Do Room Acoustics Affect the Amplitude of Sound-Field Auditory Steady-State Responses?

Valentina Zapata-Rodriguez[1,2] (iD), Søren Laugesen[1], Cheol-Ho Jeong[2], Jonas Brunskog[2], and James Harte[1]

Abstract
The sound-field auditory steady-state response (ASSR) is a promising measure for the objective validation of hearing-aid fitting in patients who are unable to respond to behavioral testing reliably. To record the sound-field ASSR, the stimulus is reproduced through a loudspeaker placed in front of the patient. However, the reverberation and background noise of the measurement room could reduce the stimulus modulation used for eliciting the ASSR. As the ASSR level is heavily dependent on the stimulus modulation, any reduction due to room acoustics could affect the clinical viability of sound-field ASSR testing. This study investigated the effect of room acoustics on the level and detection rate of sound-field ASSR. The study also analyzed whether early decay time and an auditory-inspired relative modulation power model could be used to predict the changes in the recorded ASSR in rooms. A monaural auralization approach was used to measure sound-field ASSR via insert earphones. ASSR was measured for 15 normal-hearing adult subjects using narrow-band CE-Chirps® centered at the octave bands of 500, 1000, 2000, and 4000 Hz. These stimuli were convolved with simulated impulse responses of three rooms inspired by audiological testing rooms. The results showed a significant reduction of the ASSR level for the room conditions compared with the reference anechoic condition. Despite this reduction, the detection rates for the first harmonics of the ASSR were unaffected when sufficiently long recordings (up to 6 min) were made. Furthermore, the early decay time and relative modulation power appear to be useful predictors of the ASSR level in the measurement rooms.

Keywords
sound-field auditory steady-state responses, hearing-aid validation, audiometric testing rooms, early decay time, modulation power

Date received: 8 February 2020; revised: 31 August 2020; accepted: 18 September 2020

Introduction

Effective early diagnosis and intervention of pediatric hearing loss at the age of 6 months, or even before, is crucial for the development of speech to a level comparable to normal-hearing infants (Moeller, 2000; Yoshinaga-Itano et al., 1998). This has led to the implementation of universal newborn hearing screening programs in many countries around the world, for example, most developed countries had implemented such hearing screening programs by 2015 (Morton & Nance, 2006; Naumann et al., 2015; Neumann et al., 2019; Ptok, 2011; Singh, 2015). Such screening programs aim at identifying hearing loss in infants at the earliest possible age, leading to the early treatment of their hearing impairment. The primary goal in early intervention is to ensure that a child has access to speech sounds, due to early critical windows for language and brain development (Sharma et al., 2002; Sininger et al., 2010). A successful early intervention of hearing loss relies on appropriate adjustments of the hearing-aid

[1]Interacoustics Research Unit, Lyngby, Denmark
[2]Department of Electrical Engineering, Technical University of Denmark, Lyngby, Denmark

Corresponding author:
Valentina Zapata-Rodriguez, Interacoustics Research Unit, c/o: Technical University of Denmark, Ørsteds Plads, Building 352, 2800 Lyngby, Denmark.
Email: valr@iru.interacoustics.com

amplification, which is called hearing-aid fitting. Hearing-aid fitting validation then becomes a critical procedure for ensuring that the infant receives adequate auditory stimuli and for avoiding potential delays in language development (Marcoux & Hansen, 2003). However, the validation of the prescribed hearing-aid gain is challenging in prelingual infants because standard behavioral tests are highly unreliable. For this reason, some researchers have suggested alternative objective procedures based on auditory evoked potentials, such as cortical auditory evoked potentials (e.g., Punch et al., 2016) and auditory steady-state responses (ASSRs; e.g., Picton et al., 1998). These electrophysiological measurements are promising because they can verify that the brain is receiving and processing the auditory input without the need for a voluntary response from the patient. In this study, an approach using the sound-field ASSR is considered for hearing-aid fitting validation. The study focused on investigating the potential challenges in sound-field ASSR recordings associated with the room in which the test is carried out.

The ASSR is an auditory evoked potential in response to repeated transient stimuli or sinusoidally amplitude-modulated tones. It is elicited by specific groups of neurons firing phase locked to the modulation envelope of the auditory signal (Picton et al., 2003; Rance, 2008). The sound-field ASSR involves acoustic stimulation through a loudspeaker instead of presenting the signal via insert earphones as it is traditionally done in the clinic (Picton et al., 1998). The loudspeaker stimulation allows the inclusion of the hearing aid into the stimulation path. In one of the first reported such studies, Picton et al. (1998) estimated physiological hearing thresholds using sound-field ASSR which were not significantly different from behavioral hearing thresholds measured via insert earphones. The majority of research to date has focused on the validation of the sound-field ASSR as an accurate tool for hearing-aid fitting validation, demonstrating a good agreement between physiological and objective thresholds measured in aided and unaided conditions (Damarla & Manjula, 2007; Hernandez-Perez & Torres-Fortuny, 2013; Park et al., 2013; Picton et al., 1998; Sardari et al., 2015; Selim et al., 2012; Shemesh et al., 2012; Stroebel et al., 2007). However, the potential effect of the room on the sound-field ASSR measurement has received little scientific attention. In fact, the studies that investigated sound-field ASSR for hearing-aid fitting validation did not assess the effect of the testing room on the ASSR. Most of these studies only reported the listening environments to be sound-treated-rooms (e.g., Hernandez-Perez & Torres-Fortuny, 2013; Park et al., 2013; Picton et al., 1998; Sardari et al., 2015; Shemesh et al., 2012) with low background noise according to the standard ANSI S3.1 (ANSI S3.1., 2003; e.g., Park et al., 2013; Picton et al.,

1998; Sardari et al., 2015), and no information about the reverberation time of the rooms was provided. To understand the effect of the room on sound-field ASSR is crucial, as the reverberation and background noise of the room in which the test is carried out can distort the acoustic stimulus modulation (Houtgast et al., 1980; Plomp, 1983). As the ASSR amplitude is heavily influenced by the modulation of the acoustic signal (e.g., John et al., 2001; Picton et al., 1987), the influence of room acoustics could present a barrier to the future clinical implementation of the sound-field ASSR test for infants' hearing-aid fitting validation.

Previous studies investigated the ASSR amplitude when the modulation depth for sinusoidally amplitude-modulated stimuli was systematically varied, and they reported that the ASSR amplitude reduces as the modulation depth of the acoustic signal decreases (Bharadwaj et al., 2015; Boettcher et al., 2001; Dimitrijevic et al., 2001; John et al., 2001; Kuwada et al., 1986; Lins et al., 1995; Picton et al., 1987; Rees et al., 1986; Roß et al., 2000; Rønne, 2012). When the ASSR is recorded with stimuli with equal root-mean-square value but different modulation depths, the ASSR amplitude reaches its maximum when it is recorded with 100% amplitude-modulated tones (Dimitrijevic et al., 2001; John et al., 2001). When instead the peak to peak value of the envelope remains equal for different modulation depths, a maximum ASSR amplitude is obtained with a 50% amplitude-modulated tone (Lins et al., 1995; Picton et al., 1987). The reduction in the ASSR amplitude is approximately linear for modulation depth represented in a logarithmic scale (Rees et al., 1986; Roß et al., 2000). The modulation of the ASSR stimulus can be easily controlled when eliciting the neural response through insert earphones (Kuwada et al., 1986; Picton et al., 1987). In sound-field ASSR, however, the resulting stimulus modulation will depend on the acoustics of the room and the loudspeaker and listener position. The stimulus modulation (at the eardrum) could then serve as a potential predictor of the effect of room acoustics on sound-field ASSR.

The influence of room acoustics on the stimulus modulation has been widely investigated for speech intelligibility. It has been demonstrated that the reverberation and the background noise attenuate the natural fluctuations of the speech signal which are necessary for speech comprehension, which leads to poorer speech intelligibility for longer reverberation times and high noise levels (Bradley et al., 1999). The reverberation time (T) is defined as the time it takes for a sound to decrease by 60 dB in a room after an abrupt termination of the sound source (ISO 3382-1, 2009). This can be quantified by standard room parameters, such as the early decay time (EDT) and T20, which use different decay ranges, from 0 to −10 dB and −5 to −25 dB for EDT and T20,

respectively (ISO 3382-1, 2009). Due to the close proximity between the loudspeaker and listener position in sound-field ASSR measurements, it is expected that the early reflections, which have larger energy, will have a greater influence on the stimulus modulation. Thus, EDT could potentially be a good predictor to estimate the ASSR level in sound-field ASSR measurements.

The primary objective of this study was to determine whether sound-field ASSR measurements would be affected by the acoustic condition of the measurement room in terms of the ASSR level and detection rate (the proportion of detected responses out of all conditions tested). Only the effect of the reverberation on the stimulus modulation was evaluated, whereas the effect of the background noise in the room was not considered in this study. It was then hypothesized that the response amplitude would be reduced due to the degradation of the stimulus modulation, resulting from the loudspeaker presentation in the measurement room. Consequently, the detection rate will likely also be reduced. This hypothesis was based on two facts: (a) The modulation of any acoustic signal in a room is distorted by its reverberation and background noise (Houtgast et al., 1980; Plomp, 1983); (b) ASSR amplitude reduces as the stimulus modulation decreases (Bharadwaj et al., 2015; Boettcher et al., 2001; Dimitrijevic et al., 2001; John et al., 2001; Kuwada et al., 1986; Lins et al., 1995; Picton et al., 1987; Rees et al., 1986; Roß et al., 2000; Rønne, 2012). The hypotheses were tested with an auralization approach using insert earphones, implemented to mimic sound-field ASSR. This consisted of the convolution of CE-Chirp® (Elberling & Don, 2010) stimuli with three simulated room impulse responses. The ASSR measurements were carried out in normal-hearing adult test subjects who were presented monaurally with the resulting auralized signals. The study also investigated whether it is possible to predict the ASSR level in any given room to determine its suitability for sound-field ASSR measurements. For this purpose, it was analyzed whether ASSR level could be estimated in a room by the EDT and the resulting stimulus modulation, which was here quantified with an auditory-inspired relative modulation power model.

Material and Methods

Participants

Fifteen young adult normal-hearing subjects (seven females, mean age 24 ± 3 years) participated in the test. Their audiological status was verified by means of otoscopy, wide-band tympanometry using the Interacoustics Titan, and air-conduction audiometry using an Interacoustics AC40 audiometer with ER-3A insert phones. All participants had pure-tone threshold equal or better than 20 dB hearing level at 125, 250, 500, 1000, 2000, 4000, and 8000 Hz. They provided written informed consent and were financially compensated with gift cards. The experiment was approved by the Science-Ethics Committee for the Capital Region of Denmark.

Stimuli and Room Acoustic Simulations

The narrow-band (NB) CE-Chirps® for ASSR recording consist of four one-octave-wide NB chirp trains, with center frequencies of 500, 1000, 2000 and 4000 Hz (Elberling & Don, 2010). These chirp trains are complex tones properly designed to compensate for the travelling wave delay in the basilar membrane. Each NB CE-Chirp® is composed by a harmonic series of cosines tones within the specific octave band, which are presented with slightly different time delays such that they all excite the basilar membrane at the same time. Hence, a broader region of the basilar membrane is synchronously excited, resulting in stronger neural responses and in turn, ASSRs with higher amplitudes that can be detected faster. In this study, each of the four commercial NB CE-Chirps® were presented through insert earphones at slightly different repetition rates around 90 Hz, all within a range from 88 to 98 Hz, as used in clinical practice with the Interacoustics Eclipse platform for ASSR recordings. A monaural room auralization approach was used to simulate sound-field ASSR and consisted of the convolution in real time of the NB CE-Chirps® with simulated monaural room impulse responses based on the acoustic Green's function (Jacobsen & Juhl, 2013). A total of 16 conditions were tested, corresponding to the combination of the four NB CE-Chirps® (presented individually) and the four acoustic conditions (three simulated rooms and one unmodified anechoic stimulus condition that served as reference).

As sound-field ASSR is not a standard clinical test at the moment, there are no specific room acoustic guidelines. The three room acoustic conditions were defined such that their acoustic environments were representative of realistic clinic rooms for other audiological sound-field testing, for example, speech in noise test, sound-field audiometry, fitting of hearing aids. These rooms were expected to be small and have short to medium reberveration times. However, the characteristics of real rooms can vary greatly among clinics due to the lack of regulations for the dimensions and specific acoustic requirements for audiological testing rooms. To the best of our knowledge, only a few standards and guidelines provide recommendations on the test environment for sound-field audiometry. One of such relevant standards is the ISO 8253-2, which defines the adequate acoustic environment in the room for sound-field audiometry based on the variation of the sound-pressure

level around the measurement point. This standard establishes a maximum allowable sound pressure level (SPL) variation of $\pm 2.5\,\text{dB}$ between the measurement position and a position located at $0.15\,\text{m}$ to the front, back, left, right, up, and down from the measurement position (ISO 8253-2, 2009). Another relevant guideline is the practice guidance on the acoustics of sound-field audiometry developed by the British Society of Audiology (BSA, 2019), which recommends a maximum reverberation time of $0.25\,\text{s}$ across all frequencies, and minimum dimensions of $6 \times 4\,\text{m}$ ($w \times l$) for the testing rooms (Health Technical Memorandum 2045, 1996). The implementation of such recommended values, however, has not been broadly adopted due to their non-mandatory nature.

The simulated rooms mimic a medium-size single-walled audiology testing booth (ATB), a standardized loudspeaker listening room (IEC) according to the standard IEC 268-13 (1985), and a room recommended by the BSA (2019) for sound-field audiometry for pediatric assessment (Health Technical Memorandum 2045, 1996). Rooms ATB and IEC were based on real rooms located at the Technical University of Denmark. In all simulations, a distance of $1\,\text{m}$ between the loudspeaker and the patient was used at an approximate height of a person sitting on a chair. Table 1 lists the dimensions, the reverberation time in one-octave bands, and the simulated source and receiver position for each of the simulated room conditions. The estimated Schroeder frequency (f_{Sch}; Schroeder & Kuttruff, 1962) for each of the rooms is also listed in the table. For the three simulated rooms, the Schroeder frequencies were below the lowest frequency limit of the one-octave-wide NB CE-Chirp® stimuli ($355\,\text{Hz}$, for the NB CE-Chirp® with center frequency of $500\,\text{Hz}$), indicating that no strong modes would affect the ASSR stimuli. The acoustic environments were simulated with a cosine room acoustic model using a modal approach that estimated the frequency response of the rooms based on a truncated Green's function (Equation 1). The Green's function

is an analytical solution to the wave equation with the boundary conditions imposed by rigid walls (Jacobsen & Juhl, 2013):

$$G(r, r_0) = -\frac{1}{V} \sum_{m=f_l}^{f_u} \frac{\Psi_m(r)\Psi_m(r_0)}{k^2 - k_m^2 - jk/(\tau_m c)} \quad (1)$$

The implemented room acoustic model accurately calculates the modal behavior of rectangular rooms, which is an important feature of small rooms below the Schroeder frequency (Schroeder & Kuttruff, 1962). The model was implemented in a custom MATLAB script that simulated the monaural frequency response of the rooms from $f_l = 10\,\text{Hz}$ to $f_u = 10\,\text{kHz}$. The room impulse responses were then obtained by calculating the inverse Fourier transform of the simulated frequency responses (Equation 1). Each term in Equation 1 represents a mode in the cartesian coordinate system,

$$\Psi_m(x, y, x) = \sqrt{\varepsilon_{n_x}\varepsilon_{n_y}\varepsilon_{n_z}}\cos\left(\frac{n_x \pi x}{l_x}\right)\cos\left(\frac{n_y \pi y}{l_y}\right)\cos\left(\frac{n_z \pi z}{l_z}\right) \quad (2)$$

where ε_{n_x}, ε_{n_y}, and ε_{n_z} are normalization constants equal to 1 for $n = 0$ and 2 for $n \neq 0$, respectively. The volume of the room is $V = l_x l_y l_z$, and the source is located at $r_0 = (x_0, y_0, z_0)$ and the receiver at $r = (x, y, z)$. The time constant is given by $\tau_m = T/13.8$, and was calculated using the reverberation time, T, in one-third-octave bands measured for rooms ATB and IEC, and the specified reverberation time for the BSA room. The wave-number corresponding to the mth natural frequency of the room is given by

$$k_m = (1 + 0.01\nu_m)\sqrt{\left(\frac{n_x\pi}{l_x}\right)^2 + \left(\frac{n_y\pi}{l_y}\right)^2 + \left(\frac{n_z\pi}{l_z}\right)^2} \quad (3)$$

Table 1. Characteristics of the simulated rooms. Input data for the simulations: dimensions of the room, source and receiver positions and reverberation time in one-octave bands. Estimated Schroeder frequency (fSch) for each individual simulated room, which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology for sound-field audiometry for pediatric assessment (BSA), and a standardized listening room (IEC).

Room	Dimensions $l_x \times l_y \times l_z$ (m)	Source position x_0, y_0, z_0 (m)	Receiver position x, y, z (m)	125	250	500	1000	2000	4000	8000	f_{Sch} (Hz)
ATB	$2.6 \times 2.0 \times 2.1$	0.9, 1.0, 1.0	1.9, 1.0, 1.0	0.14	0.06	0.04	0.04	0.06	0.05	0.06	153.5
BSA	$6.0 \times 4.0 \times 2.8$	2.5, 1.8, 1.0	3.5, 1.8, 1.0	0.25	0.25	0.25	0.25	0.25	0.25	0.25	122.0
IEC	$4.7 \times 7.5 \times 2.8$	3.2, 5.5, 1.4	1.8, 5.2, 1.4	0.27	0.24	0.31	0.29	0.34	0.30	0.29	108.7

(Reverberation time, T20 (s), Frequency band (Hz) spanning 125–8000 columns)

Note. Input data for the simulations: dimensions of the room, source and receiver positions and reverberation time in one-octave bands. Estimated Schroeder frequency (f_{Sch}) for each individual simulated room, which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology (BSA) for sound-field audiometry for pediatric assessment, and a standardized listening room (IEC).

where the speed of sound *c* takes a value of 343 m/s. A small random factor $v_m \sim N(0,1)$ was added to the wavenumbers given by Equation 3 to produce a more natural auralized sound.

ASSR Measurements

The Interacoustics Eclipse platform was used to generate the standard NB CE-Chirps®, as well as to record and process the ASSR responses. Each generated NB CE-Chirp® was sent to an external computer through an RME Fireface UCX sound card. They were then convolved online with the simulated room impulse responses using the virtual studio technology plugin SIR v1.011 running on the free version of the LiveProfessor v1.2.5 software. The resulting stimuli were sent through the Tucker-Davis Technologies HB7 headphone driver, which was connected to the ER-3A insert earphone used to present the stimuli to the subjects.

The ASSRs were measured using a standard clinical four-electrode montage. The four surface electrodes were placed on the high forehead (reference), cheeck (ground), and on each mastoid (left and right, ipsi- and contra-lateral mastoids active). The electrode impedances were kept as equal as possible across the four electrodes and never exceeded 3 kΩ. The signal was pre-amplified by the ERA preamplifier and was recorded with the commercial Interacoustics Eclipse system. The ASSRs were recorded using the setting accuracy-test method priority and the adult sleeping protocol in the Eclipse software. The accuracy-test method priority determines the significance level used in the ASSR detector algoritm, which corresponds to $p = .01$. The adult sleeping protocol establishes the repetition rates of the stimuli used (fast repetition rates, around 90 Hz). The electroencephalogram (EEG) response was recorded with a sampling frequency of 30 kHz for a total of 6 min of continuous recording with an artifact rejection level of ±30 μV, which was increased to ±40 μV for restless subjects (higher EEG noise). The recordings were carried out in a darkened, single-walled acoustically treated and electrically shielded booth. During testing, the participants lay on a comfortable bed and were instructed to relax and sleep if possible. The experiment consisted of two sessions of 1.5 h each. The 16 testing conditions were presented once in a random order to each test subject. The ASSRs were recorded for only one ear that was randomly chosen while the non-test ear was blocked with a foam earplug.

To calibrate the tested acoustic conditions, each stimulus was presented via the insert earphones connected to an ear simulator B&K 4157. The presentation levels were then measured with the B&K 4157 ear simulator, which simulates the presentation level at the eardrum. The 16 stimuli (four acoustic conditions and four NB CE-

Chirps®) were calibrated individually. The stimuli were calibrated to the target values 68.0, 62.6, 68.0, and 58.7 dB SPL for the NB CE-Chirps® with center frequencies of 500, 1000, 2000, and 4000 Hz, respectively. The levels were defined from the standard method for computing the speech intelligibility index (ANSI S3.5, 1997). These values were selected such that when the stimuli are combined, the resulting overall level matches the long-term spectrum of speech in one-octave wide frequency bands and with a broad-band level of 72 dB SPL at the eardrum position.

Data Analysis

ASSR Postprocessing. The EEG recordings were analyzed per block (*epoch*) of 65,536 samples, corresponding to 2.18 s each. Only recordings with 162 blocks and an artifact rejection level of ±30 μV were used to ensure consistent ASSR detection. On this basis, 93.75% of the total data collected was used (only 15 of 240 recordings were discarded). Table 2 shows the number of recordings included in the analysis for each condition. The ASSR is typically analyzed in the frequency domain by epochs with a duration relative to the periodicity of the ASSR. The EEG spectrum is then composed by the EEG noise produced and the ASSR, which is found at the frequency bin of the stimulus repetition rate and its harmonics (Picton et al., 2003; Rance, 2008). Here, the ASSR data were analyzed offline with the weighted averaging method (John et al., 2001) and an *F*-ratio test with a strict error rate of 1% (Dobie & Wilson, 1996). ASSR detection was individually evaluated for each of the first four response harmonics, and without making use of the multiharmonic detector of the standard Eclipse (Cebulla et al., 2006). The multiharmonic detector uses the amplitude and phase of the fundamental frequency and higher harmonics (12 harmonics in the commercial Eclipse system) of the ASSR to determine whether the response is present. Although a multiharmonic detector is more

Table 2. Number of Measurements per Condition Included in the Analysis After Postprocessing With an Artifact Rejection Level of ±30 μV.

Acoustic condition	Analysis band			
	500 Hz	1000 Hz	2000 Hz	4000 Hz
REF	14	14	14	15
ATB	14	13	13	15
BSA	14	15	14	12
IEC	14	14	15	15

Note. The tested acoustic conditions correspond to the reference anechoic (REF), and three simulated rooms which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology (BSA) for sound-field audiometry for pediatric assessment, and a standardized listening room (IEC).

sensitive than the *F*-ratio test, this was used to analyze the detection rates of the individual ASSR harmonics (Dobie & Wilson, 1996). The noise-corrected ASSR amplitude (in dB reference to 1 nV, hereafter, referred to as ASSR level) for the different acoustic conditions tested (Dobie & Wilson, 1996) was the primary variable of interest in this study. The ASSR level is then calculated by subtracting the estimated noise power from the response power. The response power was estimated at the frequency bin of the repetition rate, while the noise power was estimated by averaging the noise power across 20 evenly distributed frequency bins around the response bin. The frequency bins of the harmonics of 50 Hz line noise (e.g., 50, 100, 200, . . . Hz) were excluded from the noise power calculation to avoid the interference from the AC power supply noise. The frequency bins corresponding to any other stimuli repetition rate harmonics were also removed from the noise power calculation. The ASSR and noise amplitudes were estimated for each harmonic. With 20 harmonics used for the noise estimate and a 1% error rate, the critical value for the *F*-test is $F(2, 40)_{0.99} = 5.18$, corresponding to a signal-to-noise ratio of 7.1 dB. The detection rate was calculated based on the total number of recordings included in the analysis for each condition listed in Table 2. The ASSR was analyzed for each individual harmonic to determine whether there could be a correlation between the ASSR level and the stimulus modulation for each individual harmonic. It is noteworthy, however, that if any harmonic was detected, the stimulus was heard.

Acoustic Descriptors of Simulated Rooms

Early Decay Time. The EDT is a reverberation time measurement estimated from the first 10 dB level drop of the decay curve, thus quantifying the early part of the decay curve. It is known to be closely related to the subjective impression of the reverberation in the room (ISO 3382-1, 2009). Considering that for sound-field ASSR measurements a source to listener distance of 1 m is used, the stimulus modulation is expected to be most affected by the early reflections of the room, which are more important for shorter source to receiver distances. The EDT was derived from the decay rate of the simulated impulse responses for each room condition, as described in the standard (ISO 3382-1, 2009). For the reference condition, the EDT was set to 0 s for the analysis. Figure 1A shows the EDT calculated per octave band from the simulated room impulse responses of the tested acoustic conditions.

Auditory-Inspired Relative Modulation Power Model. The modulation of amplitude-modulated tones is well described by the modulation depth (*m*), which is defined as the ratio of the maximum ($y_{\max}$) and minimum ($y_{\min}$)

amplitudes of the waveform's envelope, $m = (y_{\max} - y_{\min})/(y_{\max} + y_{\min})$, assuming a sinusoidal envelope. However, when signals are presented in non anechoic room conditions, their envelopes are distorted and thus the modulation depth is not well defined. Instead, the stimulus modulation can be estimated using the discrete Fourier transform of its envelope (Houtgast et al., 1980; Schroeder, 1981). Considering this, a simple modulation power model was designed to estimate the efficiency of the stimulus modulation in eliciting an ASSR. The model takes any input signal and extracts the changes in the modulation due to the acoustic conditions of the room in relation to the reference anechoic signal. The input signal can be either recorded in the room or simulated by convolution with the room impulse response. The model builds on a previous model that characterized the stimulus waveform based on its envelope power (Laugesen et al., 2018). The model is also inspired by the modulation transfer function for the speech transmission index calculation (Houtgast & Steeneken, 1985), as well as similar approaches used to estimate speech intelligibility based on the envelope power (Relaño-Iborra et al., 2016), and to characterize the degradation of amplitude-modulated stimuli due to reverberation (Slama & Delgutte, 2015). The overall structure of the proposed relative modulation power model is shown in Figure 2.

The first phase of the model uses a linear filter bank of 12 gammatone filters (Johannesma, 1972) uniformly spaced 1/12th octave apart over the stimulus frequency band of interest to simulate the frequency specificity of the human basilar membrane. The envelope of the output of each gammatone filter is then extracted using the Hilbert transform. The temporal envelopes are normalized by subtracting their respective DC component. The envelopes are split into blocks as described in the ASSR postprocessing section. The discrete Fourier transform is then calculated per block and averaged across all blocks and all filter bands to obtain the stimulus envelope power. These steps are conducted for both the reference and the reverberant signals.

The modulation power is then estimated at the frequency bin of the repetition rate of the stimulus frequency band and its harmonics. Finally, the relative modulation power is calculated in dB referenced to the modulation power of the first harmonic of the reference anechoic signal. This quantifies the changes in the stimulus modulation brought about by the acoustical properties of the measurement room relative to the reference signal for each of the stimulus harmonics. Figure 1B shows the relative modulation power for each of the stimulus band center frequencies of the tested acoustic conditions, where a reduced relative modulation power of the room conditions compared with the reference signal is observed for the harmonics of interest.
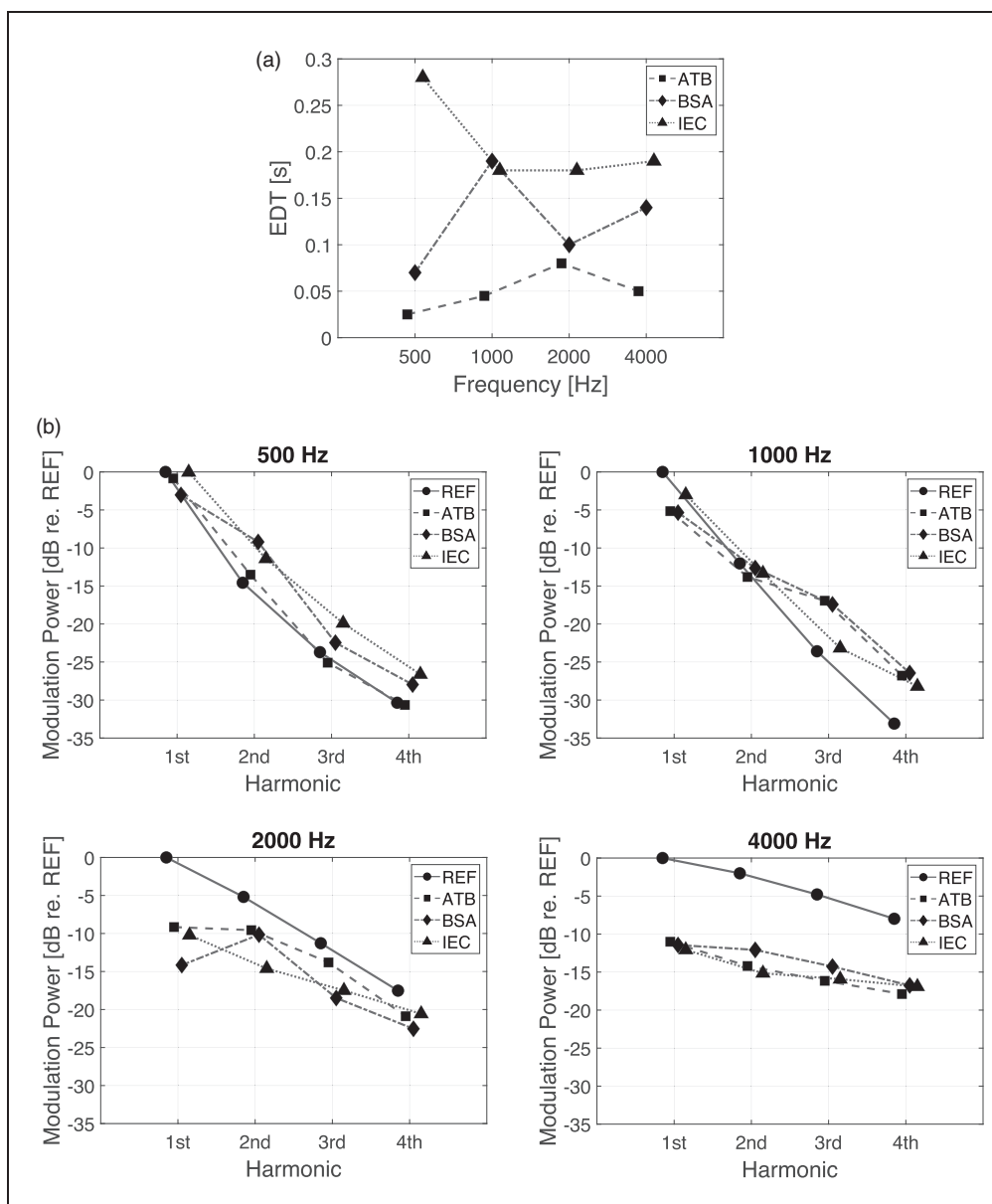
Figure 1. EDT (Top Panel) and Relative Modulation Power (Bottom Panel) of the Acoustic Conditions Tested. Panel A: The EDT estimated from the acoustic conditions for each octave frequency bands (500, 1000, 2000, and 4000 Hz). Panel B: The relative modulation power for the first four harmonics of the acoustic conditions per stimulus frequency. The tested acoustic conditions correspond to the reference anechoic (REF), and three simulated rooms which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology (BSA) for sound-field audiometry for pediatric assessment, and a standardized listening room (IEC).

Statistical Analysis

The statistical analysis was carried out using linear mixed-effects models fitted to the ASSR level, considering the participants as a random effect (*Test Subjects*: 1, 2, 3 . . . 15). All analyses were performed in the software R version 3.5 with the lme4 library (Bates et al., 2015). To analyze the effect of the room on the ASSR level, a first model was estimated including the fixed effects of the room (*Room*: REF, ATB, BSA, and IEC), stimulus frequency (*Frequency*: 500, 1000, 2000, and 4000 Hz)

and ASSR harmonic (*Harmonic*: 1, 2, 3, and 4). The predictable (*ASSR level*) and explanatory (*Room, Frequency, and Harmonic*) variables were defined as a continuous and categorical variables, respectively. Significance was evaluated for all main effects and their interactions, and those that were not significant were removed. Post hoc analysis was conducted to determine significant differences between the reference condition and each room condition (ATB, BSA, and IEC) across frequencies and harmonics. For this analysis,
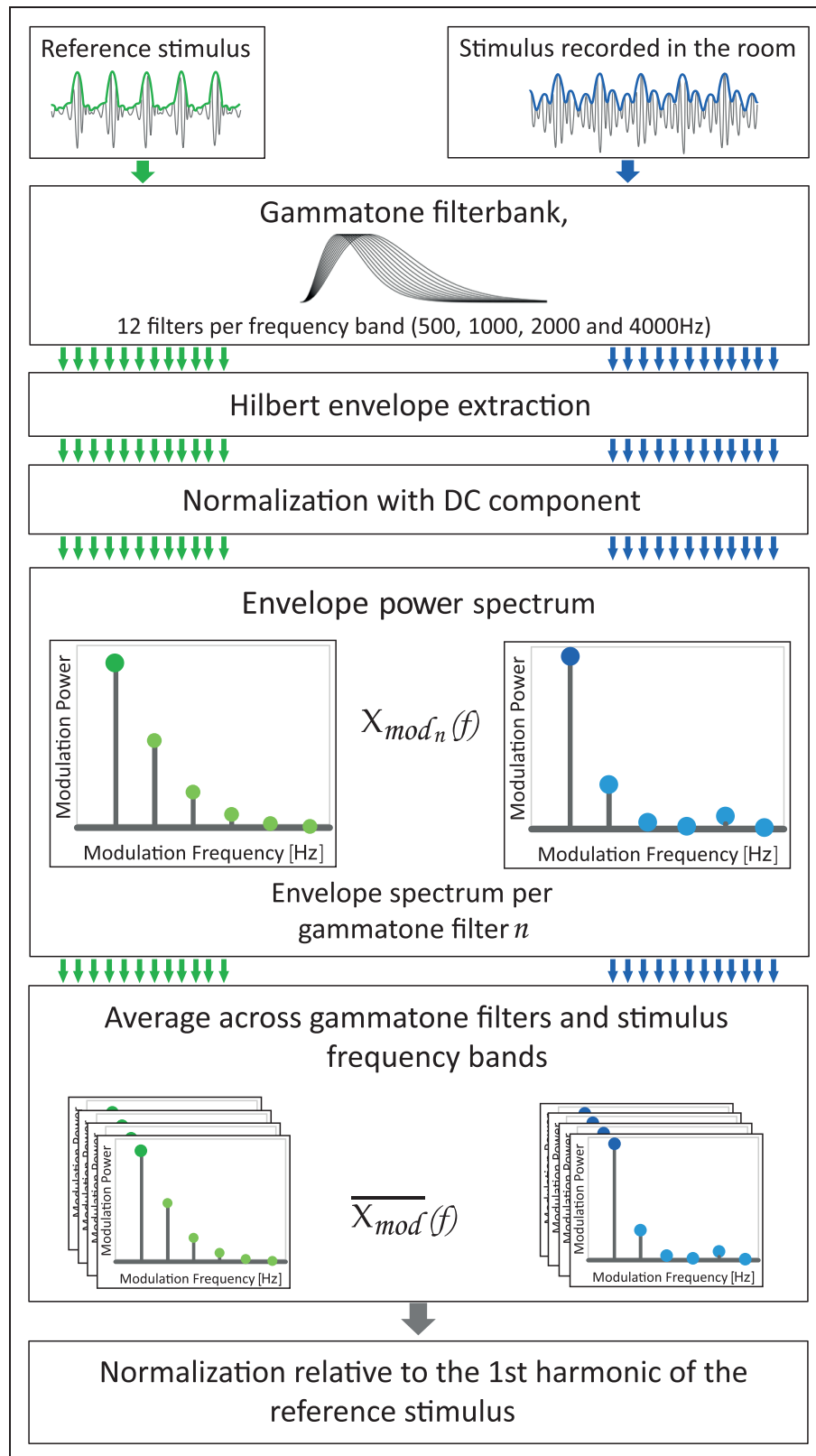
Figure 2. Diagram of the Proposed Auditory-Inspired Relative Modulation Power Model. The reference stimulus and the stimulus recorded in the room are processed through a gammatone filter bank. The envelope of the output of each filter is extracted by applying a Hilbert transform, and are normalized by their respective DC component. The envelope spectrum is calculated per blocks for each filterbank output. The stimulus modulation power is then calculated by averaging the envelope spectrum across all blocks and filter bands. The relative modulation power of the stimulus recorded in the room is then calculated in dB referenced to the modulation power of the first harmonic of the reference stimulus.

the estimated marginal means (Searle et al., 1980) with the Tukey method was used (Tukey, 1949).

Two additional linear mixed-effects models were computed to determine whether the ASSR level could be predicted using either EDT or the relative modulation power, which are measurable properties inherent to the rooms. For these analyses, only the post hoc comparisons that turned out significant in the first statistical model for all combinations of reference and room conditions were considered. Instead of the categorical variable Room, the models included the continuous fixed effects of either the *EDT* (time in seconds) or the relative modulation power (*RModP*: in

dB). Nonsignificant main effects and interactions were removed from the models, which were evaluated with analysis of variances (ANOVAs) and comparing the Akaike information criterion (Akaike, 1974).

Results

Effect of the Room on ASSR Level

Figure 3 shows the distribution of the ASSR level for the tested acoustic conditions across stimulus band center frequencies (columns) and harmonics (rows). The statistical model showed that the main effects (*Room*,
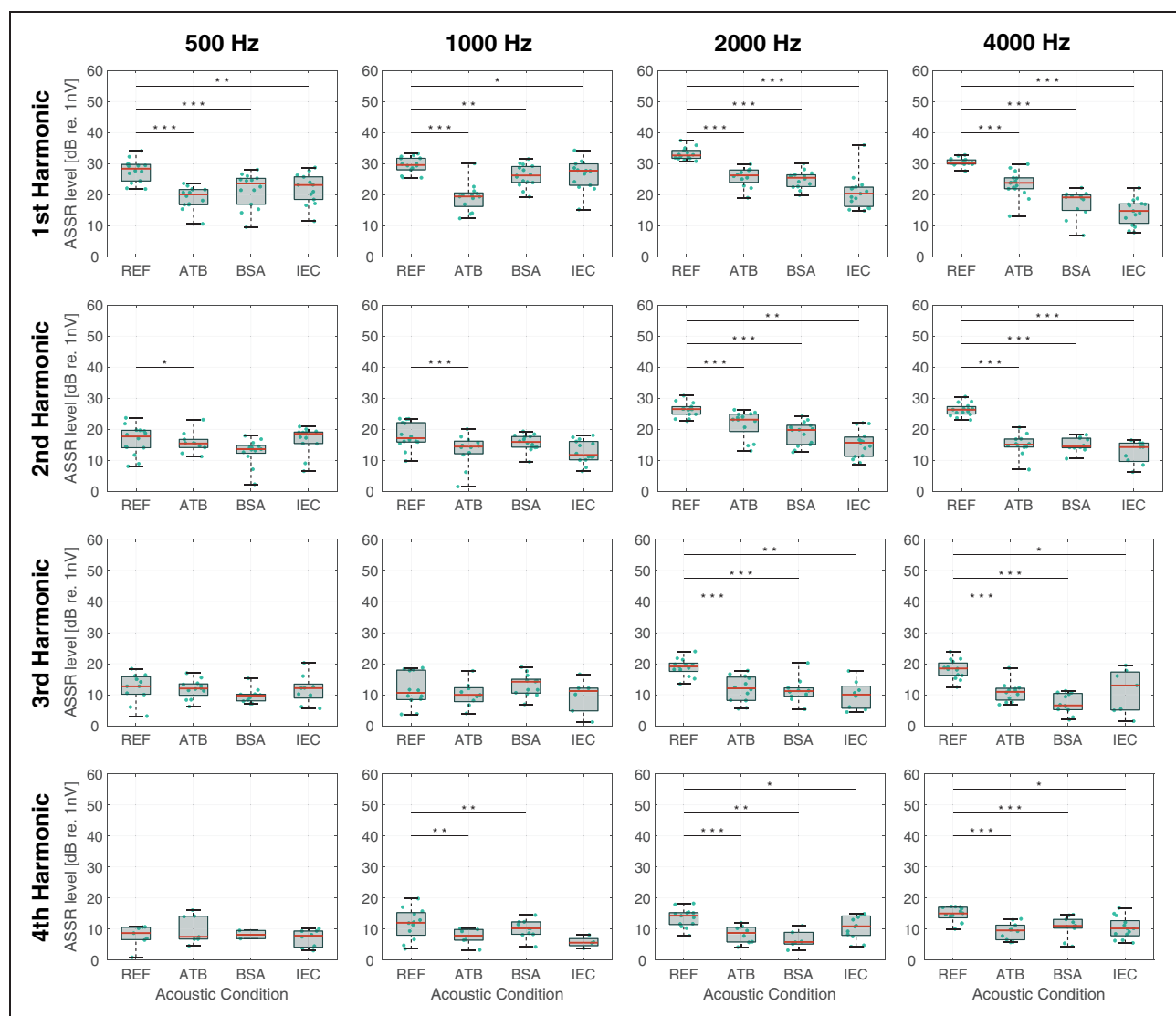

Figure 3. ASSR Levels as a Function of the Tested Acoustic Conditions for Each Stimulus Center Band (Columns) and Harmonics (From First Harmonic in the Top Row to the Fourth Harmonic in the Bottom Row). The whiskers of the boxplot indicate the minimum and maximum data points of the distribution, boxes show the 25th (bottom edge) and 75th (top edge) percentiles as well as the median (red line). The green points indicate the individual measurements. The tested acoustic conditions correspond to the reference anechoic (REF), and three simulated rooms which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology (BSA) for sound-field audiometry for pediatric assessment, and a standardized listening room (IEC).

Frequency, and *Harmonic*), as well as all two- and three-way interactions were significant. The summary of the ANOVA is shown in Table 3. In general, the ASSR results showed a reduction in the mean response amplitude for the room conditions in comparison with the reference condition. The effect was more prominent for the first harmonic, for which the ASSR level decreased between 4 and 12 dB for the room conditions. This indicates, as expected, a significant effect of the acoustic conditions of the room on the ASSR level, which could be due to the degradation of the stimulus modulation in the reverberant conditions.

Subsequently, the post hoc comparisons between the reference condition and the room conditions across all stimulus frequencies and harmonics were analyzed. The rooms for which the ASSR level were significantly different from that obtained in the reference condition are indicated in Figure 3 by horizontal lines with the corresponding significance levels. It is noteworthy that for the stimuli at 500 and 1000 Hz, all three post hoc comparisons between the reference and room conditions were significantly different only for the first ASSR harmonic. In contrast, in the case of the 2000 and 4000 Hz stimuli, the paired comparisons revealed significant differences for all tested harmonics. These results suggest that the effect of the room condition on the stimulus modulation depends on the frequency and harmonics of the ASSR stimuli. Moreover, only at high frequencies, the higher harmonics seem to be informative of the acoustic influence of the room on the obtained ASSR level.

Detection Rate of Simulated Sound-Field ASSR

Figure 4 shows the detection rate (in %) for each individual tested condition across frequencies and harmonics. For this analysis, the detection rates were calculated based on the total number of remaining measurements after the postprocessing procedure (see Table 2). For the first harmonic, a detection rate of 100% was obtained for all acoustic conditions across all frequencies, except for the ATB room at 2000 Hz that had a detection rate of 92%. For the higher harmonics, the detection rates were mostly higher or equal in the reference condition than in the room conditions across all frequencies. The lowest detection rate (14%) was obtained with the BSA room for the fourth harmonic of the 500 Hz ASSR stimulus. However, the room condition with the fewest successful detections overall was the IEC room. Importantly, the pattern of detection rates for all acoustic conditions varied across harmonics and frequencies.

ASSR Level and Early Decay Time

Figure 5 shows the mean ASSR level obtained for the tested acoustic conditions as a function of the EDT calculated for each stimulus band center frequency and harmonic. For this analysis, EDT was added as a continuous predictor, and the analysis included only the harmonics in which all post hoc comparisons between the reference and the room conditions showed a significant difference. The linear mixed-model revealed significant main effects of *EDT*, *Frequency*, and *Harmonic*. The analysis also showed significant two-way interactions between *EDT* and *Frequency* and between *EDT* and *Harmonic*. In contrast, the two-way interaction between *Frequency* and *Harmonic*, as well as the three-way interaction were not significant. The outcome of the ANOVA is summarized in Table 3.

To determine whether the ASSR level can be predicted by the EDT, linear regression models were fit to the data, in terms of the slope and coefficient of
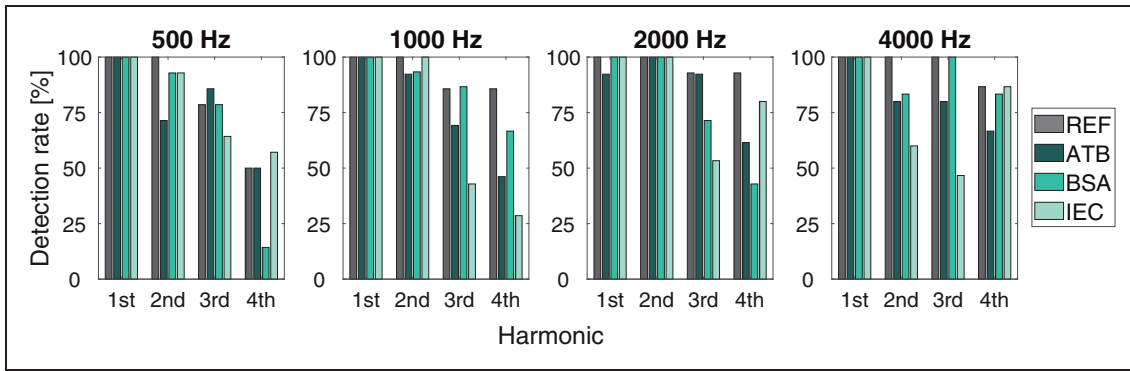
Table 3. Summary Results of the Mixed-Model Analyses of Variance.

Model 1. Effect of the room. AIC = 879.26

Factor	ASSR level	
	F statistic	p
Room	$F(3, 729.4) = 131.9$	<.0001***
Freq.	$F(3, 729.5) = 25$	<.0001***
Harm.	$F(3, 729.4) = 603.9$	<.0001***
Room × Freq.	$F(3, 729.9) = 17.5$	<.0001***
Room × Harm.	$F(3, 729.7) = 5.7$	<.0001***
Freq. × Harm.	$F(3, 729.9) = 10$	<.0001***
Room × Freq. × Harm.	$F(3, 729.6) = 4.6$	<.0001***

Model 2. Effect of the EDT. AIC = 796.93

Factor	ASSR level	
	F statistic	p
EDT	$F(1, 486.9) = 101.6$	<.0001***
Freq.	$F(3, 486.1) = 7.6$	<.0001***
Harm.	$F(3, 485.9) = 14.5$	<.0001***
EDT × Freq.	$F(3, 486.3) = 5.3$	<.0001***
EDT × Harm.	$F(3, 486.5) = 5.3$	.0014**
Freq. × Harm.	$F(3, 485.6) = 2.6$	.0526
EDT × Freq. × Harm.	$F(3, 486.4) = 0.1$	.9812

Model 3. Effect of the relative modulation power. AIC = 845.25

Factor	ASSR level	
	F statistic	p
RModP	$F(1, 485.5) = 11.5$	.0007***
Freq.	$F(3, 485.4) = 23.1$	<.0001***
Harm.	$F(3, 485.9) = 38.5$	<.0001***
RModP × Freq.	$F(3, 485.6) = 11.3$	<.0001***
RModP × Harm.	$F(3, 486.2) = 53.6$	<.0001***
Freq. × Harm.	$F(3, 485.8) = 18.2$	<.0001***
RModP × Freq. × Harm.	$F(3, 486.1) = 11.1$	<.0001***

Note. AIC = Akaike information criterion; ASSR = auditory steady-state response; EDT = early decay time; RModP = relative modulation power. * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Figure 4. Detection rates for each tested acoustic condition, across harmonics and frequencies. The tested acoustic conditions correspond to the reference anechoic (REF), and three simulated rooms which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology (BSA) for sound-field audiometry for pediatric assessment, and a standardized listening room (IEC).
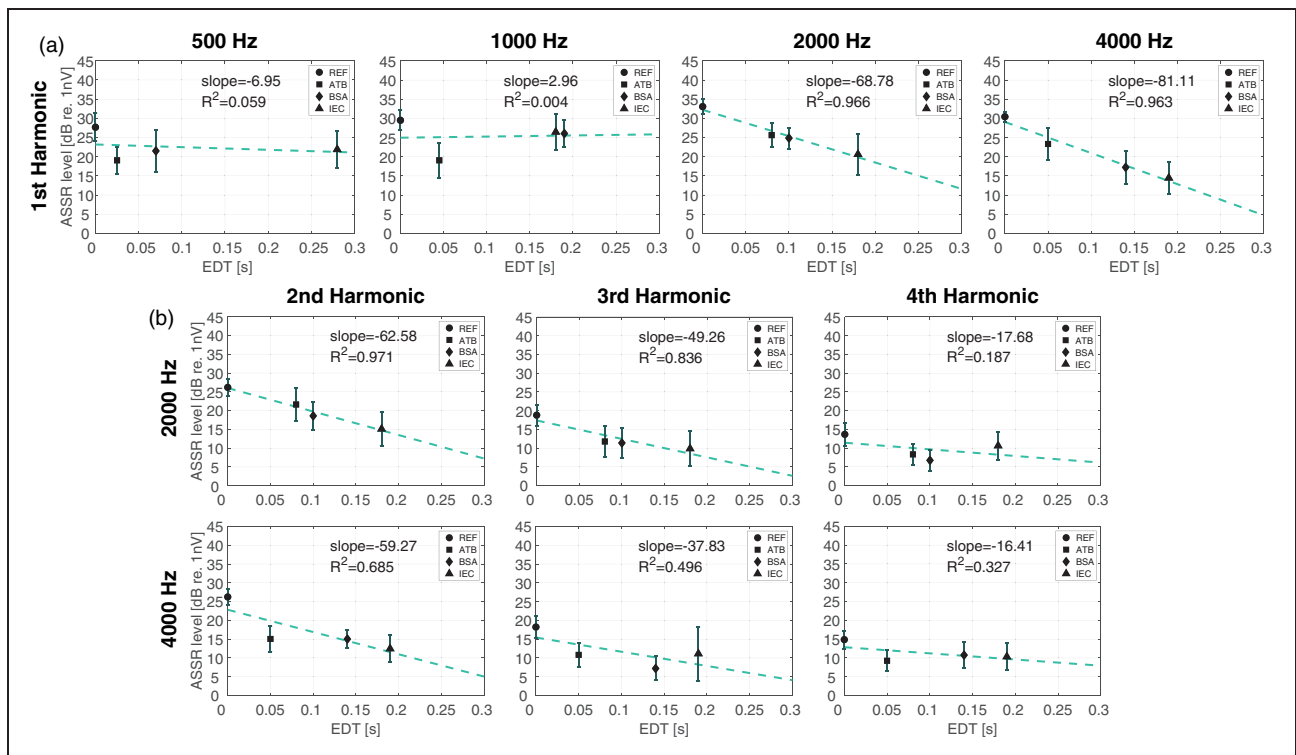

Figure 5. Mean ASSR Level as a Function of *EDT* for Each *Room*, *Frequency*, and *Harmonic*. Panel A: Results of the first harmonic for all stimulation frequencies. Panel B: Results for the included higher harmonics (second, third, and fourth) for 2000 and 4000 Hz. Error bars indicate standard deviation. Estimated regression lines (- - - -) are added to each panel. The tested acoustic conditions correspond to the reference anechoic (REF), and three simulated rooms which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology (BSA) for sound-field audiometry for pediatric assessment, and a standardized listening room (IEC).

determination (R^2). Panel A of Figure 5 shows the results for the first response harmonic for each stimulus frequency. A low correlation between the EDT and ASSR level was obtained for 500 ($R^2 = .06$) and 1000 Hz ($R^2 = .004$) for which the ASSR level did not decrease substantially with increasing the EDT. In contrast, a high correlation was found between the EDT and the ASSR response amplitude for 2000 ($R^2 = .97$) and 4000 Hz ($R^2 = .96$). In general, high correlations were also obtained for the higher harmonics, as

shown in panel B of Figure 5. The regression models showed that more than 60% of the variation in ASSR level can be explained by the EDT for the second and third harmonics for 2000 Hz, as well as the second harmonic for 4000 Hz. In the case of the fourth harmonic for 2000 Hz, and third and fourth harmonics for 4000 Hz, the linear regression models predicted approximately 20% of the variance of the ASSR data. These results indicate that EDT could be a useful predictor of the ASSR level in the rooms for 2000 and 4000 Hz. In addition, the different slopes obtained for the regression models further support that the effect of EDT on the ASSR level is frequency and harmonic dependent, as reflected in the significant two-way interactions.

ASSR Level and Relative Modulation Power

Figure 6 shows the relation between mean ASSR level and the relative modulation power across the tested acoustic conditions for each stimulus band center frequency and harmonic. The model revealed significant main effects (*RModP*, *Frequency*, and *Harmonic*), as well as all significant two- and three-way interactions.

As in the analysis of the *EDT*, only harmonics with all significant post hoc comparisons between the reference condition and the three room conditions were included. The summary of the ANOVA is shown in Table 3. As expected, the stimulus modulation was degraded for all three room conditions compared with the reference condition, resulting in lower ASSR levels as the relative modulation power decreased. This effect is observed across all analyzed stimulus band center frequencies and harmonics.

Linear regression models were fit to the data measured for each combination of stimulus band center frequency and harmonic. This was done to test whether the proposed relative modulation power model could account for the changes in the ASSR level due to the acoustics of the room. The linear regression models with their respective slopes and R^2 are shown in Figure 6. Panel A depicts the ASSR level as a function of the stimulus relative modulation power for the first harmonic of all stimulus frequencies, and Panel B depicts the results for the higher harmonics for 2000 and 4000 Hz. The regression model showed a good correlation between the relative modulation power and the ASSR
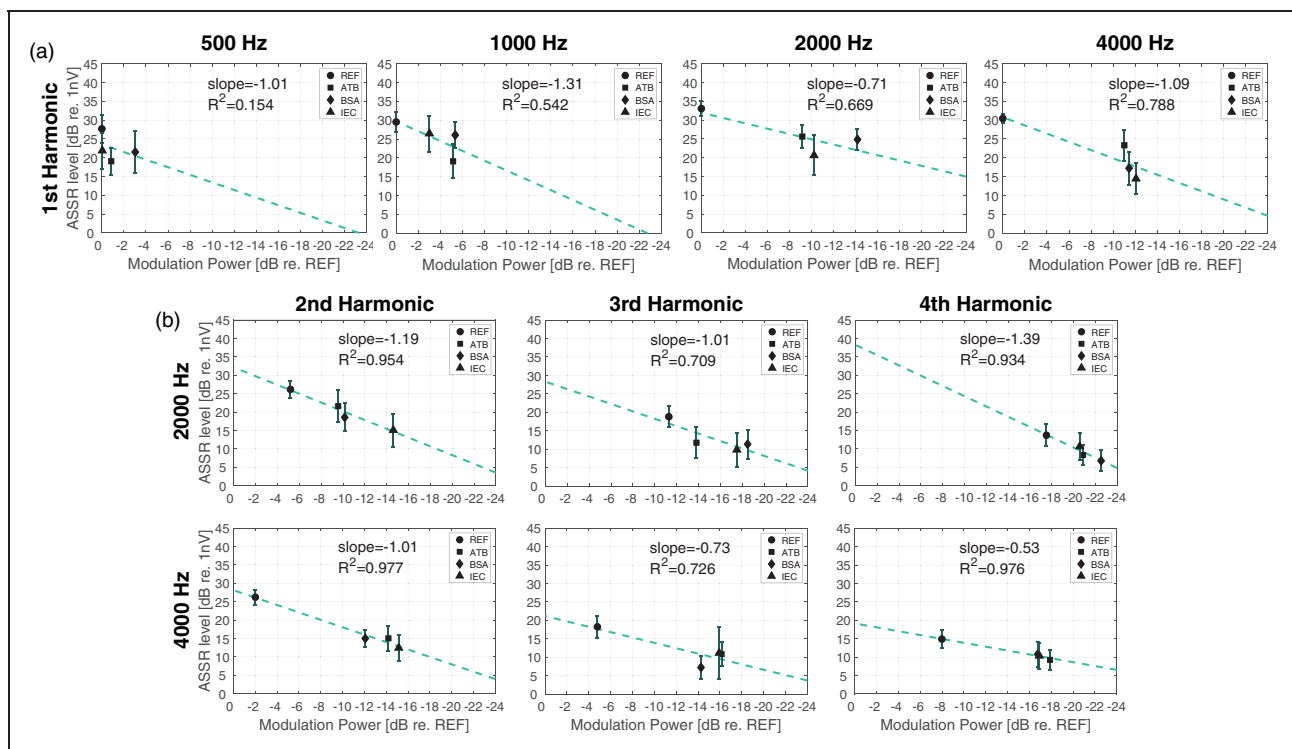

Figure 6. Mean ASSR Level as a Function of the Relative Modulation Power (*RModP*) for Each *Room*, *Frequency*, and *Harmonic*. Panel A: Results of the first harmonic for all stimulation frequencies. Panel B: Results for the higher harmonics (second, third, and fourth) for 2000 and 4000 Hz. Error bars indicate standard deviation. Estimated regression lines (- - - -) are added to each panel. The tested acoustic conditions correspond to the reference anechoic (REF), and three simulated rooms which mimic an audiometric testing booth (ATB), a room recommended by the British Society of Audiology (BSA) for sound-field audiometry for pediatric assessment, and a standardized listening room (IEC).

level, with varying slopes across stimulus frequencies and harmonics. For the first ASSR harmonic, a high correlation was found for 1000, 2000, and 4000 Hz, with R^2 values of .54, .67, and .79, respectively. In contrast, the correlation was low for the stimulus band center frequency of 500 Hz, with an R^2 value of .15. For the higher harmonics of 2000 and 4000 Hz, high correlations were obtained with coefficients of determination between .71 (for 2000 Hz, third harmonic) and .98 (for 4000 Hz, third and fourth harmonic). The results suggest that the ASSR level in the room can be partly predicted by the relative modulation power model for all considered frequencies and harmonics.

Discussion

Effect of the Room on the ASSR Level

The main finding that emerged from this study is that the ASSR level indeed was reduced for the nonanechoic room conditions compared with the reference anechoic condition. This is a novel finding as the effect of room acoustics on ASSR level has not been systematically investigated in previous studies, which have been mostly focused on testing the viability of ASSR measurements in sound field for hearing assessment and hearing-aid fitting validation (Damarla & Manjula, 2007; Hernandez-Perez & Torres-Fortuny, 2013; Park et al., 2013; Picton et al., 1998; Sardari et al., 2015; Shemesh et al., 2012; Stroebel et al., 2007; Selim et al., 2012). The reduction in the ASSR level for the room conditions was ascribed to smaller stimulus modulations for the measured nonanechoic conditions (see Figure 2B). This is also consistent with the fact that the modulation of an acoustic signal can be degraded by the reverberation and background noise of the room in which it is reproduced (Houtgast et al., 1980; Plomp, 1983). The reduction in the ASSR level could directly lead to an increase in the measurement time. This is because a longer recording time would be needed for lower ASSR levels to reach the signal-to-noise ratio required for the detection of the response, as demonstrated in earlier studies (Cebulla et al., 2006; Dobie & Wilson, 1996; Laugesen et al., 2018). This could pose a challenge to the clinical implementation of sound-field ASSR, where minimizing the testing time is crucial, especially when testing infants and hard-to-test patients.

Detection Rate of Simulated Sound-Field ASSR

Despite the reduction in the ASSR level, the ASSR was detected in all simulated acoustic conditions tested. For the reference condition, the detection rate analysis showed a reduction in the detected responses toward the higher harmonics. This is in agreement with previous investigations in which the ASSR was measured with traditional insert earphone stimulation (Cebulla et al., 2006; Laugesen et al., 2018). Interestingly, this pattern was not observed consistently across the simulated tested rooms for which the detection rate did not consistently reduce for the higher harmonics. For instance, for the IEC room condition, the percentage of successful detections for the fourth harmonic was higher than for the second and third harmonics of the 4000 Hz frequency band. Considering this, a multiharmonic detector, such as the q-sample detector, might provide higher successful detection rates for sound-field ASSR measurements compared with a one-sample detector that only analyzes a single harmonic (Cebulla et al., 2006). Furthermore, a multiharmonic detector might be particularly useful to compensate for the potential longer detection times caused by the reduced response amplitude obtained in the room conditions.

ASSR Level and Early Decay Time

In this study, it is investigated whether EDT can be used as a predictor of the sound-field ASSR level measured in a room. The results showed high correlations between the EDT and ASSR level for the frequencies of 2000 and 4000 Hz for all harmonics: A reduction in the ASSR level as the EDT increases was clearly observed. In the case of the frequencies of 500 and 1000 Hz, it was surprising that there was no correlation between the ASSR level and EDT as the EDTs observed were similar to those for 2000 and 4000 kHz. Further investigation is required to clarify whether the observed frequency-dependent effect of the EDT on the ASSR level generalizes when considering a larger sample of rooms.

ASSR Level and Relative Modulation Power

The relation between the ASSR level and the stimulus modulation in a room was analyzed. The stimulus modulation was quantified using an auditory-inspired relative modulation power model, which correlated well with the ASSR level. In general, it was observed that as the relative modulation power decreased, ASSR level was also reduced, as expected. However, a direct comparison between this study and the literature is challenging due to the lack of systematic investigations of the effect of stimulus modulation on ASSR measurements. Many studies have reported the effect of the stimulus modulation on the ASSR level as a function of the modulation depth for amplitude-modulated sinusoidal signals presented to normal-hearing subjects through insert earphones (Bharadwaj et al., 2015; Boettcher et al., 2001; Dimitrijevic et al., 2001; John et al., 2001; Kuwada et al., 1986; Lins et al., 1995; Picton et al., 1987; Rees et al., 1986; Roß et al., 2000; Rønne, 2012). Although in these

studies the researchers employed different measurement parameters (e.g., stimulus level, carrier and modulation frequency), all of them consistently showed an increase in the ASSR level as the modulation depth increased for the first harmonic of the response.

To compare the modulation-growth functions reported in the literature with the one obtained in this study, linear regression lines were fitted to each data set from the literature. For all studies, the ASSR level and modulation depth values were transformed to dB relative to 1 nV and a 100% modulation depth, respectively. The slopes obtained for each study as well as the measurement parameters used are reported in Table 4. The modulation-growth functions (physiological input/output curves) of the current investigation for the first harmonic were in general steeper than those obtained in the literature. For instance, Rees et al. (1986), Lins et al. (1995), and John et al. (2001) measured IO curves for 1000 Hz and repetition rates around 85 Hz, with estimated slopes of −0.44, −0.34, and −0.66 (dB/dB), respectively. In the present experiment, a slope of −1.31 (dB/dB) was obtained for

the measured IO curve for 1000 Hz. It is important to highlight that for fast repetition rates, only modulation-growth functions for 1000 Hz have been previously reported in the literature.

Implication and Limitations

The findings of this study highlighted the importance of the evaluation of room acoustics for the implementation of sound-field ASSR measurements. However, some consideration should be taken into account before generalizing the results to realistic clinic environments: (a) The room acoustic model implemented in this study was limited to a monaural point-to-point simulation, and hence, it did not include the effect of the patient's head and torso on the local sound field. In addition, the model is most accurate for lightly damped rooms with evenly distributed absorption on the surfaces, which is far from realistic clinic rooms. (b) This investigation only focused on three room conditions, which is a limited sample compared with the variety of audiological testing

Table 4. Modulation-Growth Functions Reported in Literature.

Report	Subjects	Stimulus rate (Hz)	Carrier frequency (Hz)	Level[a]	Tested modulation	Slope
40 Hz range						
Roß et al. (2000)	8	39	250	70 dB SL	100%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, and 5%	−0.51
Boettcher et al. (2001)	10	40	520	65 dB SPL	100%, 80%, 70%, 50%, 40%, 20%, 10%, 5%, and 0%	−0.63
Kuwada et al. (1986)	4	50	1000	60 dB SPL	90%, 80%, 70%, 60%, 50%, 30%, 10%, and 1%	−0.70
Picton et al. (1987)	5	39.1	1000	70 dB HL	90%, 70%, 50%, 30%, and 10%	−0.61
Picton et al. (1987)	8	∼40	500	76.5 dB SPL	50%, 30%, and 10%	−0.76
Picton et al. (1987)	8	∼40	1000	76.5 dB SPL	50%, 30%, and 10%	−0.81
Picton et al. (1987)	8	∼40	2000	76.5 dB SPL	50%, 30%, and 10%	−0.55
Picton et al. (1987)	8	∼40	4000	76.5 dB SPL	50%, 30%, and 10%	−0.54
Rønne (2012)	10	40	1000	55 dB SPL	0, −4, −8, −12 dB	−0.78
Boettcher et al. (2001)	10	40	4000	65 dB SPL	100%, 80%, 70%, 50%, 40%, 20%, 10%, 5%, and 0%	−0.41
90 Hz range						
Dimitrijevic et al. (2001)	10	80.1	750	50 dB SPL	100% and 50%	−0.98
Rees et al. (1986)	10	80	1000	55 dB SL	100%, 80%, 50%, 20%, 10%, and 5%	−0.44
Lins et al. (1995)	5	91	1000	60 dB SPL	100%, 75%, 50%, and 25%	−0.34
John et al. (2001)	8	82.3	1000	60 dB SPL	100%, 50%, 20%, 10%, and 5%	−0.66
Dimitrijevic et al. (2001)	10	85	1500	50 dB SPL	100% and 50%	−0.99
Dimitrijevic et al. (2001)	10	89.8	3000	50 dB SPL	100% and 50%	−1.00
Dimitrijevic et al. (2001)	10	94.7	6000	50 dB SPL	100% and 50%	−0.99
Bharadwaj et al. (2015)	26	100	4000	75 dB SPL	0, −4, −8, −12 dB	−0.97

Note. For the Picton et al.'s (1987) data presented at 76.5 dB SPL, only six subjects participated in the recording with 30% modulation depth. For the Bharadwaj et al.'s (2015) data, the stimulus used was an SAM tone in notched noise. For the Lins et al.'s (1995) data, the stimuli were calibrated based on a constant peak-to-peak value.
[a]Level: SPL, sound pressure level; HL, hearing level; SL, sensation level; SAM = sinusoidally amplitude-modulated.

rooms. It would thus be beneficial to expand the room sample in future studies to consider a broader range of acoustic scenarios that can be found in clinics. (c) Only one measurement point was considered for the analysis of the sound field. In reality, during the sound-field ASSR measurement, it is expected that the patients move their heads, producing local changes in the sound field. (d) The effect of the background noise of the test room on the sound-field ASSR measurements was not considered in this study, which is important due to the high ambient noise levels in audiometric testing rooms (Frank & Williams, 1994; Siegenthaler, 1981). An additional degradation on the ASSR level must be expected due to the background noise of the room as this also reduces the stimulus modulation (Houtgast et al., 1980; Plomp, 1983). (e) This investigation only considered monaural stimulation, which agrees with the preferred approach for hearing-aid (HA) fitting validation: to test each ear separately. However, in consideration of testing time, some audiologists may choose to test both ears simultaneously. Watson et al. (2019) considered binaural stimulation for hearing-aid fitting validation with ASSR, but only in simulated anechoic conditions. Their investigations of effects of head shadow, interaural time differences, and asymmetrical hearing-aid fittings showed slightly stronger responses with binaural versus monaural stimulation, except in an extreme case of an interaural delay that set the stimulus envelope in antiphase between the two ears. In real sound-field ASSR, there will additionally be binaural effects of room reverberation. While the perceptual benefits of binaural over monaural listening are well documented, we are not aware of any published studies of this aspect of sound-field ASSR. These aspects should be further explored for a better understanding of the effect of room acoustics on sound-field ASSR measurements that could lead to the successful implementation in clinics.

Conclusions

This study provides a first step toward understanding the effect of room acoustics on sound-field ASSR measurements. Using a simple room-acoustic model to simulate three rooms inspired by audiological testing rooms, it was shown that room acoustics indeed affects the level of sound-field ASSR measurements. This was evident in the general reduction of ASSR level for all harmonics obtained across the tested room conditions compared with the anechoic reference. This reduction in the ASSR level is likely to be attributed to the degradation of the stimulus modulation due to the nonanechoic reproduction. Although the ASSR level was reduced for all tested room conditions, ASSRs were almost always detected for the first harmonic across all acoustic conditions tested. For the room conditions, the detection rates did not

consistently decrease with increasing harmonic number. In contrast, for the reference condition, the number of detected ASSRs reduced toward the higher harmonics. The effect of the room on the ASSR level was characterized in terms of the EDT and relative modulation power. While EDT performed well for 2000 and 4000 Hz, the relative modulation power correlated well for all frequencies. These two parameters appear to be useful to analyze the changes in ASSR level produced by the acoustical properties of the measurement room. The relative modulation power and the EDT are acoustic parameters that can be easily recorded in any room. This will then be important in clinical practice as clinicians could measure the proposed parameters to evaluate the testing environment and determine whether it is acoustically suitable for sound-field ASSR measurements.

ORCID iD

Valentina Zapata-Rodriguez (iD) https://orcid.org/0000-0003-3175-5919

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

ANSI S3.1. (2003). American National Standard Methods for Maximum permissible ambient noise levels for audiometric test rooms (Rev. ed.)(ANSI S3. 1-1999).

ANSI S3.5. (1997). American National Standard Method for calculation of the speech intelligibility index.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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. https://doi.org/10.1523/jneurosci.3915-14.2015

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*(1–2), 32–42. https://doi.org/10.1016/s0378-5955(00)00255-0

Bradley, J. S., Reich, R. D., & Norcross, S. G. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. *The Journal of the Acoustical Society of America, 106*(4), 1820–1828. https://doi.org/10.1121/1.427932

British Society of Audiology. (2019). *Practice guidance: The acoustics of sound field audiometry in clinical audiological applications*. https://www.thebsa.org.uk/wp-content/uploads/2019/04/OD104-79-Acoustics-of-Sound-Field-Audiometry-in-Clinical-Audiological-Applications-FINAL-Feb-2019.pdf

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. https://doi.org/10.3766/jaaa.17.2.3

Damarla, V. K., & Manjula, P. (2007). Application of ASSR in the hearing aid selection process. *Australian and New Zealand Journal of Audiology, 29*(2), 89–97. https://doi.org/10.1375/audi.29.2.89

Dimitrijevic, A., John, M. S., van Roon, P., & Picton, T. W. (2001). Human auditory steady-state responses to tones independently modulated in both frequency and amplitude. *Ear and Hearing, 22*(2), 100–111. https://doi.org/10.1097/00003446-200104000-00003

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. https://doi.org/10.1121/1.417933

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), 2955–2964. https://doi.org/10.1121/1.3489111

Frank, T., & Williams, D. L. (1994). Ambient noise-levels in industrial audiometric test rooms. *American Industrial Hygiene Association Journal, 55*(5), 433–437. https://doi.org/10.1080/15428119491018871

Health Technical Memorandum 2045. (1996). *Acoustics – audiology*. NHS Estates.

Hernandez-Perez, H. & Torres-Fortuny, A. (2013). Auditory steady state response in sound field. *International Journal of Audiology, 52*(2), 139–143. https://doi.org/10.3109/14992027.2012.727103

Houtgast, T., & Steeneken, H. J. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *The Journal of the Acoustical Society of America, 77*(3), 1069–1077. https://doi.org/10.1121/1.392224

Houtgast, T., Steeneken, H. J., & Plomp, R. (1980). Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acoustics. *Acta Acustica United With Acustica, 46*(1), 60–72.

IEC 268-13. (1985). Sound system equipment-Part 13: Listening tests on loudspeakers. International Electrotechnical Commission.

ISO 3382-1. (2009). Acoustics—Measurement of room acoustic parameters—Part 1: Performance spaces. International Organization for Standardization, Geneva, Switzerland.

ISO 8253-2. (2009). Acoustics-Audiometric test methods—Part 2. Sound field audiometry with pure-tone and narrow-band test signals. International Organization for Standardization, Geneva, Switzerland.

Jacobsen, F., & Juhl, P. M. (2013). *Fundamentals of general linear acoustics*. John Wiley & Sons.

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., Dimitrijevic, A., van Roon, P., & Picton, T. W. (2001). Multiple auditory steady-state responses to AM and FM stimuli. *Audiology and Neurotology, 6*(1), 12–27. https://doi.org/10.1159/000046805

Kuwada, S., Batra, R., & Maher, V. L. (1986). Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. *Hearing Research, 21*(2), 179–192. https://doi.org/10.1016/0378-5955(86)90038-9

Laugesen, S., Rieck, J. E., Elberling, C., Dau, T., & Harte, J. M. (2018). On the cost of introducing speech-like properties to a stimulus for auditory steady-state response measurements. *Trends in Hearing, 22*. https://doi.org/10.1177/2331216518789302

Lins, O. G., Picton, P. E., Picton, T. W., Champagne, S. C., & Durieux-Smith, A. (1995). Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz. *The Journal of the Acoustical Society of America, 97*(5), 3051–3063. https://doi.org/10.1121/1.411869

Marcoux, A., & Hansen, M. (2003). Ensuring accuracy of the pediatric hearing aid fitting. *Trends in Amplification, 7*(1), 11–27. https://doi.org/10.1177/108471380300700103

Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics, 106*(3), e43–e43. https://doi.org/10.1542/peds.106.3.e43

Morton, C. C., & Nance, W. E. (2006). Newborn hearing screening—A silent revolution. *New England Journal of Medicine, 354*(20), 2151–2164. https://doi.org/10.1056/nejmra050700

Naumann, K., Euler, H. A., Knauth, M., & White, K. (2015, October 09–10). *The global status of newborn and infant hearing screening* [Paper presentation]. 6th Annual Coalition for Global Hearing Health Conference, Washington, DC, United States. http://conference.usu.edu/SYSTEM/Uploads/pdfs/15233_1964KatrinNeumann.pdf

Neumann, K., Chadha, S., Tavartkiladze, G., Bu, X., & White, K. R. (2019). Newborn and infant hearing screening facing globally growing numbers of people suffering from disabling hearing loss. *International Journal of Neonatal Screening, 5*(1), 7. https://doi.org/10.3390/ijns5010007

Park, E. S., Bahng, J., Lee, H. J., & Kim, H. J. (2013). The usefulness of sound-field auditory steady state response (SF ASSR): Comparison of hearing sensitivity and typical ASSR. *Audiology, 9*(1), 15–24. https://doi.org/10.21848/audiol.2013.9.1.15

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.

Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D. (2003). Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos. *International Journal of Audiology, 42*(4), 177–219. https://doi.org/10.3109/14992020309101316

Picton, T. W., Skinner, C. R., Champagne, S. C., Kellett, A. J., & Maiste, A. C. (1987). Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *The Journal of the Acoustical Society of America, 82*(1), 165–178. https://doi.org/10.1121/1.395560

Plomp, R. (1983, August). Perception of speech as a modulated signal. In *Proceedings of the tenth international congress of phonetic sciences* (pp. 29–40). Dordrecht, Foris.

Ptok, M. (2011). Early detection of hearing impairment in newborns and infants. *Deutsches Ärzteblatt International, 108*(25), 426. https://doi.org/10.3238/arztebl.2011.0426

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. https://doi.org/10.1055/s-0035-1570331

Rance, G. (2008). *The auditory steady-state response: Generation, recording, and clinical application*. Plural Publishing.

Rees, A., Green, G. G. R., & Kay, R. H. (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing Research, 23*(2), 123–133. https://doi.org/10.1016/0378-5955(86)90009-2

Relaño-Iborra, H., May, T., Zaar, J., Scheidiger, C., & Dau, T. (2016). Predicting speech intelligibility based on a correlation metric in the envelope power spectrum domain. *The Journal of the Acoustical Society of America, 140*(4), 2670–2679. https://doi.org/10.1121/1.4964505

Roß, B., Borgmann, C., Draganova, R., Roberts, L. E., & Pantev, C. (2000). A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America, 108*(2), 679–691. https://doi.org/10.1121/1.429600

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

Sardari, S., Jafari, Z., Haghani, H., & Talebi, H. (2015). Hearing aid validation based on 40 Hz auditory steady-state response thresholds. *Hearing Research, 330*, 134–141. https://doi.org/10.1016/j.heares.2015.09.004

Schroeder, M. R. (1981). Modulation transfer functions: Definition and measurement. *Acta Acustica United With Acustica, 49*(3), 179–182.

Schroeder, M. R., & Kuttruff, K. H. (1962). On frequency response curves in rooms. Comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing between maxima. *The Journal of the Acoustical Society of America, 34*(1), 76–80. https://doi.org/10.1121/1.1909022

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *American Statistician, 34*(4), 216–221. https://doi.org/10.2307/2684063

Selim, M. H., Mourad, M. E., El-Shennawy, A. M., & Elfouly, H. S. (2012). Comparing sound field audiometry and free field auditory steady state response in the verification of hearing aid fitting in adults. *The Egyptian Journal of Otolaryngology, 28*(3), 201. https://doi.org/10.7123/01.EJO.0000418067.42430.a3

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. https://doi.org/10.1097/00003446-200212000-00004

Shemesh, R., Attias, J., Magdoub, H., & Nageris, B. I. (2012). Prediction of aided and unaided audiograms using sound-field auditory steady-state evoked responses. *International Journal of Audiology, 51*(10), 746–753. https://doi.org/10.3109/14992027.2012.700771

Siegenthaler, B. M. (1981). A survey of hearing test rooms. *Ear and Hearing, 2*(3), 122–126. https://doi.org/10.1097/00003446-198105000-00007

Singh, V. (2015). Newborn hearing screening: Present scenario. *Indian Journal of Community Medicine: Official Publication of Indian Association of Preventive & Social Medicine, 40*(1), 62. https://doi.org/10.4103/0970-0218.149274

Sininger, Y. S., Grimes, A., & Christensen, E. (2010). Auditory development in early amplified children: Factors influencing auditory-based communication outcomes in children with hearing loss. *Ear and Hearing, 31*(2), 166–185. https://doi.org/10.1097/aud.0b013e3181c8e7b6

Slama, M. C., & Delgutte, B. (2015). Neural coding of sound envelope in reverberant environments. *Journal of Neuroscience, 35*(10), 4452–4468. https://doi.org/10.1523/JNEUROSCI.3615-14.2015

Stroebel, D., Swanepoel, D., & Groenewald, E. (2007). Aided auditory steady-state responses in infants: Respuestas auditivas de estado estable en niños con auxiliares auditivos. *International Journal of Audiology, 46*(6), 287–292. https://doi.org/10.1080/14992020701212630

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics, 5*(2), 99–114. https://doi.org/10.2307/3001913

Watson, S. D., Laugesen, S., & Epp, B. (2019). Provoking and minimising potentially destructive binaural stimulation effects in auditory steady-state response (ASSR) measurements. In A. Kressner, J. Regev, J. C.-Dalsgaard, L. Tranebjærg, S. Santurette, and T. Dau (eds), *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. The Danavox Jubilee Foundation, 2019. c The Authors. ISSN: 2596-5522.

Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of early- and later identified children with hearing loss. *Pediatrics, 102*(5), 1161–1171. https://doi.org/10.1542/peds.102.5.1161