How to compare hearing-aid processing of real speech and a speech-modified stimulus for objective validation of hearing-aid fittings?

SØREN LAUGESEN[1,*]

[1] *Interacoustics Research Unit, Kgs. Lyngby, Denmark, DK-2800 Lyngby, Denmark*

A method is proposed to evaluate whether a modern hearing aid with all automatic features enabled processes a speech-modified stimulus for objective validation of hearing-aid fittings as speech. The proposed method measures short-term coupler gains from brief snippets of steady-state probe noise, crossfaded into either the International Speech Test Signal (ISTS) or the speech-modified stimulus, which thus act as conditioning signals. For reference, the method is also applied to a steady-steady noise signal, which drives the hearing aids into noise mode. Results for a selection of hearing aids show that the method classifies the hearing aids' mode of processing according to expectations, with all three conditioning signals.

INTRODUCTION

For the purpose of validating hearing-aid fittings in prelingual infants, an objective assessment based on the auditory steady-state response (ASSR) is considered. One important aspect of an appropriate assessment is to ensure that speech-relevant gain and signal-processing features are activated in the hearing aids during the measurement. To avoid modifying the hearing-aid settings for the validation measurement, a family of speech-modified ASSR stimuli has been devised. The preferred member of this stimulus family consists of three bandlimited CE-Chirps® (Elberling and Don, 2010) presented at different repetition rates, individually modified by frequency-band specific envelopes derived from the International Speech Test Signal (ISTS; Holube *et al.*, 2010), and scaled in level to match the long-term ISTS band levels as described by Laugesen *et al.* (2018).

Prior to testing, it needs to be verified that the speech-modified ASSR stimulus in fact drives the hearing aid into speech mode. As a benign example, the Genie fitting software for e.g. the Oticon Alta and Sensei hearing aids offers 'Live Demonstration' of the current classification of the incoming soundscape, which in tests clearly indicates that the speech-modified ASSR stimuli are processed as if they were speech, whereas CE-Chirp stimuli without speech-modifications are classified as noise. However, most hearing aids' fitting tools do not offer such feedback about mode of processing and therefore a 'black-box' measurement method is proposed in this paper. Under the assumption that the ISTS properties applied to the speech-modified ASSR stimuli are used by the hearing aids for detecting speech, the method does not require any specialist hearing-aid brand knowledge and can therefore be used broadly to verify correct processing of speech-modified ASSR stimuli also in the clinic.

*Corresponding author: slau@iru.interacoustics.com

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 7: Auditory Learning in Biological and Artificial Systems, August 2019, Nyborg, Denmark. Edited by A. Kressner, J. Regev, J. C.-Dalsgaard, L. Tranebjærg, S. Santurette, and T. Dau. The Danavox Jubilee Foundation, 2019. ISSN: 2596-5522.

MATERIAL

The hearing aids tested were Oticon Alta Pro, Sensei Pro Power, and Opn 1; GN Resound LiNX 3D 9 and Quattro; and Phonak Sky B90-M and B90-P. They were all programmed according to standard audiograms from (Bisgaard *et al.*, 2010), either N3 (moderate hearing loss) or N5 (severe hearing loss) as reported below, using the respective fitting software's suggested prescription, except that WhistleBlock was enabled and SoundRecover2 was disabled for the Sky aids (initially). As already mentioned, the Genie fitting software for the Alta and Sensei aids allows for feedback about the processing mode; this was not available for the other aids.

All measurements were taken in an Interacoustics TBS25 test box. The input to the hearing aid was recorded with a G.R.A.S. 40BL ¼" microphone, located next to the hearing aid's microphone inlets at the reference position in the test box. The output was recorded with a G.R.A.S. RA0045-S1 ear simulator to which the hearing aid was connected as appropriate for the type of aid (Sensei, LiNX 3D, and Sky behind-the-ear with sound tube; Alta, Opn 1, and Quattro receiver-in-the-ear with closed 10-mm domes). Playback and recording were accomplished through an RME Fireface UC soundcard connected to a standard laptop running custom Matlab software. Microphone sensitivities were calibrated relative to a G.R.A.S. 42AB Sound Calibrator prior to recording. Furthermore, the playback chain frequency response (considering both magnitude and phase) was equalised with a 16384-tap FIR filter.

The specific stimuli used were the ISTS, a steady-state noise (SSN_{ISTS}) delivered together with the ISTS having the same spectrum, and the 3B ISTS-modified CE-Chirp®. The latter consists of a two-octave wide low-frequency chirp centred at 707 Hz, and two one-octave wide chirps centred at 2000 and 4000 Hz. The repetition rates for the three chirps were 38.1, 68.4, and 69.3 Hz, respectively.

METHOD

The family of speech-modified ASSR stimuli is based on different combinations of the narrow-band (NB) CE-Chirps® (Elberling and Don, 2010), which are one-octave wide chirps centred at 500, 1000, 2000, and 4000 Hz. Accordingly, the current method of hearing-aid evaluation compares the reference gain applied to the ISTS with the gain applied to the speech-modified ASSR stimulus in matching one-octave analysis bands. It is thus assumed that all hearing aids will classify the ISTS as speech in agreement with current standards (IEC 60118-15, 2012; IEC 61669, 2015).

To determine the time evolution of gain for each signal (and in each frequency band), inspiration was taken from Naylor and Johannesson (2009) who determined hearing-aid gain trajectories from time-aligned input and output signals in 10-ms time windows. The short windows were selected to capture the fastest gain-modifying behaviour of the hearing aids: wide dynamic-range compression or output limiting. Naylor and Johannesson compared time-averaged gain values across different settings of hearing-aid compression for speech and noise mixed at different signal-to-noise ratios. Accordingly, the original idea of the present investigation was to compare such gain trajectories measured with either the ISTS or the speech-modified ASSR stimulus passed through the hearing aid, assuming they would be very similar if the hearing aid was in speech mode. However, because the detailed waveforms of the two signals are markedly different when evaluated in 10-ms time windows, as can be seen from Figure 1 below, a direct comparison turned out to be difficult, even when the (Oticon Alta) hearing aid was in speech mode for both signals, according to the fitting software.

Evaluation by probe snippets

Instead, short probe snippets of the SSN_{ISTS} are crossfaded into the two signals. Thus, the ISTS and the speech-modified ASSR stimulus are merely used as conditioning signals, whereas the actual gain comparison is based on the probe snippets. This approach assumes that the probe snippets are short enough and occur rarely enough that the hearing aid remains in speech mode throughout the recording. This assumption is justified as all modern hearing-aid noise-reduction systems have a built-in activation sluggishness to avoid overly rapid switching back and forth between processing modes. For the same reason, the conditioning signal will be allowed to run without probe snippets for a while to ensure the hearing aid has settled into a stable processing mode. Thus, the method is characterised by four parameters, where T_S is the settling time (with unmodified conditioning signal), T_P is the duration of the probe snippets, and T_I is the time interval between successive probes, see Figure 1. The fourth parameter, N_T, denotes the total number of probe snippets included. Raised-cosine gates with 1-ms rise and fall times were used to crossfade between conditioning signal and probe snippets. At each probe interval, the exact same probe snippet is used with each conditioning signal, whereas different snippets are used at successive intervals. The selection of the parameter values (T_S, T_P, T_I, and N_T) is a compromise among (i) obtaining enough probe-signal recording time for a reliable gain evaluation, (ii) ensuring that the hearing aid remains in speech mode irrespective of the probe snippets, and (iii) total measurement time. This compromise will be explored below.

For further reference, measurements were also made with the SSN_{ISTS} as the conditioning signal. It was verified that this signal drove the Oticon Alta and Sensei aids into noise mode, allowing the effects on hearing-aid gain to be observed.

Time alignment

In contrast to Naylor and Johannesson (2009), who examined mild non-linearities related to dynamic-range compression, most modern hearing aids pose an additional challenge as they comprise highly non-linear signal processing such as frequency-shifting for suppressing acoustic feedback. This means that classical linear cross-correlation methods for time alignment of input and output signals from the hearing aid are useless, typically from around 1000 Hz and upwards in frequency where the frequency-shifting is employed. Consequently, the time alignment for the present method is based solely on the octave-band filtered signals centred at 500 Hz. In practice, the time alignment is done between the digital stimulus and the recorded input and output from the hearing aid, respectively. In this way, the probe snippets can be isolated from both the recorded input and output signals for analysis.

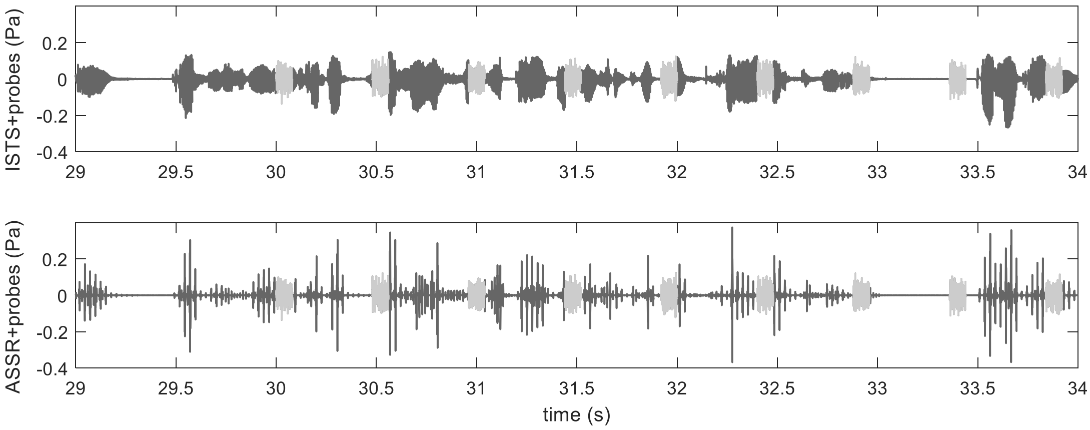

Fig. 1: Measurement signal excerpts with probe snippets indicated by light grey crossfaded into the ISTS (top) and the 3B ISTS-modified CE-Chirp ASSR stimulus (bottom panel). The first probe snippet occurs at $T_S = 30$ s, the probe duration is $T_P = 80$ ms, and the between-probe interval is $T_I = 0.4$ s.

Gain comparison metric

The respective sequences of probe snippets in each analysis frequency band and at the input and output of the hearing aid are first concatenated. Then, coupler-gain trajectories are computed as successive ratios of output and input RMS values taken across 10-ms rectangular windows in each frequency band, see Figure 2 for examples measured in the Oticon Alta and the Phonak Sky B90-M aids. Note that the examples in Figure 2 deliberately were measured with the settling time set to zero and a high number of probes in order also illustrate the course of noise-reduction activation. In addition, an extreme setting with noise-reduction parameters set to maximum values was chosen for the Sky aid to illustrate the possible range of gain reduction.

There are several important observations to make from Figure 2. First, the gain trajectories for the ISTS and the speech-modified ASSR stimulus overlap in all cases, as expected. The gain trajectories for the SSN_{ISTS} signal shows rather different activation profiles for the two aids: the Alta aid's noise reduction is fully activated after $n = 160$ analysis windows, which translates to 9.6 s (8 analysis windows per 80-ms probe and 0.4 s between-probe intervals, $\frac{160}{8}(0.08\ \mathrm{s}+0.4\ \mathrm{s})=9.6\ \mathrm{s}$), while the Sky noise reduction is activated in two stages with full activation after $n = 200 \sim 12$ s. When fully activated, the Alta aid provides about 5 dB of gain reduction at 500 Hz and almost none at 4000 Hz, while the Sky aid provides about 17 and 10 dB gain reduction at the two analysis frequencies. The exact numbers will possibly depend on the precise orientation of the hearing aid in the test box, since adaptive directionality may be activated by the SSN_{ISTS} signal. (This was, however, not the case for the Alta aid as observed through the fitting software.)

To form a single-valued metric of gain comparison between two signals A and B (with added probe snippets and appropriate settling time), median coupler gain values in dB are determined in each frequency band and for each of the gain trajectories, $g_{Ak}(n)$ and

$g_{Bk}(n)$, where n denotes the analysis time-window index and k denotes the frequency-band index. Thus, the proposed gain comparison metric is the maximum absolute difference in median gain across the four frequency bands:

$$C = \max_k |\mathrm{med}\{g_{Ak}(n)\} - \mathrm{med}\{g_{Bk}(n)\}|. \qquad \text{(Eq. 1)}$$

Inspired by the results presented in Tables 1 and 2 below, the proposed criterion value is C_{crit} = 1 dB. That is, if C is below 1 dB, the two signals A and B are considered as being processed similarly and *vice versa*. A gain error of 1 dB is small compared with the expected variation from practical sound-field measurements.

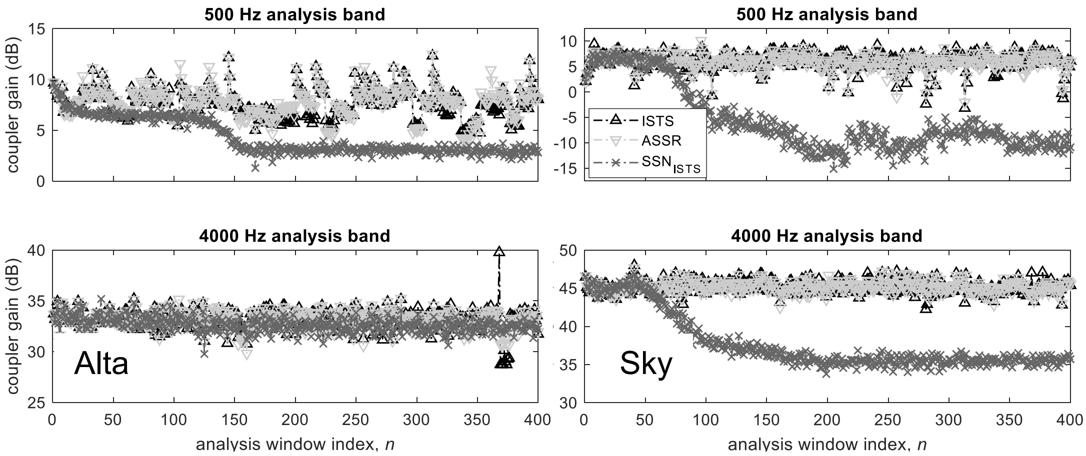

Fig. 2: Gain trajectories for the 500 and 4000-Hz octave bands and 10-ms analysis windows indexed by n, measured in the Oticon Alta (left panels, N3 hearing loss) and Phonak Sky B90-M (right panels, N3 loss). Probe parameters as in Figure 1, except the settling time which was set to $T_S = 0$. Note the different ordinate-scale ranges in all panels.

RESULTS

The results in Figure 2 as well as data from Bentler and Chiou (2006) suggest that a fail-safe choice of settling time is T_S = 30 s. In this way, the proposed test method will be able to fulfil its purpose, by allowing the hearing aid under test enough time to fully activate noise reduction, if – contrary to expectations – the speech-modified ASSR stimulus is classified as noise by the hearing aid. Next, the selection of the remaining analysis parameters is considered. The results across different parameter selections are compiled in Table 1, in terms of the gain comparison metric, C from eq. 1, for comparisons between ISTS and the 3B ISTS-modified CE-Chirp® ASSR stimulus, as well as between the ISTS and SSN_{ISTS}. In addition, the total recording time for one signal is stated. The measurements for Table 1 were all done with the Oticon Alta aid, which allowed the ground truth speech/noise classification to be examined through the fitting software, as described above. The results in Table 1 show remarkable robustness of the method towards variation in the analysis parameters. With all parameter combinations, the classification of the ASSR and SSN_{ISTS} signals is correct according to the suggested 1-dB criterion value. Varying the probe length, T_P, from

40 to 160 ms has no effect, which is to be expected from the results in Figure 2 where the onset of noise reduction is observable only from about $n = 40 \sim 2.4$ s. This assumes that the switching mechanism is reset as soon as the probe snippet stops. This seems to be the case for the Alta aid, since lowering the probe interval, T_I, from 4 s down to as little as 40 ms had only marginal effect on the results in Table 1. For the two conditions with $T_I = 40$ ms, the fitting software would very occasionally and briefly change the classification from 'Speech' to 'Speech in noise' during the ISTS and ASSR recordings. In all the other conditions the classification remained stable at 'Speech'. The SSN_{ISTS} was always classified as 'Noise'. Finally, the number of probe snippets, N_P, was varied with no observable effect to the results. Further reduction of N_P was not considered, since the total measurement time was already dominated by the 30-s settling time with $N_P = 30$. The results in Table 1 allows for selecting parameters with a broad safety margin in consideration of hearing aids potentially keener to switch to noise mode. Thus, probe duration was set to $T_P = 80$ ms, not exceeding typical syllable duration, and the between-probe interval was set to $T_I = 0.4$ s, (as used in Figures 1 and 2). Using these parameters, additional hearing aids and alternative settings were tested, with the results shown in Table 2.

T_P	T_I	N_P	T_{total}	C (ISTS vs. ASSR)	C (ISTS vs. SSN_{ISTS})
40 ms	4 s	120	511 s	0.8 dB	4.2 dB
80 ms	1 s	60	94 s	0.5 dB	4.1 dB
80 ms	0.4 s	30	44 s	0.4 dB	4.5 dB
80 ms	40 ms	60	37 s	0.4 dB	3.2 dB
160 ms	40 ms	30	36 s	0.2 dB	3.2 dB

Table 1: Total recording time for one signal, T_{total}, and the gain comparison metric, C, for two signal comparisons and various combinations of probe-method parameters. All measurements were done with the Oticon Alta hearing aid and a settling time of $T_S = 30$ s.

The results in Table 2 show the expected classification in all cases considered, that is, the gain differences between the ISTS and the speech-modified ASSR stimulus is below criterion value in all cases, whereas gain reductions above criterion value are observed for SSN_{ISTS} in all cases. The Opn 1, LiNX 3D, and Sky aids allow considerable user-defined changes to the noise-reduction (NR) parameters. Thus, in addition to measuring with the default settings described above, the NR parameters were set to maximum values, denoted NR max in Table 2. For the Opn 1 aid, settings were additionally brought way beyond what was achievable through the Genie 2 fitting software (Zaar *et al.*, 2020); this setting is denoted NR++. The grossly non-linear frequency-shifting features, Speech Rescue (SR) in Opn1, Sound Shaper (SS) in LiNX 3D, and SoundRecover2 (SR2) in Sky, were also enabled for separate measurements. In all these extreme cases, the classification was correct according to the results in Table 2. Finally, repeat recordings were made for the Sky B90-M aid showing very minor variations in C-values within ±0.1 dB, which further testifies to the robustness of the method (results not shown in Table 2).

Hearing aid	**Setting**	***C*** (ISTS vs. ASSR)	***C*** (ISTS vs. SSN_{ISTS})
Oticon Sensei	N5	0.1 dB	4.7 dB
Oticon Opn 1	N3	0.5 dB	5.3 dB
Oticon Opn 1	N3, NR max	0.9 dB	13.2 dB
Oticon Opn 1	N3, NR++	0.6 dB	10.9 dB
Oticon Opn 1	N3, SR on	0.4 dB	1.5 dB
GN Resound LiNX 3D	N5	0.1 dB	4.3 dB
GN Resound LiNX 3D	N5, NR max	0.4 dB	16.6 dB
GN Resound LiNX 3D	N5, SS on	0.1 dB	4.2 dB
GN Resound Quattro	N3	0.2 dB	12.3 dB
GN Resound Quattro	N5	0.1 dB	11.0 dB
Phonak Sky B90-P	N5	0.3 dB	4.6 dB
Phonak Sky B90-P	N5, NR max	0.3 dB	10.0 dB
Phonak Sky B90-P	N5, SR2 on	0.3 dB	4.6 dB
Phonak Sky B90-M	N3	0.2 dB	7.3 dB
Phonak Sky B90-M	N3, NR max	0.3 dB	16.5 dB

Table 2: The two gain comparison metrics measured with the proposed method for a selection of hearing aids and feature settings, see text for details.

GENERAL DISCUSSION

The evidence presented above indicates that the proposed method for evaluating whether a speech-modified ASSR stimulus is processed as speech by a hearing aid works robustly and as intended, across a selection of modern hearing aids, and with no need for prior knowledge about the settings of the hearing aid. In this way, it will potentially be possible to assess in the clinic whether a given hearing aid can be used with an aided ASSR protocol for hearing-aid validation in infants, without any modifications to the hearing aid's settings. This will ensure clinical expedience and add to the face validity of the aided ASSR test.

The proposed method will continue to be verified against so far untested hearing-aid brands as well as new hearing-aid models brought to the market. One critical factor is expected to be the time-alignment, which currently assumes quasi-linear processing within the 500-Hz octave band. In this regard it should be noted that the frequency-warped filter-bank design used in the GN Resound aids created no problems despite its varying delay across frequency. Another potential challenge is future generations of hearing aids using soundscape classification based on deep neural networks, for which the methods of detection will be opaque.

Besides demonstrating the robustness of the proposed classification method, as well as the effectiveness of the speech-modified ASSR stimulus in driving the tested hearing aids into speech mode, the results in Figure 2 and Table 2 serve to illustrate the potential gain-measurement errors if due care is not taken to bring the hearing aid into speech mode for an aided ASSR recording. Thus, gain errors up to 16.5 dB were observed, which would seriously confound a measurement intending to validate a hearing-aid fitting. According to the results presented, this can be avoided by using the speech-like 3B ISTS-modified CE-Chirp® stimulus for aided ASSR.

REFERENCES

Bentler, R. and Chiou, L.-K. (**2006**). "Digital noise reduction: an overview," Trends Amplif., **10**, 67–82.

Bisgaard, N., Vlaming, M. S. M. G., and Dahlquist, M. (**2010**). "Standard audiograms for the IEC 60118-15 measurement procedure," Trends Amplif., **14**, 113–120.

Elberling, C. and Don, M. (**2010**). "A direct approach for the design of chirp stimuli used for the recording of auditory brainstem responses," J. Acoust. Soc. Am., **128**, 2955-2964.

Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (**2010**). "Development and analysis of an International Speech Test Signal (ISTS)," Int. J. Aud., **49**, 891–903.

IEC 60118-15. (**2012**). "Electroacoustics – Hearing aids – Part 15: Methods for characterising signal processing in hearing aids with a speech-like signal," IEC International Standard.

IEC 61669. (**2015**). "Electroacoustics – Measurement of real-ear acoustical performance characteristics of hearing aids," IEC International Standard.

Laugesen, S., Rieck, J. E., Elberling, C., Dau, T., and Harte, J. M. (**2018**). "On the cost of introducing speech-like properties to a stimulus for auditory steady-state response measurements," Trends Hear., **22**, 1-11.

Naylor, G. and Johannesson, R. B. (**2009**). "Long-term signal-to-noise ratio at the input and output of amplitude-compression systems," J. Am. Acad. Audiol., **20**, 161–171.

Zaar, J., Simonsen, L. B., Behrens, T., Dau, T., and Laugesen, S. (**2020**). "Investigating the relationship between spectro-temporal modulation detection, aided speech perception, and directional noise reduction preference in hearing-impaired listeners," Proc. ISAAR, **7**, 181-188.