Evaluation of a transient noise reduction algorithm in cochlear implant users

Karl-Heinz Dyballa,[1] Phillipp Hehrmann,[2] Volkmar Hamacher,[2] Waldo Nogueira,[1] Thomas Lenarz,[1] Andreas Büchner[1]

[1]Department of Otolaryngology, Hannover Medical School and Cluster of Excellence Hearing4all; [2]European Research Center Advanced Bionics, Hannover, Germany

Abstract

Dealing with environmental noises presents a major issue for cochlear implant (CI) users. Hence, digital noise reduction (DNR) schemes have become important features of CI systems. Many noises like for example clinking glasses or slamming doors, have impulsive onsets and decay quickly. Common DNR algorithms cannot handle this type of noise in an appropriate way. In this study, we investigated the effect of an algorithm specially designed for such noises with 12 CI users (age range: 45 to 75 years). Speech scores in noise and quiet as well as subjective ratings of speech clarity, comfort and overall preference were measured. The main finding was a significant improvement of up to 1.7 dB of the speech reception threshold in noise as well as increased speech clarity. Speech in quiet was not negatively affected by the algorithm. The study revealed that the tested algorithm has the potential to improve CI listening. However, further research is needed regarding the effectiveness and suitability of the algorithm in daily use.

Correspondence: Karl-Heinz Dyballa, Hannover Medical School, Karl-Wiechert-Allee 3, 30625 Hannover, Germany.
Tel.: +49.511.532.6835 - Fax: +49.511.532.6833.
E-mail: Dyballa.Karl-heinz@mh-hannover.de

Key words: transient, impulsive noise, noise reduction, cochlear implant, speech intelligibility.

Acknowledgments: the authors would like to thank all participants in this study.

Funding: this research was supported by Advanced Bionics GmbH, European Research Center, Hannover, Germany.

Contributions: KD, PH, TL, VH, AB, conceptual idea; KD, PH, study design; KD, data collection; KD, PH, data analysis; KD, PH, VH, AB, WN, manuscript preparation.

Conflict of interest: PH and VH are employees of Advanced Bionics and provided writing assistance. The authors alone are responsible for the content and writing of the paper.

Received for publication: 31 July 2014.
Revision received: 27 March 2015.
Accepted for publication: 19 April 2015.

This work is licensed under a Creative Commons Attribution NonCommercial 3.0 License (CC BY-NC 3.0).

©Copyright K.-H. Dyballa et al., 2015
Licensee PAGEPress, Italy
Audiology Research 2015;5:116
doi:10.4081/audiores.2015.116

Introduction

Transient noises such as dishes clattering, door slams or even typing on a computer keyboard are often annoying for hearing impaired people.[1,2] Transient noise reduction (TNR) algorithms have been developed by some hearing aid (HA) companies in order to mitigate this issue.[3,4] Until now, no dedicated TNR algorithm is available for cochlear implant (CI) users, despite the fact that they typically have greater difficulty in understanding speech in noise.[5] In this paper we investigate the potential use of one such algorithm with CI users. Noises which we consider as transient have three basic characteristics: an impulsive onset (potentially even below 1 millisecond), a fast decay (tens of milliseconds) and a short overall duration (hundreds of milliseconds up to one second).[3,6] In contrast, common single-channel digital noise reduction (DNR) algorithms are designed for stationary or near-stationary noise, where the spectral and temporal signal characteristics are stable over hundreds of milliseconds or seconds at least. Thus, those algorithms have time constants, which are too slow to detect and suppress transient noises in an appropriate way, and shortening them is likely to create an intolerable amount of distortion.[7]

Beamformers are only effective if the transient noise originates from behind or from the side of the listeners while the target signal is situated in front.[8] Peak clipping creates severe signal distortions, which can be mitigated to some degree by the use of fast-acting automatic gain control (AGC) algorithms. Both methods, however, suffer fundamentally from the fact that their behavior is determined by the absolute signal level only, whereas transient noises are characterized not so much by absolute level as by fast level changes.[7] Thus, the task of a TNR algorithm is to react to fast-rising signal levels rapidly and attenuate them with minimal distortion of the target signal.

Previous studies with HA users have investigated the subjective annoyance of transient noises[1,2] and the effect of TNR algorithms on a variety of outcome measures including speech intelligibility, sound quality and listening comfort.[7,9-11]

In all but one of these studies, a benefit of TNR with regard to listening comfort could be shown for at least some of the interferer signals used. Out of two studies measuring speech intelligibility, one showed a small improvement of 4.4% due to TNR,[9] whereas the other one showed no effect.[10] Considering these partly inconsistent results obtained with HA users, we asked whether CI users can benefit from TNR and how this compares to the aforementioned studies. In this paper, we tested a commercially-available TNR algorithm with customized parameter settings. The evaluation comprised speech tests in noise and subjective ratings of

Non-commercial use only

speech clarity, comfort and overall preference. All experiments were conducted in compliance with the declaration of Helsinki and approved by the Freiburg Ethics Commission (permission number 010/1315).

Materials and Methods

Participants

Twelve experienced CI users participated in this study. The average age of the participants was 61 years, ranging from 45 to 75 years. They were implanted with a HiRes90k or CII implant of the CI manufacturer Advanced Bionics on at least one side and had a listening experience with their CI of 6 years on average, ranging from 3 to 12 years. All of them were wearing a Harmony behind-the-ear speech processor in their everyday life, fitted with the coding strategy HiRes F120.[12] Some of them were using the DNR algorithm ClearVoice which is designed to reduce stationary noise.[5]

Settings

The TNR algorithm under study was provided by HA manufacturer Phonak. It has been available on HA models since 2006 and was primarily developed to make transient noise less annoying. It continuously monitors the slope and level of the broadband signal envelope, triggering a short-term gain reduction if these features exceed a certain threshold criterion. The detector is situated in a low-delay signal path in parallel to the remaining, slower-acting signal processing steps, which effectively allows transients to be attenuated in the output instantaneously at the time of occurrence.

For the purpose of this study, we used an implementation of the algorithm running on a laptop. This enabled us to choose arbitrary settings for the internal algorithm parameters outside the restricted fitting range of the commercial product. All test signals were stored on the laptop and preprocessed token by token during testing. A Roland Cakewalk UA-1G sound card was used for D/A conversion. The generated waveforms were delivered to the speech processor electrically using a Direct Connect cable and ear hook.

Testing was conducted unilaterally. The contralateral ear was obstructed with an earplug, if the participant had residual hearing. If the participant was bilaterally implanted, testing was conducted on the better side with the second speech processor turned off. All participants were provided with the same speech processor for testing which was programmed with each participant's individual everyday program. Where applicable, ClearVoice noise reduction was turned off and the input sensitivity adjustment set to 0 in order to avoid confounding effects from these settings. All participants were tested using the standard clinical AGC. It consists of a slow and a fast acting compressor with linear gain below knee point and a compression ratio of 12:1 above (Table 1).[13] The choice of the active compressor is input-dependent and designed to maintain the full amplitude modulation depth of speech signals.

We tested the algorithm in the two-parameter settings TNR_{high} and TNR_{low} in addition to the unprocessed reference condition TNR_{off}. The settings were identical for all subjects and noises. They were determined in advance to yield near-optimal transient detection on a test signal in terms of the number of false alarms given their total hit rates. The test signal was hammering noise at a fixed rate of 3 Hz mixed with speech where both signals were varied in amplitude to cover a wide range of possible speech and noise levels. TNR_{high} was chosen to have a higher hit rate than TNR_{low}, *i.e.* it would detect and attenuate more transients overall (Figure 1).

Table 1. Technical data of integrated automatic gain control in cochlear implant speech processor.

	Attack time (ms)	Release time (ms)	Knee point (dB SPL)
Compressor fast	3	80	71
Compressor slow	240	1500	63

SPL, sound pressure level.

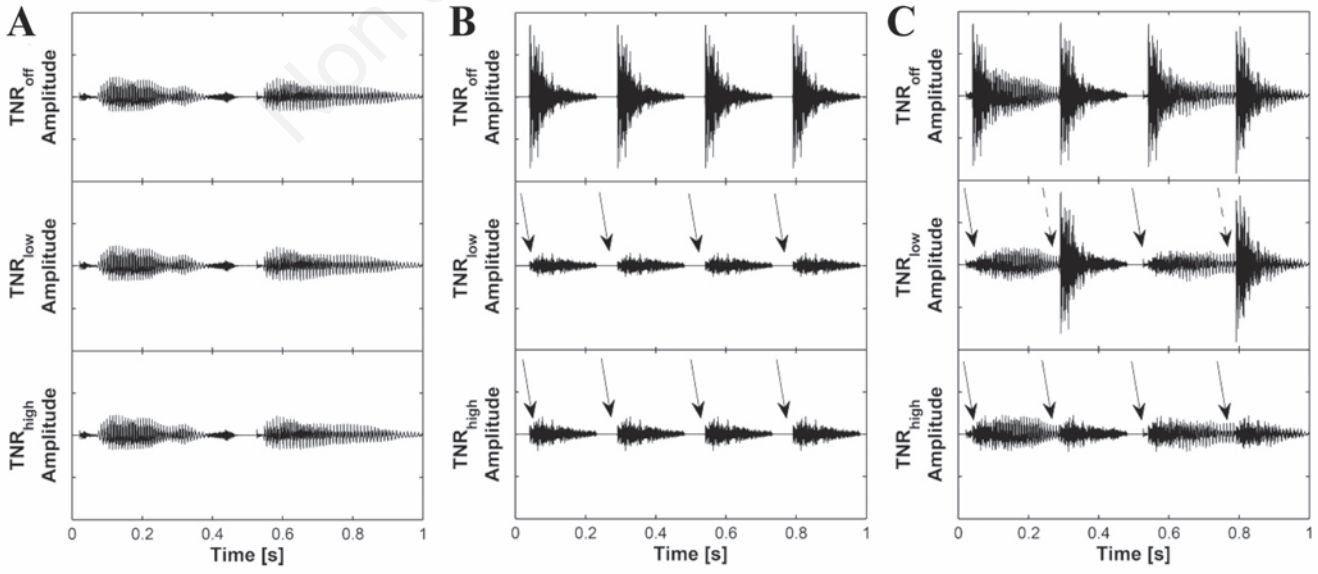


Figure 1. Example of three signals processed with transient noise reduction TNR_{off}, TNR_{low} and TNR_{high}. A) shows speech only; B) hammering noise only; and C) both mixed. Solid arrows indicate instances where the TNR algorithm triggered correct short-term signal attenuations, whereas dashed arrows indicate false negatives.

Non commercial use only

Test procedure and stimuli

One appointment per participant was scheduled for the study, each comprising three sets of tests: speech test in noise, speech test in quiet and subjective preference rankings (Table 2). The order of stimulus conditions (noises and program settings) and test lists was randomized across participants to rule out order effects. Breaks were given in between sets. All participants signed a consent form before the testing session.

Speech tests in noise were conducted for two types of transient noise: repetitive hammer blows and dishes (clinking cups and spoons). The hammer blows occurred at a rate of 4 Hz, regularly and with equal amplitude. Following onset, the amplitude of each individual blow decayed approximately exponentially with a time-constant of 50 ms. Dishes clinking provided a more realistic noise, consisting of transients of varying durations and amplitudes. The recorded signal had a total duration of 9 s and comprised a series of individual transients of 100 to 300 ms in duration, which occurred at an average rate of approximately 5 Hz. On each test trial, playback of the noise started from a different, random initial position within the signal. Both noises were scaled to a peak level of 90 dB sound pressure level (SPL), corresponding to a RMS level of approximately 70 dB SPL for both types. Noise levels were chosen with the intention of obtaining realistic conversational speech levels at the 50% speech reception threshold (SRT). This could not be achieved without also triggering the AGC due to its threshold level of 63 dB SPL (Table 1). The noise levels were comparable to levels used by Keidser *et al.*[10] and softer than used by DiGiovanni *et al*.[9]

The speech testing in noise was conducted with the Oldenburg sentence test (OLSA),[14] which determines the SRT adaptively, *i.e.* the signal-to-noise ratio at which a subject can correctly identify 50% of the spoken words. For this purpose speech level is altered in gradually decreasing steps (approaching the SRT) dependent on the number of correctly recognized words, while the noise level stays unchanged. In the OLSA test, each sentence has a length of 5 words with identical grammatical structure (name, verb, numeral, adjective, object noun). Ten different words are available for each position resulting in a total size of 50 words for the entire test corpus. Twenty grammatically correct but semantically unpredictable sentences of this kind are combined into one test list. Each combination of program settings and noises (3 settings x 2 noises) was tested with two lists. Two to four training lists were performed before the first test. For each sentence the onset of the noise preceded that of the speech signal by 1 s. This allowed the AGC to settle after the preceding silent interval and provided a cue for the onset of the next sentence to the subject.

The speech test in quiet was done to find out if the algorithm had an influence on sentences at a typical conversation loudness level. For that we used the Hochmair-Schulz-Moser sentence test (HSM),[15] which consists of real-life open-set sentences of different lengths. Each test list contains 20 sentences separated by short, fixed-duration pauses during which subjects have to repeat the previous sentence. The speech presentation level was equivalent to 65 dB SPL. White noise at a level equivalent to 25 dB SPL was added to the signal in order to emulate the speech processor microphone noise floor. Each program setting was tested with two sentence lists. The second list was skipped if a participant was close to ceiling performance with over 90% correct words.

In the subjective rankings the participants were asked to rank-order the three program options (TNR_{off}, TNR_{low} and TNR_{high}) for different signals according to the criteria speech clarity, listening comfort and overall preference. For this purpose, a hand-held touch screen was used by the subjects with which they could control the playback of the sound tokens. These tokens corresponded to the three program options for each combination of input signal and subjective criterion at a time. The same speech and noise signals were used as in the preceding speech tests (OLSA sentences in noise, HSM sentences in quiet). Noise was presented at a peak level of 90 dB SPL as before and the speech was scaled individually to an RMS level between 65 dB SPL and 75 dB SPL, at which self-reported speech understanding was barely acceptable and effortful. Comfort was assessed for sentences in noise and noise only (hammering and dishes). Speech clarity was assessed for sentences in quiet and sentences in noise and the overall preference was assessed only for sentences in noise.

Statistical analysis

For a statistical evaluation, differences between the three program settings TNR_{off}, TNR_{high} and TNR_{low} for all sets of tests were analyzed using Friedman's test (a non-parametric analog to a repeated-measures analysis of variance) followed by Conover's post-hoc test[16] for pair-wise contrasts whenever the Friedman test indicated a significant difference in the medians (α=0.05). Non-parametric tests were chosen because the data within the groups of measurements to be compared were not all normally distributed (Shapiro-Wilk-Test, α=0.05).

Results

Speech test in noise

Friedman tests showed a significant effect of processing on SRTs for both hammering noise (P<0.013) and dishes noise (P<0.028). For hammering noise, Conover's post-hoc test revealed a significant improvement (*i.e.,* reduction) in median SRTs over TNR_{off} of 1.7 dB for TNR_{high} and 0.6 dB for TNR_{low}. For clattering noise, TNR_{high} performed 0.4 dB better than TNR_{off}. All other pair-wise contrasts were not significant (Figure 2).

Speech test in quiet

Friedman's test did not show a significant difference between programs for speech in quiet (P=0.59; Figure 3). Thus, speech was not negatively influenced by TNR_{high} or TNR_{low}.

Subjective rating

Amongst the subjective preference rankings regarding speech clarity, only the Friedman test for speech with hammering noise showed a significant difference between the three program options (P<0.006). No effect was observed for speech only or speech with dishes noise. Post hoc tests for speech with hammering noise showed a significantly better average ranking for both TNR_{low} and TNR_{high} compared to TNR_{off} (P<0.05; Figure 4A). Friedman's test revealed no significant effect of program options on either comfort or overall preference (Figure 4B and C).

Table 2. Overview of the conducted tests and associated stimuli.

Test	Stimuli
Speech test in noise	OLSA speech + hammering/dishes
Speech test in quiet	HSM speech
Subjective rating: clarity, comfort, overall	OLSA speech + hammering/dishes, HSM speech

OLSA, Oldenburg sentence test; HSM, Hochmair-Schulz-Moser sentence test.

Non-commercial use only

Discussion

We tested a TNR algorithm with CI users regarding its effect on speech intelligibility as well as subjective ratings of speech clarity, comfort and overall preference. Two parameter settings of the algorithm (TNR_{low} and TNR_{high}) were compared with the unprocessed baseline setting TNR_{off}. The two types of noise used in this evaluation were hammering and clattering dishes.

Any result, positive or negative, obtained for the hammering noise can be regarded as an extreme case due to the artificial nature of the signal which provided near-optimal conditions for the detection mechanism: all noise transients have the same rapid onset and they occur at a rate that is high enough to affect intelligibility while at the same time being sufficiently separated in time to be recognized as separate events. It should be noted that the TNR algorithm does not inherently depend on or benefit from the periodicity of the signal as such. The second type of noise, clattering dishes, is more real-life-like and varied in terms of onset and decay times, peak amplitude, spectral content and event rate. It therefore provides a much harder test-case.

For the hammering noise, both TNR_{high} and TNR_{low} improved median speech intelligibility by 1.7 dB and 0.6 dB respectively. Consistently, subjective speech clarity was also ranked higher on average with either TNR setting than with TNR_{off}. When applied to the noise of clattering dishes, only the TNR_{high} setting was able to provide a small intelligibility improvement of 0.4 dB over TNR_{off}, while TNR_{low} showed no effect. No subjective difference in speech clarity was observed for either setting. Our data show no significant difference of TNR_{high} and TNR_{low}, for either type of noise or test. Furthermore, we found no effect of TNR on subjective ratings of listening comfort and overall preference.

We can at this point only speculate why comfort was not improved despite a clear effect of TNR on the input signal and speech intelligibility. We suggest that any potential effect on comfort may have been obscured in part by our technical test setup, in which the TNR algorithm and clinical AGC were put in series. Thus, any reduction in long-term signal power (a major determinant of perceived loudness) affected by TNR was counter-acted to a substantial degree by the subsequent compression mechanism. If this was the case, then a different coupling mechanism between these two components (such as letting them act in parallel rather than in series) should at least facilitate an effect of TNR on comfort compared to our setup in this study. This would, however, require more profound changes to the core CI signal processing chain and remains for future investigation.

Two studies have previously examined the effect of TNR algorithms on both intelligibility and subjective ratings in HA wearers. DiGiovanni *et al.*[9] tested the TNR algorithms Antishock (Unitron) and SoundSmoothing (Siemens) and achieved a slight improvement in speech intelligibility across all noise conditions of 4.4% (4.7% Siemens, 4.2% Unitron) but not in subjective ratings. Conversely, Keidser *et al.*[10] found that SoundSmoothing did not affect speech intelligibility, but could show a benefit in subjective annoyance and preference ratings. DiGiovanni *et al.* hypothesized that this difference in findings stemmed from a difference in the amplification settings between the two studies. While they tested subjects with their own, compressive HA AGC setting, Keidser *et al.* used a linear amplification setting instead. With linear amplification, the loudness of the HA output for high-amplitude input sounds is higher than with compression. This is very likely to make it more annoying and gives the TNR a better chance to improve comfort.[9] Another aspect of AGC potentially influencing the effect of TNR is the speed at which it acts. An AGC with fast (syllabic) time constants is less likely to overamplify sudden loud sounds for a significant amount of time than a slow-acting AGC, thus providing better comfort in transient noise. On the other hand, fast AGC is also prone to compress amplitude modulations in speech which is likely to decrease speech performance

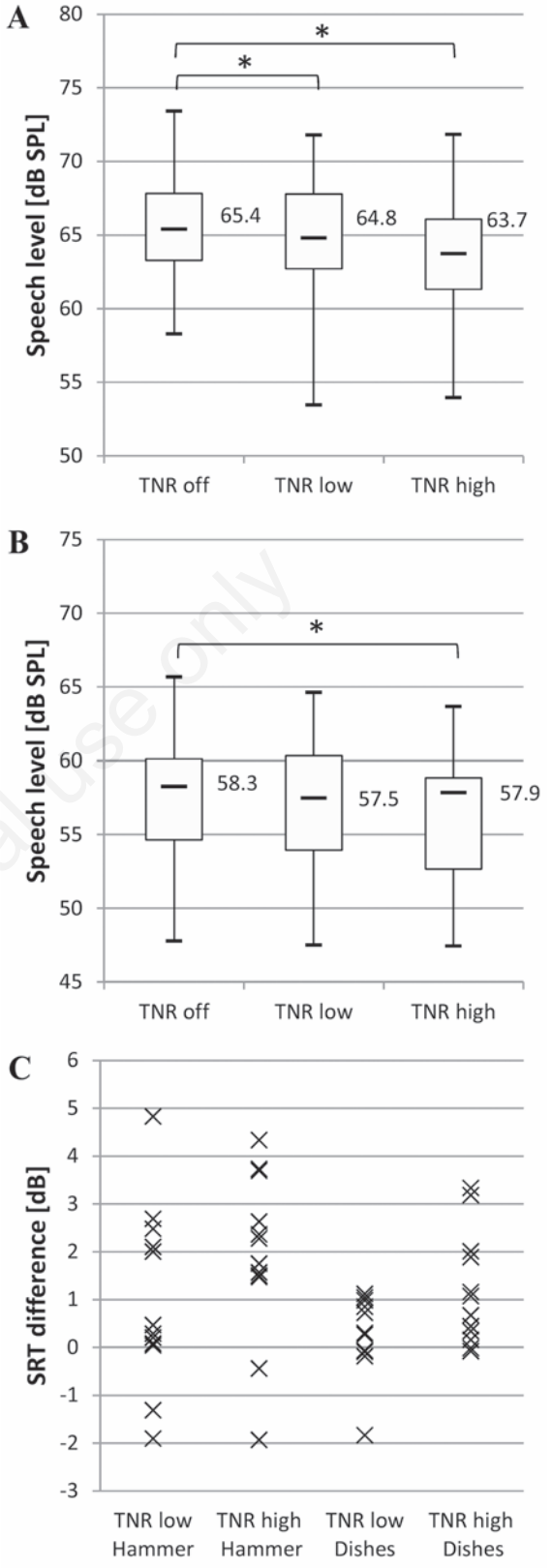


Figure 2. Speech reception threshold (SRT) results of the speech test in hammering (A) and dishes (B). Values and horizontal lines represent the median speech level at the SRT. Boxes and whiskers represent the interquartile range and minimum/maximum respectively. Significant differences in median (Conover post hoc test, P<0.05) are marked with brackets and asterisks. c) Shows the distribution of individual SRT differences for TNR_{high} and TNR_{low} relative to TNR_{off}. TNR, transient noise reduction.

Non-commercial use only

in CI users.[13] This creates a trade-off in the AGC design which may be avoided using TNR where the stimulus rise time itself is a major determinant of the attenuation, rather than absolute level as in the case of classical AGC. Furthermore, the results of DiGiovanni *et al*., obtained with syllabic multi-channel AGC, indicate that TNR can in principle improve speech performance even when the AGC acts fast. DiGiovanni *et al.* further argued that the improvements in speech intelligibility found in their own study might be related to specific differences in spectral properties of the noises. Theirs had a higher energy in the dominant speech frequency range below approximately 2 kHz than those of Keidser *et al.* It is a plausible hypothesis that attenuating such low-frequency transients could reduce non-simultaneous masking of speech components, resulting in a greater effect of the TNR algorithm. Other studies have also shown improvements in subjective preference ratings using TNR algorithms in HA users for a range of transient noises including hammering, pen tapping, clanking mugs, door slamming and others without assessing speech intelligibility altogether.[7,11]

Despite the large range of results obtained under different test conditions, previous studies have shown that TNR algorithms can lead to improved listening in noise for HA users. They also demonstrated that even the qualitative effect of TNR algorithms depends strongly on the exact characteristics of the test signals being used. These basic findings from the literature are also reflected in our results obtained with CI users. Nevertheless, our study is the first to show benefits both in speech intelligibility and in subjective ratings (speech clarity) using the same algorithm and test material on one group of subjects. In addition to methodological reasons, fundamental differences between the processes of electric and acoustic hearing may further influence the effects of TNR. Firstly, CI users are more seriously challenged by environmental noise than normal-hearing or hearing-impaired listeners owing, for example, to the decreased spectral and temporal resolution of the electrical stimulation pattern which results in a poorer ability to perceptually segregate multiple ongoing streams of acoustic events. Therefore we may, *a priori*, expect noise reduction techniques to yield

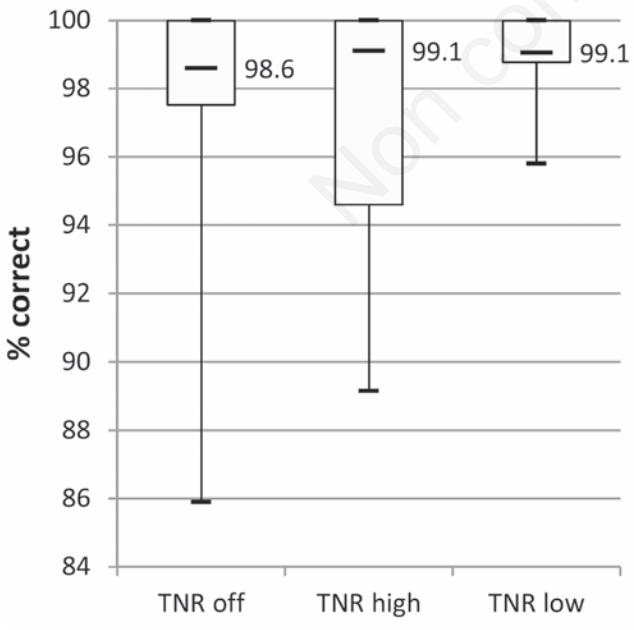


Figure 3. Percentage of correct answers of the speech test in quiet. Values and horizontal lines represent the median. Boxes and whiskers represent the interquartile range and minimum/maximum, respectively. TNR, transient noise reduction.

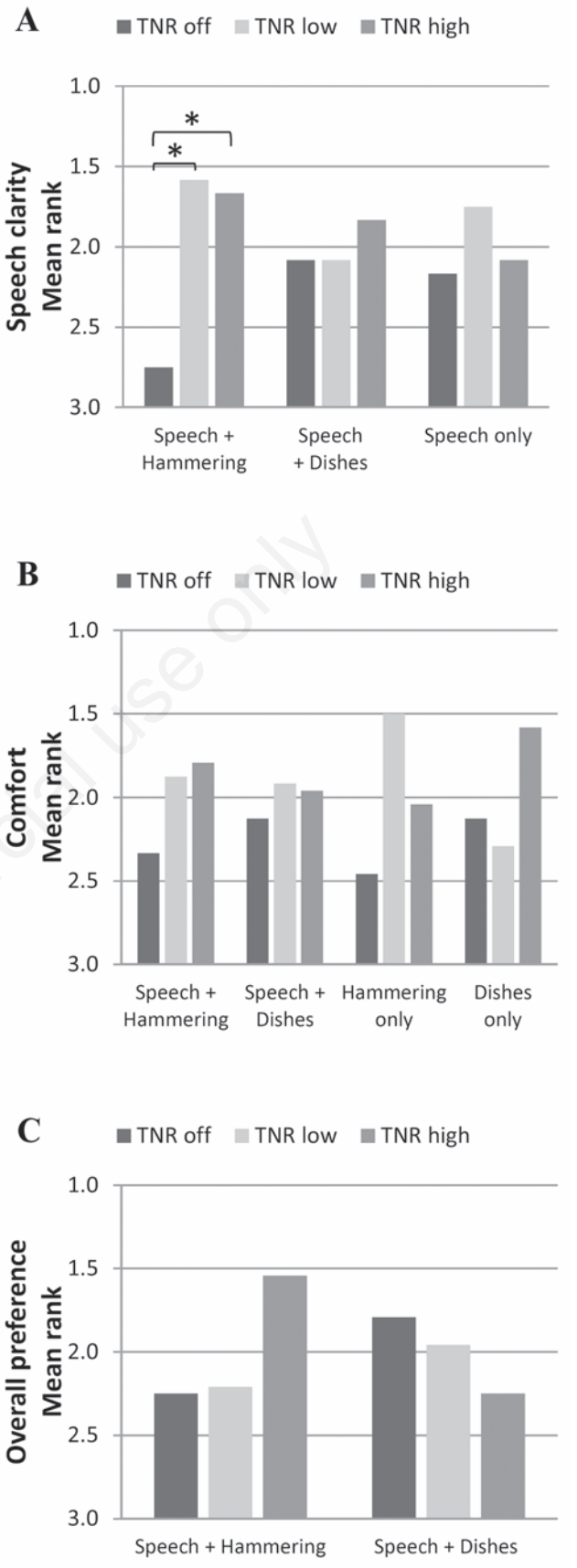


Figure 4. Mean ranks of subjective ratings regarding speech clarity (A), comfort (B) and overall preference (C). Significant effects between settings (Conover post hoc test, P<0.05) are marked with brackets and asterisks. TNR, transient noise reduction.

Non commercial use only

larger benefits in electrical hearing. Secondly, most CI users perceive sound exclusively through their speech processor, whereas HA users typically perceive a mixture between the processed output of the HA and the direct sound reaching their ear drum unprocessed through a vent or other acoustic leaks. Hence, the effect of any kind of signal processing is larger in electrical hearing, as illustrated for example by the comparatively larger effect of beam forming on CI listening performance.[8]

Conclusions

This study was the first step to determine whether CI users can benefit from a TNR algorithm, which is currently available in acoustic HAs. Our results show that speech intelligibility and speech clarity in hammering noise improved. Small improvements were also observed for speech intelligibility in dishes clattering. Further studies will be required in order to assess the applicability of the algorithm in prolonged everyday use. Another direction for future investigation is to determine a suitable fitting range of the algorithm parameters in order to efficiently accommodate the varying demands of individual users.

References

1. Keidser G, Convery E, Kiessling J, Bentler R. Is the hearing instrument to blame when things get really noisy? Hear Rev 2009;16:12.
2. Hernandez AR, Chalupper J, Powers TA. An assessment of everyday noises and their annoyance. Hear Rev 2006;13:16-20.
3. Chalupper J, Powers TA. New algorithm is designed to take the annoyance out of transient noise. Hear J 2007;60:42-4.
4. Hayes D. Transient impulse control for hearing aids. Hear Rev 2006;13:56.
5. Buechner A, Brendel M, Saalfeld H, Litvak L, Frohne-Buechner C, Lenarz T. Results of a pilot study with a signal enhancement algorithm for hires 120 cochlear implant users. Otol Neurotol 2010;31:1386-90.
6. Henderson D, Hamernik. RP. Impulse noise: critical review. J Acoust Soc Am 1986;80:569-84.
7. Korhonen P, Kuk F, Lau C, Keenan D, Schumacher J, Nielsen J. Effects of a transient noise reduction algorithm on speech understanding, subjective preference, and preferred gain. J Am Acad Audiol 2013;24:845-58.
8. Buechner A, Dyballa KH, Hehrmann P, Fredelake S, Lenarz T. Advanced beamformers for cochlear implant users: acute measurement of speech perception in challenging listening conditions. PloS One 2014;9:e95542.
9. DiGiovanni JJ, Davlin EA, Nagaraj NK. Effects of transient noise reduction algorithms on speech intelligibility and ratings of hearing aid users. Am J Audiol 2011;20:140-50.
10. Keidser G, O'Brien A, Latzel M, Convery E. Evaluation of a noise-reduction algorithm that targets non-speech transient sounds. Hear J 2007;60:29-32.
11. Liu H, Zhang H, Bentler RA, Han D, Zhang L. Evaluation of a transient noise reduction strategy for hearing aids. J Am Acad Audiol 2012;23:606-15.
12. Brendel M, Buechner A, Krueger B, Frohne-Buechner C, Lenarz T. Evaluation of the harmony soundprocessor in combination with the speech coding strategy hires 120. Otol Neurotol 2008;29:199-202.
13. Boyle P, Buechner A, Stone M, Lenarz T, Moore BC. Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants. Int J Audiol 2009;48:211-21.
14. Wagener K, Kühnel V, Kollmeier B. Development and evaluation of a German sentence test part I: design of the Oldenburg sentence test. Audiol Acoust 1999;38:4-15.
15. Hochmair-Desoyer I, Schulz E, Moser L, Schmidt M. The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. Am J Otolaryngol 1997;18:S83.
16. Conover WJ. Practical nonparametric statistics. 3rd ed. New York: Wiley; 1999.

Non commercial use only