On the Cost of Introducing Speech-Like Properties to a Stimulus for Auditory Steady-State Response Measurements

Søren Laugesen[1], Julia Eva Rieck[1,2], Claus Elberling[3], Torsten Dau[4], and James M. Harte[1]

Abstract

Validating hearing-aid fittings in prelingual infants is challenging because typical measures (aided audiometry, etc.) are impossible with infants. One objective alternative uses an aided auditory steady-state response (ASSR) measurement. To make an appropriate measurement, the hearing aid's signal-processing features must be activated (or deactivated) as if the ASSR stimulus was real speech. Rather than manipulating the hearing-aid settings to achieve this, an ASSR stimulus with speech-like properties was developed. This promotes clinical simplicity and face validity of the validation. The stimulus consists of narrow-band CE-Chirps®, modified to mimic the International Speech Test Signal (ISTS). This study examines the cost of introducing the speech-like features into the ASSR stimulus. Thus, 90 to 100 Hz ASSRs were recorded to the ISTS-modified stimulus as well as an equivalent stimulus without the ISTS modification, presented through insert phones to 10 young normal-hearing subjects. Noise-corrected ASSR magnitudes and clinically relevant detection times were estimated and analyzed with mixed-model analyses of variance. As a supplement, the observed changes to the ASSR magnitudes were compared with an objective characterization of the stimuli based on modulation power. The main findings were a reduction in ASSR magnitude of 4 dB and an increase in detection time by a factor of 1.5 for the ISTS-modified stimulus compared with the standard. Detection rates were unaffected given sufficient recording time. For clinical use of the hearing-aid validation procedure, the key metric is the detection time. While this varied considerably across subjects, the observed 50% mean increase corresponds to less than 1 min of additional recording time.

Keywords

auditory steady-state response, hearing-aid validation, speech-like stimulus, modulation power

Date received: 29 January 2018; revised: 20 June 2018; accepted: 21 June 2018

Introduction

The success of newborn hearing screening leads to very young infants being fitted with hearing aids, down to the age of about 2 months (Wood, Sutton, & Davis, 2015). Standard tools for validation of these fittings (aided audiometry, questionnaires, etc.) are either not possible or highly unreliable in these very young infants. At the age of 8 to 9 months, visual reinforcement audiometry (Bamford & McSporran, 1993; Day, Bamford, Parry, Shepherd, & Quigley, 2000) is recommended, for example, by the British Society of Audiology (2014). However, given the critical importance of providing early access to sound stimulation (Sharma, Dorman, & Spahr, 2002), a hearing-aid validation method for the 2- to 9-month age-group is needed. Therefore, objective validation methods based on electrophysiology have been investigated (e.g., Punch, Van Dun, King, Carter, & Pearce, 2016). Here, an approach using the auditory steady-state response (ASSR) is considered (Picton et al., 1998).

Aided-ASSR measurements are associated with several challenges, some of which are related to the intended

[1]Interacoustics Research Unit, Lyngby, Denmark
[2]Faculty of Sciences, Free University Amsterdam, the Netherlands
[3]Virum, Denmark
[4]Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, Lyngby, Denmark

Corresponding author:
Søren Laugesen, Interacoustics Research Unit, Ørsteds Plads, Building 352, 2800 Lyngby, Denmark.
Email: slau@iru.interacoustics.com

sound field presentation of the stimulus, which is introduced to bring the hearing aid into the stimulation path. Furthermore, as the ASSR stimulus passes through the hearing aid, it must be ensured that the correct gain is applied and that the correct signal processing features relevant for speech are activated. This can be achieved by manipulating the settings of the hearing aid, for example, by turning off helping systems such as noise reduction or directionality (Billings, Tremblay, Souza, & Binns, 2007; Carter, Dillon, Seymour, Seeto, & Van Dun, 2013; Easwar, Purcell, Aiken, Parsa, & Scollie, 2015). This is necessary because any standard ASSR stimulus, such as modulated tones or noise bands, will be classified as noise by the hearing aid due to these stimuli's lack of speech-like features. For example, with noise reduction activated in the hearing aid, the validation measurement will be misleading because of the gain reductions imposed by the noise-reduction signal processing. However, manipulating the settings of the hearing aid weakens the ecological argument that the hearing aid is in a normal mode of operation. An alternative approach is to construct an ASSR stimulus with sufficiently speech-like properties that the hearing aid automatically classifies the ASSR stimulus as speech. The benefit in terms of strengthening the counseling of the infant's parents suggests the latter approach because in that case the hearing aid can be fitted to the infant and tested in the exact same setting in which it will be used in daily life. In addition, a speech-like stimulus will corroborate the relevance of the measurement for both parents and clinicians (Mehta et al., 2017). Finally, not having to reprogram the hearing aid for the validation measurements saves time and reduces inconvenience in the clinic.

For this purpose, narrow-band (NB) CE-Chirps® (Elberling & Don, 2010) were modified to have speech-like properties, and it was verified that this stimulus indeed is classified as speech by the hearing aids tested and by observation in the fitting software. However, as the NB CE-Chirps® were designed for optimal efficiency, it is expected that any change, such as adding speech-like features to the chirps, will come at a cost of reduced ASSR magnitudes and increased detection times. For clinical purposes, the time taken to obtain detection of the ASSR is particularly important. The changes to ASSR magnitudes and detection times were investigated in the experiment presented later. In addition, the observed changes were compared with an objective characterization of the speech modified versus the standard chirps based on an auditory model-inspired analysis of modulation power.

Materials and Method

The experiments were approved by the Science-Ethics Committee for the Capital Region of Denmark. All participants provided written informed consent and received compensation for their participation in terms of gift cards.

Participants

To isolate the effects of the stimulus modifications, the experiment was carried out with young adult normal-hearing test subjects ($N = 10$, mean age 25 years) who were tested on both ears. The hearing status of the participants was examined by means of a questionnaire (ISO 389-9, 2009, Annex A), otoscopy, wide-band tympanometry using the Interacoustics Titan, and air-conduction audiometry at 250, 500, 1000, 2000, 4000, and 8000 Hz using an Interacoustics AC40 audiometer with ER-3A insert phones. All participants had hearing threshold levels of 20 dB hearing level (HL) or better at all tested frequencies and no other abnormalities that would preclude testing.

Stimuli and Apparatus

The NB CE-Chirps® consist of 4 one-octave-wide chirps centered at 500, 1000, 2000, and 4000 Hz, and the speech modifications (patent pending) were derived from the International Speech Test Signal (ISTS; Holube, Fredelake, Vlaming, & Kollmeier, 2010), as described later. First, the ISTS was filtered through one-octave-wide band-pass filters with center frequencies corresponding to the NB CE-Chirps. Second, the envelope of each ISTS band was determined by taking the absolute value and by convolution with a 16.6-ms Hamming window (Payton & Braida, 1999). The individual envelopes were then imposed on the respective chirp trains. Finally, the root-mean-square levels of the individual ISTS-modified one-octave-band chirps were scaled to match the one-octave band levels of the ISTS, when the ISTS in turn was scaled to its nominal broadband level of 65 dB sound pressure level (SPL). The latter step was also completed for the standard unmodified NB CE-Chirps, see Figure 1 (right). The chirp-band-specific repetition rates were 90.8, 94.7, 102.5, and 96.7 Hz at 500, 1000, 2000, and 4000 Hz, respectively, and the chirps in each chirp-train were presented with alternating polarity, as seen in Figure 1 (left). A section of the final ISTS-modified stimulus is shown in Figure 1 (middle). The discrepancies in the one-octave band levels at 250 and 8000 Hz occur because the chirp stimuli do not comprise stimulus bands at these center frequencies.

In a further effort to focus on the effects of the stimulus modifications, the stimuli were delivered by insert phones (Etymotic Research ER-1 fed by a Tucker-Davis HB7 headphone driver). To mimic the sound field listening conditions to be used in the ultimate
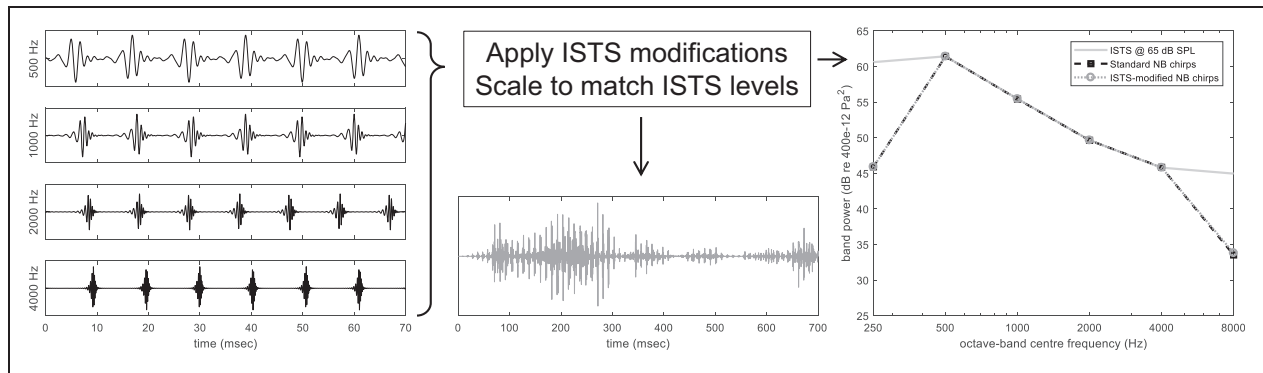
Figure 1. Sketch of the construction of the stimuli. Left: individual one-octave-band CE-Chirps. Middle: excerpt of the resulting waveform of the ISTS-modified chirps. Right: resulting one-octave band spectra compared with that of the ISTS. ISTS = International Speech Test Signal; NB = narrow-band; SPL = sound pressure level.

aided ASSR measurements, the stimuli were shaped according to real-ear measurements (REMs) in terms of the real-ear unaided gain. The REMs were performed in a $1.75 \times 1.95 \times 1.97\,\mathrm{m}^3$ sound-treated audiometry booth using an Interacoustics Affinity system. The stimulus conditioning comprised the following steps. First, the overall sensitivity of the ER-1 insert phones and the headphone driver was determined by supplying a one-octave wide noise signal to the insert phone and recording the SPL in a GRAS RA0045-S1 ear simulator, which was calibrated with a GRAS 42AB Sound Calibrator. Second, the insert phone was positioned in the ear of the test subject together with the same Affinity probe microphone used previously for the REM, taking advantage of the marks made on the probe tubes to indicate insertion depth, determined during the REM. The response of the insert phone was then recorded with the probe microphone in the individual ear using a broadband pink noise signal presented at a nominal level of 65 dB SPL. Finally, the insert phone response was corrected for the probe microphone response, which had been measured previously in an Interacoustics TBS25 test box against a calibrated GRAS 40BL reference microphone. Note that the frequency response of the ER-1 insert phones is designed to mimic an average open-ear gain. This means that the stimulus conditioning strategy only addresses the difference between the individual real-ear unaided gain and the ER-1 average open-ear gain.

ASSR Measurements

ASSR recordings were made using standard clinical four-electrode montages (high forehead ground, ipsi- and contra-lateral mastoids active, and cheek reference), with 15 min of recording time per condition, presented in each ear sequentially. The order of the test conditions was balanced across subjects. Test and retest recordings

were made for the two main conditions (ISTS modified and standard), and in addition the standard stimulus was measured at 55 and 75 dB SPL nominal levels (results not reported here). Each ear was tested in individual sessions separated by at least 1 day. With a recording time of 15 min per condition, each session involved 1½ h of recordings. Detection was evaluated online by means of a simple F-test detector on only the first harmonic of each repetition rate, see later for further details. The noise levels used for the online F tests were estimated as simple running averages across the same noise bins as described later. A 1% error rate was used, which was Bonferroni corrected for successive testing as the recording went along. A 40 μV artifact rejection level was applied in the online evaluation, while all recorded epochs were stored to disk. The test subjects were lying comfortably on a bed in a darkened, sound-treated, and electrically shielded room. The test subjects were instructed to relax and sleep if possible. Most subjects fell asleep already in the first recording, and while the subjects initially were offered breaks between conditions, only a few subjects took a break within any session. Part of the Interacoustics Eclipse EP25 ABR system was used as a front-end, with line-level signals accessed from test pins on the Eclipse unit's circuit board and routed to an RME Fireface UC soundcard that also delivered the stimuli to the insert phones. Recording and analysis of the electroencephalogram as well as the playback of the stimuli were handled by custom Matlab software. Except for the low-cut and antialiasing filters built into the RME soundcard, no filtering was applied. The sampling frequency was 32 kHz, and the analysis block ("epoch") length was 65,536 samples, corresponding to 2.048 s.

Postprocessing

The stored recordings were analyzed with a 40 μV artifact rejection level, weighted averaging (John, Dimitrijevic, &

Picton, 2001), and *F*-test detection (Dobie & Wilson, 1996) for each of the first six response harmonics separately. Note that a multiharmonic detector (Cebulla, Stürzebecher, & Elberling, 2006) as used with the standard Eclipse has yet to be developed for the ISTS-modified stimulus. The outcomes considered were as follows:

- Noise levels estimated by averaging noise power across 20 frequency bins distributed evenly around each response harmonic (fundamental as well as higher harmonics) excluding frequency bins close to harmonics of 50-Hz line noise, global system for mobile communications (GMS) interference, and any of the other repetition rate harmonics.
- Noise-corrected ASSR magnitudes (Dobie & Wilson, 1996). The estimated noise power specific to each response harmonic was subtracted from the response-bin power to yield the noise-corrected power. This was then used to compute the ASSR magnitude, which was finally converted to dB to allow analysis of the relative changes in response with stimulus type. Note that both the noise-corrected ASSR magnitude and the noise level described earlier were computed from the entire available recording, and that data were included for further analysis based on a one-shot *F* test with a 5% error rate based on the weighted averages taken across the whole recording. This was done in order to obtain the best possible estimate of ASSR magnitude.
- Detection times evaluated as the first time a response was detected according to the *F*-test detector using a 5% criterion in successive weighted averages, ignoring any corrections for repeated testing. This more lax criterion than that used during recording was used as representative of a typical clinical criterion for a threshold determination. In this way, the computed detection times and ASSR magnitudes are not directly related, as the determination of ASSR magnitudes and noise levels potentially includes data recorded after detection occurred, and thus after a clinical recording would have been terminated. Detection times were $\log_{10}$ transformed (relative to 1 s.) before statistical analysis. The log transformation was preferred over no transform based on inspection of residuals in the statistical modeling. Note that in Laugesen, Rieck, Elberling, Dau, and Harte (2017), the detection-time analysis of variance (ANOVA) was made using the untransformed data leading to slightly different results but similar conclusions in terms of which effects were significant and vice versa.

Statistical Analysis

The outcomes were analyzed with mixed-model ANOVAs with *Test ear* (a unique identifier for each ear tested) as a random effect. In a preliminary analysis, it was established that there were no systematic differences between the left and right ears in any of the outcome measures considered, and therefore the test ear rather than the test subject was chosen as the random effect variable. In addition, the statistical model included fixed effects of *Stim type* (stimulus type: standard or ISTS modified), *Stim freq* (stimulus frequency: 500, 1000, 2000, or 4000 Hz), and *Harmonic* (response harmonic: first, second,...sixth). The Statistica software package Version 13 was used.

Experimental Results

Detection Rates

Figure 2 shows the number of successful detections in terms of percentages (detection rates), evaluated from the results that were deemed detected based on the one-shot *F* test across the whole recording, as described earlier for the ASSR magnitude determination. The upper panel shows results for the first six harmonics individually, whereas the lower panel shows the detection rates accumulated across harmonics, meaning that an ASSR is deemed detected if there was a detection in either of the first *n* harmonics, $n = 1,...6$. Considering only the dominant first harmonic, the detection rates are very similar for the two types of stimuli. Individually, the higher harmonics provide fewer detections for the ISTS modified than for the standard stimulus; nevertheless, the accumulated percentages are also very similar.

ASSR Magnitude and Detection Time

Figure 3 displays ASSR magnitudes (top panels) and $\log_{10}$ transformed detection times (lower panels) for each harmonic, stimulus band center frequency, and stimulus type. Both three-way interactions in the statistical models (*Stim type × Stim freq × Harmonic*) were statistically significant, see Table 1 for a summary of the two ANOVAs. Of the 1,920 potential data points (20 Ears × 2 Repetitions × 2 Stimulus Types × 4 Stimulus Frequencies × 6 Harmonics), 1,190 ASSR magnitudes and 1,625 detection times were available for analysis. This mismatch is a consequence of the postprocessing discussed earlier and specifically occurs because in some cases, the detection-time criterion was fulfilled early in the recording, yielding a valid detection time, whereas evaluation of the complete recording resulted in a nondetection, leading to a missing ASSR-magnitude-harmonic data point.

There are several interesting observations to make from Figure 3. With increasing harmonic number, the ASSR magnitudes are generally reduced while detection times are increased, as expected. Particularly for the
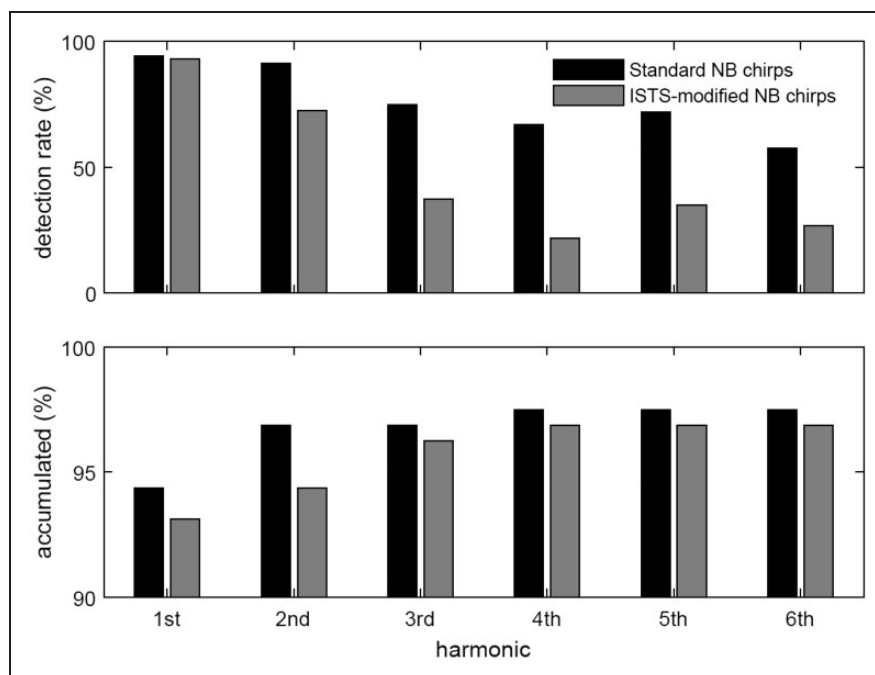
Figure 2. Detection rates across *Test ear* and *Stim freq* for each *Harmonic* separately (top) and accumulated (bottom), for each *Stim type*. Note that in order to better visualize the accumulated results, the range of the ordinate in the bottom plot is zoomed in to the 90% to 100% range. ISTS = International Speech Test Signal; NB = narrow-band.
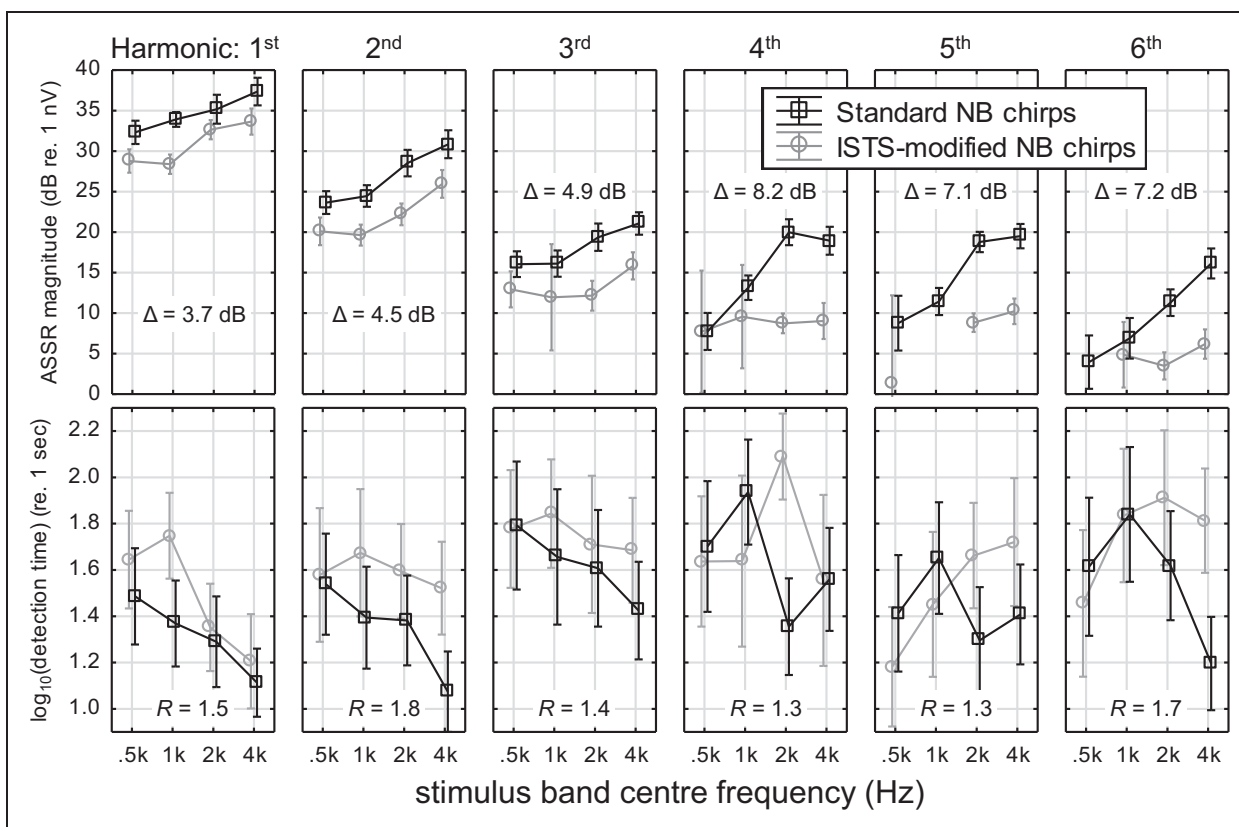

Figure 3. Noise-corrected ASSR magnitudes estimated across the entire recordings (top) and $\log_{10}$ transformed estimated clinical detection times (bottom) for each *Harmonic*, *Stim freq*, and *Stim type*, averaged across all cases in which respective detections were obtained (see text for details). Error bars indicate 95% confidence intervals. Mean *Stim type* magnitude differences, Δ, and response time ratios, *R*, are indicated in each panel. ASSR = auditory steady-state response; ISTS = International Speech Test Signal; NB = narrow-band.

Table 1. Summary Results of the Mixed-Model Analyses of Variance.

Factor	ASSR magnitude		Log_{10} (detection time)		Noise level$_{100}$	
	F statistics	p	F statistics	p	F statistics	p
Test ear (random)	$F(19, 1125) = 25$	$<.0001*$	$F(19, 1558) = 2.2$	$.0018*$	$F(19, 1159) = 36$	$<.0001*$
Stim type	$F(1, 1125) = 334$	$<.0001*$	$F(1, 1558) = 18.4$	$<.0001*$	$F(1, 1159) = 19.9$	$<.0001*$
Stim freq	$F(3, 1125) = 93$	$<.000*$	$F(3, 1558) = 8.2$	$<.0001*$		
Harmonic	$F(5, 1125) = 934$	$<.0001*$	$F(5, 1558) = 9.9$	$<.0001*$	$F(5, 1159) = 839$	$<.0001*$
Stim type × *Stim freq*	$F(3, 1125) = 7.0$	$.0001*$	$F(3, 1558) = 6.0$	$.0005*$		
Stim type × *Harmonic*	$F(5, 1125) = 4.0$	$.0014*$	$F(5, 1558) = 0.7$	$.6$	$F(5, 1159) = 2.1$	$.06$
Stim freq × *Harmonic*	$F(15, 1125) = 3.6$	$<.0001*$	$F(15, 1558) = 1.8$	$.03*$		
Stim type × *Stim freq* × *Harmonic*	$F(13, 1125) = 3.4$	$<.0001*$	$F(13, 1558) = 2.4$	$.002*$		

Note. Note that for the analysis of noise level$_{100}$, the *Stim freq* factor was disregarded.
ASSR = auditory steady-state response.
$*p < .05$.

ISTS-modified stimulus, the number of detected responses decreases toward the higher harmonics, which implies that the estimated mean values and their patterns are less reliable. Considering the difference between the two stimulus types, the ISTS-modified stimulus produces lower magnitudes and longer detection times than the standard stimulus, which again was expected. The patterns of magnitude versus stimulation frequency seem stable up to the third harmonic in terms of a constant difference between the stimuli, while greater mean differences (the inserted Δ values) and differences in slopes between stimuli are observed at the higher harmonics. The detection-time data are more variable, as indicated by the wider error bars, but it is striking that the relative increase in detection time between stimuli (the inserted R values) is almost constant across harmonics, in agreement with the nonsignificant *Stim type* × *Harmonic* interaction for detection time. This is a distinctly different pattern than the observed increase in the Δ values with harmonic number for the ASSR magnitude.

Noise Levels

Figure 4 shows the estimated noise levels across the 1,190 conditions where detection of ASSR magnitude was achieved. For this analysis, all noise levels were adjusted to the levels that would have been found at a testing time of 100 s, denoted noise level$_{100}$. This was accomplished by adding $10 \log_{10}(t/100)$ to the full-recording noise levels (in dB), where t is the full-recording time in seconds, counting only accepted epochs. Note that the full-length recording time not necessarily was 15 min because of the varying number of epochs rejected by the 40 μV artifact rejection. The corresponding ANOVA summary is included in Table 1. Note that for the noise-level$_{100}$ ANOVA, the *Stim freq* factor was disregarded because there was no reason to expect any differences related to
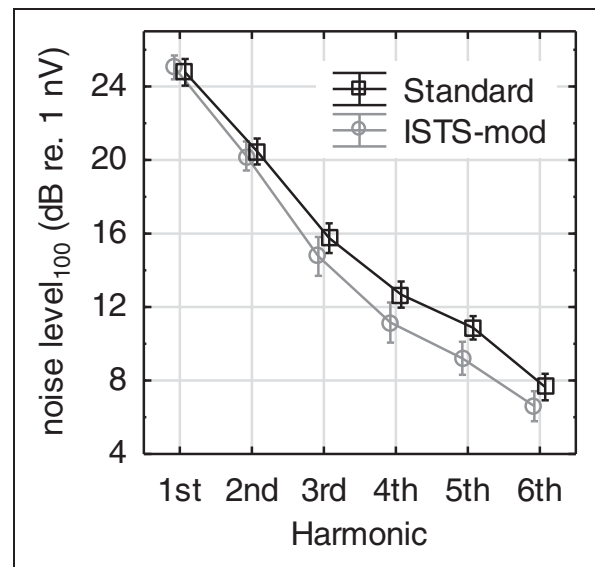

Figure 4. Mean noise level$_{100}$ across *Test ear* and *Stim freq*, including data from all cases in which an ASSR magnitude was detected. See text for details. ISTS = International Speech Test Signal.

that variable, even if separate sets of noise harmonics were used for the different repetition rates. The results in Figure 4 and Table 1 show a highly significant reduction in noise level$_{100}$ with increasing harmonic number, which was expected because biological noise typically has a $1/f$-shaped spectrum toward low frequencies (Pritchard, 1992). As seen in Table 1, the main effect of *Stim type* is also highly significant, while the plotted *Stim type* × *Harmonic* interaction just fails to meet the 5% criterion for statistical significance ($p = .06$). Finally, note that the noise level$_{100}$ values shown in Figure 4 should not be used as a detection baseline for the ASSR magnitudes in Figure 3 (top panels). This is because in most cases, the recordings extended well

beyond 100 s (up to 15 min) leading to a substantially lower noise level than those displayed in Figure 4.

Stimulus Analysis

To characterize the stimulus waveforms, an unnormalized modulation spectrum analysis was applied. This analysis was introduced under the assumption that the ASSR is driven by a nonlinear representation of the stimulus (such as rectification), here modeled in terms of envelope power. This unnormalized approach requires that the stimuli under comparison are scaled to the same level, in this case the nominal one-octave-band levels of the ISTS, as described earlier and illustrated in Figure 1.

The two stimuli (both consisting of all four NB chirps) were first passed through a bank of gammatone filters (Johannesma, 1972) corresponding to the stimulus frequency band of interest, to mimic the frequency specificity of the auditory system. Thus, for example, the analysis of the 500-Hz stimulus bands included 24 gammatone filters with center frequencies spaced 1/24th octave apart and covering the one-octave wide frequency band around 500 Hz. In addition, in order to remove epoch edge effects for the unperiodic speech-modified stimulus, all 65,536-sample epochs of each Hilbert envelope were multiplied with a raised-cosine window (using 5% at either end of each epoch for the cosine ramps) before taking the discrete Fourier transform and averaging the envelope power across epochs as well as across gammatone filter bands. Note that the steps of windowing the envelope epochs and averaging envelope power across a bank of gammatone filters were introduced after the publication of Laugesen et al. (2017). Therefore, the modulation spectra in that article are different from those shown here, while the overall conclusions are unchanged. The results are displayed in Figure 5 for the two stimuli and the two extreme stimulus band center frequencies: 500 and 4000 Hz.

For the standard stimulus (left-hand panels), the modulation power is almost entirely represented at the response harmonics, that is, the repetition rate of the respective stimulus band and its harmonics. Slight spectral splatter is observed around the dominant peaks due to the raised-cosine windowing operation. Smaller modulation power components are seen at frequencies not belonging to any stimulus repetition rate; these occur because of nonlinear interactions among the four different repetition rates that are present at the same time in the stimulus. For the ISTS-modified stimulus, the modulation power is clearly smeared out around the respective repetition rate and its harmonics. This is a consequence of the additional temporal envelope fluctuations imposed from the ISTS.

The modulation power at the response harmonics still dominates in the right-hand panels in Figure 5, but their magnitude is reduced compared with the standard stimulus. Table 2 lists the change in modulation power for all four stimulation bands and the first three harmonics. These results show that the estimated change in modulation power is very similar across the first response harmonics and among the stimulus bands. The mean reduction is $\Delta_{\mathrm{mod.power}} = 4.2\,\mathrm{dB}$.

Discussion

Stimulus Modulation Power Versus ASSR Magnitude

Two examples of data similar to those presented in Figure 3 and 5 ("physiological input/output curves") are shown in Figure 6. Note that the data sets shown in Figure 6 relate ASSR magnitude to modulation depth (converted to dB), whereas this study relates ASSR magnitude to modulation power (in dB). The relation between modulation depth and modulation power for different types of stimuli is currently under investigation. Also note that the left-hand panel in Figure 6 presents data for 40-Hz ASSR, whereas the right-hand panel shows data for 100 Hz, which is within the range of repetition rates used in this study.

Both sets of results in Figure 6 indicate a slope of about $s = 0.8$ between dB measures of modulation depth and response magnitude. Applying this to the dominant first-harmonic data from the present experiment yields,

$$\widehat{\Delta_{\mathrm{ASSR}}} = \Delta_{\mathrm{mod.power}} \times s = 4.2\,\mathrm{dB} \times 0.8\,\mathrm{dB/dB} \approx 3.4\,\mathrm{dB}$$

which agrees well with the observed $\Delta_{\mathrm{ASSR}} = 3.7\,\mathrm{dB}$ from Figure 3 (top-left panel).

In addition, the modulation power analysis reproduces the trend that the ASSR magnitudes drop more rapidly across harmonics for the lower stimulation frequencies than for the higher (Figure 3, top panels), at least considering the standard NB chirps. This is consistent with the successively fewer stimulus line-spectrum components present within an auditory (or gammatone) filter toward lower stimulus band center frequencies. For the ISTS-modified chirps, the higher order harmonic responses do not show the same trend. This may be because the higher order responses drown in the noise floor such that only the largest responses lead to detections, which creates an upward bias to the mean ASSR magnitudes.

The Cost of Stimulus Modifications

By comparing the changes to the ASSR magnitude and the detection time between the standard and the ISTS-modified stimuli, it is seen that the observed reduction in ASSR magnitudes is out of proportion with the increase
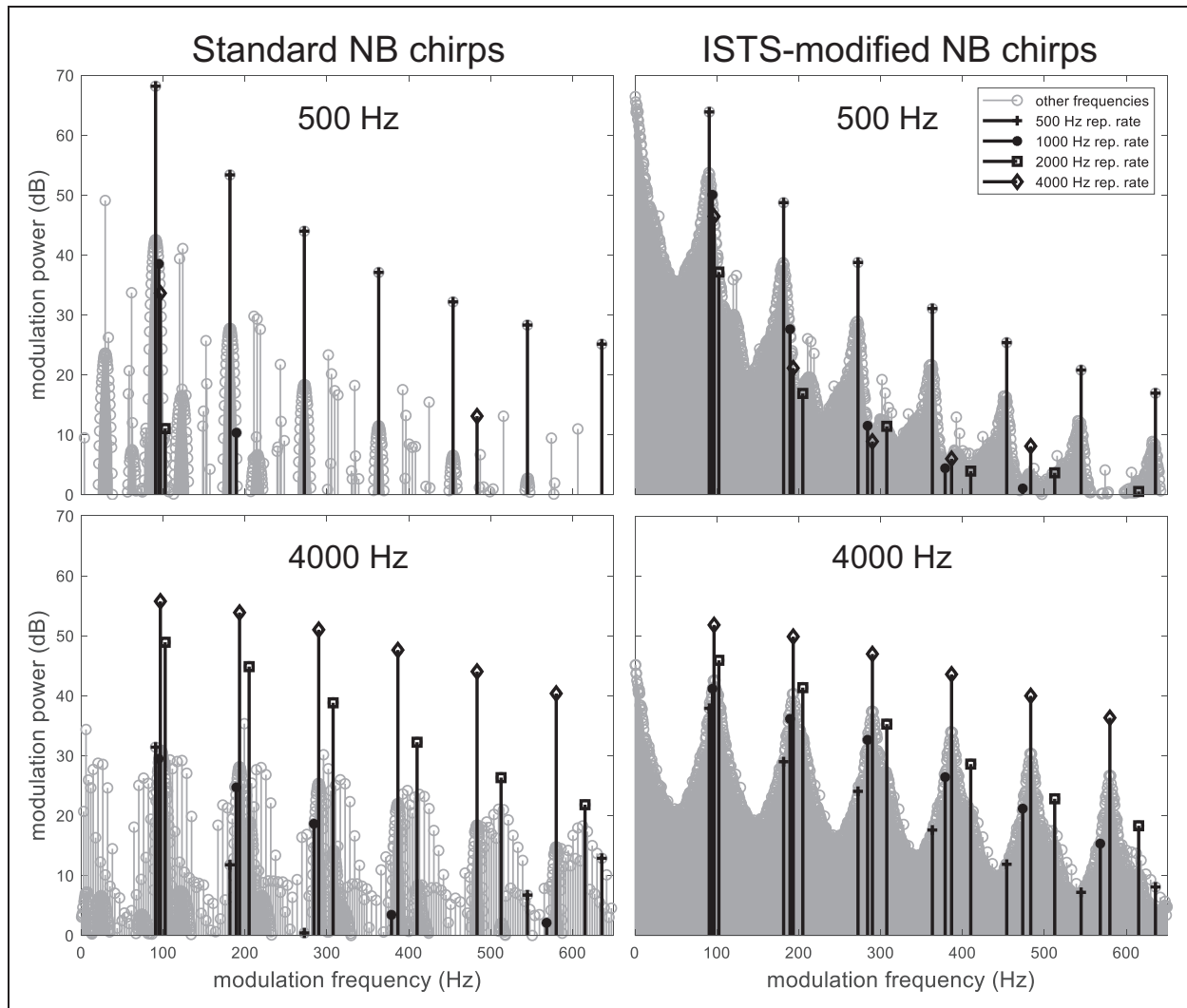
Figure 5. Unnormalized modulation power spectra of the two stimuli, computed as averages across gammatone filter banks covering the 500 Hz (top) and 4000 Hz (bottom) stimulus frequency bands, respectively. The stimulus harmonics for each stimulus frequency band are highlighted, see legend. ISTS = International Speech Test Signal; NB = narrow-band.

in detection time, particularly at the higher harmonics. For example, at the first harmonic, $\Delta_{ASSR} = 3.7\,dB$ suggests—if all else were equal—an increase in detection time by a factor of $R = 2.3$, where $R = 1.5$ was observed. Similarly, at the sixth harmonic $\Delta_{ASSR} = 7.2\,dB$ suggests $R = 5.2$, with $R = 1.7$ observed. The (albeit nonsignificant) difference in the noise level$_{100}$ patterns seen in Figure 4 hints toward a partial explanation of why the detection-time ratios increase less with harmonic number, namely, that the lower noise levels observed for the ISTS-modified chirps toward the higher harmonics partly offset the effect of lower ASSR magnitudes on detection time. Recall, however, that the detection times were determined from successively averaged weighted spectra, whereas the ASSR magnitudes were determined from weighted-average spectra across all accepted epochs of the full 15-min recordings.

Table 2. Reduction in Modulation Power (in dB) at the Response Harmonics due to the International Speech Test Signal Modifications to the Chirp Stimulus.

Response harmonic	Analysis band			
	500 Hz	1000 Hz	2000 Hz	4000 Hz
First	4.3	3.8	3.9	3.9
Second	4.6	4.0	4.0	4.0
Third	5.2	4.1	4.0	4.0

Detection Rates

Very similar detection rates are observed for the two stimuli (Figure 2). The reduced response contribution from successively higher harmonics seen for the ISTS
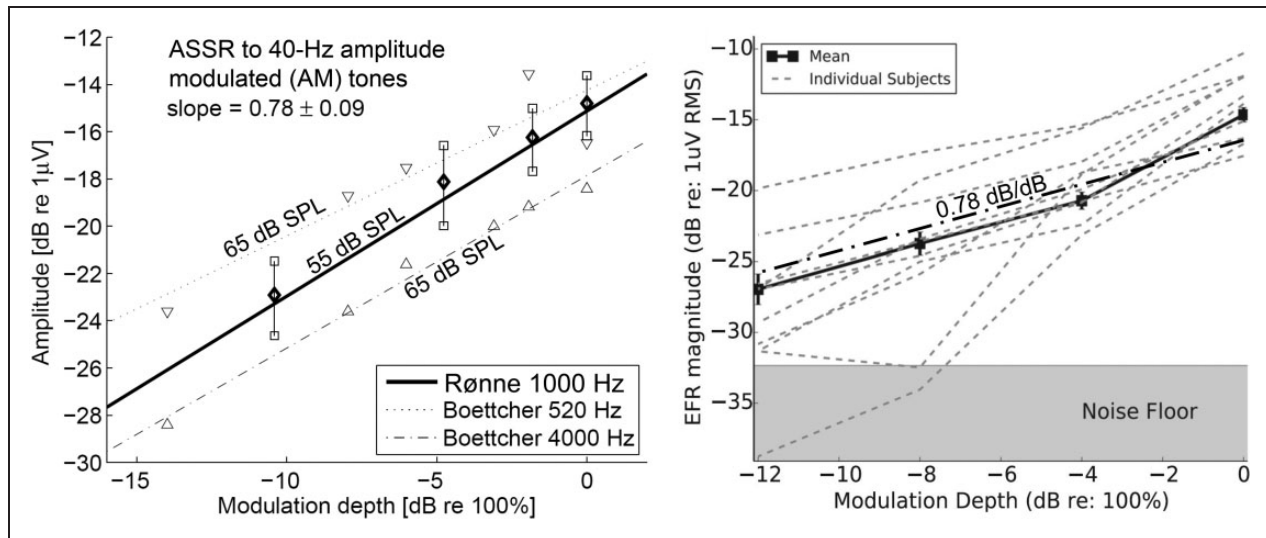
Figure 6. Examples of physiological input/output curves. Left: 40-Hz ASSR measurements (Rønne, 2012) and (Boettcher, Poth, Mills, & Dubno, 2001), as indicated. Right: 100-Hz envelope-following response measurements (Bharadwaj, Masud, Mehraei, Verhulst, & Shinn-Cunningham, 2015). The regression line from the left-hand panel is superimposed on the right-hand panel as the dash-dotted bold line. AM = amplitude modulated; ASSR = auditory steady-state response; RMS = root mean square; SPL = sound pressure level.

modified versus the standard NB chirps, observed in Figure 2, top panel, indicates that a multiharmonic detector, such as the q-sample detector (Cebulla et al., 2006) may provide less benefit for the ISTS-modified stimulus compared with what has been found for standard stimuli. On the other hand, the accumulated detection rates in Figure 2, bottom panel, show bigger improvement from including more harmonics for the ISTS modified than for the standard stimulus. This is not intuitive. An increase in the accumulated detections from one harmonic number to the next depends on how many new recordings can be deemed detected from the added analysis. So, for example, when the accumulated detection rates grow by a similar amount from including four compared with three harmonics in Figure 2, this is because the—very different—number of individual detections of the fourth harmonic contains equally many new detections. In any case, the subject of a multi-harmonic detector is subject to further research.

Test Subject Population

The experiments reported earlier were conducted with young adult normal-hearing test subjects, and therefore it needs to be considered whether the results can be generalized to the prelingual infant target population. Evidence from the literature suggests similar responses in infants compared with young adults after about 2 to 3 months of age (John, Brown, Muir, & Picton, 2004), which corresponds well with the expected age at which hearing-aid fittings are completed in infants. On the other hand, more variability must be expected due to

less compliance in the infant population. Additional performance degradation must be expected between normally hearing and hearing-impaired aided infants based on the existing evidence from similar recordings of cortical responses in adults (Van Dun, Kania, & Dillon, 2016) and infants (Punch et al., 2016). This is, again, subject to further research.

Conclusions

The consequences of adding speech-like properties to the NB CE-Chirps® for ASSR recordings—for the purpose of hearing-aid validation in infants—were investigated. The main findings were as follows:

- Detection rates were very similar for both the speech-modified and the standard stimuli, given sufficient recording time.
- ASSR magnitude decreased by about 4 dB (for the dominant first response harmonic) due to the speech modifications.
- Detection times increased relatively less by a factor of 1.5.

For hearing-aid validation in the clinic, a key metric is the detection time. While the results presented earlier suggest considerable between-subjects variation in detection time, the observed 50% mean increase in detection time for the first harmonic corresponds to less than 1 min of additional recording time. This seems acceptable, given the potential for allowing aided ASSR recordings to be carried out with hearing aids in their daily-life

mode of operation. The proposed procedure is expected to have benefits in terms of clinical simplicity, as the hearing aids will not need to be reprogrammed for the validation measurement, and in terms of the face validity of the validation measurement toward clinicians and the parents of the infant in question.

Finally, the unnormalized modulation power spectrum including preprocessing through a bank of gammatone filters appears to be a useful tool for characterizing the efficacy of complex stimuli for ASSR measurements.

Declaration of Conflicting Interests

Funding

References

Bamford, J., & McSporran, E. (1993). Visual reinforcement audiometry. In B. McCormick (Ed.), *Paediatric audiology, 0–5 years* (2nd ed., pp. 124–154). London, England: Whurr.

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *Journal of Neuroscience*, *35*(5), 2161–2172. doi:10.1523/JNEUROSCI.3915-14.2015

Billings, C. J., Tremblay, K. L., Souza, P. E., & Binns, M. A. (2007). Effects of hearing aid amplification and stimulus intensity on cortical auditory evoked potentials. *Audiology and Neurotology*, *12*(4), 234–246. doi:10.1159/000101331

Boettcher, F. A., Poth, E. A., Mills, J. H., & Dubno, J. R. (2001). The amplitude-modulation following response in young and aged human subjects. *Hearing Research*, *153*(1), 32–42. doi:10.1016/S0378-5955(00)00255-0

British Society of Audiology. (2014). *Recommended procedure: Visual reinforcement audiometry*. Retrieved from http://www.thebsa.org.uk/wp-content/uploads/2014/04/BSA_VRA_24June2014_Final.pdf

Carter, L., Dillon, H., Seymour, J., Seeto, M., & Van Dun, B. (2013). Cortical auditory-evoked potentials (CAEPs) in adults in response to filtered speech stimuli. *Journal of the American Academy of Audiology*, *24*(9), 807–822. doi:10.3766/jaaa.24.9.5

Cebulla, M., Stürzebecher, E., & Elberling, C. (2006). Objective detection of auditory steady-state responses: Comparison of one-sample and q-sample tests. *Journal of the American Academy of Audiology*, *17*(2), 93–103.

Day, J., Bamford, J., Parry, G., Shepherd, M., & Quigley, A. (2000). Evidence on the efficacy of insert earphone and sound field VRA with young infants. *British Journal of Audiology*, *34*(6), 329–334. doi:10.3109/03005364000000148

Dobie, R. A., & Wilson, M. J. (1996). A comparison of *t* test, *F* test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids. *The Journal of the Acoustical Society of America*, *100*(4), 2236–2246.

Easwar, V., Purcell, D. W., Aiken, S. J., Parsa, V., & Scollie, S. D. (2015). Evaluation of speech-evoked envelope following responses as an objective aided outcome measure: Effect of stimulus level, bandwidth, and amplification in adults with hearing loss. *Ear and Hearing*, *36*(6), 635–652.

Elberling, C., & Don, M. (2010). A direct approach for the design of chirp stimuli used for the recording of auditory brainstem responses. *The Journal of the Acoustical Society of America*, *128*(5), 2955. doi:10.1121/1.3489111

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology*, *49*(12), 891–903. doi:10.3109/14992027.2010.506889

ISO 389-9. (2009). *Acoustics—Reference zero for the calibration of audiometric equipment—Part 9: Preferred test conditions for the determination of reference hearing threshold levels* (ISO European Standard).

Johannesma, P. I. (1972). The pre-response stimulus ensemble of neurons in the cochlear nucleus. In *Symposium on Hearing Theory* (pp. 58–69). Eindhoven, Holland: Institute for Perception Research.

John, M. S., Brown, D. K., Muir, P. J., & Picton, T. W. (2004). Recording auditory steady-state responses in young infants. *Ear and Hearing*, *25*(6), 539–553. doi:10.1097/01.AUD.0000148050.80749.AC

John, M. S., Dimitrijevic, A., & Picton, T. W. (2001). Weighted averaging of steady-state responses. *Clinical Neurophysiology*, *112*(3), 555–562.

Laugesen, S., Rieck, J. E., Elberling, C., Dau, T., & Harte, J. M. (2017). On the cost of introducing speech-like properties to a stimulus for auditory steady-state response measurements. In S. Santurette, T. Dau, J. C. Dalsgaard, L. Tranebjærg, T. Andersen, & T. Poulsen (Eds.),, *Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 6: Adaptive Processes in Hearing, August 2017, Nyborg, Denmark* (pp. 223–230). Ballerup, Denmark: The Danavox Jubilee Foundation.

Mehta, K., Watkin, P., Baldwin, M., Marriage, J., Mahon, M., & Vickers, D. (2017). Role of cortical auditory evoked potentials in reducing the age at hearing aid fitting in children with hearing loss identified by newborn hearing screening. *Trends in Hearing*, *21*, 233121651744409. doi:10.1177/2331216517744094

Payton, K. L., & Braida, L. D. (1999). A method to determine the speech transmission index from speech waveforms. *The Journal of the Acoustical Society of America*, *106*(6), 3637–3648.

Picton, T. W., Durieux-Smith, A., Champagne, S. C., Whittingham, J., Moran, L. M., Giguère, C., & Beauregard, Y. (1998). Objective evaluation of aided thresholds using auditory steady-state responses. *Journal of the American Academy of Audiology*, *9*, 315–331.

Pritchard, W. S. (1992). The brain in fractal time: 1/f-like power spectrum scaling of the human electroencephalogram. *International Journal of Neuroscience*, *66*(1–2), 119–129.

Punch, S., Van Dun, B., King, A., Carter, L., & Pearce, W. (2016). Clinical experience of using cortical auditory evoked potentials in the treatment of infant hearing loss in Australia. *Seminars in Hearing*, 37(1), 36–52. doi:10.1055/s-0035-1570331

Rønne, F. M. (2012). *Modeling auditory evoked potentials to complex stimuli* (PhD thesis). Retrieved from Department of Electrical Engineering, Technical University of Denmark: http://orbit.dtu.dk/fedora/objects/orbit:127704/datastreams/file_250c6e0a-10de-405b-bfa2-2ab780406324/content

Sharma, A., Dorman, M. F., & Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear and Hearing*, 23(6), 532–539.

Van Dun, B., Kania, A., & Dillon, H. (2016). Cortical auditory evoked potentials in (un)aided normal-hearing and hearing-impaired adults. *Seminars in Hearing*, 37(1), 9–24. doi:10.1055/s-0035-1570333

Wood, S. A., Sutton, G. J., & Davis, A. C. (2015). Performance and characteristics of the Newborn Hearing Screening Programme in England: The first seven years. *International Journal of Audiology*, 54(6), 353–358. doi:10.3109/14992027.2014.989548