membrane and outer hair cells) along with the mechanoelectrical transduction of the inner hair cells, so that travel-

---

*Corresponding author: hzs@idmt.fraunhofer.de

Proceedings of ISAAR 2013: Auditory Plasticity – Listening with the Brain. 4th symposium on Auditory and Audiological Research. August 2013, Nyborg, Denmark. Edited by T. Dau, S. Santurette, J. C. Dalsgaard, L. Tranebjærg, T. Andersen, and T. Poulsen. ISBN: 978-87-990013-4-7. The Danavox Jubilee Foundation, 2014.

ling-wave delays and several psychoacoustic phenomena are accounted for inherently. The produced stimulation patterns differ greatly from that of the wide-spread Cochlear ACE™ (advanced combination encoder) strategy. Although the computation of SAM requires considerably more operations than that of ACE, the current C/C++ implementation of SAM can run in real-time on a state-of-the-art desktop computer.

At ISAAR 2011 we showed the outline of the algorithm along with first simulation results concerning speech reception thresholds (Harczos *et al.*, 2012b) and horizontal-plane localization abilities using SAM (Harczos *et al.*, 2012a). We also presented a real-time visualization of the strategy and a vocoder algorithm making SAM stimuli audible (Chilian *et al.*, 2012). In the meantime we did first tests with CI users to explore benefits with SAM. In this paper, we present these results.

METHODS

Participants

Five post-lingually deafened adult CI users participated in the study. They were all native speakers of German and had at least two years of CI experience at the commencement of the study. Every subject had a Nucleus® Freedom™ implant with a Contour Advance™ electrode together with a Freedom™ sound processor from Cochlear®. More detailed demographic information is presented below.

Subject	**Age** (yr)	**Deaf** (yr)	**CI** (yr)	**Most probable cause of hearing impairment**	**Laterali-zation**	***CSR*** (pps)	***N***
S1	37	3	4	Circulatory disorder	Bimodal	900	11
S2	70	1	5	Genetic	Unilateral	900	9
S3	69	15	2	Diphtheria	Bimodal	900	8
S4	50	1	5	Genetic / Traumatic	Bilateral	1200	10
S5	27	3	13	Meningitis	Bilateral	900	8

Table 1: Detailed demographic information. The ACE parameters *CSR* and *N* mean channel stimulation rate and number of spectral peaks, respectively.

Assessment procedure

Cochlear-implant users' performance was measured in various ways with a number of tests as listed below.

(1) Testing of speech intelligibility in quiet with the *Freiburg monosyllabic test* (see Hahlbrock, 1953). Corresponding results did not appear to be particularly meaningful and were not listed in this paper.

(2) Testing of speech intelligibility in speech-shaped noise with the *Oldenburg sentence test* (*OLSA*, see Wagener *et al*., 1999).

(3) Testing of speech intelligibility in simulated reverberant environments using clean but reverberated OLSA sentences with four distinct magnitudes of reverberation. Sentences were played back and the subject was asked to repeat them. The percentage of correctly repeated words was computed.

(4) Testing of pitch discrimination thresholds for pure tones and sung vowels in an adaptive three-interval three-alternative forced-choice (‘3I-3AFC’, or ‘odd-one-out’) procedure using the 1-up 2-down paradigm (see Levitt, 1971). Without having to tell which tone was lower or higher in frequency, the subject was asked to identify the tone that was different in pitch. The discrimination threshold was then computed for each test-tone type.

(5) Testing of discrimination ability of consonant pairs (b/p, m/n, n/l, and k/t) using minimal-pair words. In each trial, either two similarly sounding (e.g., bark/park) or two identical words were played back in a sequence. The subject was asked to tell if the two words were the same or not. The percentage of correct answers was computed for each consonant pair.

(6) Subjective quality rating of speech and music via direct comparison of snippets processed with either SAM or ACE.

Reverberation conditions as used in assessment method (3) are listed in Table 2. RT_{60} and STI mean reverberation time and speech transmission index (Steeneken and Houtgast, 1980), respectively. The former is the time required for the sound level to decrease by 60 dB, while the latter is a measure of speech transmission quality. STI is a well-established objective measurement predictor of how well a listener may understand speech using the given transmission channel. STI values may vary between 0 (bad) and 1 (excellent). STI values presented in Table 2 were calculated for the same (randomly selected) OLSA sentence for 65 dB SPL presentation level.

	Reverb-1	**Reverb-2**	**Reverb-3**	**Reverb-4**
Simulated environment	living room	empty office	train station	stairwell (concrete walls)
RT_{60}	935 ms	1440 ms	1380 ms	2700 ms
STI	0.988	0.897	0.745	0.467

Table 2: Summary of reverberation conditions and related parameters.

Conditions as used for the pitch discrimination tests described in assessment method (4) are listed in Table 3. Each presented sequence consisted of three tones (two identical reference and an odd one, each 600 ms long) separated by a 400-ms pause. The spectral distance between the differing and the reference tones was varied adaptively with a quantization of one semitone. Frequencies (or fundamental frequencies in the case of sung vowels) of the tones, as expressed in notes, were determined to be symmetrical around the centre of the valid range for the given test variant (see Table 3). The initial distance was six semitones. The intensity of each tone was randomized by +/− 3 dB to reduce any unwanted effects of loudness variations on the subjects' ranking of pitches. Subjects were instructed to ignore loudness variations, if they perceived any. The task was to identify the tone that was different in pitch.

	Pure tones (C5)	**Pure tones (C6)**	**Pure tones (C7)**	**Female sung "A" and "I"**	**Male sung "A" and "I"**
Range	C4 (262 Hz) – C6 (1046 Hz)	C5 (523 Hz) – C7 (2093 Hz)	C6 (1046 Hz) – C8 (4186 Hz)	C4 (262 Hz) – F5 (698 Hz)	G2 (98 Hz) – A#3 (233 Hz)
Centre of range	C5 (523 Hz)	C6 (1046 Hz)	C7 (2093 Hz)	G#4 (415 Hz)	D#3 (156 Hz)

Table 3: Summary of conditions for pitch discrimination tests.

Within five sessions (each 2 × 45 minutes plus breaks) all tests were conducted with each participant using both the ACE and the SAM strategies. The latest individual clinical map was used with ACE. With SAM a new map was created and fitted for each CI user. Subjects were provided an excerpt of 6 to 10 minutes of an audio book prior to testing with SAM to get accustomed to the new strategy.

Except for the duration of fitting and initial practice with SAM, subjects were blinded to the choice of processing strategy used in any test.

Stimulation

The Nucleus Implant Communicator (NIC) version 2 from Cochlear® (see Irwin, 2006; Swanson and Mauch, 2006) was employed to directly stream the stimuli from the PC to the CI. All computer programs developed and used during this study were able to process sounds by both the ACE and the SAM strategy. This way, the switch between the strategies was easy for the operator and without attracting subjects' attention.

RESULTS

Speech intelligibility

The standard OLSA test revealed that implant users S4 and S5, being already high-performers with the ACE strategy (i.e., OLSA SRT < 0 dB), could not benefit from switching to the SAM strategy in terms of speech intelligibility in speech-shaped noise. For the other three subjects (having about 10-15 dB worse speech reception thresholds using ACE than S4 and S5) the switch to SAM manifested itself in better SRTs (on average 2.44 dB better).

Results based on the reverberant OLSA corpus showed similar trends: SAM showed no benefit in S4 and S5, while the other three subjects achieved slightly better scores on average. For detailed results, please see Table 4.

Speech intelligibility test type	S1		S2		S3		S4		S5	
	SAM	ACE	SAM	ACE	SAM	ACE	SAM	ACE	SAM	ACE
OLSA (Standard)	1.9 dB	4.9 dB	5.2 dB	6.3 dB	7.6 dB	10.9 dB	-3.6 dB	-5.9 dB	-2.3 dB	-4.1 dB
OLSA (Reverb-1)	92 %	85 %	87 %	87 %	80 %	55 %	97 %	100 %	100 %	100 %
OLSA (Reverb-2)	80 %	83 %	82 %	76 %	49 %	52 %	100 %	100 %	100 %	100 %
OLSA (Reverb-3)	73 %	70 %	70 %	56 %	37 %	36 %	93 %	100 %	94 %	100 %
OLSA (Reverb-4)	65 %	20 %	9 %	16 %	4 %	18 %	60 %	71 %	93 %	100 %

Table 4: Results of speech-intelligibility tests with the OLSA corpus. The first row shows speech reception thresholds (in dB SNR) measured with the standard OLSA test procedure (speech-shaped noise). The other rows show percentage of correctly identified words of reverberant OLSA sentences at four fixed reverberation magnitudes. Cells with grey background colour denote cases where the ACE strategy performed better.

Pitch discrimination

Since the SAM strategy was designed to provide a considerable amount more temporal pitch information than ACE does, cochlear-implant users were expected to perform better (in terms of pitch-discrimination performance) with SAM than with ACE. Test results showed that this expectation was reasonable: except for isolated cases, all tests delivered much better scores with the proposed new signal-processing strategy.

Signal type in pitch test	S1		S2		S3		S4		S5	
	SAM	ACE	SAM	ACE	SAM	ACE	SAM	ACE	SAM	ACE
Pure tones (C5)	2.3	8.5	4.6	2.0	2.2	3.9	1.4	2.3	1.5	2.5
Pure tones (C6)	3.3	8.7	3.3	2.5	1.8	2.5	1.5	1.7	1.3	1.0
Pure tones (C7)	2.8	4.1	1.5	2.7	1.5	3.0	3.0	3.5	1.8	2.3
Female sung A	10.3	6.0	6.4	5.0	6.2	7.4	4.3	6.6	5.9	7.5
Female sung I	7.8	10.7	2.5	3.3	3.4	6.5	2.0	3.8	1.8	4.0
Male sung A	6.0	6.5	6.0	12.5	3.5	6.4	6.1	7.4	6.2	6.3
Male sung I	4.5	7.7	7.7	13.5	4.8	10.4	4.8	6.6	6.0	7.0

Table 5: Pitch-discrimination thresholds (in semitones) measured using the adaptive 3-AFC procedure (with 1-up 2-down rule) that targeted 70.7% ($p = 1/\sqrt{2}$) correct discrimination level. Cells with grey background colour denote cases where the ACE strategy performed better.

Table 5 shows discrimination thresholds (in semitones) of all subjects for various signal types. Tests with pure tones seem to be much easier for all subjects: The number of semitones (ST) for the 70.7% discrimination threshold averages to 2.83 ($\sigma = 1.8$), while the same measure for the sung vowels yields 6.28 ST ($\sigma = 2.57$). Vowel 'I' sung by the male singer proved to be the most difficult signal: Subjects needed a pitch difference of 7.29 ST ($\sigma = 2.81$) on average (i.e., an interval larger than a perfect fifth!) to correctly identify the difference (with $p = 1/\sqrt{2}$).

The results listed in Table 5 clearly indicate that the tested CI listeners can utilize the additional temporal information provided by the new strategy. The benefit with SAM averages to 1.16 ST ($\sigma = 2.17$), 1.02 ST ($\sigma = 2.27$), and 2.86 ST ($\sigma = 2.36$) for pure tones, female sung vowels, and male sung vowels, respectively.

Consonant discrimination

Results of the consonant-discrimination tests did not deliver clear trends, as shown in Table 6. CI users' performance seems to be at about the same level with both strategies.

Consonant pair	**S1**		**S2**		**S3**		**S4**		**S5**	
	SAM	**ACE**	**SAM**	**ACE**	**SAM**	**ACE**	**SAM**	**ACE**	**SAM**	**ACE**
b / p	100 %	87 %	100 %	100 %	93 %	87 %	73 %	67 %	100 %	100 %
m / n	27 %	33 %	40 %	67 %	73 %	40 %	87 %	80 %	72 %	67 %
n / l	67 %	40 %	67 %	60 %	53 %	67 %	80 %	80 %	87 %	60 %
k / t	73 %	93 %	80 %	80 %	87 %	87 %	80 %	80 %	80 %	100 %

Table 6: Percentages of correct answers in the consonant pairs test. Cells with a grey background denote cases where the ACE strategy performed better.

Subjective quality

At first sight, subjective quality ratings yielded mixed strategy preferences (see Table 7). However, the preferences of the two bimodal users (S1 and S3) of the test group were remarkable. Since these subjects still had a more or less natural contralateral auditory perception to compare with (hearing aid in the contralateral ear), results would suggest that stimulation via SAM elicits more natural sensation.

Signal type	**S1**	**S2**	**S3**	**S4**	**S5**
Speech	SAM better	ACE better	SAM better	ACE better	SAM better
Music	SAM better	no preference	SAM better	no preference	no preference

Table 7: Results of subjective quality rating after direct comparisons. Cells with a grey background denote cases where the ACE strategy was preferred.

DISCUSSION

SAM is a novel speech-processing strategy for implantable auditory prostheses that we have tested in a pilot study with five cochlear-implant users. Even though the number of testees was very low (and hence the variance of the results considerably high), we were able determine some trends of benefit with SAM: Better speech reception thresholds in speech-shaped noise of CI users performing poorly with ACE, and much better pitch-discrimination performance of all testees were the most prominent quantifiable results. These were also predicted by the simulation study published in Harczos *et al.* (2012b).

Another important outcome was that bimodal users rated the quality of sensation higher with SAM than with ACE. This indicates that the firing patterns of the auditory nerve elicited by the SAM stimulation are more similar to the natural ones. Investigations with an acoustical simulation tool (see Chilian *et al.*, 2012) also showed strong preference for SAM (over ACE) in normal-hearing subjects.

Tests of the presented study have also shown that no subject was stressed or disturbed by SAM. Furthermore, knowing that a successful switch from one CI strategy to another may take weeks or months, the fact that all participants understood speech immediately after switching to SAM was an astonishing outcome by itself.

Unfortunately, the NIC v2 tool provided by Cochlear® did not support continuous real-time streaming, which had two important implications. First, there was an unavoidable delay (ranging from seconds to minutes, depending on the duration of the test signal) between sending a stimulus signal from the PC and perceiving it via the CI. Second, to be able to communicate with the CI users, their everyday processor (using ACE) needed to be placed back and turned on again, which might have interfered with the learning processes involved in extracting information from the SAM stimulation patterns.

Preparations are currently underway in our lab to be able to provide a longer uninterrupted habituation and testing period with SAM. Furthermore, we plan to run a longer study including at least 20 CI users to yield more statistically relevant results.

Finally, as the simulation study in Harczos *et al.* (2012a) indicates huge improvements in horizontal plane localization with binaural SAM configurations over ACE, this issue should also be investigated with cochlear-implant users.

ACKNOWLEDGMENTS

We thank volunteer subjects for participating in the pilot study and the staff of the Cochlear-Implant Rehabilitationszentrum Thüringen for their support.

The authors would also like to thank Thomas Vandahl and Akos Kitta for their help with programming, Gero Szepannek for his help with statistics, and Cochlear Limited for access to implant communication software and hardware. We are also grateful for the talks with Horst Hessel, Thomas Hocke, and Jan Poppeliers.

This research was supported in part by grant B514-09020 of the Thuringian Ministry of Education, Science and Culture, Germany, and by Cochlear Ltd.

REFERENCES

Chilian, A., Braun, E., and Harczos, T. (**2012**). "Acoustic simulation of cochlear implant hearing," in *Speech perception and auditory disorders*. 3rd International Symposium on Auditory and Audiological Research. Nyborg, Denmark. Edited by T. Dau, M.L. Jepsen, T. Poulsen, and J. C.-Dalsgaard. ISBN: 978-87-990013-3-0. (The Danavox Jubilee Foundation, Copenhagen), pp. 425-432.

Hahlbrock, K.-H. (**1953**). "Über Sprachaudiometrie und neue Wörterteste," Arch. Ohren Nasen Kehlkopfheilkd., **162**, 394-431.

Harczos, T., Chilian, A., and Katai, A. (**2012a**). "Horizontal-plane localization with bilateral cochlear implants using the SAM strategy," in *Speech perception and auditory disorders*. 3rd International Symposium on Auditory and Audiological Research. Nyborg, Denmark. Edited by T. Dau, M.L. Jepsen, T. Poulsen, and J. C.-Dalsgaard. ISBN: 978-87-990013-3-0. (The Danavox Jubilee Foundation, Copenhagen), pp. 339-345.

Harczos, T., Fredelake, S., Hohmann, V., and Kollmeier, B. (**2012b**). "Comparative evaluation of cochlear implant coding strategies via a model of the human auditory speech processing," in *Speech perception and auditory disorders*. 3rd International Symposium on Auditory and Audiological Research. Nyborg, Denmark. Edited by T. Dau, M.L. Jepsen, T. Poulsen, and J. C.-Dalsgaard. ISBN: 978-87-990013-3-0. (The Danavox Jubilee Foundation, Copenhagen), pp. 331-338.

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**, 414-425.

Irwin, C. (**2006**) "NIC v2 Software Interface Specification E11318RD (Technical Report)," Lane Cove NSW, Australia, Cochlear Ltd.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am., **49**, 467-477.

Steeneken, H.J.M., and Houtgast, T. (**1980**). "A physical method for measuring speech-transmission quality," J. Acoust. Soc. Am., **67**, 318-326.

Swanson, B.A., and Mauch, H. (**2006**). "Nucleus Matlab Toolbox 4.20 software user manual," Lane Cove NSW, Australia, Cochlear Ltd.

Wagener, K., Kühnel, V., and Kollmeier, B. (**1999**). "Entwicklung und Evaluation eines Satztests in deutscher Sprache I: Design des Oldenburger Satztests (Development and evaluation of a sentence test in German language I: Design of the Oldenburg sentence test)," Z. Audiol., **38**, 4-15.

Wilson, B.S., Lopez-Poveda, E.A., and Schatzer, R. (**2010**). "Use of auditory models in developing coding strategies for cochlear implants." in *Computational Models of the Auditory System*. Edited by R. Meddis, R.R., Fay, E.A., Lopez-Poveda, and A.N. Popper (Springer Science+Business Media LLC, Boston, Massachusetts, USA), pp. 237-260.