Effect of room acoustics on sound-field auditory steady-state response (ASSR) measurements

Zapata Rodriguez, Valentina

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Acoustic Technology
Department of Electrical Engineering

Effect of room acoustics on sound-field auditory steady-state response (ASSR) measurements

Valentina Zapata-Rodríguez

Ph. D. Thesis
2020

IRU
Interacoustics Research Unit

Interacoustics

Effect of room acoustics on sound-field auditory steady-state response (ASSR) measurements

Author
Valentina Zapata-Rodríguez

Supervisors
Cheol-Ho Jeong, Associate Professor, Technical University of Denmark
Jonas Brunskog, Associate Professor, Technical University of Denmark
James M. Harte, Senior Director, Interacoustics Research Unit
Søren Laugesen, Senior Research Engineer, Interacoustics Research Unit

Technical University of Denmark
Acoustic Technology, Department of Electrical Engineering
Ørsteds Plads, Building 352
2800 Kgs. Lyngby
Denmark

Interacoustics Research Unit
c/o Technical University of Denmark
Ørsteds Plads, Building 352
2800 Kgs. Lyngby
Denmark

Interacoustics A/S
Audiometer Allé 1
5500 Middelfart
Denmark

Education Ph.D.

Field Engineering Acoustics and Technical Audiology

Class Public

Edition 1. edition

Preface

This thesis was submitted to the Department of Electrical Engineering in partial fulfillment of the requirements for obtaining the degree of Doctor of Philosophy (Ph.D.) at the Technical University of Denmark (DTU). It is the product of an industrial Ph.D. project carried out at the Acoustic Technology group in DTU and the Interacoustic Research Unit (IRU) from Interacoustics A/S. The work was completed between July 2016 and March 2020 under the supervision of Associate Professor Cheol-Ho Jeong and Associate Professor Jonas Brunskog from DTU, along with Senior Director James Harte and Senior Research Engineer Søren Laugessen from IRU. The project was funded by Interacoustics A/S, the William Demant Foundation and the Innovation Fund Denmark, Grant. No. 5189-00070B.

Abstract

The main purpose of universal newborn hearing screening programs is to identify hearing loss in infants at the earliest stage possible, which can lead to clinical intervention at a very young age (< 3 months). Accordingly, many pre-lingual infants are fitted with hearing aids which allow them to develop communication skills. The validation of the hearing aid fitting is a key procedure, through which it is ensured that the infant is receiving an appropriate amplification, and hence, has access to speech sounds. For infants younger than 8 months, standard audiological measurements (e.g. visual reinforcement audiometry) are highly unreliable since they require active participation of the patient. Sound-field auditory steady-state responses (ASSRs), have been suggested as a potential clinical test for the objective validation of the hearing-aid fitting. This test measures the evoked potential elicited by auditory neurons, which follow the envelope of a periodically repeated stimulus presented through a loudspeaker. The loudspeaker presentation allows a clinician to validate the fitting of the hearing aid under its normal mode of operation. However, when the stimulus is presented in a sound field, its envelope and modulation could be affected due to the reverberation of the test room. This, in turn, can potentially affect the ASSR level that is highly dependent on the stimulus modulation. At the start of this project, a proper characterization of the effect of the room on sound field ASSR measurements was lacking, which is crucial for assessing the clinical viability of the test for hearing aid fitting validation.

This PhD project explored how room acoustics affects sound field ASSR measurements, and the practical implications for its potential clinical implementation. Throughout this thesis, it is shown that the room has a detrimental effect on the ASSR level. This reduction in the level leads to longer detection times and could result in lower detection rates when the testing time is limited. The assessment of 31 state-of-the-art audiometric testing rooms conducted in this project revealed a large variability of the acoustical conditions across rooms as well as between testing positions within the rooms. The effect of the room acoustics was systematically investigated in loudspeaker-based virtual acoustic environments using accurate acoustic simulation. Moreover, the inclusion of a nearby loudspeaker within the reproduction system to faithfully account for the short source to receiver distance (1 m) required for sound field ASSR measurements was validated. The results of the project showed that despite the large acoustic variability across audiometric testing rooms, sound field ASSR measurements could be generally conducted in the existing acoustic environments. However, care should be taken when selecting the room and testing positions for sound field ASSR measurements, since the detrimental effect of the room on the ASSR level can compromise, in some cases, the clinical viability of the test. It was also demonstrated that the room-induced variation

in ASSR level could be predicted using two room acoustic descriptors: the early decay time and a proposed auditory-model-inspired relative modulation power metric. Both descriptors can be measured in situ and could be used for determining whether the room is suitable for sound field ASSR. This, in the long term, could be implemented as part of a calibration protocol for sound field ASSR in the clinics.

Resumé

Det primære formål med universel undersøgelse af nyfødtes hørelse er at identificere et eventuelt høretab så tidligt som muligt, således at en klinisk behandling kan igangsættes allerede inden 3 måneder efter fødslen. I overensstemmelse hermed bliver mange spædbørn tilpasset med høreapparater før de begynder at tale, hvilket gør det muligt for dem at udvikle deres kommunikationsevner. Validering af høreapparattilpasningen er en vigtig procedure gennem hvilken det sikres at barnet faktisk får passende forstærkning således at talelydende er tilgængelige for barnet. For børn under 8 måneder er de almindelige audiometrisk målemetoder (for eksempel VRA (visual reinforcement audiometry)) særdeles upålidelige idet metoderne kræver patientens aktive deltagelse. Såkaldt lydfelts auditory steady-state responser (ASSR), er blevet foreslået som en potentiel objektiv metode til validering af høreapparattilpasninger. Denne test måler det elektriske potentiale fra neuroner i hørebanen genereret i takt med en periodisk gentaget lydstimulus der afspilles over en højttaler. Ved at præsentere lyden fra en højttaler (altså i lydfeltet) kan klinikeren validere tilpasningen af høreapparatet under normale brugsbetingelser. Ulempen er dog at når stimulussen afspilles i lydfeltet kan indhyllingen og modulationen blive påvirket af efterklangen i rummet hvor der testes. Dette kan endvidere påvirke det målte ASSR niveau, som er stærkt afhængigt af netop modulationen af den anvendte stimulus. Der fandtes ved projektets start ikke en grunding karakterisering af effekten af testrummet på lydfelts ASSR målinger, hvilket er afgørende for at kunne vurdere den kliniske brugbarhed af den foreslåede test til validering af høreapparattilpasninger.

Dette PhD projekt har undersøgt hvordan rumakustik påvirker lydfelts ASSR målinger samt de praktisk konsekvenser for en potential klinisk implementering. Igennem afhandlingen vises det at testrummet har en reducerende virkning på ASSR niveauet. Denne reduktion i niveau medfører længere detektionstider og kan resultere i lavere detektionsrater når den til rådighed værende testtid er begrænset. En undersøgelse af 31 nyere audiometriske testrum foretaget som en del af dette projekt afslørede en stor variation i de akustiske forhold mellem rum og ligeledes imellem forskellige måle-positioner i de enkelte rum. Effekten af de rumakustiske forhold blev systematisk undersøgt i højttaler-baserede virtuelle akustiske miljøer ved brug af omhyggelig rumakustisk simulering. For troværdigt at kunne gengive den korte afstand (1 m) mellem kilde og modtager, som bruges til lydfelts ASSR, tilføjedes en ekstra højttaler til det anvendte simuleringssystem, og denne fremgangsmåde blev valideret. Projektets resultater viser at på trods of den store variation i de akustiske forhold i audiometriske testrum, kan lydfelts ASSR målinger gennemføres i de eksisterende akustiske omgivelser. Man bør dog være omhyggelig med at udvælge rum såvel som målepositioner til lydfelts ASSR målinger, idet den reducerende effekt af rummet på ASSR niveauet i visse tilfælde kan hindre den

kliniske gennemførlighed af testen. Det er også vist at den rum-inducerede variation i ASSR niveau kunne forudsiges ved hjælp af to rumakustiske parametre: EDT (early decay time) og et mål for relativ modulationseffekt inspireret af auditiv modellering. Begge parametre kan måles in situ og kan således bruges til at afgøre om et givet rum er velegnet til lydfelts ASSR. På længere sigt kan dette implementeres som en del af en klinisk kalibreringsprotokol for lydfelts ASSR.

Resumen

El propósito principal de los programas universales de exploración auditiva para recién nacidos es identificar la pérdida de audición en los bebés en la etapa más temprana posible, lo cual puede llevar a una intervención clínica a una edad muy temprana (< 3 meses). En consecuencia, muchos niños prelocutivos reciben auxiliares auditivos que les permiten desarrollar habilidades de comunicación. La validación de la adaptación del auxiliar auditivo es un procedimiento clave, mediante el cual se garantiza que el niño recibe una amplificación adecuada y, por lo tanto, tiene acceso a los sonidos del habla. En el caso de los bebés menores de 8 meses, las mediciones audiológicas estándar (por ejemplo, la audiometría de refuerzo visual) son muy poco fiables, ya que requieren de la participación activa del paciente. Se ha sugerido que las respuestas auditivas de campo sonoro de estado estacionario (ASSR), son una prueba clínica potencial para la validación objetiva de la adaptación de los auxiliares auditivos. Esta prueba mide el potencial evocado que provocan las neuronas auditivas, al seguir la envolvente de un estímulo repetido periódicamente y presentado a través de un altavoz. La presentación por altavoz permite al audiólogo validar la adaptación del auxiliar auditivo en su modo normal de funcionamiento. Sin embargo, cuando el estímulo se presenta en un campo sonoro, su envolvente y su modulación podrían verse afectadas debido a la reverberación de la sala de pruebas. Esto, a su vez, puede afectar potencialmente el nivel del ASSR que depende en gran medida de la modulación del estímulo. Al comienzo de este proyecto, faltaba una caracterización adecuada del efecto de la sala en las mediciones del ASSR de campo sonoro, lo cual es crucial para evaluar la viabilidad clínica de la prueba de validación de la adaptación del auxiliar auditivo.

Este proyecto de doctorado exploró cómo la acústica de la sala afecta a las mediciones de ASSR de campo sonoro, y las implicaciones prácticas para su posible aplicación clínica. A lo largo de esta tesis, se demuestra que la sala tiene un efecto perjudicial en el nivel del ASSR. Esta reducción del nivel conduce a tiempos de detección más largos y podría resultar en tasas de detección más bajas cuando el tiempo de prueba es limitado. La evaluación de 31 salas de pruebas audiométricas de última generación realizada en este proyecto reveló una gran variabilidad de las condiciones acústicas entre las salas, así como entre las posiciones de prueba dentro de las salas. Se investigó sistemáticamente el efecto de la acústica de la sala en entornos acústicos virtuales basados en altavoces utilizando una simulación acústica precisa. Además, se validó la inclusión de un altavoz cercano dentro del sistema de reproducción para tener en cuenta fielmente la corta distancia entre la fuente y el receptor (1 m) requerida para las mediciones del ASSR de campo sonoro. Los resultados del proyecto demostraron que, a pesar de la gran variabilidad acústica entre las salas de pruebas audiométricas, las mediciones del ASSR de campo

sonoro podían realizarse en general en los entornos acústicos existentes. Sin embargo, se debe tener cuidado al seleccionar la sala y las posiciones de prueba para las mediciones del ASSR de campo sonoro, ya que el efecto perjudicial de la sala sobre el nivel del ASSR puede comprometer, en algunos casos, la viabilidad clínica de la prueba. También se demostró que la variación del nivel del ASSR inducido por la sala podía predecirse utilizando dos descriptores acústicos de la sala: el tiempo de decaimiento temprano y una medición relativa de la potencia de modulación inspirada en modelos auditivos que se propuso. Ambos descriptores pueden medirse in situ y podrían utilizarse para determinar si el recinto es adecuado para el ASSR de campo sonoro. Esto, a largo plazo, podría implementarse como parte de un protocolo de calibración para el ASSR de campo sonoro en las clínicas.

Acknowledgements

Firstly, I would like to thank my supervisors James Harte, Cheol-Ho Jeong, Søren Laugesen, and Jonas Brunskog, for their unconditional support all these years. Thanks for the trust, working with all of you has been truly inspiring. Without your encouragement and guidance, this project would not have been possible.

I would like to express my gratitude to Gerd Marbjerg, Axel Ahrens, Johannes Zaar, Efren Fernandez, Marton Marshall, and Mario Cebulla for the productive discussions. Thanks to all the representatives from the hospitals for their assistance in organizing the measurement sessions. Your contributions have been valuable for the project.

l want to thank all my colleagues at the Acoustic Technology Group and the Hearing System for the good moments. It was great to work in such a supportive environment.

Special thanks go to the Interacoustic Research Unit: Sinnet Kristensen, Lisbeth Simonsen, Bue Kristensen, and Claus Elberling. It has been a pleasure to be part of the team.

This project would not have been possible without the technical assistance from Jørgen Rasmussen, Tom Arent Petersen, Henrik Hvidberg, Dana Gilbert, and Andrew King. Thanks for your help all these years.

I would like to express my gratitude to Interacoustics A/S, the William Demant Foundation, and the Innovation Fund Denmark DTU for their financial.

Finally, I want to thank my family here and in Colombia for the unconditional love and understanding. Especially to Camilo, for all the support during this long journey. You really helped me to make this possible.

Valentina Zapata-Rodríguez, March 2020

Related publications

Journal papers

- Zapata-Rodríguez, V., Laugensen S., Jeong, C.-H., Brunskog, J., and Harte, J. M. (**2020a**). "Does room acoustics affect the amplitude of sound-field auditory steady-state responses?," *Trends in Hearing*, under review.
- Zapata-Rodríguez, V., Jeong, C.-H., Zaar J., Laugensen S., Brunskog, J., and Harte, J. M. (**2020b**). "Evaluation of acoustical characteristics of audiological clinic rooms for sound field audiometric testing," *International Journal of Audiology*, in preparation.
- Zapata-Rodríguez, V., Laugensen S., Marbjerg G., Jeong, C.-H., Brunskog, J., and Harte, J. M. (**2020c**). "Reproduction of nearby sources using loudspeaker-based virtual acoustic environments," *The Journal of the Acoustical Society of America Express Letters*, submitted.
- Zapata-Rodríguez, V., Laugensen S., Cebulla M., Jeong, C.-H., Brunskog, J., and Harte, J. M. (**2020d**). "Effect of room acoustics on sound-field auditory steady-state responses for hearing aid fitting validation," *Trends in Hearing*, in preparation.

Additional papers

The Ph.D. candidate has also contributed during the project to the additional publications listed below. These publications do not constitute an essential part of the dissertation, but they are added as appendices since some minor aspects described in these publications are mentioned throughout the text.

- Marbjerg G., Zapata-Rodríguez, V., Brunskog, J., and Jeong, C.-H. (**2020**). "Comparing loudspeaker array reproduction techniques using a phased combination of the image source method and acoustical radiosity," *Manuscript. Appendix A*.
- Narayanan, S. K., Laugensen S., Zapata-Rodríguez, V., Brunskog, J., and Jeong, C.-H. (**2019**). "Effect of head-movement on sound-field auditory steady state response measurement," *Proceedings of 23rd International Congress on Acoustics*, 855–62. *Appendix B*.

- Marbjerg G., Brunskog, J., Jeong, C.-H., and Zapata-Rodríguez, V. (2018). "The influence of overlapping band filters on octave band decay curves," *Acta Acustica united with Acustica*, 104(6). *Appendix C.*

Published abstracts

- Zapata-Rodríguez, V., Marbjerg, G. H., Brunskog, J., Jeong, C.-H., Laugesen, S., and M. Harte, J. (2017). "Evaluation of a loudspeaker-based virtual acoustic environment for investigating sound-field auditory steady-state responses," *Journal of the Acoustical Society of America*, 141(5), 3997–3997.

- Marbjerg, G. H., Brunskog, J., Jeong, C.-H., and Zapata-Rodríguez, V. (2017). "Auralizations with loudspeaker arrays from a phased combination of the image source method and acoustical radiosity," *Journal of the Acoustical Society of America*, 141(5), 3783–3783.

- Zapata Rodríguez, V., Laugesen, S., Jeong, C.-H., Brunskog, J., and Harte, J. M. (2018a). "Towards Predicting Room Acoustical Effects on Sound-Field ASSR from Stimulus Modulation Power," *41st Annual MidWinter Meeting organized by Association for Research in Otolaryngology, ARO, San Diego, USA.*

- Zapata-Rodríguez, V., Laugesen, S., Jeong, C.-H., Brunskog, J., and Harte, J. M. (2018b). "Acoustic Conditions in Clinic Rooms for Measuring Sound-Field ASSR," *Dansk Teknisk Audiologisk Selskab, DTAS, Stouby, Denmark.*

- Narayanan, S. K., Laugesen, S., Brunskog, J., Jeong, C.-H., and Zapata-Rodríguez, V. (2018). "Effect of Head-Movement on Sound-Field ASSR Measurements," *Dansk Teknisk Audiologisk Selskab, DTAS, Stouby, Denmark.*

- Zapata Rodríguez, V., Laugesen, S., Jeong, C.-H., Brunskog, J., and Harte, J. M. (2019). "Investigating Room Acoustical Effects of Simulated Clinic Rooms on Sound-Field ASSR Measurements Using a Virtual Acoustic Environment," *XXVI International Evoked Response Audiometry Study Group Biennial Conference, IERASG 2019, Sidney, Australia.*

Contents

Preface i

Abstract iii

Resumé v

Resumen vii

Acknowledgements ix

Related publications xi

Contents xiii

1 Introduction 1
- 1.1 Motivation and background . . . 1
- 1.2 Overview of the thesis . . . 5
- 1.3 Main contributions of this thesis . . . 7

2 Does room acoustics affect the amplitude of sound-field auditory steady-state responses? 9
- 2.1 Introduction . . . 10
- 2.2 Material and Methods . . . 12
  - 2.2.1 Participants . . . 12
  - 2.2.2 Stimuli and room acoustic simulations . . . 12
  - 2.2.3 ASSR measurements . . . 14
  - 2.2.4 Data analysis . . . 15
  - 2.2.5 Statistical analysis . . . 18
- 2.3 Results . . . 18
  - 2.3.1 Effect of the room on ASSR level . . . 18
  - 2.3.2 Detection rate of simulated sound-field ASSR . . . 20
  - 2.3.3 ASSR level and Early Decay Time . . . 20
  - 2.3.4 ASSR level and Relative Modulation Power . . . 23
- 2.4 Discussion . . . 24
  - 2.4.1 Effect of the room on the ASSR level . . . 24
  - 2.4.2 Detection rate of simulated sound-field ASSR . . . 25
  - 2.4.3 ASSR level and Relative Modulation Power . . . 25
  - 2.4.4 Implication and Limitations . . . 26

2.4.5 Conclusions 27

3 Evaluation of acoustical characteristics of audiological clinic rooms for sound field audiometric testing 29
3.1 Introduction 30
3.2 Methods 33
3.2.1 Audiometric clinic rooms 33
3.2.2 Objective room acoustic parameters 34
3.2.3 Speech transmission index 35
3.2.4 Sound pressure level distribution of warble tones 36
3.2.5 Sound-field auditory steady-state response measurements 36
3.3 Results 39
3.3.1 Objective room acoustic parameters 39
3.3.2 Speech transmission index 41
3.3.3 Sound pressure level of warble tones 42
3.3.4 Sound-field auditory steady-state response measurements 43
3.4 Discussion 45
3.4.1 Objective room acoustic parameters 45
3.4.2 Speech transmission index 47
3.4.3 Sound pressure level distribution for warble tones 47
3.4.4 Sound-field auditory steady-state response measurements 49
3.5 Summary and conclusions 49

4 Reproduction of nearby sources using loudspeaker-based virtual sound environments 51
4.1 Introduction 51
4.2 Material and methods 53
4.2.1 Test rooms 53
4.2.2 Acoustic scene generation 53
4.2.3 Virtual sound environment 55
4.2.4 Calibration procedure 55
4.2.5 Objective evaluation of the VSE 56
4.2.6 Results and discussions 57
4.3 Conclusions 58
4.4 Supplementary 59

5 Effect of Room Acoustics on Sound-Field Auditory Steady-State Responses for Hearing Aid Fitting Validation 63
5.1 Introduction 64
5.2 Methods 66
5.2.1 Participants 66
5.2.2 Stimuli 66
5.2.3 Virtual room acoustic enviroments 67
5.2.4 ASSR Measurements 70

5.2.5 ASSR postprocessing . . . 72
5.2.6 Statistical analysis . . . 73
5.3 Results . . . 73
5.3.1 Sound-field ASSR measurements in different acoustic conditions . 73
5.3.2 Relation between ASSR level and the early decay time and the relative modulation power . . . 77
5.3.3 Detection time and detection rate of sound-field ASSR . . . 79
5.4 Discussion . . . 80
5.4.1 Effect of the room on the ASSR level . . . 80
5.4.2 Predicting the ASSR level with the EDT or the relative modulation power . . . 81
5.4.3 Detection rate and detection time in sound-field ASSR measurements . . . 82
5.4.4 Acoustic descriptors . . . 82
5.5 Conclusion . . . 83
5.6 Supplementary Material . . . 84

6 General discussion **87**
6.1 Summary . . . 87
6.2 Room acoutsical effect on sound-field ASSR in the clinics . . . 89
6.3 Outlook and future work . . . 91

Appendices **93**

A Comparing loudspeaker array reproduction techniques using a phased combination of the image source method and acoustical radiosity **95**

B Effect of head-movement on sound-field auditory steady state response measurement **97**

C The influence of overlapping band filters on octave band decay curves **99**

Bibliography **101**

CHAPTER 1

Introduction

1.1 Motivation and background

Hearing loss is one of the most predominant congenital diseases present at birth, affecting one to three out of every 1000 newborns (Erenberg et al., 1999). In children, permanent hearing loss has huge negative effects on cognitive and speech development (Moeller et al., 1986; Ruben, 1991). Early diagnosis and intervention are critical for successful auditory rehabilitation, which highly depends on the access to speech signals at the youngest possible age (Yoshinaga-Itano et al., 1998b; Sharma et al., 2002). The implementation of universal newborn hearing screening programs in most developed nations, has led to infants being fitted with hearing aids at a very young age (Wood et al., 2015). The prescribed gain of the hearing aid should then be validated to ensure that each infant receives adequate sound stimulation (Marcoux and Hansen, 2003). At present, once a child is fitted with a hearing aid, the most common approach to validation is to wait until the child is old enough (6 to 8 months) that standard behavioral audiology tests, such as visual reinforcement audiometry (Bamford and McSporran, 1993; Day et al., 2000) can be conducted. The development of objective hearing aid fitting validation tests is an important emerging application in technical audiology, where several methods based on electrophysiological measurements are being developed. The advantage of these methods is that they do not require an active participation from the patient (e.g. Picton et al., 1998; Easwar et al., 2015; Punch et al., 2016; Uhler et al., 2018), which make them particularly useful for the hearing aid fitting validation in infants and hard-to-test adults. In this project, auditory steady-state responses (ASSR) measured in sound field have been considered (Picton et al., 1998). This is an objective clinical test of the status of the auditory system, which consist in recording the electrical activity from the brain in response to an acoustic signal. In the test, the acoustic stimulus is presented through a loudspeaker and thus, the patient can wear the hearing aid as they would do every day, as seen in Figure1.1.

ASSRs are periodic auditory evoked potentials elicited by a group of neurons synchronously firing in a phase-locked manner to the envelope of an acoustic signal (Picton et al., 2003; Rance, 2008). Figure 1.2 presents an overview of the ASSR test from the stimulus presentation to the analysis of the response. ASSRs can be evoked by sinusoidally amplitude modulated tones or repeated transient stimuli, such as the CE-Chirps® (Elberling and Don, 2010). Throughout this project, the narrow band CE-Chirp® stimulus has been considered for the ASSR measurements. The family of narrow bands CE-Chirps® consists of four one octave band chirps with center frequencies of 500, 1000, 2000 and 4000 Hz, all with different repetition rates (Elberling et al., 2007),

Figure 1.1: Photo of sound-field ASSR measurement for hearing aid fitting validation.

as illustrated in Figure 1.2.A. These stimuli were designed to compensate for the wave traveling delay in the basilar membrane within an octave band, resulting in higher across-neuron synchronization, and hence higher ASSR levels. The CE-Chirp® is commercially available in the Interacoustics Eclipse platform, and it is routinely used in the clinics. ASSRs offer the possibility to measure frequency specific responses that can be recorded simultaneously. Once the signal is presented to the subject, the basilar membrane is stimulated in different relatively independent regions due to its tonotopic organization, as represented in Figure 1.2.B). Afterwards the signal is transformed into an electrophysiological signal (EEG signal) processed by the brain and recorded from surface electrodes. The EEG signal is a signal in the time domain (see Figure 1.2.D), which contains the periodic ASSR and EEG noise produced by all brain activity. The ASSR is analyzed in the frequency domain by "epochs," which are normally around 2 s long, by convention. The ASSR spectrum has the information of the EEG noise and the ASSR found at the frequency bin of the repetition rate and the higher harmonics, as illustrated in Figure 1.2.E. The EEG noise is normally larger than the ASSR level. However, since the noise is uncorrelated, the noise contribution can be decreased by averaging across epochs as $\sqrt{N}$, with N number of included epochs. The ASSR is detected using automated statistical tests, such as the F ratio test, which determines whether the ASSR signal is significantly different from the noise floor calculated from the surrounding frequency bins around the ASSR (Dobie and Wilson, 1996). The detection time will then depend on both the ASSR level and the EEG noise, as well as the number of epochs required to obtain the signal to noise ratio necessary to detect the ASSR.

It has been widely reported that the stimulus modulation influences the ASSR level (e.g., Rees et al., 1986; Kuwada et al., 1986; Picton et al., 1987), which reduces as the modulation depth decreases. The modulation depth is defined as the difference between the minimum and maximum values of the stimulus envelope and is commonly used to quantify the modulation for periodic amplitude modulated signals. Stimuli with higher modulation depths will then have envelopes with larger amplitude fluctuations, which

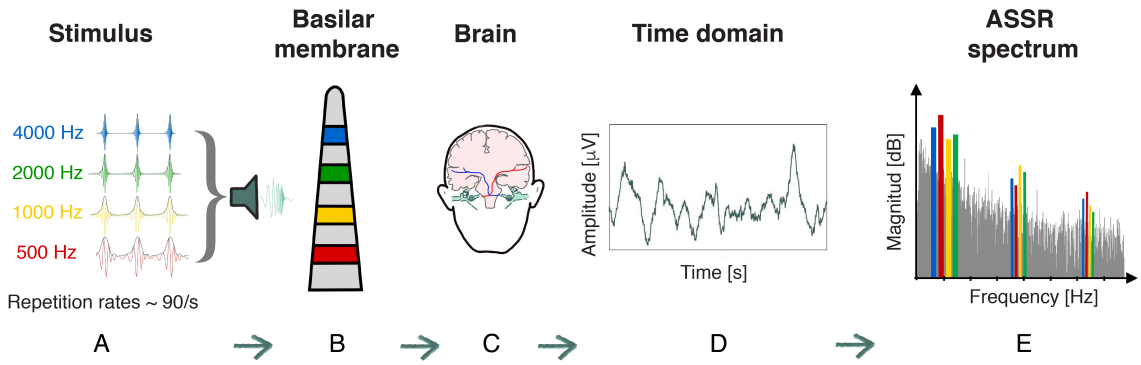

Figure 1.2: Overview of ASSR measurements from stimulus presentation to ASSR analysis.

lead to higher ASSR levels. The relation between the ASSR level and the modulation depth for amplitude modulated tones has been demonstrated in several studies in which the stimulus was presented to the subject through insert earphones (e.g., Rees et al., 1986; Kuwada et al., 1986; Picton et al., 1987; Lins et al., 1995; Dimitrijevic et al., 2001). Given the mode of presentation used in these previous studies, the stimulus arrived relatively unaltered to the ear. For sound-field ASSR, however, the signal presented through the loudspeaker will be affected to some extent by the reverberation and background noise of the room. In fact, it was demonstrated that the modulation of any signal reduces in a room due to the sound reflections from the reverberation, which "fill" the gaps of the stimulus modulation (Plomp, 1983; Houtgast et al., 1980). This has been a topic of extensive research in classroom acoustics, where an excellent speech intelligibility is required. Many investigations have reported that speech intelligibility is reduced in reverberant rooms, as a result of the reduced speech modulations (which provide cues to speech comprehension) produced by the reverberation in the rooms (Bradley et al., 1999; Bistafa and Bradley, 2000; Hodgson and Nosal, 2002). Thus, the loudspeaker presentation for ASSR measurements could result in lower ASSR levels, and hence, complicate the ASSR detection.

The majority of research in sound-field ASSR has focused on evaluating the discrepancies between hearing thresholds estimated via behavioral pure-tone audiometry and via ASSR, measured with earphones as well as in a sound field (e.g., Picton et al., 1998; Stroebel et al., 2007; Damarla and Manjula, 2007; Hernández-Pérez and Torres-Fortuny, 2013). Overall, the studies showed good correlations between the measurements and suggested that sound-field ASSR is a promising tool for the objective hearing aid fitting validation. In contrast, the potential effect of the room has received little scientific interest, even though for other audiological tests (e,g., sound field audiometry), the sound field presentation is a limiting factor of the measurement (Dillon and Walker, 1982). For audiological sound field testing (e.g., visual reinforcement audiometry, speech in noise test and real ear measurements), such limitation comes from the fact that the sound field in a room is not completely homogeneous, and therefore, any change in the mea-

surement positions (produced by head movements or the loudspeaker placement in the room) can lead to substantial changes in the sound pressure level (Shaw and Greenwood, 2012). Consequently, in the case of sound-field ASSR, it is also important to investigate the potential changes in the stimulus modulation and how these could affect the ASSR measurement.

A better characterization of the effect of the room on sound-field ASSR measurements is also necessary due to the lack of specific acoustic regulations for audiometric testing rooms, for which a large variability of the dimensions and acoustic properties could then be expected. This could limit the repeatability of the measured ASSR within clinic rooms and the reproducibility across clinics. Figure 1.3 shows some exemplary pictures of real audiometric testing rooms, with marked differences in materials, dimensions and measurement setups. The available acoustic regulation and standards for healthcare facilities usually address only the patient rooms, waiting areas and consultation rooms, where high levels of speech privacy and speech intelligibility are a priority. Most of the available acoustic regulations provide recommendations on minimum allowable sound insulation levels (e.g., Rasmussen, 2018a; Machimbarrena et al., 2019; Fausti et al., 2019), and only a few include recommendations about the reverberation time (e.g., Rasmussen, 2018b; NBA-CA-82, 1982). In Australia and New Zealand, for instance, the reverberation time in consultation rooms should be within 0.4 s and 0.6 s (AS/NZ 2107, 2000), whereas in Denmark, the regulation establishes reverberation times shorter than 0.6 s for hospital bedrooms, and examination and treatment rooms (DBR, 2018). However, audiometric testing rooms may require higher acoustic quality than other type of rooms in healthcare facilities, due to the acoustic requirements necessary to conduct the audiological test in the rooms. In the case of sound field audiometry, which is one of the most commonly conducted tests in audiology practice, the Bristish Society of Audiology recommends a maximum reverberation time of 0.25 s (BSA, 2019a). Furthermore, the ISO 8253-2 (2010) has established a calibration procedure based on differences in sound pressure level around the measurement point, which vary depending on the type of sound field in the test room. It is unknown, however, whether the recommendations provided for sound field audiometry could be sufficient to ensure that sound-field ASSR measurements can be conducted. Thus, it is important to systematically investigate the effect of the room to determine the minimum acoustic requirements for standardized ASSR measurements.

Virtual sound environments created with room acoustical simulations and loudspeaker-based reproduction techniques provide a valuable tool to systematically investigate the effect of different acoustic environments under controlled laboratory conditions. These can be used to investigate the influence of the room on sound-field ASSR measurements, by reproducing realistic clinic acoustic environments. Such a system can provide a more dynamic environment where ASSR can be recorded in a sound field, implicitly incorporating the natural head movements expected during the clinical test. Here, to create virtual sound environments for sound-field ASSR measurements, the Phased Acoustical Radiosity and the Image Source Method (PARISM; Marbjerg et al., 2015) was used for the simulation of the audiological testing rooms. This model was selected due to its

Figure 1.3: Exemplary photos of the audiometric clinic rooms used for audiological sound field measurements.

potential for simulating small rooms with absorbing surfaces (such as classrooms and clinic rooms), by accurately calculating their characteristic modal behavior. The simulated rooms were reproduced using the nearest loudspeaker method (NLM; Seeber et al., 2010). This method was implemented based on an early investigation that compared different reproduction techniques for PARISM simulations, which showed that the reproduction with the NLM was more accurate than reproductions based on higher order Ambisonics and vector-base amplitude panning methods (Marbjerg et al., 2020). The proposed virtual acoustic framework can correctly represent the acoustic characteristics of the rooms where ASSR can be conducted, which can lead to a proper characterization of the effect of the room on sound-field ASSR measurements.

1.2 Overview of the thesis

The primary goal of the present Ph.D. project was to conduct a thorough analysis of the effect of room acoustics on sound field ASSR measurements, considering the potential implications for the clinical implementation of the test for hearing aid fitting validation. Moreover, a thorough examination of the acoustic conditions in a set of real audiometric testing rooms was conducted to determine the characteristics of the rooms where sound-field ASSR could be recorded in the clinics. The project seeks to answer the following main research questions:

1. Is there an effect of the room on sound-field ASSR measurements?
2. Can the impact of room acoustics on ASSR amplitudes be predicted using easily measured acoustic parameters?
3. Are the detection time and detection rate affected by room acoustics in sound-field ASSR measurements?

4. Are the acoustic characteristics similar across the standard audiometric testing rooms in the clinics?

Chapter 2 provides the initial contribution toward understanding how the room acoustics influences sound-field ASSR measurements. To answer the first two research questions, ASSR was measured via insert earphones, using a simplistic monaural auralization approach with three different simulated room impulse responses. The rooms were simulated using a cosine acoustic model based on audiometric testing rooms. The effect of the room on sound-field ASSR measurements was analyzed in terms of the changes in the ASSR level and the detection rate for the fundamental frequency and higher harmonics of the ASSR spectrum. This paper also explores the relation between the ASSR level and two acoustic descriptors: the early decay time and an auditory-inspired modulation power model. It was also analyzed whether the acoustic descriptors could be used as a tool for predicting the ASSR level in any given room. This chapter is based on Zapata-Rodríguez et al. (2020a).

Chapter 3 addresses the fourth research question by presenting a thorough investigation of the acoustic characteristics of 31 state-of-the-art audiometric clinic rooms where sound-field testing is carried out. The room acoustic measurements were conducted for different positions in the clinic rooms. The uniformity of the sound field in each of the rooms, as well as the variability of the acoustic characteristics across rooms were evaluated based on three analyses: (1) comparison of objective acoustic parameters with the available acoustic standards for healthcare facilities; (2) speech transmission index recommended for the type of rooms; (3) analysis of the uniformity of the sound field in the room for sound field audiometry measurements. Moreover, sound-field ASSR was measured in a room adapted to have acoustic conditions similar to the real audiometric rooms to analyze whether the acoustics of these rooms is sufficiently good for measuring sound-field ASSR. This study provided crucial information about the acoustic conditions of the existing clinic rooms, which served as input to systematically investigate the effect of the room based on realistic acoustic conditions. This chapter is based on Zapata-Rodríguez et al. (2020b).

Chapter 4 describes a measurement and a calibration procedure to improve the reproduction of nearby sources in loudspeaker-based reproduction systems. This is important for the investigation of sound-field ASSR, which requires a relatively short source-to-receiver distance of 1 m. The paper presents the validation of the proposed methodology based on objective acoustic parameters, using virtual sound environments created for small rooms. This paper represents the validation of the simulation and reproduction system implemented in *Chapter 5* for the systematic investigation of the room effect on sound-field ASSR measurements. This chapter is based on Zapata-Rodríguez et al. (2020c).

Chapter 5 describes an extended analysis of the effect of the room on sound-field ASSR measurements, which is based on the initial investigation presented in *Chapter 2*. Here, sound field-ASSR were measured using virtual sound environments created

as described in *Chapter 4*. The tested acoustic conditions were designed based on the real audiometric testing rooms evaluated in *Chapter 3*. This study addresses the third research question by evaluating how the room influences sound-field ASSR measurements in terms of the ASSR level, detection rate and detection time, and discusses its clinical implications. Moreover, the models based on the early decay time, as well as the relative modulation power, were updated with the new dataset collected in the virtual acoustic enviroment to assess their suitability for predicting the ASSR level in any given room. This chapter is based on Zapata-Rodríguez et al. (2020d).

Finally, *Chapter 6* summarizes the main findings of each chapter, and further discusses the clinical implications of the effect of the room for sound-field ASSR measurements for hearing aid fitting validation. Additionally, perspectives for further research on sound-field ASSR measurements are suggested.

1.3 Main contributions of this thesis

- A frequency-dependent detrimental room effect on the ASSR level was demonstrated, which in turn, affects the detection times and detection rates of the ASSR measurements.
- The ASSR level has been shown to reduce by either increasing the early decay time or decreasing the relative modulation power.
- An extensive analysis of the acoustic conditions of state-of-the art audiometric testing rooms was provided.
- An experimental framework for the reproduction of nearby sources in loudspeaker-based virtual sound environments has been provided.
- Acoustic validation of the experimental framework for the evaluation of room acoustic conditions of small rooms for ASSR measurements.
- Prediction models of the ASSR levels as a function of either the early decay time or the relative modulation power were made, which can serve as a groundwork for the implementation of a calibration tool that evaluates whether a given room is appropriate for sound field ASSR measurements.

CHAPTER 2

Does room acoustics affect the amplitude of sound-field auditory steady-state responses?[1]

Abstract

Sound-field auditory steady-state response (ASSR) is a promising tool for the objective validation of hearing aid fitting in patients who are unable to respond to behavioral testing reliably. In 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 (EDT) 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 (NB) 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 to the reference 'anechoic' condition. Despite this reduction, the detection rates for the first harmonics of the ASSR were unaffected when sufficiently long recordings were made. Furthermore, the EDT 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.

[1]This chapter is based on Zapata-Rodríguez et al. (2020a).

2.1 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 (Yoshinaga-Itano et al., 1998a; Moeller, 2000). This has led to the implementation of universal newborn hearing screening programs (UNHS) in many countries around the world, e.g., most developed countries had implemented such hearing screening programs by 2015 (Morton and Nance, 2006; Ptok, 2011; Neumann et al., 2015; Singh, 2015; Neumann et al., 2019). The primary goal in such 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 amplification, which is called hearing aid fitting. Hearing aid fitting validation then becomes a critical procedure that can ensure that the infant receives adequate auditory stimuli and avoids potential delays in language development (Marcoux and Hansen, 2003). However, the validation of the prescribed hearing aid gain is challenging in pre-lingual 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 (ASSR) (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, a procedure for hearing aid fitting validation using sound-field ASSR is examined.

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). 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. estimated physiological hearing thresholds using sound-field ASSR that were not significantly different from behavioral hearing thresholds measured via insert earphones (Picton et al., 1998). The majority of research to date has focused on the validation of 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 (Picton et al., 1998; Stroebel et al., 2007; Damarla and Manjula, 2007; Selim et al., 2012; Shemesh et al., 2012; Hernández-Pérez and Torres-Fortuny, 2013; Park et al., 2013; Sardari et al., 2015). However, the potential effect of the room on the sound-field ASSR measurement has received little scientific attention. This is potentially problematic because 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 ASSR level is heavily influenced by the modulation of the acoustic signal (Rees et al., 1986; Kuwada et al., 1986; Picton et al., 1987; Lins et al.,

1995; Roß et al., 2000; Dimitrijevic et al., 2001; John et al., 2001; Boettcher et al., 2001; Rønne, 2013; Bharadwaj et al., 2015), the influence of room acoustics could present a barrier to the future clinical implementation of sound-field ASSR test for infants' hearing aid fitting validation.

Previous studies investigated ASSR level as a function of the modulation depth for sinusoidally amplitude modulated stimuli, and they reported that the ASSR level reduces as the modulation depth of the acoustic signal decreases (Rees et al., 1986; Kuwada et al., 1986; Picton et al., 1987; Lins et al., 1995; Roß et al., 2000; Dimitrijevic et al., 2001; Boettcher et al., 2001; John et al., 2001; Rønne, 2013; Bharadwaj et al., 2015). The ASSR level reaches its maximum when it is recorded with 100% amplitude modulated tones, keeping an equal RMS value across the signals with different modulation depths (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 level is obtained with a 50% amplitude modulated tone (Picton et al., 1987; Lins et al., 1995). The reduction in the ASSR level 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 upon 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 attenuates the natural fluctuations of the speech signal that are necessary for speech comprehension, which leads to poorer speech intelligibility for longer reverberation times (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 the current 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). It was 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: (1) the modulation of any acoustic signal in a room is distorted by its reverberation, and background noise

(Houtgast et al., 1980; Plomp, 1983); (2) ASSR level reduces as the stimulus modulation decreases (Kuwada et al., 1986; Rees et al., 1986; Picton et al., 1987; Lins et al., 1995; Roß et al., 2000; Dimitrijevic et al., 2001; John et al., 2001; Boettcher et al., 2001; Rønne, 2013; Bharadwaj et al., 2015). 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 and Don, 2010) stimuli with three simulated room impulse responses. The ASSR measurements were carried out in normal-hearing adult test subjects who were presented 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.

2.2 Material and Methods

2.2.1 Participants

Fifteen young adult normal-hearing subjects (7 female, mean age 24 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 HL 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.

2.2.2 Stimuli and room acoustic simulations

The NB CE-Chirps® for ASSR recording consists of four narrow-band (NB) chirp trains, with center frequencies of 500, 1000, 2000 and 4000 Hz (Elberling and Don, 2010). Each of the four stimuli were presented through insert earphones at slightly different repetition rates around 90 Hz. 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 and 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).

Since 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, e.g., speech in noise test, sound-field audiometry, fitting of hearing aids, etc. 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. These have been based on the variation of the sound-pressure level around the measurement point (ISO 8253-2, 2010), or on the maximum reverberation time and minimum dimensions of the testing room (BSA, 2019b; HTM2045, 1996). Their implementation, 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 (IEC IEC268-13, 1985), and a room recommended by the British Society of Audiology for sound-field audiometry for pediatric assessment (BSA, 2019b; HTM2045, 1996). Rooms ATB and IEC were based on real rooms located at the Technical University of Denmark. In all simulations, a distance of 1 m between the loudspeaker and the patient was used at an approximate height of a person sitting on a chair. Table 2.1 lists the dimensions, the reverberation time in one-octave bands, and the simulated source and receiver position for each of the simulated room conditions.

Table 2.1: 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.

Room	**Dimensions**	**Source position**	**Receiver position**	**Reverberation time, T20 [s]** Frequency band [Hz]						
	$l_x \times l_y \times l_z$ [m]	x_0, y_0, z_0 [m]	x, y, z [m]	125	250	500	1000	2000	4000	8000
ATB	2.6 × 2.0 × 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
BSA	6.0 × 4.0 × 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
IEC	4.7 × 7.5 × 2.8	3.2, 5.5, 1.4	1.8, 5.2, 1.4	0.27	0.24	0.31	0.29	0.34	0.3	0.29

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, Eq. (2.1). The Green's Function is an analytical solution to the wave equation with the boundary conditions imposed by rigid walls (Jacobsen and Juhl, 2013):

$$G(r, r_0) = \frac{-1}{V} \sum_{m=f_l}^{f_u} \frac{\Psi_m(r)\, \Psi_m(r_0)}{k^2 - k_m^2 - jk/(\tau_m\, c)} \,. \tag{2.1}$$

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 and Kuttruff, 1962). The model was implemented in a custom

MATLAB script that simulated the monaural frequency response of the rooms from $f_l =$ 10 Hz to $f_u = 10$ kHz. The room impulse responses were then obtained by calculating the inverse Fourier transform of the simulated frequency responses (Eq. (2.1)). Each term in Eq. (2.1) represents a mode in the cartesian coordinate system,

$$\Psi_m\ (x, y, x) = \sqrt{\varepsilon_{n_x}\ \varepsilon_{n_y}\ \varepsilon_{n_z}}\ \cos\Big(\frac{n_x\ \pi x}{l_x}\Big) \cos\Big(\frac{n_y\ \pi y}{l_y}\Big)\ \cos\Big(\frac{n_z\ \pi z}{l_z}\Big)\ , \tag{2.2}$$

where $\varepsilon_{n_x}, \varepsilon_{n_y}$,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 wavenumber corresponding to the m-th natural frequency of the room is given by

$$k_m\ = (1 + 0.01\nu_m) \sqrt{\Big(\frac{n_x\ \pi}{l_x}\Big)^2 + \Big(\frac{n_y\ \pi}{l_y}\Big)^2 + \Big(\frac{n_z\ \pi}{l_z}\Big)^2}\ , \tag{2.3}$$

where the speed of sound c takes a value of 343 m/s. A small random factor $\nu_m \sim N(0, 1)$ was added to the wavenumbers given by Eq. (2.3) to produce a more natural auralized sound.

2.2.3 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 presented to the test subject using a Tucker-Davis Technologies HB7 headphone driver and the ER-3A earphones, with foam ear tips. The stimuli were calibrated individually per frequency to eventually create a combined stimulus matching 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. The levels were defined from the standard method for computing the speech intelligibility index (ANSI, 1997). The presentation levels were measured at the eardrum position using an ear simulator BK 4157. The target values were 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.

ASSRs were measured using a standard clinical four-electrode montage (ipsi- and contra-lateral mastoids active, high forehead ground, and cheek reference). The electrode impedances were kept as equal as possible across the four electrodes and never exceeded 3 kΩ. The signal was preamplified by the ERA preamplifier and was recorded with the

Eclipse system using the accuracy-test method priority (ASSR detected for $p = 0.01$) and the adult sleeping protocol (fast repetition rates, around 90 Hz). The EEG response was recorded with a sampling frequency of 30 kHz for a total of 6 minutes of continuous recording. An artifact rejection level of 30 μV or 40 μV was used, depending on the EEG noise of the measurements, and in turn, on how relaxed the patient was. 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 one and a half hours 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.

2.2.4 Data analysis

2.2.4.1 ASSR post-processing

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 out of 240 recordings were discarded). Table 2.2 shows the number of recordings included in the analysis for each condition. 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 and Wilson, 1996). ASSR detection was individually evaluated for each of the first four response harmonics, and without making use of the multi-harmonic detector of the standard Eclipse (Cebulla et al., 2006). Although a multi-harmonic detector is more sensitive than the F-ratio test, this was used to estimate the ASSR level for the different acoustic conditions tested (Dobie and Wilson, 1996). The ASSR level was calculated at the frequency bin of the repetition rate. Noise levels were calculated by averaging the noise power across 20 evenly distributed frequency bins around the response bin, excluding the frequency bins near the harmonics of 50 Hz line noise, as well as the frequency bins corresponding to any other repetition rate harmonics. The ASSR and noise amplitudes were estimated for each harmonic. The noise-corrected ASSR level (in dB reference to 1 nV, hereafter, referred to as ASSR level) was analyzed to reduce the effect of intersubject variability produced by the electrophysiological noise. It was calculated by subtracting the estimated noise power from the response power (Dobie and Wilson, 1996). The detection rate was calculated based on the total number of recordings included in the analysis for each condition listed in Table 2.2. 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 is detected, the stimulus was heard.

Table 2.2: Number of measurements per condition included in the analysis after post-processing 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
BST	14	15	14	12
IEC	14	14	15	15

2.2.4.2 Acoustic Descriptors of Simulated Rooms

Early Decay Time (EDT)

The early decay time 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 2.1A 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, assuming that a nonlinear representation of the stimulus evokes the ASSR (Laugesen et al., 2018). The model is also inspired by the modulation transfer function for the speech transmission index (STI) calculation (Houtgast and 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 and Delgutte, 2015).

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.

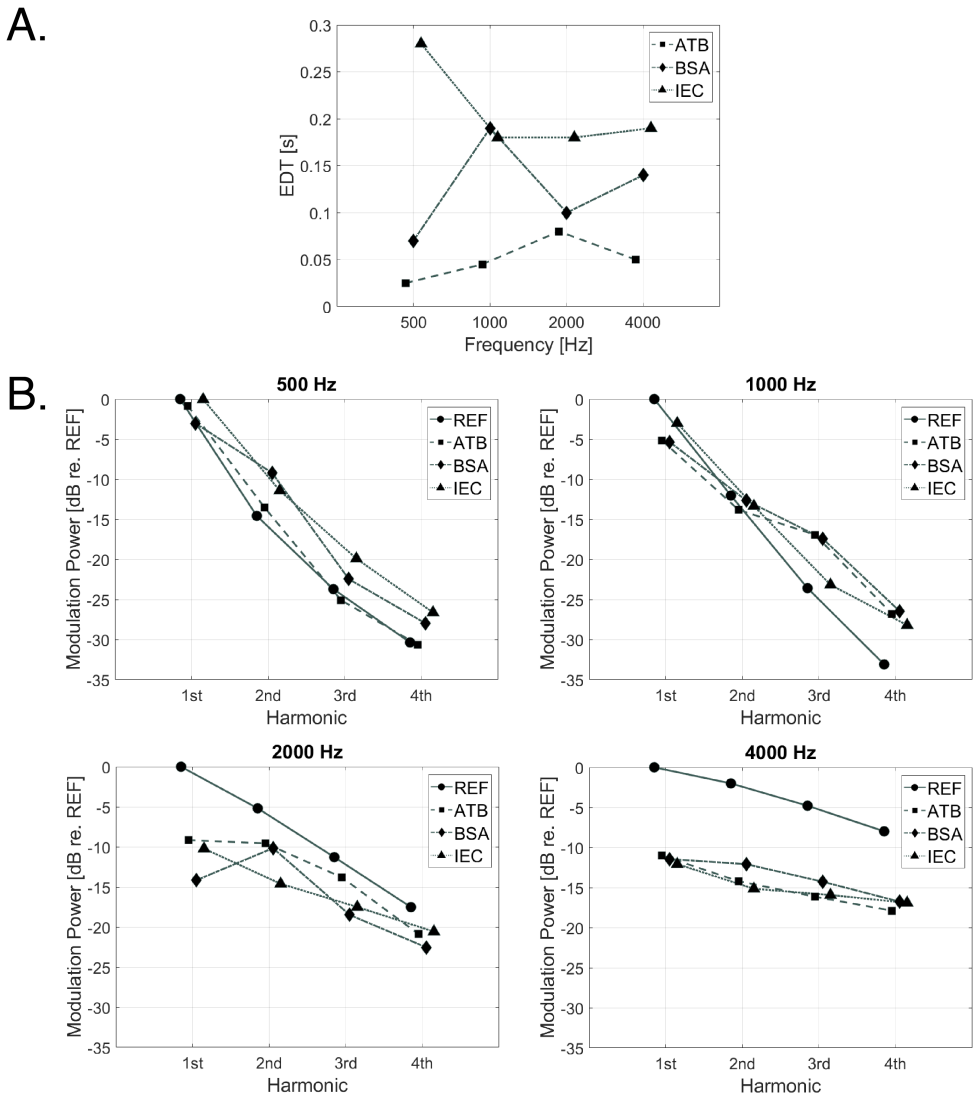

Figure 2.1: EDT (top panel) and relative modulation power (bottom panel) per stimulus harmonic of the acoustic conditions tested.

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 2.1B 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 to the reference signal is observed for the harmonics of interest.

2.2.5 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 (*TS*: 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, IEC), stimulus frequency (Frequency: 500, 1000, 2000, 4000 Hz) and ASSR harmonic (*Harmonic*: 1, 2, 3, 4). The predictable (*ASSR level*) and explanatory (*Room*, *Frequency*, *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, IEC) across frequencies and harmonics. For this analysis, 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 if 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 pairwise 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 early decay time (*EDT*: time in seconds) or the relative modulation power (*RModP*: in dB). Non-significant main effects and interactions were removed from the models, which were evaluated with analysis of variances and comparing the Akaike information criterion (AIC) (Akaike, 1974).

2.3 Results

2.3.1 Effect of the room on ASSR level

Figure 2.2 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*, *Frequency*, and *Harmonic*), as well as all two- and three-way interactions were significant. The summary of the ANOVA is shown in Table 2.3. 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 4-12 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.

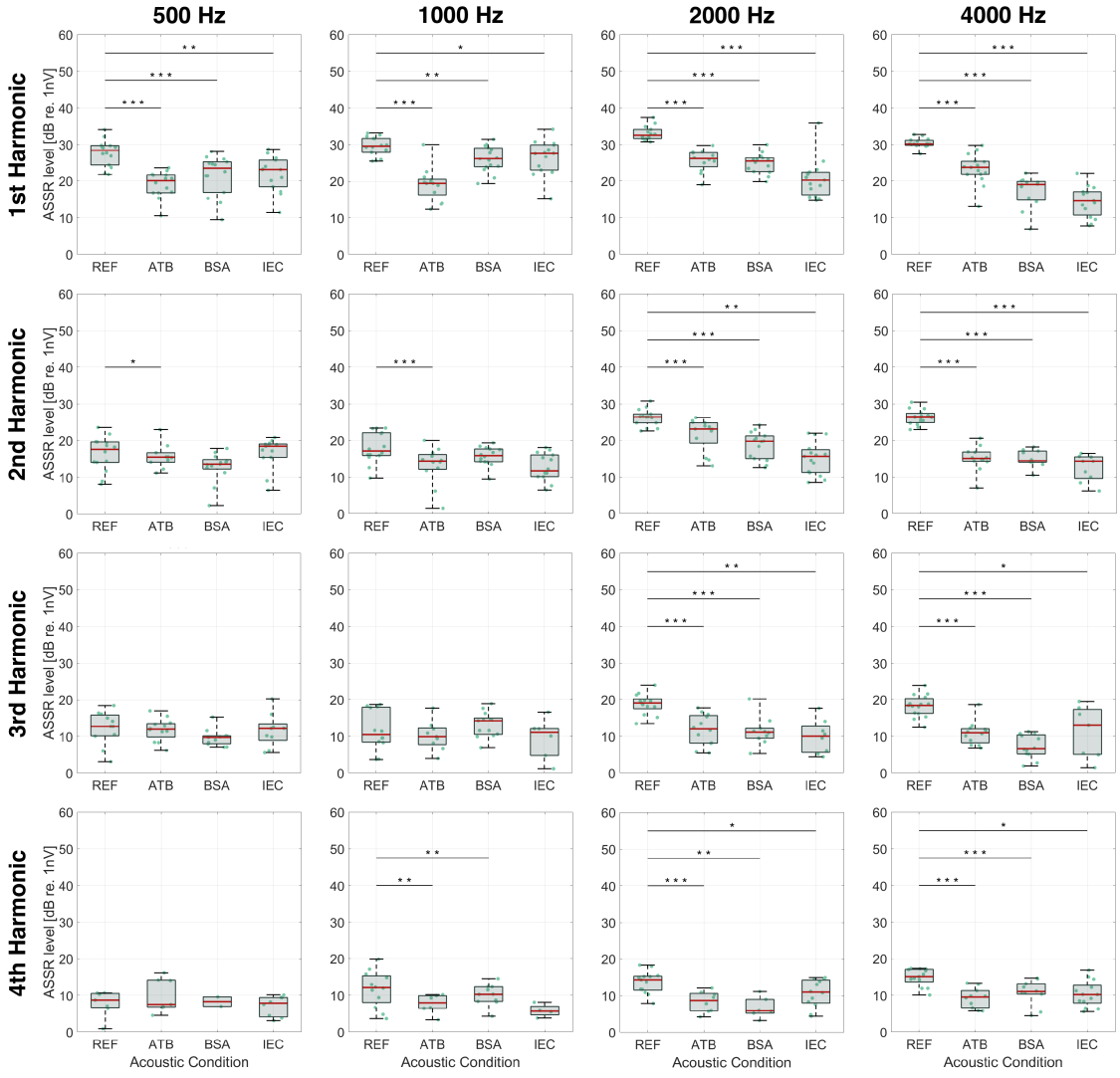

Figure 2.2: ASSR levels as a function of the tested acoustic conditions for each stimulus center band (columns) and harmonics (from 1st harmonic in the top row to the 4th harmonic in the bottom row). The whiskers of the boxplot indicate the minimum and maximum data points of the distribution, boxes show the 25th (bottom edge) and 75th (top edge) percentiles as well as the median (red line). The green points indicate the individual measurements.

Subsequently, the pairwise 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 2.2 by horizontal lines with the corresponding significance levels. It is noteworthy that for the stimuli at 500 and 1000 Hz, all three pairwise 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.

2.3.2 Detection rate of simulated sound-field ASSR

Figure 2.3 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 post-processing procedure, see Table 2.2. 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.

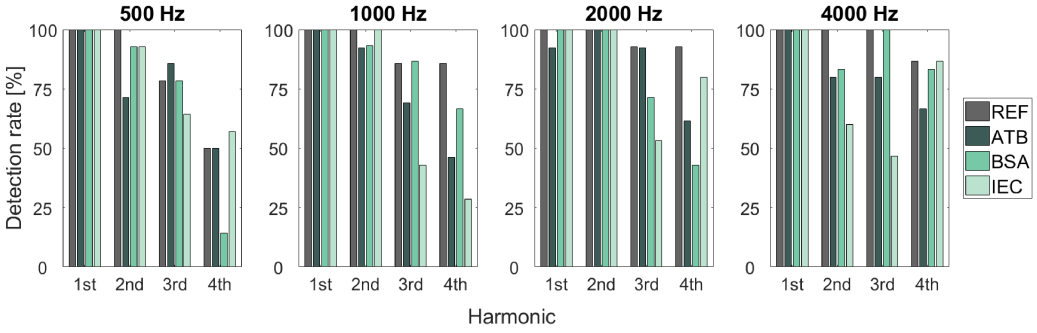

Figure 2.3: Detection rates for each tested acoustic condition, across harmonics and frequencies.

2.3.3 ASSR level and Early Decay Time

Figure 2.4 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.

Table 2.3: Summary results of the Mixed-Model Analyses of Variance.

Model 1. Effect of the room		AIC = 879.26
	ASSR level	
Factor	F statistic	ρ value
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
	ASSR level	
Factor	F statistic	ρ value
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*	$F(3, 485.6) = 2.6$	0.0526
EDT × *Freq* × *Harm*	$F(3, 486.4) = 0.1$	0.9812
Model 3. Effect of the relative modulation power		AIC = 845.25
	ASSR level	
Factor	F statistic	ρ value
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***

For this analysis, EDT was added as a continuous predictor, and the analysis included only the harmonics in which all pairwise 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 2.3.

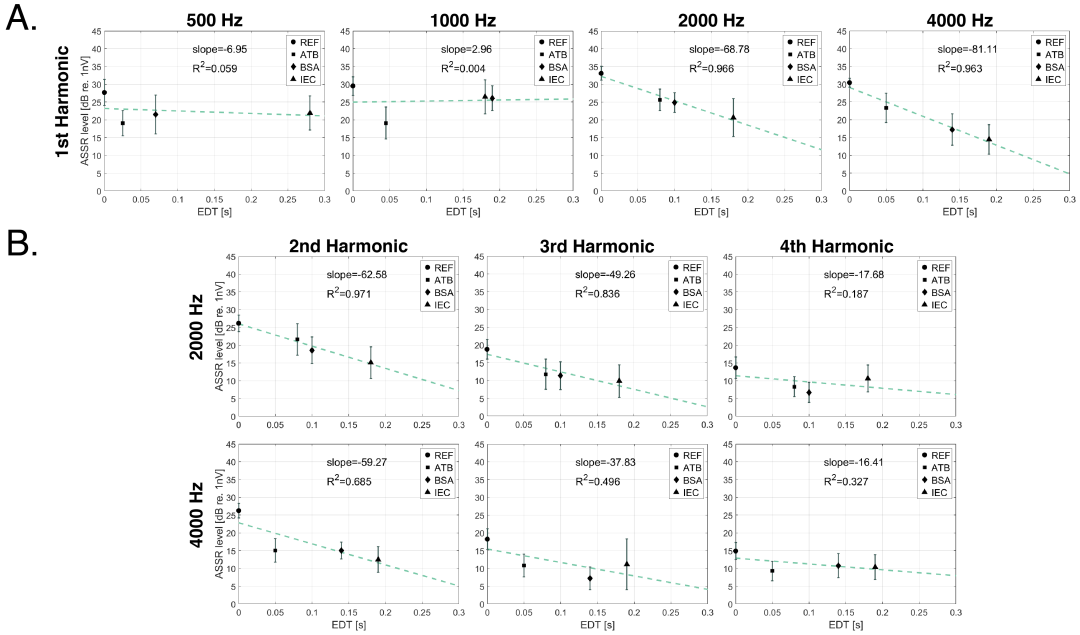

Figure 2.4: Mean ASSR level as a function of *EDT* for each *Room*, *Frequency* and *Harmonic*. Panel A shows results of the first harmonic for all stimulation frequencies. Panel B shows results for the included higher harmonics (2nd, 3rd and 4th) for 2000 and 4000 Hz. Error bars indicate standard deviation. Estimated regression lines (−−) are added to each panel.

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 determination (R^2). Panel A of Figure 2.4 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 ($R^2 = 0.06$) and 1000 Hz ($R^2 = 0.004$), 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 ($R^2 = 0.97$) and 4000 Hz ($R^2 = 0.96$). In general, high correlations were also obtained for the higher harmonics, as shown in panel B of Figure 2.4. 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. Additionally, 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.

2.3.4 ASSR level and Relative Modulation Power

Figure 2.5 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 pairwise comparisons between the reference condition and the three room conditions were included. The summary of the ANOVA is shown in Table 2.3. As expected, the stimulus modulation was degraded for all three room conditions compared to 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.

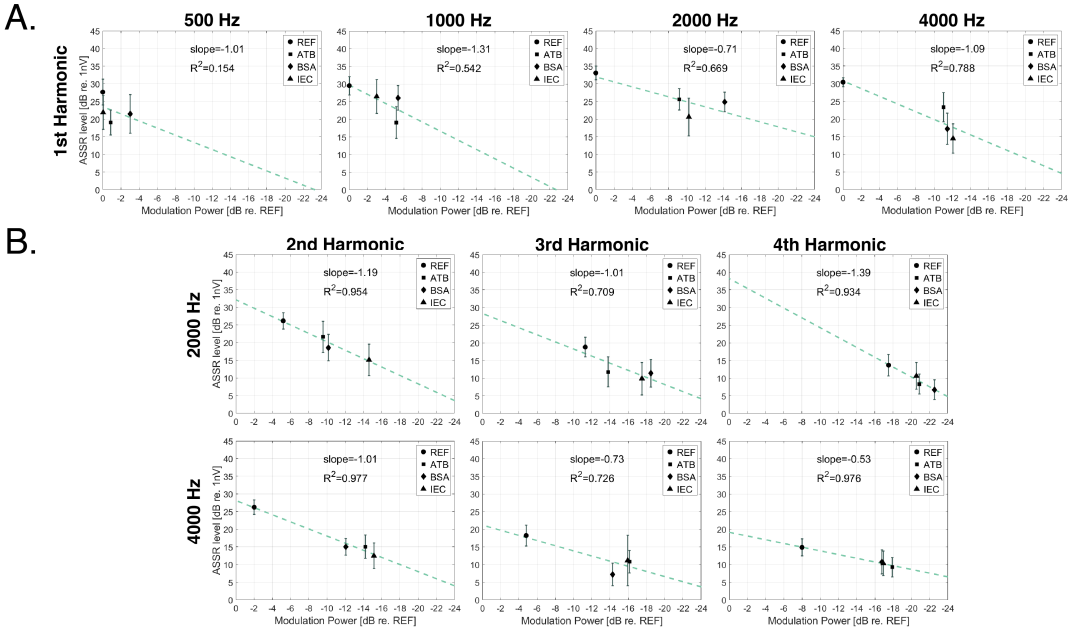

Figure 2.5: Mean ASSR level as a function of the relative modulation power (*RModP*) for each *Room*, *Frequency* and *Harmonic*. Panel A shows results of the first harmonic for all stimulation frequencies. Panel B shows results for the higher harmonics (2nd, 3rd and 4th) for 2000 and 4000 Hz. Error bars indicate standard deviation. Estimated regression lines (−−) are added to each panel.

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 2.5. 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 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 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 0.54, 0.67 and 0.79, respectively. In contrast, the correlation was low for the stimulus band center frequency of 500 Hz, with an R^2 value of 0.15. For the higher harmonics of 2000 and 4000 Hz, high correlations were obtained with coefficients of determination between 0.71 (for 2000 Hz, 3rd harmonic) and 0.98 (for 4000 Hz, 3rd and 4th 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.

2.4 Discussion

2.4.1 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 non-anechoic room conditions compared to the reference anechoic condition. This is a novel finding since 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 (Picton et al., 1998; Stroebel et al., 2007; Damarla and Manjula, 2007; Shemesh et al., 2012; Hernández-Pérez and Torres-Fortuny, 2013; Selim et al., 2012; Park et al., 2013; Sardari et al., 2015). The reduction in the ASSR level for the room conditions was ascribed to smaller stimulus modulations for the measured non-anechoic conditions (see Figure 2.1B). 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 (Dobie and Wilson, 1996; Cebulla et al., 2006; 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.

2.4.2 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 towards 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 multi-harmonic detector, such as the q-sample detector, might provide higher successful detection rates for sound-field ASSR measurements compared to a one-sample detector that only analyzes a single harmonic (Cebulla et al., 2006). Furthermore, a multi-harmonic detector might be particularly useful to compensate for the potential longer detection times produced by the reduced response amplitude obtained in the room conditions.

2.4.3 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 is also reduced, as expected. However, a direct comparison between the current 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 (Rees et al., 1986; Kuwada et al., 1986; Picton et al., 1987; Lins et al., 1995; Roß et al., 2000; Dimitrijevic et al., 2001; John et al., 2001; Boettcher et al., 2001; Rønne, 2013; Bharadwaj et al., 2015). 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.

In order to compare the modulation-growth functions reported in the literature with the one obtained in the current 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 2.4. The modulation-growth functions ('physiological IO 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.

Table 2.4: 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 (2013)	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

[a]Level: SPL, sound pressure level; HL, hearing level; SL, sensation level.
For the Picton et al. (1987) data presented at 76.5 dB SPL, only 6 subjects participated in the recording with 30% modulation depth.
For the Bharadwaj et al. (2015) data, the stimulus used was a SAM tone in notched noise.
For the Lins et al. (1995) data, the stimuli were calibrated based on a constant peak-to-peak value.

2.4.4 Implication and Limitations

The findings of the current 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: 1) 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. Additionally, the model is most

accurate for lightly damped rooms with evenly distributed absorption on the surfaces, which is far from realistic clinic rooms. 2) This investigation only focused on three room conditions, which is a limited sample compared with the variety of audiological testing 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. 3) 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. 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.

2.4.5 Conclusions

This study provides a first step towards 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 to the anechoic reference. This reduction in the ASSR level is likely to be attributed to the degradation of the stimulus modulation due to the non-anechoic 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, which will be important in potential clinical applications to evaluate the testing environment and determine whether it is acoustically suitable for sound-field ASSR measurements.

Acknowledgments

The authors would like to thank Johannes Zaar for the insightful comments and suggestions.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was funded by Interacoustics A/S, the William Demant Foundation and the Innovation Fund Denmark, Grant. No. 5189-00070B.

CHAPTER 3

Evaluation of acoustical characteristics of audiological clinic rooms for sound field audiometric testing[1]

Abstract

Sound field audiometric tests are routine procedures in audiological assessment. The present study evaluated the acoustic conditions of 31 state-of-the-art audiometric testing rooms in which sound field tests are conducted. The early decay time (EDT), reverberation time (T20) and speech definition (D50) estimated from impulse responses measured in the rooms were analyzed according to ISO 3382-1, (2009). The reverberation time was compared with acoustic regulations, standards, and guidelines for healthcare facilities. The rooms were also evaluated based on the speech transmission index (STI) and the spatial variation in sound pressure level for warble tones for sound field audiometry according to ISO 8253-2, (2010). Moreover, sound field auditory steady-state responses (ASSR) were recorded in a test room with similar characteristics to the audiometric testing rooms. The results showed a large variability of the acoustic conditions across the audiometric testing room. The comparison with the acoustic standards indicated that the rooms fulfilled the recommended limits based on reverberation time and STI, but not in terms of the variability of sound pressure level for warble tones. Moreover, lower ASSR levels were obtained in sound field compared to a reference dataset recorded with insert earphones. These findings highlight the need for specialized acoustic regulations for sound field audiometric testing rooms.

Keywords: audiometric test room, sound-field testing, hospital acoustics, speech transmission index, sound field audiometry, sound-field auditory steady-state responses

[1]This chapter is based on Zapata-Rodríguez et al. (2020b).

3.1 Introduction

Some of the regular tests in clinical audiology involve the presentation of a stimulus through loudspeakers. These sound field audiometric tests are suitable for patients who do not tolerate standard testing procedures using earphones (ASHA, 1991). For instance, the hearing assessment of infants (from 6 months onwards) is commonly carried out with sound field virtual reinforcement audiometry (VRA) (Day et al., 2000; Shaw and Nikolopoulos, 2004). Sound field testing is also valuable for the validation of hearing assistive devices, providing a natural approach for conducting hearing tests under more realistic conditions, e.g., testing patients while wearing their hearing aids, and cochlear or bone-anchored implants (Dillon and Walker, 1982; Abouchacra et al., 2011). Real-ear measurements (REM), for instance, are used for technical verification of hearing aid fitting to the prescribed target amplification (Swan and Gatehouse, 1995; BSA, 2018). The validation of the fitting is conducted using sound field audiometry or speech-in-noise tests (Rochlin, 1993; BSA, 2019a). These measurements require the active participation of the patients and are therefore highly unreliable for the hearing aid fitting validation of infants and hard-to-test adults. For such patients, sound field auditory evoked cortical potentials (Punch et al., 2016), and sound field auditory steady-state responses (ASSR) (Picton et al., 1998) have been proposed as alternative objective procedures for hearing aid fitting validation. Although sound field measurements have proven useful audiological procedures, their implementation in clinics requires consideration of the interaction of the testing signal with the acoustics and background noise of the testing room, as well as the position and movements of the patient during the test (ASHA, 1991).

Several studies have investigated various factors affecting sound field audiometric testing. The interaction between the testing room and the stimulus, for instance, has been widely studied for sound field audiometry in typical audiometric booths (Dillon and Walker, 1982). It has been established that frequency-modulated (warble) tones provide less spatial variation than pure tones of the sound pressure level (SPL) around the measurement point in the room. This is essential in sound field audiometry, in which the patient's head movements may produce small variations of the measurement position during the test. Shaw and Greenwood (2012) reported differences of approximately 2.7 dB around the measurement points in eight audiology testing rooms with quasi-free sound fields (Shaw and Nikolopoulos, 2004). This was studied considering head movements within a radius of 0.3 m, which have been suggested to be realistic for the infant population (Walker et al., 1984). To ensure reliable results in sound field testing, however, it has been recommended that the variability around the measurement position should not exceed ±2 dB approximately (ISO 8253-2, 2010).

The effect of the loudspeaker placement has also been investigated for sound field audiometry (Stream and Dirks, 1974) as well as for REM (Killion and Revit, 1987; Ickes et al., 1991). It has been shown that the placement of the loudspeaker in the room, and its position relative to the patient can affect the uniformity of the sound field produced by the stimulus around the measurement point. This is important because the placement of the loudspeaker in the room has not yet been standardized for most sound

field measurements. Consequently, large variability of testing setups can be found across clinic rooms, which will impact the reliability of sound field measurements carried out across different clinic environments.

The acoustic properties of the test room in audiology clinics represent another critical factor for sound field measurements that has received little scientific attention. A few studies have analyzed the effect of room acoustics in speech tests, e.g., speech audiometry and speech in noise. When the tests are carried out using loudspeakers, it can become more challenging for the patient due to the influence of the acoustics of the room on the test signal. This is because the sound reflections produced by the surfaces of the room (reverberation) interfere with the natural modulation of the speech (and any given signal) (Houtgast et al., 1980; Plomp, 1983), affecting its perception. Dirks et al. (1972) investigated the differences between speech audiometry measured in normal-hearing listeners via insert earphones and in sound field, both in an anechoic chamber and an audiometric room. They found differences of up to 3.6 and 7 dB for loudspeaker positions located at 0° and ±45° in azimuth, respectively (Dirks et al., 1972). Abouchacra et al. (2011) examined the effect of the reverberation time on speech comprehension in multi-talker scenarios. This was tested for normal-hearing listeners in two typical audiometric test rooms with reverberation times of 0.4 and 0.6 s, and volumes of 16.5 and 23.8 m^3, respectively. Their results indicated a detrimental effect of the reverberation on speech recognition scores when two or more independent speech distracters were presented (Abouchacra et al., 2011). While reverberation may be desirable from the point of view of test realism, these studies highlight the importance of having test rooms with appropriate acoustics to ensure reproducible and reliable sound field measurements across audiology clinics.

The acoustics of audiology clinic rooms is also vital for the communication between the audiologist and the patients, who commonly suffer from hearing disabilities, and therefore require rooms with acoustic conditions that facilitates excellent speech intelligibility (Iglehart, 2019; Blamey et al., 2001; Boothroy, 1984). In general, the current architectural standards and regulations are mostly focused on the minimum background noise levels required for healthcare facilities. The only existent recommendations are provided for the case of the reverberation time of the waiting areas and patients' rooms (Rasmussen, 2018a; Rasmussen, 2018b; Machimbarrena et al., 2019). For instance, the Danish Building Regulation recommends reverberation times below 0.6 s in the frequency range of 125-4000 Hz (with the liberation of 20% for 125 Hz) (DBR, 2018). Audiometric test rooms, however, are not particularly addressed by the standards, and hence a wide range of room acoustic conditions can be expected.

For sound field audiometry, ISO 8253-2 (2010) specifies the acoustics of test rooms based on three types of sound fields: free sound field, quasi-free sound field, and diffuse sound field. These definitions are based only on sound pressure level differences between six positions around the reference point placed at least 1 m away in front of the loudspeaker. The six positions are located on the left-right, up-down, and front-back directions from the measurement point. The recommended maximum differences

in sound pressure level depend on the type of sound field. In the case of the diffuse sound field, for instance, the sound pressure level variations should not exceed ±2.5 dB from the reference point at the positions located 0.15 m left-right, up-down, and front-back (ISO 8253-2, 2010). The acoustic requirements for sound field audiometry in clinical applications have also been considered in a best practice report by The British Society of Audiology (BSA, 2019b). The guidelines state that test rooms should have reverberation times below 0.25 s (HTM2045, 1996). Furthermore, the dimensions of the test rooms are suggested to be larger than 4 × 6 m for pediatric assessment, with the further recommendation of having an independent control room. These recommendations are defined for sound field audiometry measured using either pure tones, warble tones, or narrow-band noise. Therefore, it is unknown whether available acoustic guidelines could be suitable for other types of sound field measurements that rely more on the modulation of the stimulus, such as sound field ASSR.

ASSR is an electrophysiological measurement evoked by a periodically varying continuous auditory stimulus. In sound field ASSR, the acoustics of the testing room play an important role since the level of the ASSR is highly dependent on the stimulus modulation and reduces as the modulation decreases (Picton et al., 1987; Rees et al., 1986). The sound field presentation of the stimulus can, therefore, reduce the stimulus modulation. Indeed, a pioneering study has shown that room acoustics has an effect on the ASSR level due to the degradation of the stimulus modulation (Zapata-Rodríguez et al., 2020a). The study was conducted measuring ASSR through insert earphones using three simulated rooms inspired in audiometric clinic rooms. Although the study by Zapata-Rodríguez et al. (2020a) provided the initial assessment of the interaction between the room and the ASSR level, it remained unclear whether the acoustics of standard real audiometric rooms is sufficient for sound field ASSR, or whether it is necessary to establish specific acoustic standards for this type of measurements.

The present study investigates the acoustic characteristics of various state-of-the-art audiometric clinic rooms where sound-field testing is carried out. For this purpose, acoustic measurements were performed in 31 clinic rooms located in different facilities in Denmark and Germany. The rooms selected were either standard audiometric booths (for sound field audiometry, speech in noise test and VRA) or audiology office-type rooms (for fitting hearing aids). Four analyses were carried out to assess the rooms: (1) analysis of the objective room acoustic parameters early decay time (EDT), reverberation time (T20), and definition (D50); (2) analysis of the speech transmission index (STI); (3) analysis of spatial sound pressure level distribution with warble tones; and (4) analysis of sound field ASSR. The main goal of the study was to gain insight into the acoustic conditions of a set of audiological testing rooms, and to estimate the variability across these environments. The study also aims to evaluate the potential effect of the acoustic conditions of the rooms on different sound field audiometric tests, such as sound field audiometry, sound field ASSR, and speech tests.

3.2 Methods

3.2.1 Audiometric clinic rooms

A total of 31 rooms were considered in the present study. All were audiometric testing rooms in which sound field audiological tests were regularly carried out, e.g., sound field audiometry, VRA, REM, etc. Most of the rooms were located in Denmark, at Aarhus Universitetshospital, Bispebjerg Hospital, Gentofte Hospital, Odense Universitetshospital, and the Interacoustics A/S headquarters. The other two rooms considered were located In Würzburg, Germany, at the ENT Clinic of the Julius Maximilians-University, and the Karl-Kroiss-Schule. The rooms were categorized into two types: sound booth (SB) or office-type (OT). Figure 3.1 shows example photos of the two types of rooms studied.

Figure 3.1: Exemplary photos of the audiometric clinic rooms. Sound booths in A and B. Office-type rooms in C and D.

Table 3.1: Summary of volumes of audiometric clinic rooms.

Room Type	N	Mean $[m^3]$	STD $[m^3]$	Minimum $[m^3]$	Maximum $[m^3]$
Sound booth (SB)	15	27.3	12.8	13.9	55.1
Office-type (OT)	16	46.4	10.9	26.5	58.2

In general, the SBs were acoustically sound-treated audiometric booths with absorbing material on all surfaces of the rooms, except for the typical window for monitoring the patient during the test. In these rooms, VRA for infants as well as sound field audiometry and speech in noise tests for adults were carried out. Most of the SB rooms had relatively small dimensions (mean volume of 27.4 m^3, see Table 3.1 and usually had a separate control room associated. The OT rooms were audiology consultation rooms that had installed sound field reproduction systems. These rooms were commonly used for REM and sound field speech tests for cochlear or bone-anchored implants. The OT rooms had generally different materials (e.g., concrete, wood, drywall, etc.) and were

not fully treated acoustically. The OT rooms were mostly larger than the SB rooms, with a mean volume of 46.4 m^3. The loudspeaker setups varied across rooms and depended on the type of sound field test for which they were employed. The loudspeakers, in some cases, were not always on axis with respect to the listening position due to space limitations in the rooms. Moreover, most SB rooms had fixed loudspeaker setups, which were normally located close to reflecting surfaces, such as doors, tables and windows, as shown in Figure 3.1.A.

3.2.2 Objective room acoustic parameters

The acoustic conditions of the audiological testing rooms were analyzed in terms of the EDT, T20 and D50. These objective acoustic parameters were estimated from measured room impulse responses (RIRs) using custom MATLAB code. The parameters were selected based on their frequent use in relevant standards as well as their importance for speech intelligibility measures (ISO 3382-1, 2009). Reverberation time is the most well-known room acoustic parameter. This is defined as the time interval (in seconds) necessary for a 60-dB sound pressure level reduction in a room after it has been excited by a steady state sound abruptly turned off (ISO 3382-1, 2009). Both EDT and T20 can be used to characterize the reverberation time from the decay curves. While EDT uses the early part of the decay curve (from 0 to –10 dB), T20 uses a longer range for the estimation (–5 to –25 dB) (ISO 3382-1, 2009). EDT, however, is known to be more closely related to the subjective perception of reverberance than T20 (ISO 3382-1, 2009). Reverberation affects speech intelligibility as it smears the temporal envelope of the speech signal (Houtgast et al., 1980; Plomp, 1983). The energy contribution within the first 50 ms of the RIR is also important for speech understanding in rooms. This is quantified by the D50, defined as the ratio between the early and total sound energy of the RIR. The just noticeable difference (JND) of these acoustic parameters is 5% for EDT and T20, and 0.05 (e.g. difference in unitless ratio) for D50 (ISO 3382-1, 2009).

For the present study, the RIRs were recorded using an exponential sweep according to ISO-3382-1 (2009). The signal was amplified via a LAB 300 Lab.Gruppen amplifier and was presented through a custom-built omnidirectional loudspeaker (dodecahedron). The sweeps were recorded with a BK 4192 ½” microphone amplified through a BK NEXUS amplifier. The RIRs were measured with the Bruël Kjær DIRAC software, which was employed for the reproduction and recording of the excitation signal, as well as for the processing of the RIRs. The number of the measured source and receiver positions depended on the dimensions of the rooms. These were carefully selected as described by ISO-3382-1 (2009). A minimum of four source-receiver positions (two sources, two receivers) were measured for each room. The measurements were conducted during holidays or during nights with no activity in the clinics. It is therefore assumed that the recordings were not influenced by the regular background noise of the clinics.

3.2.3 Speech transmission index

The audiological testing rooms were also evaluated based on the speech transmission index (STI) (Houtgast et al., 1980). This is an objective room acoustic parameter that predicts the speech intelligibility through a transmission channel between a talker and a receiver (Houtgast et al., 1980). The STI is a value between 0 to 1. It predicts speech intelligibility in acoustic conditions that contain reverberation and stationary noise (Houtgast et al., 1980; Houtgast and Steeneken, 1985; Steeneken and Houtgast, 1980). STI is calculated from the modulation transfer function and quantifies the reduction of the envelope fluctuations as the ratio between the processed reference signal and the clean reference signal at a specific location (ISO 60268-16, 2011). The relation between the intelligibility rating and the STI for normal-hearing listeners and older listeners (over 60 years old) with hearing loss are defined in ISO 9921 (2003) and ISO 22411 (2008), respectively, and are summarized in Table 3.2. For a prolonged normal person-to-person communication the minimal intelligibility rating acceptable is "Good" (ISO 9921, 2003; ISO 22411, 2008), which would be the appropriate for audiological clinic rooms.

Table 3.2: Intelligibility rating and relation between STI for normal hearing and older listeners with age-related hearing loss.

Intelligibility Rating	**STI for normal hearing listeners (ISO 9921, 2003)**	**STI for older listeners (23 dBHL of PTA of 0.5 to 2 kHz) (ISO 22411, 2008)**
Excellent	>0.75	>0.8
Good	0.60 – 0.75	0.75 – 0.80
Fair	0.45 – 0.60	0.60 – 0.75
Poor	0.30 – 0.45	0.45 – 0.60
Bad	<0.30	<0.45

The STI values were estimated from the measured impulse responses, as defined in ISO 60268-16 (2011). Unlike the reference standard, the STIs were computed using the non-gender specific weights defined in Houtgast and Steeneken (1985). Furthermore, instead of using modulated octave band noise signals, the STI was calculated with a speech sentence from the Danish version of the Hearing in Noise Test (Nielsen and Dau, 2010), similar to the speech-based STI from Payton and Braida (1999). For this analysis, RIRs were measured in all rooms according to the ISO 8253-2 for sound field audiometry (ISO 8253-2, 2010). Five measurements were conducted for the determination of the STI. One at the reference position 1 m from the source, while the other four positions were located 0.15 m to the front, back, left, and right of the reference. All recordings were made 1 m above the floor. The RIRs were measured as described in the previous section.

3.2.4 Sound pressure level distribution of warble tones

The specific effect of the acoustic conditions of the audiological testing rooms on sound field audiometry was also evaluated. For this, the expected sound pressure level of warble tones in the room was analyzed. A reference source and receiver distance of 1 m was considered, as well as the same four additional receiver positions as above located 0.15 m to the front, back, left, and right of the reference position. The analysis was conducted by convolving the measured RIRs and warble tones with nominal frequencies of 500, 1000, 2000 and 4000 Hz. The warble tones were created according to the ISO standard for sound field audiometry (ISO 8253-2, 2010), with a warble rate of 5 Hz and a ±5% frequency swing. The source and receiver positions were carefully located as closely as possible to the loudspeaker setups in each room. When this was impossible due to obstacles (furniture) or the room layout, they were located near the center of the room. The analysis of the variation in sound pressure level for the warble tones was conducted according to ISO 8253-2 (2010). Therefore, the difference in sound pressure level (ΔSPL) was calculated between the reference position and each of the four additional positions. The purpose of this analysis was to compare the maximum variation of the sound pressure level estimated from the measurements with that recommended by the standard for sound field audiometry. Table 3.3 summarizes the main reference values defined by the standard, for specific details refer to ISO 8253-2 (2010).

3.2.5 Sound-field auditory steady-state response measurements

To analyze the potential effect of typical acoustic conditions of audiometric clinic rooms on the recorded ASSR level, sound field ASSR was recorded in a test room (volume of 42.8 m^3) with similar acoustic conditions to the real audiometric rooms evaluated in the present study. These measurements were then compared with a reference ASSR dataset previously measured via insert phones in Laugesen et al. (2018).

3.2.5.1 Participants

Ten young adult normal-hearing subjects (7 female, mean age 25 years) participated in the study. All ten subjects had also participated in the reference study (Laugesen et al., 2018). Their normal-hearing status was confirmed through otoscopy, wide-band tympanometry using the Interacoustics Titan, and air-conduction audiometry using an Interacoustics AC40 audiometer with ER-3A insert phones. The subjects had pure-tone thresholds 20 dB HL from 125–8000 Hz. All subjects provided written informed consent prior to the tests and were financially compensated with gift cards. The experiment was approved by the Science-Ethics Committee for the Capital Region of Denmark.

Table 3.3: Summary of maximum allowable sound pressure level variations for the different types of sound fields defined by ISO 8253-2 (2010).

Sound field	Positions	SPL difference
Free sound field*	Reference point, and 0.15 m to the left, right, up and down from the reference	SPL ±1 dB from the SPL at the reference point (up to 4000 Hz). SPL ±2 dB from the SPL at the reference point (above 4000 Hz). SPL 3 dB between the extreme left-right positions.
Free sound field*	Reference point, and 0.15 m to the front and back of the reference	SPL ±1 dB from the theoretical value given by the inverse sound pressure distance law.
Quasi-free sound field	Reference point, and 0.15 m to the left, right, up and down of the reference	SPL ±2 dB from the SPL at the reference point
Quasi-free sound field	Reference point, and 0.10 m to the front and back of the reference	SPL ±1 dB dB from the theoretical value given by the inverse sound pressure distance law
Diffuse sound field	Reference point, and 0.15 m to the front, back, left, right, up and down of the reference	SPL ±2.5 dB from the SPL at the reference point SPL 3 dB between the extreme left-right positions

* These requirements can only be met in an anechoic room.

3.2.5.2 Stimuli

ASSR was recorded using the four narrow-band NB CE-Chirps®, which consist of one-octave-wide chirp trains centered at 500, 1000, 2000, and 4000 Hz presented at slightly different repetition rates all around 90 Hz (Elberling and Don, 2010). The stimuli were adjusted to have a spectrum similar to speech, by calibrating the individual NB-CE Chirps to 61.4, 55.5, 49.7, and 45.8 dB SPL in the sound field for the 500, 1000, 2000, and 4000 Hz chirps, respectively. The four chirps were presented simultaneously to produce a total of 65 dB SPL. These stimuli as well as the presentation levels were the same in both the reference (Laugesen et al., 2018) and the present study.

3.2.5.3 ASSR recordings and playback conditions

ASSR was recorded twice (test and retest) for each individual ear. The recordings were made using the standard clinical four-electrode montage (ipsi- and contra-lateral mastoids active, high forehead ground, and cheek reference). The impedance across electrodes did not exceed 3 kΩ. The two aforementioned playback conditions were considered.

Insert phone measurements: The reference ASSR dataset measured by Laugesen et al. (2018) were recorded using the ER-1 insert phones. Their test consisted of two sessions (one for each ear) separated by at least one day. During the test, the participants were encouraged to relax and sleep if possible, while they lay on a bed in a darkened audiometric booth that was acoustically treated and electrically shielded. The ASSRs were recorded using an artifact rejection level of 40 V. The data were analyzed online using the F-ratio test on the first harmonic implemented in a custom Matlab function. Further details of the system and the measurement procedure can be found in the reference study (Laugesen et al., 2018).

Sound-field measurements: The stimuli were delivered via a KEF loudspeaker located 1 m from the listener position. Both ears were tested in one session, with the non-tested ear blocked with a foam plug. The participants sat on a chair and were instructed to relax and sleep if possible. The measurement position was marked with a head rest that was used by the subjects who were instructed to keep their head still facing the loudspeaker. The ASSRs were recorded with an artifact rejection level of 40 V using the Interacoustics Eclipse platform, which was also used to generate the standard NB CE-Chirps®. The measurements were conducted in a small room (42.8 m^3) located at the Technical University of Denmark. The room consisted of six hard walls and was conditioned to have acoustic conditions similar to the standard audiometric clinic rooms. To achieve this, two of its walls were covered completely with absorbent material, and two diffusers were placed on the floor around the measurement position. The acoustic evaluation of the test room with respect to the real audiometric rooms is presented in the Results section. Figure 3.2 shows a sketch of the room and measurement setup.

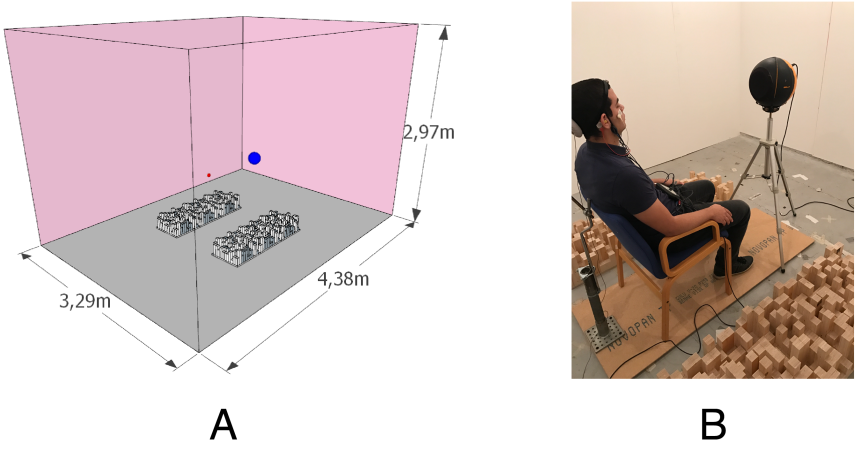

Figure 3.2: Sketch of the test room A, and exemplary photo of the sound field ASSR measurement B. The two pink colored surfaces in A are the absorbing walls.

3.2.5.4 ASSR postprocessing

The recorded ASSR for both the reference and the current study were analyzed offline, in the same way, using an artifact rejection level of ±40 V to guarantee low electrophysiological noise and weighted averaging (John et al., 2001). To minimize the effect of intersubject variability produced by the electrophysiological noise, the analysis was carried out on the noise-corrected ASSR level in dB referenced to 1 nV (hereafter, referred to as ASSR level). This was computed by subtracting the estimated noise power from the ASSR power (Dobie and Wilson, 1996). For each of the frequencies, the ASSR level was estimated at the frequency bin of the repetition rate. The noise power was averaged across 20 frequency bins that were uniformly distributed on either side of the response bin. The frequency bins belonging to any other repetition rate harmonics and those that were close to the harmonics of the 50 Hz line noise were excluded for the calculation of the noise power. ASSR detection was computed for only the first harmonic of the ASSR using the F-ratio test (Dobie and Wilson, 1996) which was calculated with a strict error-rate of 1%.

3.3 Results

3.3.1 Objective room acoustic parameters

Figure 3.3 shows the obtained mean and standard deviation (STD) for EDT, T20 and D50 across all measurement positions as a function of the room volume. These were estimated per octave frequency band (125 – 4000 Hz) for each of the clinic rooms considered in the study. The results show that the SB rooms had overall shorter EDT and T20 than the OT rooms. For the SB rooms the mean EDT and T20 (averaged across all frequencies and rooms) were 0.12 and 0.13 s, respectively, whereas for the OT rooms the mean values were 0.28 s for EDT and 0.31 s for T20. The lower EDT and T20 for the SB rooms was expected since this type of rooms are acoustically treated with more absorbing material than the OT rooms, which can further decrease the reverberation. Moreover, for both types of rooms the STD was generally larger for EDT than for T20. This result agrees with the fact that the amplitude of the early reflections is more strongly affected by the materials near the measurement point, and therefore, EDT is expected to be more sensitive to the measurement position in the room. This is further supported by the fact that the STD for the OT rooms were generally larger than for SB rooms, which had a more uniform distribution of materials on the surfaces.

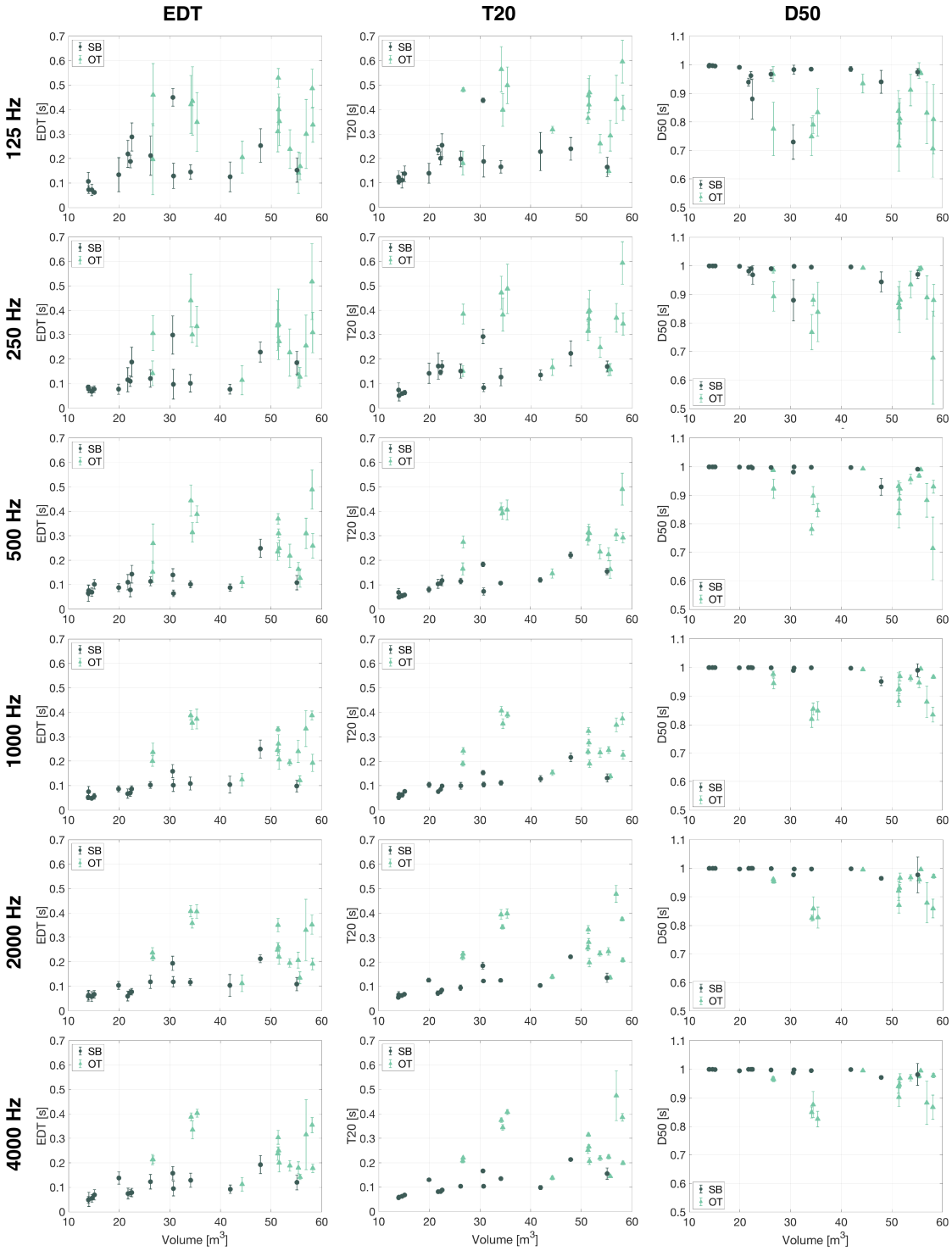

Figure 3.3: Mean EDT (left column), T20 (center column) and D50 (right column) as a function of room volume for the sound booths (SB, dark green) and office-type (OT, light green) audiometric clinic rooms. The error bars indicate the standard deviation. The means and standard deviations were calculated from the spatially distributed measurements according to ISO 3382 (2009).

For D50, in general, the results show that the OT rooms had smaller mean values and larger STD than the SB rooms. This is in line with the longer T20 obtained for the OT rooms, which indicates larger energy contributions of the late reflections (after 50 ms). The D50 values for both SB and OT rooms indicate that the audiometric clinic rooms included in the study generally have good acoustic conditions for speech communication with mean D50 values (averaged across frequencies and rooms) of 0.98 and 0.90 for the SB and OT rooms, respectively.

Overall, the objective evaluation of the rooms revealed a large variability of the acoustic parameters across the audiometric testing rooms. To estimate the variability across rooms for EDT, T20 and D50, it was calculated the mean across frequency for each of the rooms. The variability was then determined as the difference between the rooms with the minimum and maximum values for each of the acoustic parameters. For the SB rooms the difference was 0.22 s, 0.21 s and 0.1 for EDT, T20 and D50, respectively. In the case of the OT rooms, the difference was for EDT, T20 and D50 to 0.34 s, 0.36 s and 0.23, respectively. Thus, for all three acoustic parameters the variability across rooms was larger than 1 JND, and hence it could be perceptible. The room volumes also varied within a relatively large range, see Table 3.1. Interestingly, the center column of Figure 3.3 indicates that for the SB rooms, the relation between the room volume and the reverberation was fairly linear (the larger the room, the longer the EDT and T20), unlike the OT rooms that showed more irregular variability.

3.3.2 Speech transmission index

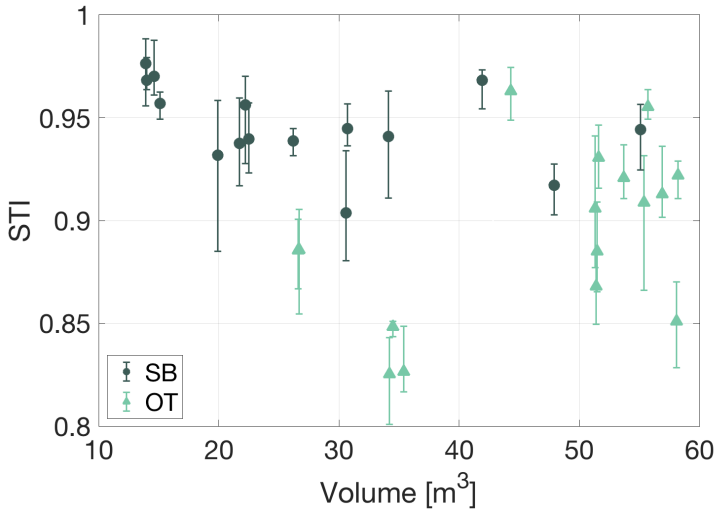

Figure 3.4: Mean STI as a function of the volume for the sound booths (SB, dark green) and office-type (OT, light green) audiometric clinic rooms. The error bars indicate the minimum and maximum STIs calculated across the reference point and the four additional measurement positions (0.15 m, left-right and front-back).

Figure 3.4 shows the mean as well as the minimum and maximum (error bars) STI as a function of the volume. The results show generally larger STI for the SB than the OT rooms, with mean values across rooms of 0.95 and 0.89, respectively. This is in agreement with the generally longer T20 obtained for the OT rooms, which is expected to affect the STI negatively as it is known that STI and T20 are inversely proportional (Houtgast et al., 1980). The variation in the mean STI, between the rooms with the lowest and highest values, was of 0.07 and 0.14 for the SB and OT rooms, respectively.

3.3.3 Sound pressure level of warble tones

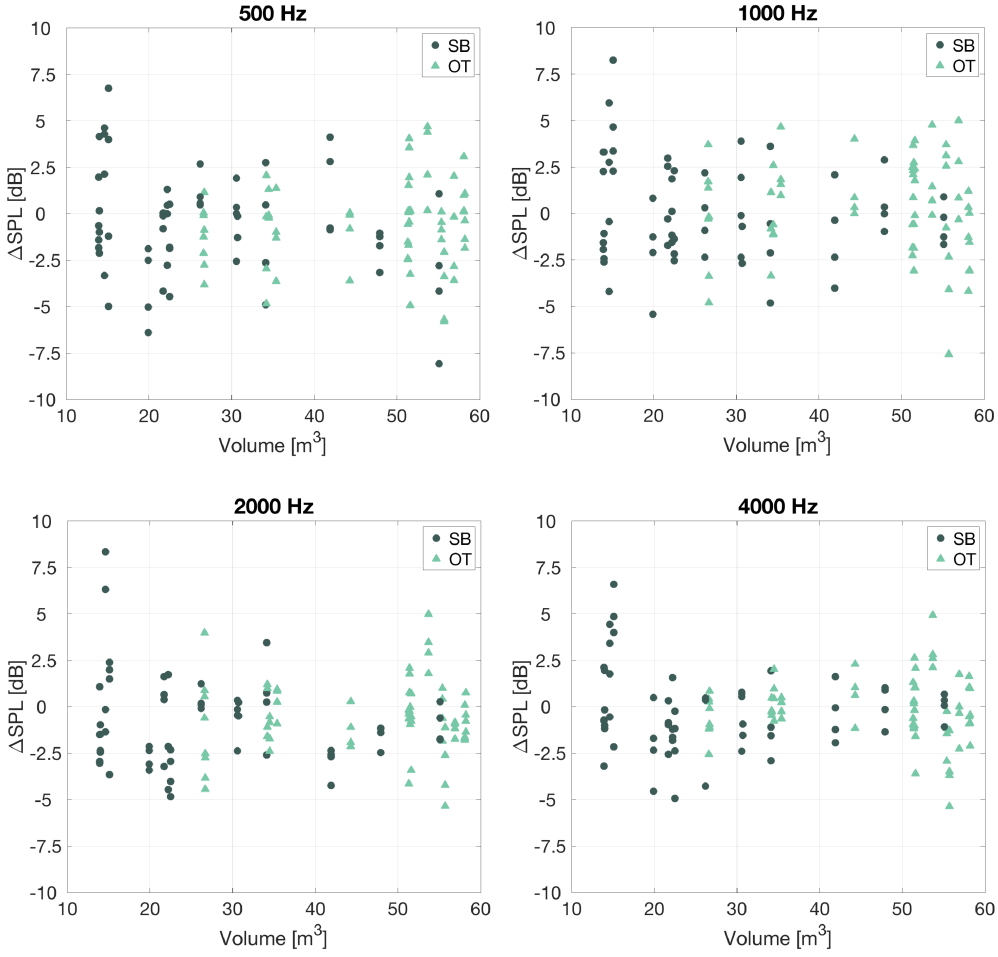

Figure 3.5: Sound pressure level difference between the reference point and each of the four additional measurement positions (0.15 m, left-right and front-back) as a function of the volume for the sound booths (SB, dark green) and office-type (OT, light green) audiometric clinic rooms. The analysis is presented for each of the four warble-tone nominal frequencies.

Figure 3.5 shows the differences in sound pressure level (ΔSPL) as a function of the room volume. The ΔSPL was calculated as the difference in the measured sound pressure level between the reference point and each of the four additional positions (left, right, front and back), as described in ISO 8253-2 (2010). The results show a large variability of ΔSPL for both the SB and OT rooms. The variability tends to decrease towards the higher frequencies. Thus, at 500 Hz the largest SPLs were 8.1 dB for the SB and 5.8 dB for the OT rooms, while for 4000 Hz, the differences were smaller with ΔSPLs of 6.6 and 5.4 dB for the SB and OT rooms, respectively. This agrees with the fact that the modal overlap is greater at high frequencies, meaning that the sound pressure level is less influenced by the position in the room. Moreover, larger spread of ΔSPLs were generally observed for the smaller rooms, which was expected since these rooms are dominated in a wider frequency range by a modal behavior.

3.3.4 Sound-field auditory steady-state response measurements

Figure 3.6 displays the distribution of the recorded ASSR levels per stimulus frequency and for the two playback conditions studied. In general, for the sound field condition, it can be observed a reduction in the ASSR level compared to the playback with insert phone. A linear mixed-effects model fitted to the *ASSR level* data was used to analyze whether the sound field presentation had an effect on the ASSR level. The playback condition (*PC*: InsertPhones, SoundField), stimulus frequency (*Frequency*: 500, 1000, 2000, 4000 Hz) and test ear (*TestEar*: Left, Right) were taken as fixed effects in the statistical model. The participants were considered as a random effect (*TS*: 1, 2, 3… 10). The ASSR level was taken as a continuous dependent variable, whereas the explanatory variables (*PC*, *Frequency*, *TestEar*) were all defined as categorical. All main effects and their interactions were analyzed, and non-significant effects were removed from the analysis, which was carried out with the lme4 library (Bates et al., 2015) in the software R version 3.5. The statistical analysis was performed including only the detected ASSR measurements. The statistical model showed significant influence of the three main effects (*PC*, $F(1, 278.6) = 72.0$, $\rho < .0001$; *Frequency*, $F(3, 277.0) = 18.2$, $\rho < .0001$; *TestEar*, $F(1, 278.0) = 10.1$, $\rho < .0016$) on the *ASSR level*. The two-way interaction between the *PC* and the *Frequency* was significant, $F(3, 277.0) = 3.7$, $\rho < .012$; whereas the two-way interactions (between *PC* and *TestEar*, $F(1, 277.0) = 3.0$, $\rho < .08$; and *Frequency* and *TestEar*, $F(3, 277.0) = 0.5$, $\rho < .69$) as well as the three-way interaction were not significant, $F(3, 277.0) = 0.9$, $\rho < .42$. The statistical analysis was performed including only the detected ASSR measurements.

The statistical analysis revealed a dependency of the *ASSR level* on the playback condition, as reflected by the reduction in the *ASSR level* for the measurements in sound-field. This is expected since sound field stimulation would naturally degrade the stimulus modulation due to the acoustics of the testing room (Houtgast and Steeneken, 1985). Despite the reduction in the *ASSR level*, the detection rates for the stimulus

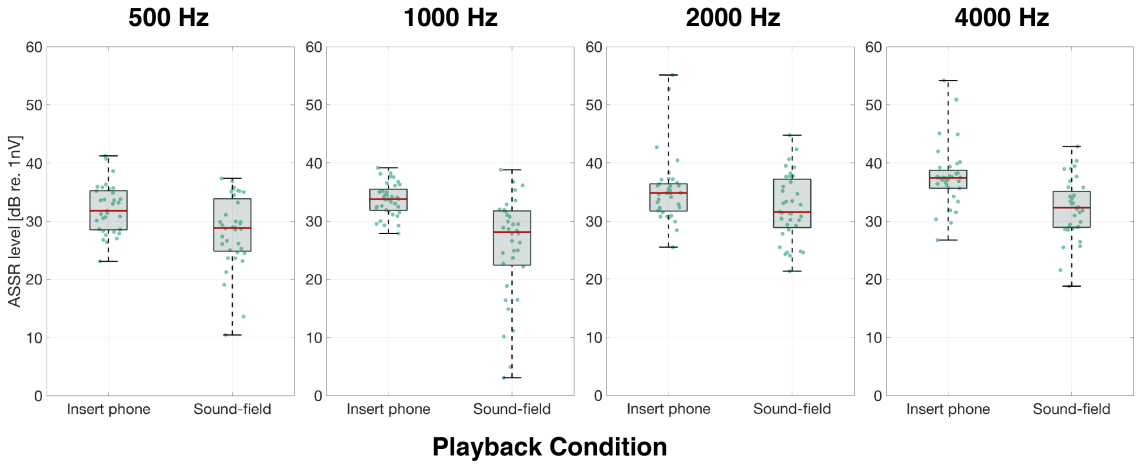

Figure 3.6: ASSR levels as a function of the playback condition for each stimulus center frequency. The whiskers indicate the minimum and maximum data points of the distribution, and the boxes show the 25th (bottom edge), 75th (top edge) percentiles and the median (red line). The green points indicate the individual measurements.

frequency bands 500, 1000, 2000 and 4000 Hz were overall very similar for both the insert phones (82.5, 92.5, 92.5 and 90 %, respectively) and sound field (90, 90, 90 and 90 %) tested conditions. Unlike the reference study, here the *TestEar* showed a significant effect on the *ASSR level*. This was surprising, considering that the two-way interaction of *PC* and *TestEar* did not reach significance. Additionally, the significant interaction between the *PC* and *Frequency* could be due to the added frequency-dependent room effect on the signal reproduced in a sound field.

To evaluate how representative the acoustic condition of the test room (categorized as a OT room) is in comparison to those of the audiometric clinic rooms, an analysis of the single number frequency acoustic parameters (EDT_{mid}, $T20_{mid}$ and $D50_{mid}$) was conducted for the reference measurement position and the four additional positions located at 0.15 m from the reference (left-right and front-back). The EDT_{mid}, $T20_{mid}$ and $D50_{mid}$ were calculated as the arithmetic mean between the frequency octave bands centered at 500 and 1000 Hz, as described in the standard ISO 3382-1 (2009). Figure 3.7 shows the averaged EDT_{mid}, $T20_{mid}$ and $D50_{mid}$ and STD across the five listener positions in the audiometric clinic rooms as a function of the volume. The acoustic parameters estimated for the test room used for sound field ASSR measurements are also presented in the figure. The results reveal that the estimated EDT_{mid}, $T20_{mid}$ and $D50_{mid}$ for the test room (0.17 s, 0.23 s, and 0.98) are within a similar range than the values obtained for the audiometric clinic rooms considered in the study, particulary for the OT rooms (0.12 – 0.44 s, 0.14 – 0.45 s, and 0.81 – 0.99). Furthermore, the STD obtained for the test room is also similar to the STD estimated for the OT rooms. This indicates that the test room is illustrative of the acoustic conditions of the audiometric clinic rooms.

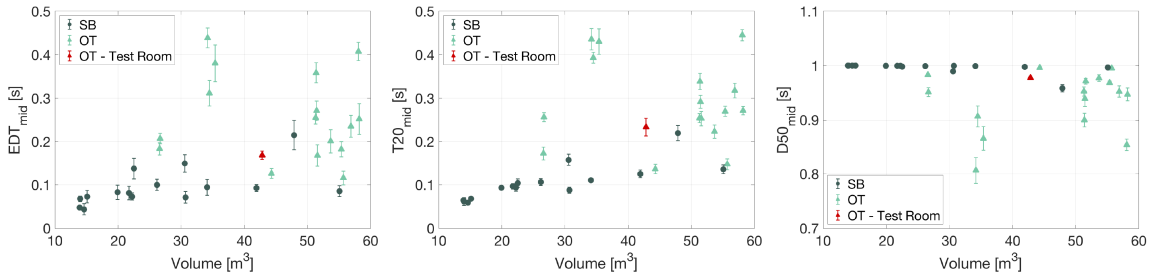

Figure 3.7: Mean EDT_{mid} (left column), $T20_{mid}$ (center column) and $D50_{mid}$ (right column) as a function of the volume for the sound booth (SB, dark green) and office-type (OT, light green) audiometric clinic rooms. The error bars indicate the standard deviation. The mean and standard deviation were calculated by averaging the estimated values across the reference point and the four additional measurement positions (0.15 m, left-right and front-back).

3.4 Discussion

3.4.1 Objective room acoustic parameters

The analysis of the estimated EDT, T20 and D50 for the audiometric testing rooms revealed a large variability of the acoustic conditions across all SB and OT rooms. The differences between the two types of rooms were also pronounced. Expectedly, the SB rooms were generally smaller with shorter reverberation times (EDT and T20) and higher speech definition (D50), compared with the OT rooms. The OT rooms showed a more irregular relation between the volume and the objective acoustic parameters than the SB rooms. This is likely due to the fact that the SB rooms are well-defined audiology facilities built with the aim of providing low reverberation and background noise for hearing screening (Hirschorn and Singer, 1989). In contrast, the OT rooms are usually adapted to the space and available facilities in the clinics, and therefore their dimensions and materials vary greatly. The substantial differences found in the acoustic conditions across the rooms also seem to reflect the lack of detailed acoustic standards, regulations and guidelines for healthcare facilities related to the needs of the audiometric testing rooms.

The acoustic regulations and standards that provide recommended limit values for healthcare facilities are often defined in terms of sound insulation (Rasmussen, 2018a; Machimbarrena et al., 2019; Fausti et al., 2019; Evans and Himmel, 2012; Zoontjens and Cockings, 2014). Some standards also highlight the importance of the acoustic comfort and acoustic privacy in healthcare facilities and provide target values for acoustic design in terms of minimum absorption area based on the type of room (Fausti et al., 2019; Van Wyk et al., 2014; DIN 18041, 2016; Nocke, 2016; HTM 08-01, 2013; Sykes et al., 2012; Clarke, 2011; UNI 11532, 2018). Although sometimes maximum recommended reverberation times are provided (Rasmussen, 2018b; NBE-CA-82, 1982;

SFS 5907, 2004; IST 45, 2016; NS 8175, 2012; SS 25268, 2007; AS/NZS 2107, 2000), these limit values are mostly defined for hospital bedrooms, treatment and consultation rooms. The audiometric clinic rooms could potentially be categorized as treatment or consultation rooms. Under this assumption, it could then be relevant to compare the estimated objective acoustic parameters with the available regulations and standards. Table 3.4 lists some of the available acoustic standards and regulations that provide maximum recommended reverberation times for healthcare facilities.

Table 3.4: List of acoustic regulations and standards that indicates reverberation time limits for healthcare facilities. Adapted from Rasmussen (2018a) and Clarke (2011).

Country	Reference	Type of room	Maximum reverberation times [s]
Denmark	DBR18 (2018)	Hospital bedrooms, examination rooms and treatment rooms	$\leq$ 0.6
Finland*	SFS 5907:2004 (2004)	Hospital bedrooms	$\leq$ 0.6 and $\leq$ 0.8
Island*	IST 45:2016 (2016)	Hospital bedrooms	$\leq$ 0.5 and $\leq$ 0.8
Norway*	NS 8175:2012 (2012)	Hospital bedrooms	$\leq$ 0.4 and $\leq$ 0.8
Sweden*	SS 25268:2007 + T1:2017 (2007)	Hospital bedrooms	$\leq$ 0.5 and $\leq$ 0.6
Australia / New Zealand	AS/NZS 2107:2000 (2007)	Casualty areas, consultation rooms, dental clinics, geriatric rehabilitation, intensive care wards	0.4 – 0.6
Spain	NBE-CA-82 (1982)	Casualty areas, consultation rooms, intensive care wards	0.8 – 1.5

*The standard includes acoustic quality classes. Detailed descriptions are found in references.

14 out of 15 SB rooms (93.3%) fulfilled the most strict recommendation (T20 $\leq$ 0.4 for all the frequency range), whereas only 7 out of 16 (43.7%) OT rooms did. Compared to the laxest requirement for reverberation (T20 $\leq$ 0.8 for all the frequency range), all SB rooms, and 15 out 16 (93.7%) OT rooms fulfilled the criterion. These set limits in the regulations and standards aim at obtaining good speech intelligibility and high acoustic comfort. However, for audiometric testing rooms it is questionable if the established recommended values are sufficiently restrictive for the purpose of the rooms, e.g., audiological assessments using sound field tests (Abouchacra et al., 2011).

As mentioned before, the British Society of Audiology (BSA) defined practice guidelines with relevant recommendations for sound field audiometry in clinical audiological applications (BSA 2019b). The guidelines recommend a reverberation time of less than 0.25 s across all frequencies. In the present study, 13 out of 15 (86.6%) SB rooms fulfilled this criterion, whereas only 2 out of 16 (12.5%) OT rooms have reverberation times below the recommended limit. The guidelines also recommend that the room dimensions for pediatric assessment should be larger than 4 x 6 m, which corresponds to a volume of around 57.6 m^3 assuming a height of 2.4 m. On this basis, only two OT rooms fulfilled the BSA criteria for pediatric assessment.

Although the reverberation time is the most common room acoustic descriptor, parameters such as D50 and STI could also be relevant and informative with regard to the

acoustic conditions of the audiometric testing rooms. This is because these parameters refer to speech intelligibility, which is crucial for audiometric tests and communication with patients who typically suffer from hearing difficulties. Unfortunately, these additional parameters are not considered in the current acoustic regulations, standards, and guidelines for health care facilities reviewed in the present study. Further support for the inclusion of additional room acoustic measures in the design and evaluation of audiometric clinic rooms comes from the fact that some of the audiologists stated (during the acoustic measurements in the clinics) that not all rooms were "good enough" for communication with the patients. Thus, even though both SB and OT rooms seemed to have overall good acoustic conditions based on the reported standards, the acoustic quality of some of the rooms might be insufficient for audiometric testing. It is therefore important to establish standards that are appropriate for the purpose of sound field audiometric testing. Another critical factor could be the uniformity of the acoustic conditions across rooms, which ensures reproducible and reliable sound field measurements across audiology clinics.

3.4.2 Speech transmission index

Across the clinic rooms, the estimated STI had a relatively large variability (around 20%). The STI values were, however, higher than 0.8 for all the SB and OT rooms. The recommended minimum intelligibility rating for rooms where prolonged face-to-face communication occurs should be "good," which corresponds to STI values in the range 0.60 – 0.75 for normal-hearing listeners (ISO 9921 2003), and 0.75 – 0.8 for older listeners with mild age-related hearing loss (ISO/TR 22411 2008). Compared to the standard, the obtained STI values for all SB and OT rooms correspond to an "excellent" speech intelligibility rating for both normal-hearing (> 0.75) and older listeners (> 0.8). This suggests that the acoustic condition of all the standard audiology clinic rooms is suitable for person-to-person speech communication within 1 m distance. This can be considered a typical distance relevant for the interaction between the audiologist and the patient, as well as for audiological sound field testing, such as speech audiometry and speech in noise tests. The speech intelligibility, as well as STI, can also be affected by the background noise in the rooms. The potential effect of the background noise, however, is not included in this study since the measurements were conducted under atypical operating conditions with fairly low background noise.

3.4.3 Sound pressure level distribution for warble tones

According to ISO 8253-2 (2010), the evaluation of the acoustics of a room for sound field audiometry should be based on the variability in the sound pressure level of the warble tones across all the frequencies considered (summarized in Table 3.3). Compared to this standard, none of the audiometric testing rooms analyzed here can be categorized as any of the established sound field types required for sound field audiometry. The obtained results in the analysis of the warble tones showed a large variability for both types of

rooms, with ΔSPL $>$ 2.5 dB for at least one frequency in all rooms. This would be critical for sound field audiometry where natural head movements are expected during the test.

The evaluation based on SPL differences could produce high uncertainties. This is because small variations in the measurement position can result in large changes in the SPL due to the low modal density at low frequencies, leading to uneven frequency responses (particularly in small rooms) (Hirschorn and Singer, 1989; Jacobsen and Juhl, 2013). This is illustrated in Figure 3.8, which shows the frequency response for the 500 Hz octave band measured at the reference and at the four additional positions in one of the audiometric clinic rooms. Here, the large frequency response differences observed across positions could explain the substantial SPL differences obtained in the analysis of warble tones. However, it is noteworthy that in the present study the warble tones were analyzed based on the recorded impulse responses, implying perfect stationarity. With a real practical measurement using a sound level meter in the room, a smaller SPL variation would be expected due to the non-stationary nature of the sound field in the rooms.

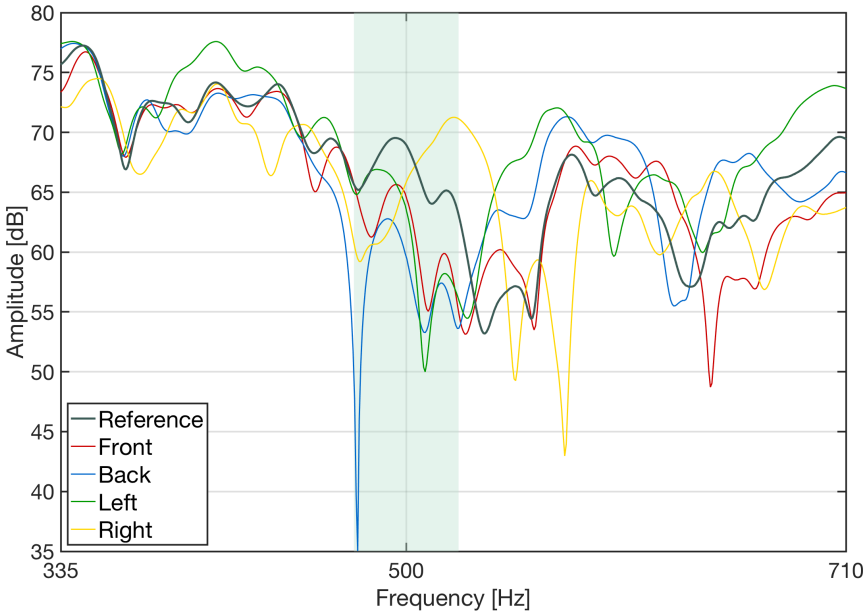

Figure 3.8: Frequency responses in the 500 Hz octave band measured in one of the audiometric testing rooms for the reference position (dark green) and the four additional positions located at the front (red), back (blue), left (green), and right (yellow). The shadowed (light green) area indicates the frequency swing of the 500 Hz warble tone.

3.4.4 Sound-field auditory steady-state response measurements

The comparison between the ASSR measurements conducted with insert earphones and in the sound field showed an effect of the playback condition, with lower ASSR levels in the case of the sound field measurements. This is likely due to the reduced stimulus modulation caused by the reverberation in the room (Houtgast et al., 1980; Plomp, 1983), which could, in turn, affect the ASSR level that is known to decrease as the stimulus modulation reduces (Picton et al., 1987; Rees et al., 1986). The effect of the room acoustics on the level of sound-field ASSR measurements has also been shown in a previous study that employed simulated rooms inspired by real audiometric testing rooms (Zapata-Rodríguez et al., 2020a). Interestingly, despite the observed reduction in the recorded ASSR level for the sound-field playback, the detection rates obtained in the present study were high and similar to those obtained through insert earphones, given sufficient recording time. Although the reduction in the ASSR level was not reflected in the detection rates, this could lead to poorer measurement performance, potentially requiring longer measurement time to detect the low ASSR levels recorded in sound field (Don and Elberling, 1994; Elberling et al., 2007).

The test room used for the sound field ASSR measurements seemed to be a representative of real audiometric clinic rooms. The room acoustic conditions of this test room were found to be appropriate based on the standards in terms of the reverberation time. However, as discussed above, an evaluation based only on the reverberation time might not be sufficiently informative, especially for more complex measurements such as ASSR that relies heavily on the properties of the stimulus being maintained, e.g., the stimulus modulation. In fact, Zapata-Rodríguez et al. (2020a) showed that EDT and an auditory inspired relative modulation power model (which quantifies the reduction in the stimulus modulation in the room) could be more accurate parameters for the evaluation of rooms for sound field ASSR measurements (Zapata-Rodríguez et al., 2020a). Considering the increasing interest in objective measurements such us sound field ASSR for hearing aid fitting validation, it is important to establish acoustic standards that can evaluate the rooms according to the requirements of this type of measurements.

3.5 Summary and conclusions

The present study assessed the acoustic conditions of 31 state-of-the-art audiometric clinic rooms where sound-field testing is regularly conducted. The evaluation was done based on the objective room acoustic parameters EDT, T20 and D50 and comparison with the available acoustic regulations, standards, and guidelines for healthcare facilities. The acoustic condition of the rooms was also evaluated in terms of the STI and the spatial variation in sound pressure level of warble tones for sound field audiometry. Additionally, the viability of sound field ASSR measurements was studied in a representative test room via comparison with insert-phone ASSR recordings. The main findings were as follows:

- Large variation in EDT, T20 and D50 was observed across audiometric testing rooms. Both the sound booth (SB) and office type (OT) rooms fulfilled the criteria of reverberation time stipulated by the acoustic regulations and standards. Considering the guideline for sound field audiometry, only the SB rooms had T20 within the recommended range.
- STI was high for all the audiometric testing rooms, indicating that all rooms provide excellent acoustic conditions in terms of speech intelligibility for listeners with normal-hearing and mild hearing loss.
- Large spatial variation in sound pressure level for warble tones was calculated in all the rooms for the four warble-tone nominal frequencies. None of the rooms met the criteria stipulated by the standard for sound field audiometry.
- A reduction in the ASSR level was found for the sound field condition compared to insert earphone measurements. Nevertheless, similar detection rates were achieved in the two conditions, given sufficient recording time.

  To conclude, the findings presented here support the need for establishing acoustic standards that define the optimal acoustic conditions for audiometric testing rooms considering all possible sound field tests that are carried out in these rooms. This could help to ensure that audiometric measurements recorded in sound field were more reliable and consistent across different clinical environments.

Acknowledgments

The authors thank Franck Michel from Aarhus Universitetshospital, Erik Kjærbøl from Bispebjerg/Gentofte Hospital, Jesper Hvass Schmidt from Odense Universitetshospital, Mario Cebulla and Ute Geiger from Universitätsklinikum Würzburg, and all the hospitals staff for their assistance in organizing the measurement sessions. The authors also thank Ecophon Group for providing the wall absorbers for the test room, Mai-Britt Beldam from Ecophon Group, and Ida Hoffmann for their assistance with the acoustic measurements in the clinics.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was funded by Interacoustics A/S, the William Demant Foundation and the Innovation Fund Denmark, Grant. No. 5189-00070B.

CHAPTER 4

Reproduction of nearby sources using loudspeaker-based virtual sound environments[1]

Abstract

This study evaluated the accuracy of the reproduction of nearby sources in virtual sound environments adding a nearby loudspeaker physically located at the simulated source position. This was investigated using the room acoustic model PARISM (Phased Acoustical Radiosity and the Image Source Method), a spherical 64-loudspeaker array and the nearest loudspeaker method. Comparison of measured impulse responses with and without the additional nearby loudspeaker showed that its inclusion produces better energy balance between the direct sound and early and late reflections, reflected in an improvement of the measured early decay time (EDT) and definition (D50) relative to the acoustic simulation.

Keywords: nearby sources, virtual sound environment, nearest loudspeaker.

4.1 Introduction

Loudspeaker-based reproduction techniques in connection with advanced room acoustic simulations can create valid virtual sound environments (VSE) adapted to diverse applications. In hearing research, for instance, VSE can be implemented for the investigation of auditory perception and cognition in challenging listening situations (Seeber et al., 2010; Ahrens et al., 2019), and for the evaluation of hearing assistive devices under their normal mode of operation (Cubick and Dau, 2016; Oreinos and Buchholz, 2015). The VSEs are limited by factors such as the reproduction method (Grimm et al., 2015), the loudspeaker array available(Ahrens and Spors, 2008), and the type of simulated acoustic scenarios. However, the primary end goal is always to create physically accurate and perceptually authentic auralizations.

The accuracy of the reproduction methods for VSEs is an important aspect that has been investigated extensively. The objective evaluation of VSEs has shown good estimations of the reverberation time (T30) (Ahrens et al., 2019; Cubick and Dau, 2016; Oreinos and Buchholz, 2015; Favrot and Buchholz, 2010), although discrepancies have

[1] This chapter is based on Zapata-Rodríguez et al. (2020c).

been reported between the early decay times (EDTs) derived from the VSEs and the real world measurements (Ahrens et al., 2019). Similarly, some discrepancies have been shown in the subjective evaluation of VSEs using speech intelligibility scores (Ahrens et al., 2019; Cubick and Dau, 2016; Favrot and Buchholz, 2009), in VSEs created with a combination of higher-order ambisonics (HOA; Grimm et al., 2015; Favrot and Buchholz, 2010) and the nearest loudspeaker method (NLM) for the direct sound and early reflections. It has been suggested that these deviations are related to a lack of agreement of EDT estimates (Ahrens et al., 2019).

The NLM approach consists of mapping each individual simulated reflection to the nearest loudspeaker in the array (Seeber et al., 2010). NLM becomes more accurate for denser loudspeaker arrays and when simulating virtual sources that match real positions of loudspeakers in the array (Seeber et al., 2010). However, when the source and receiver distances are substantially shorter than the physical distance of the array, the direct sound and reflections cannot be accurately reproduced. This includes, for instance, typical face-to-face conversations where distances are often shorter than 1 m, or audiology tests such as speech in noise test and sound field ASSR, in which the sound source is 1 m away from the listener.

The reproduction of nearby sources is a challenging task for loudspeaker-based reproduction methods, as these assume that the listener is located in the far-field (Seeber et al., 2010). To overcome this, regularization filters were developed to create near field compensated HOA (Daniel, 2003). Although theoretically this method correctly reproduces nearby sources in the sweet spot, in practice, it is inviable because it produces loudspeaker signals of extremely large amplitudes at low frequencies. To improve the reproduction of nearby sources with the NLM, here, it is proposed to include an additional nearby (NB) loudspeaker located within the array, at the exact simulated source-to-receiver distance. Westermann and Buchholz (2017), followed a similar approach to investigate the influence of nearby maskers on speech intelligibility measurements with a VSE of a cafeteria using a 41-loudspeaker array (Westermann and Buchholz, 2017). The reverberant sound field was reproduced with HOA, whereas the NLM was employed for the reproduction of the direct sound and the early reflections, as defined by the LoRA toolbox (Favrot and Buchholz, 2010). The nearby maskers were closer to the listener than the loudspeakers in the array. The direct sound component of the nearby maskers was presented from the additional nearby loudspeakers within the array, whereas the remaining part of the room impulse responses (IRs) was reproduced using the 41-loudspeaker array. Their results focused on the effect of nearby maskers on subjective measures of speech intelligibility. However, the VSE including the additional nearby sources was not acoustically validated.

The goal of the present study is to objectively assess the loudspeaker-based auralization using an array of 64 loudspeakers without and with the additional NB loudspeaker, hereafter, referred to as 64-array and 64+NB-array, respectively. A calibration procedure is also presented to include NB sources within any loudspeaker array systems. The study evaluates loudspeaker reproduction in terms of measurable objective parameters

compared to the room acoustic simulation. It is hypothesized that the reproduction of the direct sound and early reflections is more accurate with the 64+NB-array since the virtual source location will match that of the nearby loudspeaker. This would be reflected in objective acoustic measurements of EDT and D50, which are more sensitive to the fine structure of the IRs than reverberation time measures such as T20.

4.2 Material and methods

4.2.1 Test rooms

Four test rooms were considered for the analysis of the loudspeaker-based auralization system. The acoustic design of the rooms was inspired by typical meeting rooms or audiological clinic rooms. These types of rooms were selected because, in them, it is common to encounter every-day scenarios in which nearby sources are critical. The rooms were highly damped rectangular rooms with six surfaces, each of them with only one material. The acoustic properties of the rooms were defined by the scattering and absorption coefficients. For absorbing walls, the flow resistivity was used to calculate the angle-dependent impedance of porous absorbers via an empirical model (Miki, 1990), implying an extended reaction modeling (Jeong, 2011). The dimensions of the rooms, simulated source and receiver positions, as well as the T20 (averaged over the 500 Hz and 1 kHz one-octave bands) are listed in Table 4.1. Detailed information about the acoustic properties used in the simulations is provided in the Supplementary Material.

Table 4.1: Dimensions, source and receiver positions, and T20 (averaged over the 500 Hz and 1 kHz one-octave bands) for all four simulated test rooms.

Simulated room	**Dimension [m]**	**Source position [m]**	**Receiver position [m]**	**Reverberation Time, T20 [s]**
	l_X, l_y, l_z	x_0, y_0, z_0	x, y, z	
1^{st} **Room**	6.0, 3.9, 2.5	2.5, 1.6, 1.0	3.5, 1.6, 1.0	0.28
2^{nd} **Room**	2.8, 2.5, 2.0	0.5, 2.0, 1.0	1.5, 2.0, 1.0	0.52
$3rd$ **Room**	2.8, 2.5, 2.0	0.7, 1.2, 1.0	1.7, 1.2, 1.0	0.22
4^{th} **Room**	6.0, 3.9, 2.5	0.6, 3.1, 1.0	1.6, 3.1, 1.0	0.65

4.2.2 Acoustic scene generation

The test rooms were simulated with the Phased Acoustical Radiosity and the Image Source Method PARISM (Marbjerg et al., 2015), all for a source to receiver pointing at each other at a distance of 1 m. The model combines the phased image source methods (ISM) and the acoustical radiosity (AR) to simulate the specular and scattered reflections, respectively. The phase information is included in the ISM predicting the

interference between room modes accurately. This is a key feature when simulating smaller rooms since their sound fields are dominated by a modal behavior in a broader frequency range. The test rooms were simulated including the frequency response and a simplified directivity pattern of the SP90 loudspeaker (SP90, RadioEar, Middelfart, Denmark), which is also used as the NB loudspeaker in the loudspeaker-based reproduction system. The SP90 is an audiometric loudspeaker with a frequency response from 125 to 8 kHz. The directivity pattern was measured 1 m away from the SP90 with a 5-degree resolution for the axes illustrated in the Figure 4.1.A. The directivity was implemented in PARISM by choosing the directivity index value corresponding to the closest measured direction.

The loudspeaker-based reproduction uses the NLM. This consists in panning each reflection to the closest loudspeaker available in the VSE based on the direction of the reflection (Seeber et al., 2010). For PARISM, the impulse responses are obtained for each of the loudspeaker channels for both the ISM and AR parts, which were then combined to obtain the total PARISM impulse response for each of the loudspeakers in the array (h_{PARISM_i}) (Marbjerg et al., 2020). The loudspeaker impulse responses were computed by panning each individual image source for the ISM, and each surface element for the AR (Marbjerg et al., 2020). This is achieved by applying weights of one and zero for the active loudspeaker and the other loudspeakers in the array, respectively.

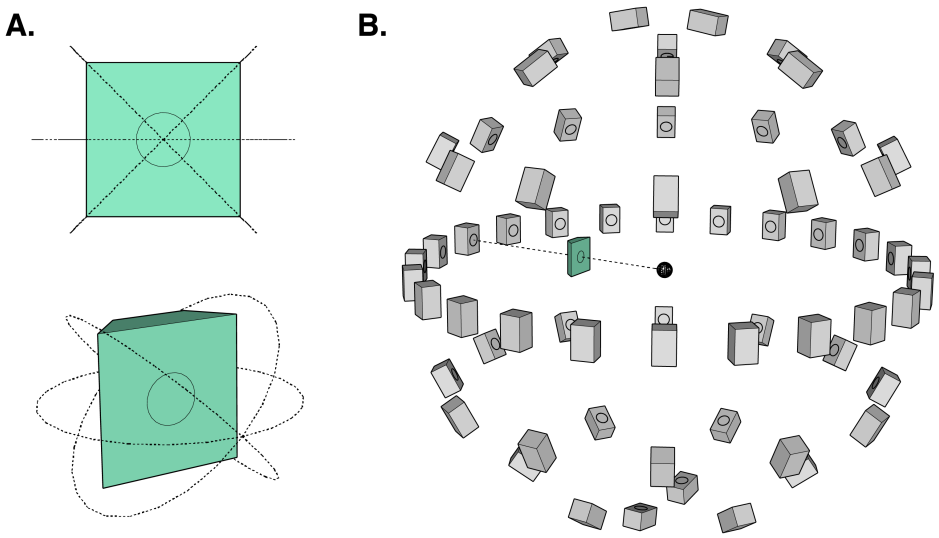

Figure 4.1: Virtual sound environment with the nearby (NB) loudspeaker. A. Sketch of the SP90 loudspeaker. The lines indicate the axes used for the measurements of the directivity pattern. B. Sketch of the 64-loudspeaker array (in gray), and the NB loudspeaker (in green) placed 1 m away from the center of the array (black sphere).

4.2.3 Virtual sound environment

The reproduction system consisted of an array of 64 loudspeakers (KEF LS50, KEF, Maidstone, United Kingdom) arranged in a full sphere with a radius of 2.4 m to the receiver position, located at the center of the array. The 64-array consisted of seven rings, with elevation angles of ±80° , ±56° ,±28° and 0° with respect to the receiver positions, as shown in Figure 4.1.B (gray). The 64 loudspeakers were calibrated to obtain a flat frequency response within a range of, ±3dB across all loudspeakers. Detailed information about the 64-loudspeaker system can be found in the paper by Ahrens et al. (2019).

In addition to the 64-array, the additional NB loudspeaker was placed 1 m away in front to position, as illustrated in Figure 4.1.B (green), which corresponded to the simulated source and receiver distance. The individual loudspeakers of the 64+NB-array were calibrated to individually generate the same sound pressure level at the center of the array (sweet spot) when reproducing pink noise. A temporal calibration was also carried out to achieve a correct acoustic time alignment at the center of the array. For this, individual delays were measured for each loudspeaker's channel and corrected to ensure that all loudspeaker's acoustic signals arrived at the same time at the center of the array. The NB loudspeaker was further delayed to compensate for its location relative to the 64-array (1.4 m).

4.2.4 Calibration procedure

The virtual acoustic scene was calibrated to obtain, during its reproduction with the 64+NB-array, the same energy balance between the direct and reverberant sound field as in the PARISM simulation. For this purpose, the simulated total sound field (TSF) was divided into the direct sound field (DSF), and the reverberant sound field (RSF). The loudspeaker impulse responses (IRs) for the TSF and DSF were obtained from the PARISM simulation with NLM for the 64-array. The RSF was computed by subtracting the DSF from the TSF ($h^{RSF}_{1:64} = h^{TSF}_{1:64} - h^{DSF}_{1:64}$). The proposed 64+NB-array system was calibrated as follows:

1. The impulse responses of the system, $h^{TSF}_{1:64}$, $h^{DSF}_{1:64}$, and $h^{RSF}_{1:64}$, were convolved with the target signal to be reproduced (s) in the VSE, i.e. $p^{TSF}_{1:64} = h^{TSF}_{1:64} * s$, generating three multichannel signals: $p^{TSF}_{1:64}$, $p^{DSF}_{1:64}$, and $p^{RSF}_{1:64}$.

2. The sound pressure level generated by each of the multichannel signals was measured at the center of the array.

3. The target signal was corrected in time considering the simulated source and receiver distance. The resulting signal (s^{DSF}_{NB}) was then reproduced through the NB-loudspeaker and calibrated to produce the same sound pressure level as that previously estimated with $p^{DSF}_{1:64}$.

4. The $p_{1:64}^{RSF}$ was recorded again and its sound pressure level was corrected ($p\prime_{1:64}^{RSF}$) to account for the effect of the placement of the NB-loudspeaker inside the array. This was done to assure that the respective $p_{1:64}^{RSF}$, with and without the NB-loudspeaker present, produced the same sound pressure level at the center of the array.

5. The multichannel signal for the 64+NB-array auralization was thus the result of the simultaneous reproduction of p_{NB}^{DSF} and $p\prime_{1:64}^{RSF}$. Finally, it was verified that the sound pressure level of the reproduction for the 64+NB-array was the same (with a ±1.5 dB tolerance) as that generated by the total sound field with the 64-array ($p_{1:64}^{TSF}$).

The proposed calibration procedure can be generalized to different loudspeaker-based auralization methods (e.g., higher-order ambisonics, vector-based amplitude panning, etc.) adapted to include nearby sources. Importantly, the NB loudspeaker considered in the VSE, should also be used in the room acoustic simulation. Here, for instance, the frequency response and directivity pattern of the SP90 loudspeaker was included in the room acoustic simulations. Therefore, additional frequency equalization was not needed for the NB-loudspeaker since it was inherently included by also selecting the SP90 as the NB loudspeaker in the reproduction system.

4.2.5 Objective evaluation of the VSE

To objectively validate the loudspeaker-based auralization, IRs were measured at the center of the array for the test case. The IRs were measured using a 10 s long exponential sine sweep from 150 to 8000 Hz, and a sample frequency of 48 kHz. The frequency range was selected based on the frequency response of the NB loudspeaker used (SP90). The evaluation was based on a comparison of both measured IRs (using the 64-array and the 64+NB-array) with the purely simulated IR by PARISM and the NLM. The latter IR was the sum of all simulated loudspeakers' IRs. EDT, T20, and D50 were calculated from the IRs in one-octave frequency bands (f_c) from 250 to 4000 Hz, as described in ISO 3382-1 (2009). The frequency band of 125 Hz was not investigated due to the low frequency limit of the SP90. A mean absolute difference was calculated across frequencies for the acoustic parameters, defined by Eq. (4.1):

$$\overline{d_X} = \frac{1}{N_{fc}} \sum_{f_c=f_{c,min}}^{f_{c,max}} \left| X_{PARISM_{fc}} - X_{array_{fc}} \right| \tag{4.1}$$

where X indicates each of the calculated acoustic parameters (EDT, T20 or D50), f_c corresponds to the frequency band (from $f_{c,min}$ = 250 Hz to $f_{c,max}$ = 4000 Hz), and N_{f_c} is the total number of frequency bands (N_{f_c}=5).

4.2.6 Results and discussions

Figure 4.2 shows the IR simulated with PARISM, and the two IRs measured with the 64-array and the 64+NB-array for the 1st simulated room. In general, the time of arrival of the reflections was similar across the three IRs. However, the energy balance between the direct sound and early reflections was not accurately reproduced in the measured IR with the 64-array. This can be seen in the early reflection at around 5 ms (highlighted gray area), which is amplified for the 64-array. In contrast, when introducing the NB loudspeaker, the energy balance substantially improves with respect to the IR measured with the 64-array.

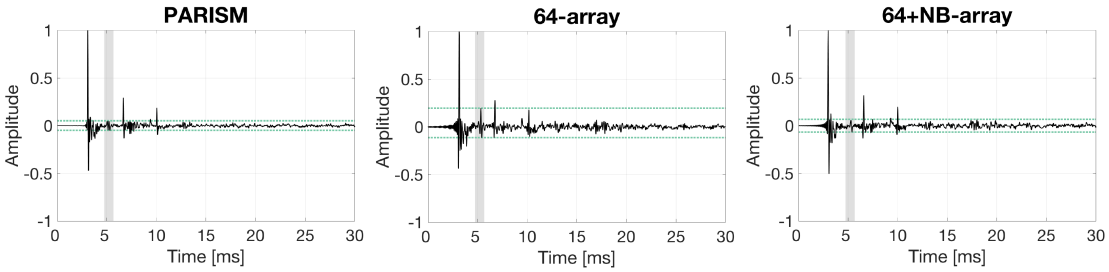

Figure 4.2: Impulse responses (IRs) of the test-case room computed with the PARISM and the NLM (left), measured with the 64-array (center), and the 64+NB-array (right).

Figure 4.3 shows the comparison of the acoustic parameters estimated from the IRs. The gray areas indicate the just-noticeable differences (JNDs) for the EDT, T20, and D50, as defined in ISO 3382-1 (2009). Overall, the evaluated acoustic parameters of the simulated room were well reproduced by both loudspeaker systems. The measured T20 and D50 agreed with the simulation across all frequency bands within one JND. However, the EDT obtained with the 64+NB-array was closer to the simulation, for which the resulting EDTs were within one JND for the frequency bands of 500, 1000 and 2000 Hz. Thus, the inclusion of the NB loudspeaker improved the reproduction of the direct sound and early reflections. This was expected since the nearby loudspeaker could reduce the positional errors of the reproduction method.

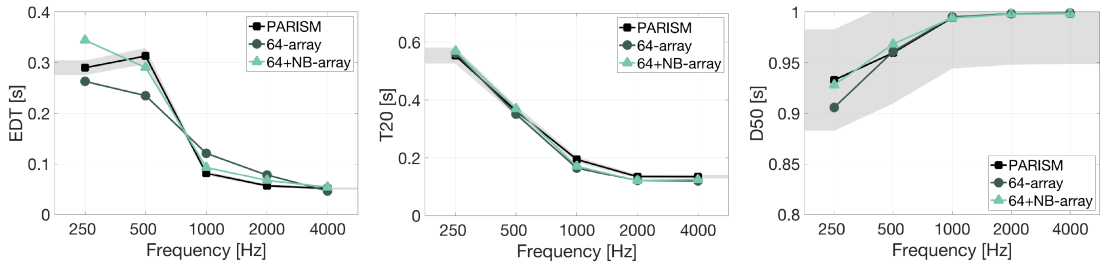

Figure 4.3: Early decay time (EDT; left), reverberation time (T20; center) and definition (D50; right) estimated from the simulated and measured IRs. The gray areas indicated the just-noticeable differences relative to the simulated IR.

Figure 4.4 presents the mean absolute error ($\overline{d}$) across frequency bands calculated for all four simulated test rooms. As observed for the results of the 1st simulated room, T20 was not substantially affected by the reproduction system. The results also revealed that the reproduction with the 64+NB-array produced more accurate EDT for three of the rooms (1st, 2nd, and 3rd), whereas D50 improved for all rooms. Objectively, this represents an improvement in the accuracy of the VSE reproduction, which could be substantially greater for arrays of fewer loudspeaker where the positional errors are larger (Seeber et al., 2010; Grimm et al., 2015). Further research is needed to determine whether this improvement can be acoustically perceived, which is expected considering that EDT correlates well with the perception of reverberance (ISO 3382-1, ISO3382-1, 2009). A good reproduction of the direct sound and early reflections could then be crucial for speech intelligibility measurements (Arweiler et al., 2009).

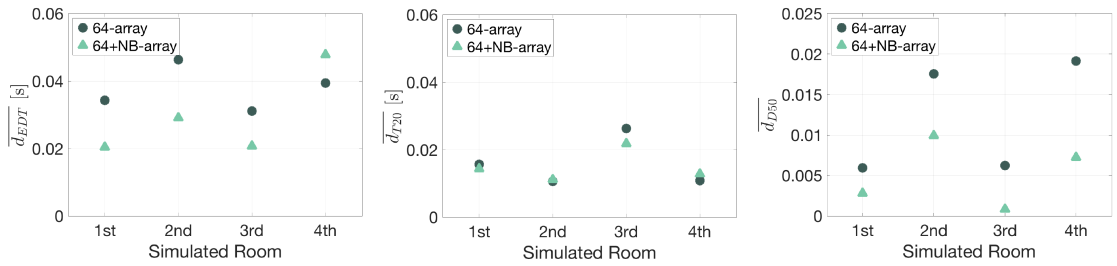

Figure 4.4: Mean absolute error calculated for the early decay time (EDT; left), reverberation time (T20; center) and definition (D50; right) for the four test rooms.

4.3 Conclusions

The results indicate that virtual acoustic reproduction of nearby sources using the nearest loudspeaker method can be improved by adding a nearby loudspeaker calibrated with the proposed procedure. The improvement is achieved by reducing the mapping errors in the reproduction method thanks to the location match between the simulated virtual source and the nearby loudspeaker. This results in a better reproduction of the direct sound and early reflections, which was specially reflected in the improvement of the measured EDT and D50 with the 64-loudspeaker array and the additional nearby loudspeaker.

Acknowledgments

The authors would like to thank Axel Ahrens, Efren Fernandez-Grande, and Marton Marschall, for the valuable discussions. This research was funded by Interacoustics A/S and the Innovation Fund Denmark, Grant. No. 5189-00070B. The William Demant Foundation supported the first and third authors' contributions to this work.

4.4 Supplementary

Includes detailed information about the simulated rooms used in the study.

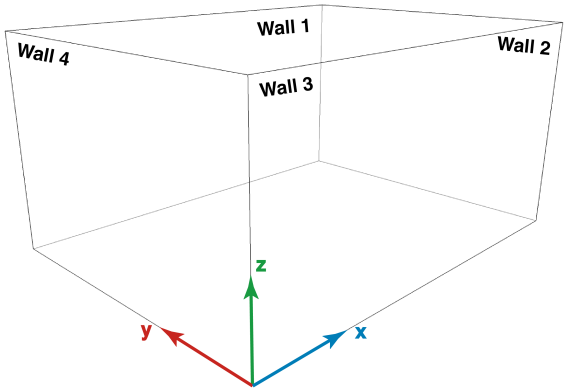

Figure 4.5: Coordinate system used for all the room acoustic simulations.

Table 4.2: Input data for the simulation of the 1^{st} Room. Absorption and scattering coefficients were defined in frequency octave bands from 63 Hz to 8 kHz. The porous absorbing walls were defined with flow resistivity (σ), thickness of the absorbers (h), and without air gap behind the absorber.

Room surface	**Frequency band [Hz]**							
	63	125	250	500	1000	2000	4000	8000
Absorption coefficients, and flow resistivity for porous absorbing walls								
Floor	0.10	0.10	0.09	0.08	0.05	0.05	0.05	0.05
Ceiling	$\sigma = 12.9\ kPas/m^2$, $h = 0.04\ m$							
Wall 1	0.20	0.10	0.15	0.05	0.06	0.08	0.08	0.09
Wall 2	$\sigma = 12.9\ kPas/m^2$, $h = 0.04\ m$							
Wall 3	$\sigma = 12.9\ kPas/m^2$, $h = 0.04\ m$							
Wall 4	0.30	0.20	0.15	0.09	0.09	0.04	0.03	0.02
Scattering coefficients								
Floor	0.43	0.43	0.44	0.44	0.46	0.48	0.48	0.51
Ceiling	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 1	0.43	0.43	0.44	0.44	0.46	0.48	0.48	0.51
Wall 2	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 3	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 4	0.43	0.43	0.44	0.44	0.46	0.48	0.48	0.51

Table 4.3: Input data for the simulation of the 2^{nd} Room. Absorption and scattering coefficients were defined in frequency octave bands from 63 Hz to 8 kHz.

Room surface	**Frequency band [Hz]**							
	63	**125**	**63**	**500**	**63**	**2000**	**63**	**8000**
Absorption coefficients								
Floor	0.10	0.10	0.09	0.08	0.05	0.15	0.15	0.15
Ceiling	0.25	0.35	0.3	0.32	0.3	0.25	0.2	0.25
Wall 1	0.20	0.10	0.15	0.05	0.06	0.18	0.18	0.19
Wall 2	0.15	0.25	0.2	0.22	0.2	0.15	0.1	0.15
Wall 3	0.15	0.25	0.2	0.22	0.2	0.15	0.1	0.15
Wall 4	0.30	0.20	0.15	0.09	0.09	0.14	0.13	0.12
Scattering coefficients								
Floor	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61
Ceiling	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 1	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61
Wall 2	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 3	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 4	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61

Table 4.4: Input data for the simulation of the 3^{rd} Room. Absorption and scattering coefficients were defined in frequency octave bands from 63 Hz to 8 kHz. The porous absorbing walls were defined with flow resistivity (σ), thickness of the absorbers (h), and without air gap behind the absorber.

Room surface	**Frequency band [Hz]**							
	63	**125**	**63**	**500**	**63**	**2000**	**63**	**8000**
Absorption coefficients, and flow resistivity for porous absorbing walls								
Floor	0.10	0.10	0.09	0.08	0.05	0.15	0.15	0.15
Ceiling	$\sigma = 12.9\ kPas/m^2$, $h = 0.04\ m$							
Wall 1	$\sigma = 12.9\ kPas/m^2$, $h = 0.04\ m$							
Wall 2	$\sigma = 12.9\ kPas/m^2$, $h = 0.04\ m$							
Wall 3	$\sigma = 12.9\ kPas/m^2$, $h = 0.04\ m$							
Wall 4	0.30	0.20	0.15	0.09	0.09	0.14	0.13	0.12
Scattering coefficients								
Floor	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61
Ceiling	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 1	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61
Wall 2	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 3	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 4	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61

Table 4.5: Input data for the simulation of the 4^{th} Room. Absorption and scattering coefficients were defined in frequency octave bands from 63 Hz to 8 kHz.

Room surface	**Frequency band [Hz]**							
	63	**125**	**63**	**500**	**63**	**2000**	**63**	**8000**
Absorption coefficients								
Floor	0.10	0.10	0.09	0.08	0.05	0.15	0.15	0.15
Ceiling	0.25	0.35	0.30	0.32	0.30	0.25	0.20	0.25
Wall 1	0.20	0.10	0.15	0.05	0.06	0.18	0.18	0.19
Wall 2	0.15	0.25	0.20	0.22	0.20	0.15	0.10	0.15
Wall 3	0.15	0.25	0.20	0.22	0.20	0.15	0.10	0.15
Wall 4	0.30	0.20	0.15	0.09	0.09	0.14	0.13	0.12
Scattering coefficients								
Floor	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61
Ceiling	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 1	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61
Wall 2	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 3	0.02	0.03	0.04	0.04	0.04	0.08	0.08	0.08
Wall 4	0.83	0.83	0.84	0.64	0.66	0.68	0.68	0.61

CHAPTER 5

Effect of Room Acoustics on Sound-Field Auditory Steady-State Responses for Hearing Aid Fitting Validation[1]

Abstract

Sound-field auditory steady-state response (ASSR) has received increasing interest given its potential application as an objective clinical procedure for hearing aid fitting validation in prelingual infants. In practice, the stimulus modulation can be degraded by the acoustics of the clinic room in which the test is carried out, which in return could reduce the ASSR level and elongate the test time. The detrimental effect of the room on the ASSR level was already reported using a monaural auralization approach through insert earphones, and the early decay time (EDT) as well as a measure of relative modulation power were proposed as predictors of the ASSR level in any given room. The present study further investigates the effect of room acoustics on sound-field ASSR measurements in an attempt to better characterize the effect under more realistic testing conditions. Here, the evaluation considered the ASSR level, the detection rates, and the detection time. ASSR was recorded from 26 normal-hearing adults using the ISTS-modified version of the narrow-band CE-Chirps. Eight room acoustic conditions were tested by reproduction in a virtual sound environment that consisted of a 64-loudspeaker array with an added nearby loudspeaker. The tested acoustic conditions were carefully selected based on a data set of 31 real audiometric testing rooms. A significant reduction of the ASSR level was found across acoustic conditions, which led to overall longer detection times and lower detection rates for the room conditions compared to the reference anechoic condition. Good correlations were established for the ASSR level predictions based on either the EDT or the relative modulation power measure, which further generalized the suitability of the models for predicting the performance of the room for sound-field ASSR measurements.

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

[1]This chapter is based on Zapata-Rodríguez et al. (2020d).

5.1 Introduction

The validation of hearing aid fitting is an essential phase in aural rehabilitation and early intervention. This process aims to verify that the patient is receiving an appropriate amplification, and hence, benefiting from the hearing aid. In pre-lingual infants, hearing aid fitting validation is even more critical since a proper amplification promotes the development of communication skills (Sininger et al., 2010; Sharma et al., 2002), allowing hearing impaired infants to reach a level of speech communication comparable to that of their normal-hearing peers (Yoshinaga-Itano et al., 1998b; Moeller, 2000). This validation, however, is a challenging procedure in infants between 2-9 months of age as they cannot respond reliably to behavioral tests. There is therefore a need for an alternative testing procedure for conducting the hearing aid fitting validation in such patients. Several studies have proposed that electrophysiological tests based on auditory steady-state responses (ASSR) measured in sound field can serve as a novel objective clinical procedure for the validation of hearing aid amplification (e.g., Picton et al., 1998). ASSR is an electrophysiological response that measures objectively the resulting signal processed by the brain in response to an auditory stimulus (Picton et al., 2003; Rance, 2008). Therefore, it bypasses the need for the patients' voluntary response, which is particularly beneficial for assessing infants and hard-to-test adults.

Sound-field ASSR measurements involve the presentation of the acoustic stimulus through a loudspeaker. This enables the inclusion of the hearing aid into the stimulation path and hence, validates its fitting under their usual mode of operation. The loudspeaker presentation, however, carries implications that can compromise the clinical implementation of the sound-field ASSR test. One critical aspect is that the ASSR is highly dependent on the stimulus modulation, which is defined as its time varying amplitude. This has been previously investigated for amplitude-modulated pure tones, for which the ASSR amplitude decreases as the stimulus modulation reduces (Dimitrijevic et al., 2001; John et al., 2001; Lins et al., 1995; Picton et al., 1987; Rees et al., 1986; Roß et al., 2000; Kuwada et al., 1986; Rønne, 2013; Boettcher et al., 2001; Bharadwaj et al., 2015). Consequently, when measuring sound-field ASSR in non-anechoic rooms, the stimulus modulation could be degraded due to the reverberation and background noise of the test room (Houtgast et al., 1980; Plomp, 1983). The detrimental effect of room acoustics on the modulation is well established for speech signals, for which the intelligibility decreases when the natural amplitude fluctuations are distorted (Houtgast et al., 1980; Plomp, 1983). Similarly, it is then expected that any reduction in the ASSR stimulus modulation due to room acoustics will be reflected as a reduction in the ASSR level. This could lead to an increase in the testing time required to reach the signal to noise ratio for detecting the low ASSR levels (Dobie and Wilson, 1996; Cebulla et al., 2006; Laugesen et al., 2018). The need for long testing sessions could compromise the clinical implementation of sound-field ASSR measurement due to the high efficiency demands in clinical testing.

The effect of the room on sound-field ASSR measurements has not been broadly studied. Most investigations have focused on establishing the feasibility of sound-field

ASSR as an accurate clinical procedure for hearing aid fitting validation (Picton et al., 1998; Stroebel et al., 2007; Damarla and Manjula, 2007; Shemesh et al., 2012; Hernández-Pérez and Torres-Fortuny, 2013; Selim et al., 2012; Sardari et al., 2015; Park et al., 2013). Picton et al. (1998), for instance, showed that similar normal-hearing physiological thresholds could be obtained for ASSR measurements conducted via earphones or in a sound field. However, the study was limited to a single acoustic environment defined as a single-walled audiometric test room with a quasi free sound field, and did not explore the consequences of the sound field stimulation in terms of the ASSR level and detection time. A clear understanding of the implications of sound field presentation in ASSR measurements is essential since the acoustic conditions of audiometric testing rooms are not standardized due to the lack of specific regulations. This has been shown in a recent study that investigated the acoustic conditions of 31 state-of-the-art audiometric testing rooms (Zapata-Rodríguez et al., 2020b). The results of Zapata-Rodríguez et al. (2020b) revealed a great variability in the estimated acoustic parameters across rooms. Moreover, the study showed a comparison of the ASSR levels measured in young adult normal-hearing subjects through either insert earphones or in sound field. Their findings also revealed a reduction for the ASSR levels measured in a sound field, even though the testing room had similar dimensions and acoustic conditions as real audiometric clinic rooms.

The investigation of Zapata-Rodríguez et al. (2020b) supported the results of a previous study that evaluated the influence of room acoustics on the level of sound-field ASSR measured in normal-hearing listeners (Zapata-Rodríguez et al., 2020a). That study employed a monaural auralization approach to measure ASSR, with insert earphones, under three simulated acoustic conditions inspired by audiometric testing rooms. A reduction of the ASSR level was obtained for the simulated room acoustic conditions in comparison to the anechoic reference. Moreover, it was suggested that this reduction can be estimated in the testing room based on two measurable acoustic predictors: (1) the early decay time (EDT), which is the time interval necessary for a 60 dB energy decay estimated from the initial 10 dB decay; and (2) an auditory-inspired relative modulation power model. The latter quantifies the changes in the ASSR stimulus modulation (induced by the room) relative to the anechoic version of the stimulus. Thus, Zapata-Rodríguez et al. (2020a), presented a first approximation for characterizing the effect of room acoustics on sound-field ASSR measurements. However, the experiment only considered monaural room acoustic simulations presented through insert earphones, and hence did not include the effect of the local sound field resulting from head movements. Moreover, the room sample size investigated did not sufficiently encompass typical room acoustic conditions of real audiometric clinic rooms due to their large variability. Thus, there is still a need for determining the effect of the room acoustics on sound-field ASSR under more representative test environments.

The current study investigates the effect of room acoustics on sound-field ASSR measurements quantified in terms of the ASSR level, detection rates, and detection time. It also analyses the potential implications for the clinical implementation of sound-field ASSR for hearing aid fitting validation. This research builds upon the study of Zapata-

Rodríguez et al. (2020a). Here, however, sound-field ASSR was measured using an extended set of room acoustic conditions that were presented through a loudspeaker-based virtual sound reproduction system. This novel framework enables a more realistic evaluation of the effect of the room on sound-field ASSR measurements by including the natural head movements as well as the effect of the listener's head, torso and pinna present in the sound field.

To isolate the effect of room acoustics, sound-field ASSR was measured in normal-hearing adult subjects for eight simulated room acoustic conditions and an additional anechoic reference condition. ASSR was elicited using a novel ISTS-modified version of the narrow-band CE-Chirps® (Watson et al., 2019). The rooms were designed based on the data set of clinic audiometric testing rooms presented in Zapata-Rodríguez et al. (2020b), and simulated with the Phased Acoustical Radiosity and the Image Source Method (PARISM) (Marbjerg et al., 2015). The virtual sound environments were reproduced using the nearest loudspeaker method (2010) through a 64-loudspeaker array with an additional nearby loudspeaker, as described in Zapata-Rodríguez et al. (2020c). It is hypothesized that the ASSR level will be reduced due to the detrimental effect of the room on the stimulus modulation. Accordingly, lower detection rates and longer measurement times are expected for the most challenging acoustic conditions (rooms with longer reverberation times). The present study also further investigates whether the models proposed in Zapata-Rodríguez et al. (2020a) can be generalized to predict ASSR levels under more realistic listening conditions.

5.2 Methods

5.2.1 Participants

Twenty-six young adult normal-hearing subjects (12 female, mean age 25) participated in the experiment. All subjects had pure-tone thresholds lower than or equal to 20 dB HL in the range 125–8000 Hz, which were measured for both ears using an Interacoustics AC40 audiometer. Ear examination via otoscopy was conducted before each test session. Informed consent was obtained prior to the experiment from all participants, who were financially compensated with gift cards. The research was approved by the Science-Ethics Committee for the Capital Region of Denmark.

5.2.2 Stimuli

The ISTS-modified version of the narrow-band CE-Chirps® was used to elicit the ASSR. This novel stimulus was developed to provide a stimulus with speech-like properties that can be identified as a speech signal by the hearing aid. For this purpose, the CE-Chirps® (Elberling and Don, 2010) were modified to have speech characteristics based on the International Speech Test Signal (IST 45, 2016 ; Holube et al., 2010). The

stimulus keeps the main properties of the CE-Chirp® as much as possible (e.g., the compensation for the wave traveling delay in the basilar membrane is maintained) to enable the successful recording of the ASSR. The stimulus was developed based on CE-Chirps® centered at 707 Hz (two-octave-band wide), as well as 2000 Hz and 4000 Hz (one-octave-band wide). Each frequency band was presented at an individual rate in the range of 40, 90 and 90 Hz, respectively. The modified CE-Chirp® stimulus also has the same one-octave band levels as the original ISTS when presented at its nominal level of 65 dB SPL. The detailed procedure for the generation of the stimulus can be found in Laugesen et al. (2018), and Watson et al. (2019).

5.2.3 Virtual room acoustic enviroments

Virtual sound environments (VSE) were created to simulate and record sound-field ASSR under eight different room acoustic conditions. An additional anechoic reference condition was also included. There are currently no specific room acoustic guidelines available for sound-field ASSR, as this is not a standardized clinical test yet. Therefore, the eight tested room acoustic conditions were designed based on the real data set of 31 state-of-the-art audiometric testing rooms presented in Zapata-Rodríguez et al. (2020b). The aim was to produce acoustic environments that resembled the conditions of the rooms where sound-field ASSR could be potentially recorded in the clinics. For the simulations, the distance between the receiver (patient) and the loudspeaker was set to 1 m. This value was also used for the height, which corresponds to the approximate height of a person sitting on the testing position, as recommended by the ISO 8253-2 for sound field audiometry (ISO 8253-2, 2010). The rooms were simulated using the novel room acoustic model PARISM (Marbjerg et al., 2015). This model was developed with a focus on simulating small rooms, and therefore, accurately calculates the room modes. It combines the acoustical radiosity (AR) and the phased image source methods (ISM) for the simulation of the scattered and specular reflections, respectively.

The VSE were reproduced using the nearest loudspeaker method (Seeber et al., 2010) and a spherical 64-loudspeaker array with a radius of 2.4 m, housed in an anechoic chamber located at the Technical University of Denmark. An additional nearby loudspeaker was used for the reproduction of the direct sound from the simulated source following the methodology proposed by Zapata-Rodríguez et al. (2020c), which showed that the reproduction of nearby sources in VSE could be improved by adding a nearby loudspeaker physically located at the simulated source position. The SP90 loudspeaker (SP90, RadioEar, Middelfart, Denmark) was used as the nearby loudspeaker in the PARISM simulations as well as in the reproduction of the VSE. This audiometric loudspeaker has a frequency response in the range 125-8000 Hz. Both the frequency response and the directivity of the SP90 loudspeaker were submitted to the model as described in Zapata-Rodríguez et al. (2020c).

The design criteria for the simulated rooms was based on the room acoustic parameters that have been shown to be potential predictors of the ASSR level for clinic rooms

(Zapata-Rodríguez et al., 2020a). As mentioned, these parameters are the EDT and the auditory-model-inspired relative modulation power measure. The EDT is a reverberation time measure estimated from the initial 10 dB decay of the sound pressure level in a room after the source has been stopped. It is known to be closely related to the subjective impression of the reverberation and quantifies the early part of the decay curve (ISO 3382-1, 2009). As described in Zapata-Rodríguez et al. (2020a), the EDT is expected to be more relevant for sound-field ASSR testing than other reverberation time measurements such as T20. This is because the early reflections of the room become increasingly more relevant when decreasing the source to receiver distance, which in the test was relatively short (1 m). The auditory inspired relative modulation power model assumes that the ASSR is elicited by a nonlinear representation of the signal, and estimates the change in the signal modulation due to the acoustics of the room. The main characteristics of the modulation power model are briefly presented in the following, for further details see Zapata-Rodríguez et al. (2020a).

First, a gammatone filter bank (Johannesma, 1972) consisting of 24 or 12 filters evenly spaced over the stimulus frequency band by 1/12th octave was used to process the anechoic and reverberant stimuli. Temporal envelopes are then calculated from each gammatone filter output and normalized based on their DC components. Next, the discrete Fourier transform is calculated for each temporal envelope divided into 65536-sample epochs, which are averaged over all epochs and all filter bands to estimate the envelope power of the signal. The relative modulation power is finally calculated in dB based on the modulation power estimated (at the frequency bin of the repetition rate for each stimulus frequency band) from the envelope power of the reverberant stimulus relative to the anechoic signal. In Zapata-Rodríguez et al. (2020a), the relative modulation power model was described for periodic signals. Here, a raised-cosine window is multiplied with each of the Hilbert envelopes per epoch before the DC component normalization to smooth the epoch edges of the resulting speech-modified stimulus. The cosine ramps were here implemented using 5% at either end of each epoch, similarly as in Laugesen et al., (2018).

The EDT was calculated for five impulse responses measured with an omnidirectional microphone in each of the 31 audiometric testing rooms at the reference position (1 m from the source) and at four additional positions (located 0.15 m to the front, back, left, and right of the reference), as presented in Zapata-Rodríguez et al. (2020b). The relative modulation power was calculated for the same positions using the ISTS-modified version of the narrow-band CE-Chirps convolved with the five measured impulse responses. The mean EDT and relative modulation power were estimated across the five positions in each of the real clinic rooms. These measurements served as the design criteria for the simulation of the eight room acoustic conditions used in the present study. Figure 5.1 shows the EDT and the relative modulation power measured for the real audiometric clinic rooms and the tested simulated room acoustic conditions for the three frequency bands of interest.

The values presented in Figure 5.1 for the tested acoustic conditions are based on

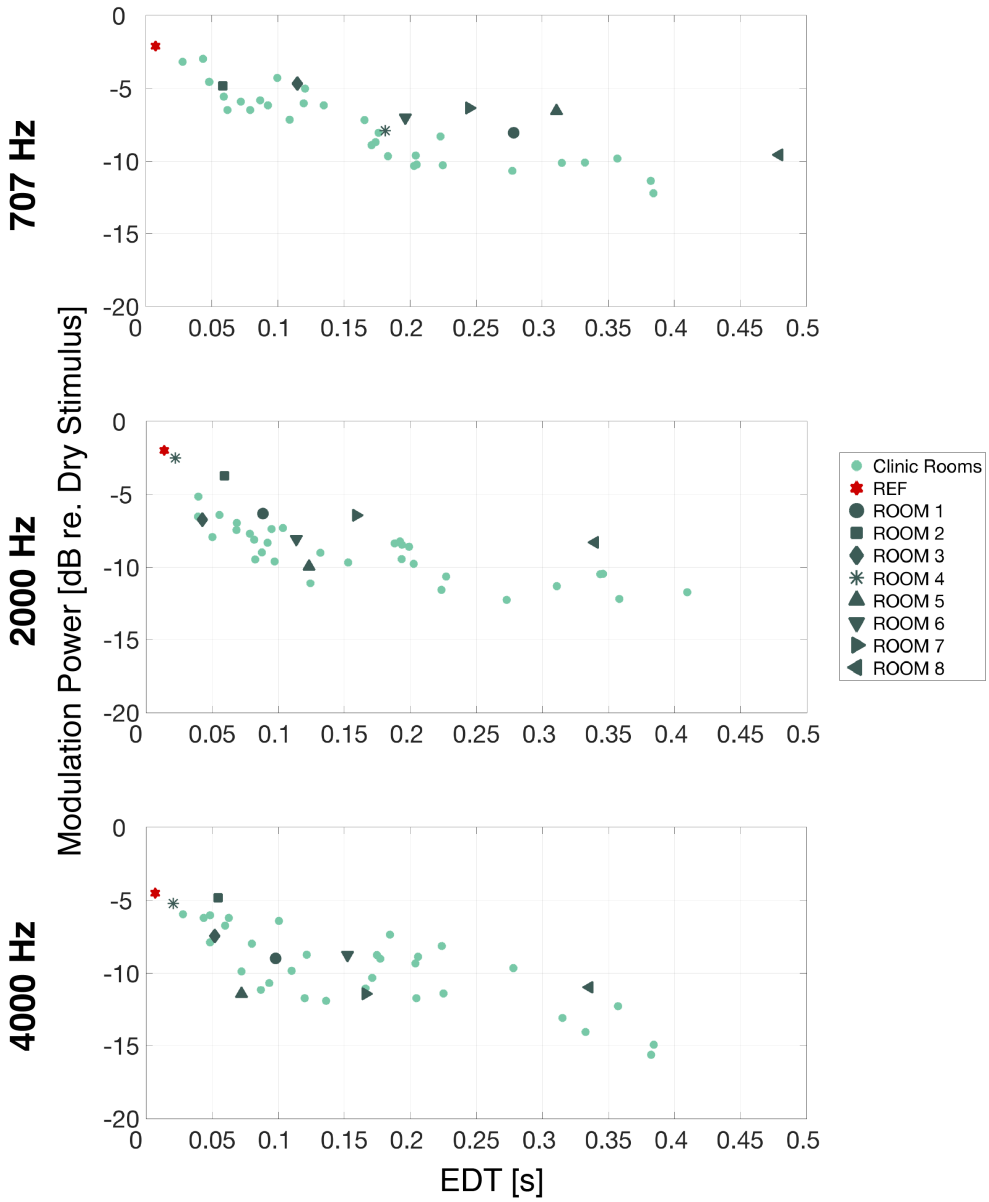

Figure 5.1: Relative modulation power model as a function of the EDT for the eight simulated acoustic conditions (red and black markers), and for the data set of real audiometric clinic rooms (light green markers) presented in Zapata-Rodríguez et al. (2020b).

the measurements of the simulated rooms in the virtual sound environment. EDT was estimated from impulse responses that were measured using a 10 s long exponential sine sweep from 150 to 8000 Hz, and a sampling frequency of 48 kHz. The relative modulation power was computed from the recorded reverberant ISTS-modified CE-Chirp stimulus for each of the eight conditions tested. The reference condition consisted of measuring the anechoic response due to the ASSR "dry" stimulus presented through the SP90

loudspeaker in the laboratory. Figure 5.1 also shows the EDT and relative modulation power measured in the anechoic chamber with the SP90 loudspeaker. The relative modulation power was calculated reference to the anechoic "dry" stimulus for all rooms, the real audiometric clinic rooms, the simulated acoustic conditions and the reference condition.

The dimensions of the simulated rooms were selected based on the smallest and largest dimensions of the real audiometric clinic rooms. Table 5.1 shows the dimensions of the rooms, the source position, the receiver position and the T20 (averaged between the one-octave-bands of 500 Hz and 1 kHz).

Table 5.1: Dimensions, source position, receiver position, and T20 averaged over the three frequency bands of interest for all eight simulated acoustic conditions.

Simulated room	**Dimension [m]**	**Source position [m]**	**Receiver position [m]**	**Reverberation Time, T20 [s]**
	l_X, l_y, l_z	x_0, y_0, z_0	x, y, z	
ROOM 1	6.0, 3.9, 2.5	2.5, 1.6, 1.0	3.5, 1.6, 1.0	0.30
ROOM 2	6.0, 3.9, 2.5	0.6, 3.1, 1.0	1.6, 3.1, 1.0	0.16
ROOM 3	2.8, 2.5, 2.0	0.7, 1.2, 1.0	1.7, 1.2, 1.0	0.11
ROOM 4	2.8, 2.5, 2.0	0.4, 2.0, 1.0	1.4, 2.0, 1.0	0.14
ROOM 5	6.0, 3.9, 2.5	0.6, 3.1, 1.0	1.6, 3.1, 1.0	0.40
ROOM 6	2.8, 2.5, 2.0	0.7, 1.2, 1.0	1.7, 1.2, 1.0	0.36
ROOM 7	2.8, 2.5, 2.0	0.4, 2.0, 1.0	1.4, 2.0, 1.0	0.35
ROOM 8	6.0, 3.9, 2.5	2.5, 1.6, 1.0	3.5, 1.6, 1.0	0.69

5.2.4 ASSR Measurements

ASSRs were recorded using a standard clinical electrode montage that consists of four surface electrodes, two placed on each mastoid (ipsi- and contra-lateral active), one on the vertex (high forehead ground), and the other on the cheek (reference). The electrode impedances never exceeded 3 kΩ and were kept as equal as possible across the four electrodes. During the measurements, the test subjects sat in a chair in front of the loudspeaker with their heads resting on a headrest. The subjects were instructed to keep their heads still facing the SP90 loudspeaker (as shown in Figure 5.2) and were encouraged to relax and sleep if possible. The position and orientation of the head were monitored during the test session and were corrected when necessary. ASSR was recorded individually for each of the ears, while the non-test ear was blocked with an earplug.

The experiment consisted of two testing sessions. In each session, the reference condition as well as four room conditions were initially measured for the test ear, and then as many room conditions as time allowed were measured for the opposite ear. The

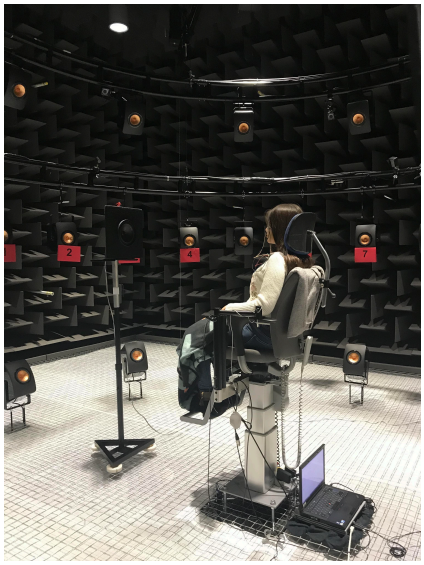

Figure 5.2: Experimental setup of sound-field ASSR measurement.

order of presentation of the conditions was randomized for all the test subjects, and the selection of the test ear was counterbalanced across subjects. Four subjects did not participate in the second session. The ASSR was recorded with a sampling frequency of 48 kHz and an analysis block length of 98304 samples (2.05 s per block). An artifact rejection level of 40 μV was used for all the measurements. Each condition was recorded for a maximum of 15 minutes (440 blocks), or less if at least 9 mins of clean blocks (264 blocks) could be collected. The stimuli were generated by custom Matlab software and were sent via the TESIRA Server digital signal processing (DSP) units connected to TESIRA SOC-4 digital-to-analog converters. Then, the stimuli were amplified by Sonible d:24 amplifiers (Sonible GmbH, Graz, Austria) before being reproduced through the loudspeaker array. The ASSR signals were preamplified by the Interacoustics ERA preamplifier connected to the Interacoustics Eclipse EP25 ABR, which was used as a front-end. The ASSR signal was then sent to the TESIRA units and was recorded in the custom Matlab software. Detection of the ASSR was evaluated online using a simple F-test detector (with a 1% error rate) on the first harmonic of each repetition rate.

The stimuli were presented at a nominal level of 65 dB. The calibration was conducted using a BK 4192 ½" microphone located at the center of the loudspeaker array. Compensation filters were implemented to slightly correct the balance of the sound pressure level across the stimulus frequency bands (induced by the room filtering) relative to the original target presentation levels. The overall nominal calibration level was verified every testing day before the measurements

5.2.5 ASSR postprocessing

All ASSR data were first pre-processed to remove blocks with high EEG noise (above 40 μV). In order to establish a fair assesment of the effect of the room on the sound-field ASSR measurements, the analysis was based on the first 254 clean epochs of each recording. On this basis, 1.6% were discarded (7 out of 428). Table 5.2 shows the number of recordings included in the analysis for each condition.

Table 5.2: Number of measurements per condition included in the analysis after post-processing with an artifact rejection level of ±40 V.

Acoustic condition	**REF**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Recordings included	59	49	39	43	48	41	49	50	43
Total recordings	59	49	40	44	48	42	51	52	43

The present study evaluated three outcomes: ASSR level, detection rate, and detection time. The ASSR level was determined using the weighted averaging method (John et al., 2001), accepting only recordings that met a frequency domain F-ratio test (Dobie and Wilson, 1996) with a 1% error rate. Thus, the ASSR level, located at the frequency bin of the repetition rate, was statistically compared with the EEG noise computed by averaging the noise power from the 20 frequency bins at each side neighboring the response bin. The frequency bins from the repetition rate harmonics, the 50 Hz line noise, and its harmonics were excluded from the noise power estimation. The ASSR level was analyzed only at the fundamental frequency of the ASSR spectrum. To minimize the variability across subjects due to the electrophysiological noise, the ASSR level was noise-corrected by subtracting the noise power from the ASSR power (Dobie and Wilson, 1996). The detection time was determined using the q-sample detector (Cebulla et al., 2006). This detection algorithm considers that the ASSR is represented in the ASSR spectrum by the fundamental frequency (located at the stimulus repetition rate) as well as its higher harmonics. The q-sample uses the amplitude and phase of the ASSR fundamental frequency and its harmonics to assess the presence of the ASSR, leading to shorter average detection times in comparison to the F-ratio test (Cebulla et al., 2006). The q-sample test provides a yes/no response in regards to the presence of the ASSR and the detection time. The exact q-sample detector clinically used in the Eclipse ASSR software was used (Cebulla and Stürzebecher, 2015). In the present study, the detection time was determined for each of the ASSR recordings collected using a q-sample detector with an error rate of 5%, and by including the ASSR fundamental frequency and the next 11 harmonics. The detection rate was evaluated using the q-sample detector and was based on the available recordings after the clean up process. The test-retest reliability of the experiment was evaluated based on the ASSR level.

5.2.6 Statistical analysis

The ASSR data were analyzed using linear mixed-model analyses of variance (ANOVA) implemented in the the software R version 3.5 with the lme4 library (Bates et al., 2015). The test ear was included as a random effect. In all the models, the main effects as well as the interactions were evaluated and removed when they were not significant.

A first model was computed to evaluate the reliability of the ASSR measurements (test-retest) based on the *ASSR level* recorded for the reference acoustic condition in the two different measurement sessions. The model included the *ASSR level* as the predictable variable, and the explanatory variables of test session (*TestSession* : 1, 2) and stimulus frequency (*Frequency*: 707, 2000, 4000 Hz) as main fixed effects. The *ASSR level* was defined as a continuous variable, whereas test session and stimulus frequency were treated as categorical variables. The effect of the room on the *ASSR level* was evaluated by fitting a linear-mixed model to the *ASSR level* considering the fixed effects of the room (*Room*: 1,2,3...8) and the stimulus frequency (*Frequency*: 707, 2000, 4000 Hz). For this model, the *ASSR level* was also taken as a continous variable, while *Room* and *Frequency* were defined as categorical variables. To assess the statistical significance across room conditions, a post-hoc analysis was implemented using the the estimated marginal means (Searle et al., 1980) with the Tukey method (Tukey, 1949). As in Zapata-Rodríguez et al. (2020a), linear mixed-effects models were fitted to the *ASSR level* to evaluate whether the early decay time (*EDT*) and the relative modulation power (*RModP*)could predict the ASSR level in a given room. These two models included, instead of the categorical variable *Room*, the continuous fixed effects for either the early decay time (*EDT*: time in seconds) or the relative modulation power (*RModP*: in dB). All the models calculated for the *ASSR level*, were also investigated for the *detection time* in order to establish its relation with the room. Before statistical analysis, detection times Log_{10} transformed (relative to 1 s).

5.3 Results

5.3.1 Sound-field ASSR measurements in different acoustic conditions

Figure 5.3 shows the boxplot distribution of the ASSR level for the tested acoustic conditions across the stimulus band center frequencies. The linear mixed-effect model revealed a highly significant effect of the two main factors (*Room*, *Frequency*) and their interaction (*Room* x *Frequency*). In addition, a significant random effect ($\rho = <.0001$) of test ear was found, indicating that much of the variability seen in the box plots is due to systematic variation between subjects/ears. Table 5.3 shows the outcome of the ANOVA. A general reduction is observed of the *ASSR level* for the room acoustic conditions compared to the reference. This effect is evident for the 707 Hz stimulus,

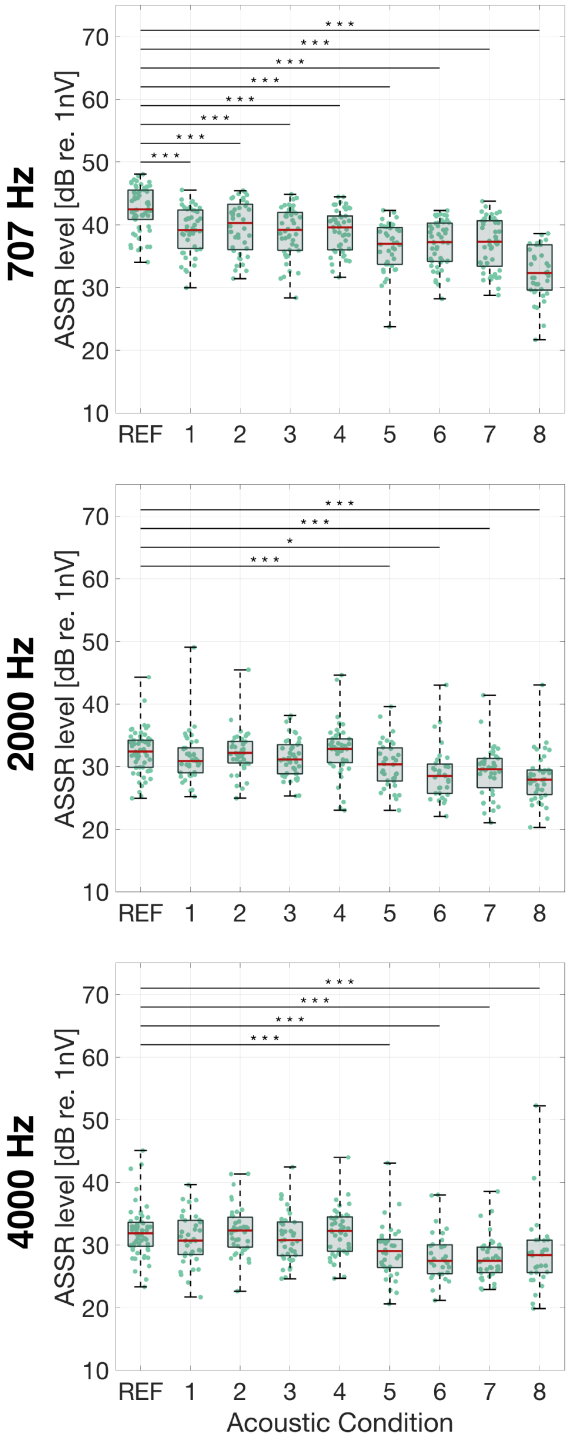

Figure 5.3: ASSR levels as a function of the tested acoustic conditions for each stimulus center frequency band (rows). The whiskers of the boxplot indicate the minimum and maximum data points of the distribution, boxes show the 25th (bottom edge) and 75th (top edge) percentiles as well as the median (red line). Room conditions significantly different from the anechoic REF condition are indicated with horizontal lines, * ρ<0.05, ** ρ<0.01, *** ρ<0.001.

where a median reduction of about 10 dB was measured. For the higher frequencies, this effect is less pronounced, even though all the main factors and interactions were significant. Also, in general, higher variability was obtained for the frequency bands

of 2000 and 4000 Hz compared to 707 Hz. These results suggest that the effect of the room is frequency-dependent, as indicated by the significant two-way interaction (*Room* x *Freq*).

Table 5.3: Summary results of the Mixed-Model Analyses of Variance.

Model 1. Effect of the room

	ASSR level		Log_{10} **(detection time)**	
		AIC = 5937.7		AIC = 2880.6
Factor	***F* statistic**	ρ **value**	***F* statistic**	ρ **value**
Room	$F(8, 1091.7) = 57.3$	<.0001***	$F(8, 1103.6) = 17.6$	<.0001***
Freq	$F(2, 1086.2) = 827.3$	<.0001***	$F(2, 1093.6) = 68.5$	<.0001***
Room × *Freq*	$F(16, 1085.7) = 4.5$	<.0001***	$F(16, 1092.7) = 2.1$	.0085***

Model 2. Effect of the EDT

	ASSR level		Log_{10} **(detection time)**	
		AIC = 6026.7		AIC = 2934.6
Factor	***F* statistic**	ρ **value**	***F* statistic**	ρ **value**
EDT	$F(1, 1090.8) = 282.4$	<.0001***	$F(1, 1101.6) = 93.8$	<.0001***
Freq	$F(2, 1085.8) = 422.2$	<.0001***	$F(2, 1092.5) = 48.4$	<.0001***
EDT × *Freq*	$F(2, 1086.9) = 1.0$	0.3647	$F(2, 1093.7) = 0.1$	0.929

Model 3. Effect of the relative modulation power

	ASSR level		Log_{10} **(detection time)**	
		AIC = 6058.8		AIC = 2899.7
Factor	***F* statistic**	ρ **value**	***F* statistic**	ρ **value**
RModP	$F(1, 1092.9) = 304.7$	<.0001***	$F(1, 1106.0) = 103.7$	<.0001***
Freq	$F(2, 1086.3) = 147.4$	<.0001***	$F(2, 1093.4) = 21.5$	<.0001***
RModP × *Freq*	$F(2, 1086.8) = 10.9$	<.0001***	$F(2, 1094.2) = 3.9$	.0231*

Model 4. Analysis of test - retest

	ASSR level	
Factor	***F* statistic**	ρ **value**
MeasSession	$F(1, 152.6) = 2.9$	0.0884
Freq	$F(2, 138.7) = 243.1$	<.0001***
MeasSession × *Freq*	$F(2, 138.7) = 0.4$	0.6545

Pairwise comparisons were subsequently conducted to assess the statistical significance across rooms for all the stimulus frequency bands. The results show that for the 707 Hz, all the room conditions were significantly different from the reference. In contrast, in the case of 2000 and 4000 Hz, only rooms 5, 6, 7 and 9 were found to be statistically different from the reference. Interestingly, all these rooms have low relative modulation power ($<$ -5 dB) and relative longer *EDTs* ($>$ 0.15 s), as shown in Figure

5.1. The horizontal lines and stars in Figure 5.3 indicate the room conditions that were significantly different from the reference. Moreover, the post-hoc analysis showed that for 2000 and 4000 Hz, the rooms with short EDT (1,2,3,4) were significantly different from the rooms with long *EDT* (6,7,8). In the case of 707 Hz, no clear trend was found between the rooms and the acoustic parameters, although 17 out of 28 pairwise comparisons between room conditions were significant. Table 5.1 in the supplementary material lists all significant differences found across the acoustic conditions. The frequency-dependent effect of the acoustic conditions on the *ASSR level* could be due to the changes in the stimulus modulation under reverberant conditions. This is supported by the fact that both parameters (*RModP* and *EDT*) are frequency-dependent as observed in Figure 5.1.

The reliability of ASSR measurements was evaluated by measuring the reference acoustic condition in two different days (test-retest). Figure 5.4 shows the *ASSR level* measured in two different sessions for each stimulus frequency band. As described above, the test–retest variability was evaluated using a linear mixed-model, which showed no statistical difference between the two testing sessions (*MeasSession*). While the main effect of Frequency on the *ASSR level* was significant, the two-way interaction (*Run* x *Frequency*) did not reach significance, see Table 5.3 for a summary of the ANOVA. Overall, the results show good reliability of the measurements, indicating that the tested effects were maintained across the two different testing sessions. The effect of *Frequency* is observed in the generally lower *ASSR levels* obtained for the 2000 and 4000 Hz stimuli than for the 707 Hz stimulus, which was previously reported in Watson et al., 2019.

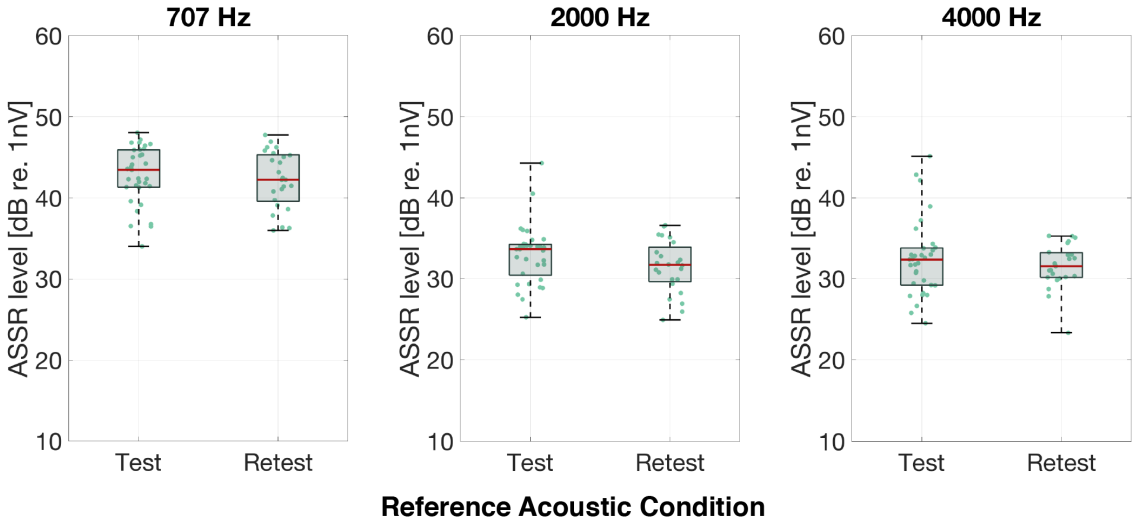

Figure 5.4: ASSR levels measured in two different test sessions for the reference condition for each stimulus center frequency band (columns). The whiskers of the boxplot indicate the minimum and maximum data points of the distribution, boxes show the 25th (bottom edge) and 75th (top edge) percentiles as well as the median (red line).

5.3.2 Relation between ASSR level and the early decay time and the relative modulation power

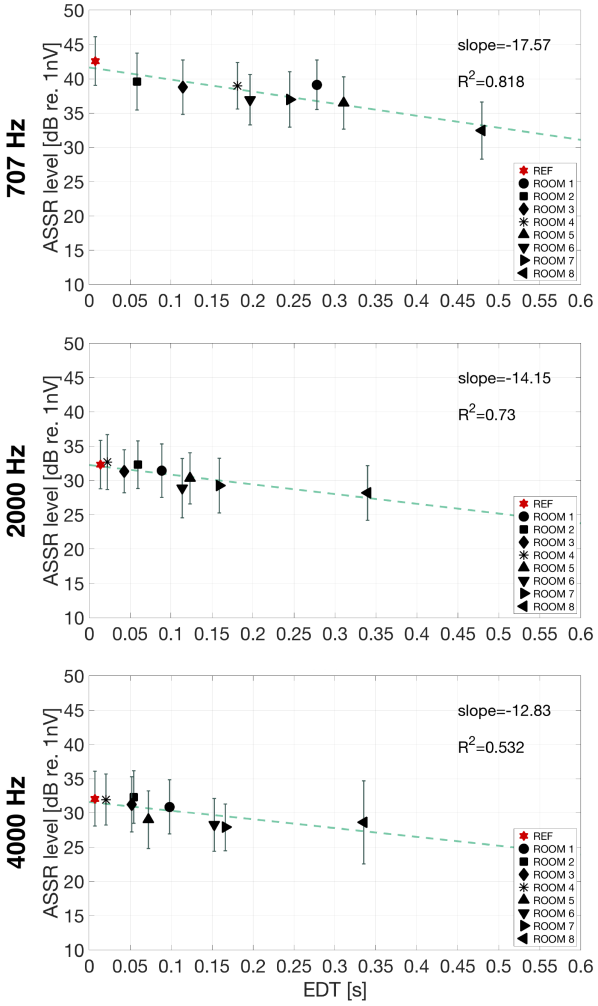

Figure 5.5: Mean ASSR level as a function of EDT for each Room and Frequency (rows). Error bars indicate standard deviation. Estimated regression lines (−−) are added to each panel.

Figure 5.5 shows the mean *ASSR level* as a function of *EDT* for each stimulus frequency band obtained with the different acoustic conditions. The results show a clear reduction of the *ASSR level* as a function of the *EDT* for all stimulus frequency bands. The statistical model revealed that both main effects (*EDT*, *Frequency*) were statistically significant, whereas their two-way interaction (*EDT* x *Frequency*) was not significant. The summary of the ANOVA is shown in Table 5.3. Linear regression models were estimated to evaluate the relation between the *ASSR level* and the *EDT* per stimulus frequency band, which are presented with their corresponding slope and coefficient of determination (R^2) in Figure 5.5. It was found that the *EDT* can predict more than 50% of the variability in the *ASSR level*, reflected in the obtained correlations for 707 Hz (R^2=0.82), 2000 Hz (R^2=0.73) and 4000 Hz (R^2=0.53). The resulting negative slopes indicate that the *ASSR level* reduces with increasing the *EDT*. The slopes only slightly

differed across frequencies, which agrees with the non-significant two-way interaction between the *EDT* and *Frequency*. These outcomes suggest that the *EDT* could be employed to estimate the *ASSR level* in the rooms for all stimulus frequency bands.

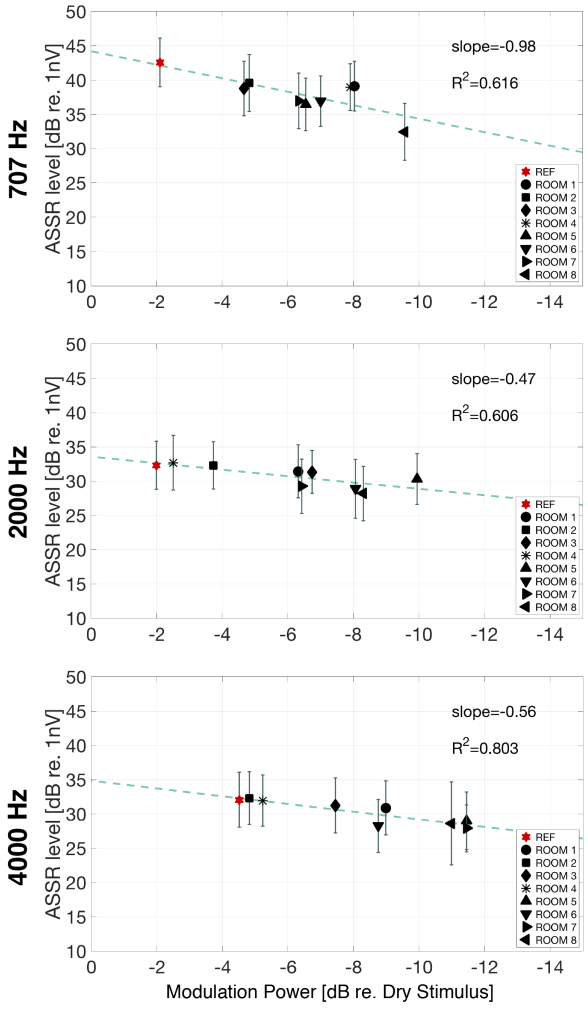

Figure 5.6: Mean ASSR level as a function of the relative modulation power (RModP) for each Room and Frequency (rows). Error bars indicate standard deviation. Estimated regression lines (——) are added to each panel.

Figure 5.6 shows the mean ASSR level as a function of the relative modulation power for each stimulus frequency band obtained with the different acoustic conditions. The results show an overall reduction of the ASSR level as the relative modulation power reduces across all stimulus frequency bands. Expectedly, the relative modulation power decreased for the room conditions compared to the reference condition. The main effect of RModP and Frequency as well as their two-way interaction on the ASSR level were all found to be significant. The outcome of the ANOVA is summarized in Table 5.3. As for the EDT, linear regression models were fitted to the ASSR data as a function of the relative modulation power per frequency band. This was done to evaluate the viability of the relative modulation power as a predictor of the ASSR level in any given room. The regression lines with their corresponding slopes and R^2 are shown in Figure 5.6. For all frequency bands, a good correlation was found between the relative modulation

power and the ASSR level, with R^2 values of 0.62, 0.61 and 0.80, for 707, 2000 and 4000 Hz, respectively. The regression models revealed a faster decrease (steeper slope) of the *ASSR level* for the frequency band of 707 Hz than for 2000 and 4000 Hz. This indicates that the relative modulation power model (*RModP*) can predict the *ASSR level* in the room for the evaluated stimulus frequency bands.

5.3.3 Detection time and detection rate of sound-field ASSR

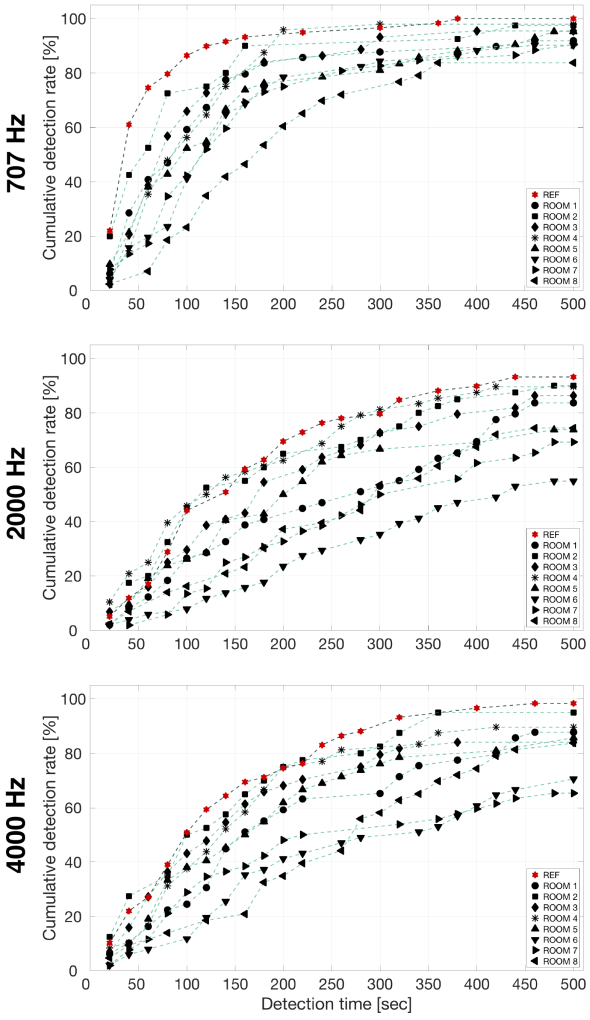

Figure 5.7: Cumulative detection rate as a function of the detection times across tested acoustic conditions estimated for each stimulus center band (rows).

Figure 5.7 shows the cumulative detection rate as a function of the *detection time* estimated using the q-sample detector for each of the tested acoustic conditions. The highest detection rates were obtained for the reference condition in all frequency bands. The ASSRs were detected faster for the 707 Hz frequency band. The linear mixed model confirmed the significant effects of the *Room*, *Frequency* and their interaction on

the *detection time*. The summary of the ANOVA analysis is presented in Table 5.3. The results also showed an overall increase in the *detection time* for the room conditions compared to the reference. For instance, in the case of 707 Hz, about 90% of responses across subjects were detected for the reference condition within the first 100 seconds, whereas for Rooms 6 and 7, only 40% of the responses were detected in the same time interval. In the case of 2000 and 4000 Hz, ASSRs were also detected generally faster for the reference condition, Rooms 2, and 4, as these two rooms also had similar acoustic conditions to the reference for the 2000 and 4000 Hz frequency bands (see Figure 5.1). This indicates a relation between the detection time and the room in which the sound-field ASSR is measured. This is further supported by the fact that the rooms with the poorest performance in terms of detection rate and detection time had substantially different acoustic conditions than the reference, as shown in Figure 5.1.

As for ASSR level, the relations between detection times, EDT and RModP were investigated. The ANOVA summaries in Table 5.3 indicate a similar pattern of results, however, with the detection time results broadly speaking being the mirror images of the ASSR level results: high ASSR level being associated with short detection time and vice versa (results not shown).

5.4 Discussion

5.4.1 Effect of the room on the ASSR level

The present study showed that the ASSR level decreases for the room conditions compared to the anechoic reference, supporting the early work presented in Zapata-Rodríguez et al. (Zapata-Rodríguez et al., 2020a). The results are in agreement with the fact that the ASSR level is known to be negatively affected by the decreased stimulus modulation (Houtgast et al., 1980; Plomp, 1983). A substantial reduction of the ASSR level could have a practical impact on the viability of sound-field ASSR test. For low ASSR levels, longer measurement time would be needed for detecting the ASSR. In a clinical environment, where the testing time is limited, this could lead to misleading results. For instance, in some extreme cases, the acoustics of the room could substantially reduce the ASSR level, making detection difficult during the testing interval, even though the subject could clearly hear the stimulus and/or the hearing aid was fitted correctly.

The results also show a frequency-dependent room effect on the ASSR level, indicating that for some frequency bands the room effect was more pronounced. While in the 707 Hz frequency band, the ASSR level for all rooms considered was significantly different from the reference, in the case of 2000 and 4000 Hz, only four of the rooms differed significantly, as shown in Figure 5.3. Moreover, the ASSR level also varied across room conditions with increasing EDT and decreasing relative modulation power, even though the acoustics of the simulated rooms was inspired by real audiometric testing rooms with relatively short reverberation times (EDT $<$ 0.5 s). The selection of the testing room

will then play an important role in sound-field ASSR measurements in terms of ensuring the reliability and repeatability of hearing aid fitting validations across clinic rooms. Consequently, these results also highlight the need for establishing acoustic standards for such measurements, as previously suggested (Zapata-Rodríguez et al., 2020b). It is noteworthy, however, that in the present study no changes were obtained on the mean ASSR level for the reference condition between testing sessions. The same robustness would be expected when the sound-field ASSR is measured using the same positions in the same room, assuming limited head movements (Narayanan et al., 2019).

5.4.2 Predicting the ASSR level with the EDT or the relative modulation power

Regression models based on the EDT and the relative modulation power were previously proposed to quantify the relation between the testing room and ASSR level (Zapata-Rodríguez et al., 2020b). Here, the proposed models are further investigated with an extended data set of room acoustic conditions. The obtained results showed, as expected, a reduction of the ASSR level when either the EDT increased or the relative modulation power decreased. The corresponding updated regression models indicated that more than 50% of the variability of the ASSR level for the first response harmonic could be explained by either of the two acoustic predictors (EDT or the relative modulation power). Taking the descriptor with the best performance for each frequency could provide a prediction of more than 70% of the variance, that is the EDT for the frequencies of 707 (slope = -17.57 dB/s, R^2 = 0.82) and 2000 Hz (slope = -14.15 dB/s, R^2 = 0.73), and the relative modulation power model for the 4000 Hz (slope = -0.56 dB/dB, R^2 = 0.80).

The updated models presented here are expected to have a better performance than the proposed models presented in (Zapata-Rodríguez et al., 2020a). This is because the tested acoustic conditions presented here were carefully simulated to represent the real audiometric clinic rooms reported in (Zapata-Rodríguez et al., 2020b). Additionally, in the present study a more realistic experimental framework was used to evaluate the effect of the room on the ASSR level. Unlike the monaural auralization approach used in (Zapata-Rodríguez et al., 2020a), here a loudspeaker-based auralization technique was implemented for the reproduction of the simulated clinic rooms. Therefore, sound-field ASSR was measured by presenting the stimulus through a loudspeaker array, including the effect of the local sound field produced by the patient's head and torso. Thus, the updated models proposed here, could play an important role in the clinical implementation of sound-field ASSR measurements since they can serve to determine whether a room (or the position in the room) would be suitable for the test.

5.4.3 Detection rate and detection time in sound-field ASSR measurements

Although the ASSR level is reduced for the room conditions compared to the reference, the ASSR was detected in all rooms for more than 50% of the participants with an analysis time of about 500 s. As in the case of the ASSR level, the detection rate also depended on the room condition and the stimulus frequency band. Two main observations can be made from the analysis of detection times. First, generally shorter detection times were needed for the reference anechoic condition than for the room conditions across frequencies. Second, a large variability of the detection times was also found across room conditions. Whereas some room conditions (e.g., Rooms 2 and 4) followed a similar trend as the reference, the detection time in other rooms (e.g., Rooms 6 and 8) was significantly lower. This analysis was conducted with a q-sample detector which increases the likelihood of the detection, since it uses both ASSR magnitude and phase of the fundamental frequency as well as eleven harmonics of the ASSR spectrum. Therefore, even longer detection times could be expected if the detection is performed with a one-sample detector such as the F-ratio test (Cebulla et al., 2006).

The extended detection time for the room conditions is consistent with the finding that lower ASSR levels were also obtained for the simulated rooms compared to the reference condition, which implies that the ASSR should be recorded for a longer period to reach the signal to noise level required for the detection. Prolonged measurements are not always feasible in clinical settings, and therefore, the sound field presentation could pose a challenge when performing the hearing aid fitting validation. While some clinic rooms could provide "ideal" acoustic conditions for a fast ASSR detection, other rooms could have a higher influence on the stimulus modulation, and hence, compromise the efficiency of the sound-field ASSR measurements.

5.4.4 Acoustic descriptors

The models presented in the current study for predicting the ASSR level in a given room were based on room acoustic descriptors measured with a single microphone located at the center of the head without the subject present. However, it is well known that the sound field in a room varies across positions and therefore, the acoustic parameters measured at the ear positions will be different to some extent. The work presented in Zapata-Rodríguez et al. (2020b), highlighted the variability of the room acoustic parameters across positions in the clinical audiometric testing rooms, which can be problematic for sound field testing. This variability is considered in the standard for sound field audiometry, which establishes a calibration procedure for sound field audiometry based on the sound pressure level differences between the measurement point and 6 additional positions arranged around this point 0.15 m away. Although for practical reasons, it would be more convenient to evaluate the clinic rooms based on acoustic measurements in a single point, it might be relevant to further investigate whether room acoustic de-

scriptors measured at the ear positions would add substantial value to the predictive model for sound-field ASSR.

5.5 Conclusion

This study investigated the effect of room acoustics on sound-field ASSR measurements. The study expanded the work presented in Zapata-Rodríguez et al. (2020a) by including a more accurate room acoustic model, as well as a more realistic reproduction technique, and a larger number of realistic room acoustic conditions. The main findings were as follows:

- A frequency-dependent room effect on the ASSR level was found. The ASSR level is reduced with increasing early decay time and decreasing the relative modulation power.
- The effect of the room on the ASSR level was characterized in terms of the EDT and the relative modulation power. Both parameters were found to be good predictors of the level of the ASSR fundamental frequency. The EDT performed particularly well for 707 and 2000 Hz, whereas the relative modulation power was better for 4000 Hz.
- The reduced ASSR level for the room conditions was also reflected in reduced detection rates and extended detection times obtained for the room conditions compared to the reference.

  Overall, this study showed that although a reduction in the ASSR level is expected for sound-field ASSR measurements, it is still possible to measure sound-field ASSR in standard audiometric testing rooms. Nevertheless, special care should be taken when selecting the test room since a substantial reduction in ASSR level (produced by the room acoustics) could lead to unpractical detection times. Quantitative models, based on the EDT and the relative modulation power, such as those proposed here can help in determining whether a given room and/or a specific set of source and receiver positions could be appropriate for conducting sound-field ASSR measurements.

Acknowledgments
The authors would like to thank Dana Gilbert and Henrik Hvidberg for valuable technical assistance.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
This research was funded by Interacoustics A/S, the William Demant Foundation, and the Innovation Fund Denmark, Grant. No. 5189-00070B.

5.6 Supplementary Material

Table 5.4: Summary of the statistical pairwise comparisons between tested acoustic conditions for the 707 Hz stimulus.

Acoustic conditions comparison	Estimate	SE	df	t ratio	p value
REF – 8	10.20638068	0.6408261	1087.73	15.927	<.0001
REF – 5	6.03698177	0.6255786	1088.15	9.65	<.0001
REF – 6	5.82595943	0.5944611	1088.56	9.8	<.0001
REF – 7	5.80810019	0.5847863	1088.42	9.932	<.0001
REF – 3	3.89506995	0.6090327	1088.57	6.396	<.0001
REF – 4	3.65598765	0.5938842	1087.9	6.156	<.0001
REF – 1	3.54559555	0.5975473	1087.86	5.934	<.0001
REF – 2	3.00670754	0.6246676	1087.07	4.813	0.0001
5 – 8	4.16939891	0.691387	1086.43	6.03	<.0001
6 – 8	4.38042125	0.6647946	1087.8	6.589	<.0001
7 – 8	4.3982805	0.6560729	1087.61	6.704	<.0001
3 – 8	6.31131073	0.6756703	1086.22	9.341	<.0001
3 – 5	2.14191182	0.6610312	1085.51	3.24	0.0335
4 – 8	6.55039304	0.6640419	1087.09	9.864	<.0001
4 – 5	2.38099413	0.6502065	1088.07	3.662	0.008
4 – 6	2.16997179	0.6179838	1085.85	3.511	0.0137
4 – 7	2.15211254	0.6086047	1085.7	3.536	0.0126
1 – 8	6.66078513	0.6669677	1086.94	9.987	<.0001
1 – 5	2.49138622	0.6535053	1088.05	3.812	0.0046
1 – 6	2.28036388	0.6214876	1086.1	3.669	0.0078
1 – 7	2.26250464	0.6121349	1085.93	3.696	0.0071
2 – 8	7.19967315	0.6905759	1085.5	10.426	<.0001
2 – 5	3.03027424	0.6768351	1086.01	4.477	0.0003
2 – 6	2.8192519	0.6504678	1087.85	4.334	0.0005
2 – 7	2.80139265	0.6412906	1087.54	4.368	0.0005

Table 5.5: Summary of the statistical pairwise comparisons between tested acoustic conditions for the 2000 Hz stimulus.

Acoustic conditions comparison	Estimate	SE	df	t ratio	p value
REF – 8	4.60709454	0.6546194	1087.49	7.038	<.0001
REF – 5	4.26162562	0.7015658	1088.2	6.074	<.0001
REF – 6	3.94099842	0.6547805	1087.47	6.019	<.0001
REF – 7	2.71776458	0.6704077	1088.48	4.054	0.0018
3 – 8	2.97887692	0.6801665	1088.04	4.38	0.0004
3 – 5	2.63340799	0.7248144	1088.3	3.633	0.0089
3 – 6	2.31278079	0.6777483	1086.13	3.412	0.0192
4 – 8	3.18769689	0.673546	1086.15	4.733	0.0001
4 – 5	2.84222796	0.7237154	1090.9	3.927	0.0029
4 – 6	2.52160076	0.6764913	1088.5	3.727	0.0063
1 – 8	4.23642296	0.6856304	1085.7	6.179	<.0001
1 – 5	3.89095403	0.7353509	1090.51	5.291	<.0001
1 – 6	3.57032683	0.6890532	1088.15	5.181	<.0001
1 – 7	2.347093	0.7002639	1086.37	3.352	0.0234
2 – 8	4.31412256	0.6326601	1087.72	6.819	<.0001
2 – 5	3.96865363	0.684067	1091.3	5.802	<.0001
2 – 6	3.64802644	0.6328105	1088.02	5.765	<.0001
2 – 7	2.4247926	0.6487212	1088.61	3.738	0.0061

Table 5.6: Summary of the statistical pairwise comparisons between tested acoustic conditions for the 4000 Hz stimulus.

Acoustic conditions comparison	Estimate	SE	df	t ratio	p value
REF – 8	4.59338706	0.7327462	1088.33	6.269	<.0001
REF – 5	4.57183322	0.6938364	1088.5	6.589	<.0001
REF – 6	4.23305622	0.7374967	1086.74	5.74	<.0001
REF – 7	3.32972416	0.7003019	1086.42	4.755	0.0001
3 – 8	2.73035051	0.7365183	1087.86	3.707	0.0068
3 – 5	2.70879667	0.6976852	1087.84	3.883	0.0035
3 – 6	2.37001966	0.7440063	1088.2	3.185	0.0396
4 – 8	3.20954603	0.736337	1088.28	4.359	0.0005
4 – 5	3.18799219	0.6986739	1089.41	4.563	0.0002
4 – 6	2.84921519	0.7414751	1086.84	3.843	0.0041
1 – 8	4.24922849	0.7079588	1088.69	6.002	<.0001
1 – 5	4.22767466	0.6654762	1086.75	6.353	<.0001
1 – 6	3.88889765	0.7149347	1088.44	5.44	<.0001
1 – 7	2.98556559	0.6767268	1088.56	4.412	0.0004
2 – 8	4.50642001	0.6851804	1089.95	6.577	<.0001
2 – 5	4.48486617	0.642562	1089.94	6.98	<.0001
2 – 6	4.14608917	0.691561	1088.92	5.995	<.0001
2 – 7	3.24275711	0.6506621	1088.43	4.984	<.0001

CHAPTER 6

General discussion

This PhD project investigated the effect of room acoustics on sound-field ASSR measurements with a view to the potential clinical implementation of the test for the objective hearing aid fitting validation in infants and hard-to-test adults. A thorough analysis of the acoustic conditions of audiometric clinic rooms was also carried out to gain insight into the variability of the sound field across and within the rooms. Moreover, an experimental framework for electrophysiological measurements in virtual sound environments was devised and validated for the investigation of sound-field ASSR in the laboratory. The key points of consideration were to obtain accurate room acoustic simulations and to reproduce nearby sources within a loudspeaker-based reproduction system. The analysis of the clinic rooms, as well as the development of the experimental framework for the reproduction of nearby sources allowed a systematic investigation of different acoustic environments based on realistic simulations of audiometric testing rooms. The thesis also describes the role of the loudspeaker presentation in sound-field ASSR measurements and discussed its potential limitations for the clinical implementation. Additionally, the relation between the ASSR and objective acoustic parameters was investigated to quantify the effect of the room. In the following, the main findings of the project are summarized and discussed.

6.1 Summary

The study presented in *Chapter 2* served as the first step into the investigation of the room's effect on sound-field ASSR measurements. The main results showed that reverberation, indeed, has a detrimental effect on the ASSR level. This effect was observed for the fundamental frequency and harmonics of the ASSR spectrum. The decreased ASSR levels observed for the room conditions were produced by the reduction in the stimulus modulation due to the acoustical properties of the simulated rooms. This is consistent with the previous literature that reported a reduction in the ASSR level as the modulation depth decreases (e.g., Rees et al., 1986; Kuwada et al., 1986; Picton et al., 1987). This study also presented two models to quantify the effect of the room on the ASSR level, based on measurable room acoustic parameters. The first model used the early decay time as an independent variable. This parameter is related to the early reflections in a room, which has been shown to influence the changes in the stimulus modulation in the room. The second model was defined based on an auditory-inspired relative modulation power model. This model was implemented to determine the variation in the stimulus modulation relative to the anechoic reference stimulus and was evaluated in terms of the efficiency in evoking the ASSR. The results showed that ASSR

level reduces either as the EDT increases or as the relative modulation power decreases. Even though only three acoustic conditions were tested in the study presented in *Chapter 2*, a strong correlation was found between the ASSR level and both acoustics parameters. Thus, it was expected that these measurements could be potentially used for predicting the ASSR levels measured in the clinic rooms.

The first study evaluated the effect of the room using a monaural auralization approach with single point simulated impulse responses. These were simulated based on rooms available at the Technical University of Denmark which were expected to have similar acoustic conditions to standard audiometric testing rooms found in the clinics. However, defining the acoustic characteristics of audiometric testing rooms is a challenging task since no specific acoustic regulations for this type of rooms currently exist. The study presented in *Chapter 3* investigated this problem by an extensive evaluation of 31 audiometric rooms in clinical use located in Denmark and Germany. In general, the results showed a great variability in the acoustic conditions across rooms, suggesting the need for determining acoustic standards according to the acoustic requirements necessary for conducting sound field tests in clinical audiometric rooms. Moreover, the study also analyzed the level differences in ASSRs measured via insert earphones and in a sound field. The findings showed that the recorded ASSR level was significantly lower for the sound field presentation compared to the measurements via insert earphones, even though the room was adapted to have similar acoustic conditions to the real audiometric clinic rooms. This study provided a clearer overview of the existing acoustic environments in the clinics where sound-field audiometry testing is conducted.

The detrimental effect of the room on the ASSR level was reported in *Chapter 2*, demonstrating the importance of considering the room acoustics in sound-field ASSR recording. However, this study constituted only a first approximation to the phenomenon, since the monaural room acoustic model implemented did not include the diffraction caused by the torso, head and pinna, present in real environments. Moreover, the measurements were conducted by presenting the stimuli via insert earphones, without including the natural head movements into the simulation, which limited the generalization of the findings. Therefore, to further investigate the room effect with a more realistic approach, an experimental framework was developed and validated using virtual sound environments created with loudspeaker-based reproduction systems. The main objective of this experimental framework was to develop a dynamic system that could be used to accurately represent the acoustic scenarios in which sound-field ASSR measurements were expected to be carried out. On this basis, small rooms and a source-to-receiver distance of 1 m were considered in the simulation. Careful considerations were then taken for the selection of the room acoustic model as well as the reproduction system. The rooms were simulated using the simulation tool PARISIM (Phased Acoustical Radiosity and Image Source Method; Marbjerg et al., 2015) and the nearest loudspeaker method (Seeber et al., 2010), which were previously reported to have the lowest errors compared to other auralization methods Marbjerg et al., 2020. Special effort was put into the reproduction of a relatively short source-to-receiver distance of 1 m, which is used for sound-field ASSR as well as standard sound-field audiometry. The

reproduction of nearby sources can be problematic in loudspeaker-based reproduction systems, as the methods used for the reproduction generally assume that the listener is located in the far-field of the simulated sources. This problem was addressed in the study presented in *Chapter 4*, where the inclusion of an additional loudspeaker located at the exact distance used in the simulation was proposed. The results showed that a better energy balance between early and late reflection is obtained with the proposed method, which translates into a more accurate estimation of the EDT. This was particularly relevant for the project since EDT was one of the acoustic descriptors proposed for the estimation of the ASSR levels in any given room.

Finally, in *Chapter 5*, a more systematic investigation of the room's effect on sound field ASSR measurements was conducted using the methodology proposed in *Chapter 4*. This approach allowed ASSR to be measured under much more realistic acoustic conditions in the laboratory. In this study, eight room acoustic conditions were investigated. These were simulated based on the findings of the second study, which were important to define the range of existing acoustic conditions for evaluating the real audiometric testing rooms. The results of the study presented in *Chapter 5* confirmed the reduction of the ASSR level for the room conditions compared to the anechoic reference conditions observed in the first study. Moreover, the investigation further revealed other consequences of room acoustics, demonstrating that the reduction of the ASSR level affected the detection time and detection rates of the sound-field ASSR measurements in clinical conditions. Interestingly, the study showed substantial differences in the ASSR levels obtained across room conditions, as a result of the changes in the stimulus modulation across rooms. The initially proposed models for the prediction of the ASSR level based on the EDT and relative modulation power were also updated in the study presented in *Chapter 5*, and the results also supports the use of both parameters as predictors of the ASSR level measured in a room.

6.2 Room acoutsical effect on sound-field ASSR in the clinics

Throughout this project, it was demonstrated that the room has a frequency-dependent effect on the ASSR level. A systematic reduction of the ASSR level was observed across room acoustic conditions relative to the anechoic reference. Despite the level reduction, detectable ASSRs were recorded in all the room conditions evaluated in the project, which were representative of standard audiometric rooms. The reduction in the ASSR level, however, led to overall longer detection times. For the clinical implementation of sound-field ASSRs, the measurement time is a key metric for several reasons. First, in clinical applications, fast measurements are always desirable,. Thus, a simpler and faster HA fitting validation protocol is expected to be more rapidly adopted in clinic environments. Second, for ASSR measurements (and electrophysiological measurements in general), the patient should be relaxed during the test to avoid high electrophysiologi-

cal noise that can occlude the ASSR, and hence, long measurement sessions are avoided as much as possible. Third, the detection rate of ASSRs can also be affected by the reduction in the ASSR level in spite of extended detection times. If the acoustics of a room reduces extensively the ASSR level, a longer detection time will be required to reach the signal-to-noise-ratio needed for detection. This can increase the probability of missed responses within a given testing interval in the clinics, potentially resulting in misdiagnosis in relation to the validation of the hearing aid fitting. Thus, special care should be taken when selecting the test room to avoid significant reductions in ASSR level due to room acoustics.

The prediction models for the ASSR level in a room based on the EDT and relative modulation power are a valuable contribution to the field of technical audiology, as they could be used to evaluate the suitability of any given room for sound-field ASSR measurements. These models can also be used to determine the most suitable position within a room for sound field-ASSR recording. This is an important factor because the EDT was found to vary substantially across space in the real audiometric rooms. The changes in the local sound field could also have an implication on the clinical viability of the test due to head movements, which are expected to occur during the measurement. An early study investigated the potential influence of changes in the head orientation on sound field ASSR measurements and reported a minimal effect for dynamic head movements, which were estimated as the combined measurement of three different head orientations with the same duration each. However, the effect became significant for abrupt head orientations during most of the measurement for 4 kHz, which could be attributed to the head-shadow effect Narayanan et al., 2019. Consequently, to minimize potential detrimental effects produced by changes in head orientation, it would be important to determine the listener position/confined area in the room where the sound field is most uniform. The proposed prediction models are also expected to provide benefits in this regard. Thus, the models could pave the way for the clinical implementation of sound-field ASSR for hearing aid fitting validation and could be used as a part of the calibration protocol for sound field ASSR measurements.

The large variability of the acoustic conditions across existing state-of-the-art audiometric testing rooms (observed in the study presented in *Chapter 3*), indirectly represents a potential limitation of the test, as this could affect the reliability and repeatability of sound-field ASSRs across clinic rooms. It is, therefore, important to establish acoustic guidelines specialized for audiometric testing rooms, addressing all the different sound-field tests performed in these rooms. For sound field ASSR measurements, the proposed model could help lay the foundations for the development of the acoustic recommendations based on EDT and the relative modulation power, which were shown to correlate well with the ASSR measurement in the room.

6.3 Outlook and future work

In conclusion, the results presented throughout this thesis indicate that sound-field ASSR can be recorded in the average audiometric testing rooms existing in the clinics. However, the loudspeaker presentation will in fact reduce the ASSR level due to the changes in the stimulus modulation produced by the reverberation in the room. Therefore, it is important to select an adequate room and measurement positions to avoid substantial reductions of the ASSR level, which could lead to unacceptable long detection times for clinical testing. These decisions could be made using the outcomes of the proposed quantitative models based on EDT and the relative modulation power, which estimated the systematic reduction of the ASSR level as a function of the room conditions. Further validation of the models would be valuable in order to verify that the estimated ASSR levels are representative of real sound-field ASSR measurements. To do so, the sound-field ASSR recorded in the real rooms could be compared with the simulated and auralized virtual acoustic version of the rooms, to determine whether the differences between virtual and real rooms are significant.

The scope of the present project was to investigate only the effect of the room on sound-field ASSR measurements, therefore, all experiments described in this project were conducted in young adult normal-hearing test subjects. Further work may be focused on validating the reported room effect for sound field ASSR measurements in hearing aid fitting validation, including testing scenarios in which the infants wear hearing aids. In the case of infants, the effect of the room is expected to be shifted towards higher frequencies, since infant's heads have small dimensions. Moreover, it is important to investigate any potential interaction between the hearing aid processing and the room. The non-linear amplification strategies of modern hearing aids (mainly dynamic range compression) could have a detrimental effect on the ASSR stimulus. Therefore, in sound field presentation, the stimulus modulation is expected to be affected by both the hearing aid compression and the reverberation of the room. However, since the hearing aid compression acts on the signal already affected by reverberation, the room is then expected to have a larger effect than the hearing aid on the stimulus modulation, and hence on the ASSR measurement.

Appendices

APPENDIX A

Comparing loudspeaker array reproduction techniques using a phased combination of the image source method and acoustical radiosity

Comparing loudspeaker array reproduction techniques using a phased combination of the image source method and acoustical radiosity

Gerd Marbjerg, Jonas Brunskog,[a] and Cheol-Ho Jeong
Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Valentina Zapata-Rodriguez
Interacoustics Research Unit, 2800 Kgs. Lyngby, Denmark.

For hearing and audiological research, auralisations from room acoustic simulations of small rooms can be useful. Small rooms can be simulated with phased geometrical acoustic methods. This can, e.g., be with the tool PARISM (Phased Acoustical Radiosity and the Image Source Method). To reproduce the sound field simulated by PARISM with a loudspeaker array, three widely used reproduction methods have been compared numerically: higher order Ambisonics, vector-base amplitude panning and nearest loudspeaker panning. The implementation of these methods are described. For simulations of a small rectangular room, conventional objective room acoustic parameters are used as performance indicators. For the studied parameters, the method of panning to the nearest loudspeaker seems most robust with a dense array of 64 loudspeakers.

PACS numbers:

I. INTRODUCTION

With recent technological advances in computational acoustics and 3D sound reproduction techniques, sound field reproduction has become more physically accurate and perceptually authentic with many engineering applications, e.g. audiological research, hearing research and building design. Many current listening tests in audiological and hearing research are however still conducted with simplified sound environments that cannot represent realistic situations. Creating sound environments with room acoustical simulations is a flexible method to create more ecologically valid tests, while still maintaining a high level of control of the conditions. For such experiments, it is often of interest to include hearing devices, which is most conveniently done with loudspeaker auralisations.

When choosing a method for simulating room acoustics, it is important to consider the room type that is to be modelled, because some methods rely on assumptions that are only valid for certain room types. Currently there are methods such as ODEON[1] with the LoRA toolbox[2] and MCRoomSim[3] available for creating loudspeaker-based auralisations with higher order Ambisonics. They belong to the family of geometrical acoustic methods. In this family there are methods which include phase information, but ODEON and MCRoomSim do not fully do so. ODEON is fully energy-based and is therefore best suited for simulating the acoustics of larger spaces with high sound scattering. MCRoomSim includes the propagation phase in specular reflections, but cannot model phase shifts on reflections. It is furthermore limited to only modelling shoebox-shaped rooms. Phased geometrical acoustical methods include phased beam tracing[4,5], phased image source method[6,7] and PARISM (Phased Acoustical Radiosity and the Image Source Method).[8] PARISM is so far different from the other methods[4–7] in that it can include scattering. PARISM is a method that is especially developed for simulating the acoustics of small rooms with absorbing surfaces,[9] e.g. classrooms and clinic rooms. It includes a phased implementation of the image source method and the ability to include sound scattering with acoustical radiosity. The present work concerns which reproduction technique performs best together with PARISM for a small room with absorbing surfaces in terms of acoustical parameters. Three different reproduction techniques are tested: panning with nearest loudspeaker (NL), vector-base amplitude panning (VBAP) and higher order Ambisonics (HOA).

Early decay time EDT, reverberation time T_{20} and sound strength G are used to check that the overall acoustic properties of the simulated room are maintained after the processing with the reproduction techniques. Speech intelligibility is often of high interest in hearing research, so the metrics deutlichkeit D_{50} and speech transmission index STI are also regarded. Finally, interaural cross correlation IACC is regarded, because the spatial impression is of interest for auralisations. This set of investigated parameters is assumed to cover the most important aspect of the sound field for hearing research.

The parameters for the reproduced sound field are determined numerically, so no measurements are included. This is done to isolate possible errors of the reproduction that will be transferred to the proceeding steps. The errors here are thus conservative and do not include measurement errors, e.g. due to loudspeaker responses and reflections from the physical setup of a loudspeaker array, and reflections from the room in which the loudspeakers are placed.

The outline of this paper is that a general introduction to the implementation of loudspeaker auralisation reproduction with PARISM is given, then the implementations of three reproduction methods with PARISM are outlined, and lastly these reproduction techniques are compared in terms of the mentioned parameters for a small rectangular room with absorbing surfaces.

[a]Electronic address: jbr@elektro.dtu.dk

II. LOUDSPEAKER AURALISATIONS FROM PARISM

PARISM combines the image source method (ISM) with acoustical radiosity (AR) with the phase information included in the image source method,[8] and it is an extension of the unphased CARISM.[10] The ISM models the specular reflections and AR models the scattered reflections. Image sources are terminated adaptively by comparing the acoustic energy in the specularly reflected sound field and the scattered sound field. To create loudspeaker auralisations, the impulse responses for the two parts of PARISM, ISM and AR, are found separately and then summed. The response, $S_i(t)$, for loudspeaker channel i is thus found as

$$S_i(t) = S_{i,ISM}(t) + S_{i,AR}(t). \tag{1}$$

A. The image source method (ISM)

The basic assumption of the image source method (ISM) is that a reflection of a source by a surface can be described by mirroring the source in the surface.[11] In PARISM, the ISM is implemented in the frequency domain, so a contribution $P_q(f)$ is added to the frequency response for each image source q. This contribution is panned to a loudspeaker array with gain factor, $g_{q,i}$ for each loudspeaker i in the array that can be determined by HOA, VBAP or NL. The resulting frequency response for loudspeaker i is thus found as the sum of the contributions from the N_q image sources, and the impulse response for each loudspeaker, $S_{i,ISM}(t)$, is found by inverse Fourier transform.

$$P_{i,ISM}(f) = \sum_{q=1}^{N_q} g_{q,i}\, P_q(f), \;\; S_{i,ISM}(t) = \mathscr{F}^{-1}\left[P_{i,ISM}(f)\right]. \tag{2}$$

The gain factors are normalised such that $\sum_i g_{q,i} = 1$.

B. Acoustical radiosity (AR)

AR is an energy-based method in which surface elements exchange acoustic energy.[8,12] These energy contributions are collected at the receiver, which results in energy impulse responses in octave bands. For implementation of the reproduction techniques, an energy impulse response at the receiver due to each AR surface element is determined in each octave band. For element j and octave band k this is denoted $w_{k,j}(t)$. A pressure impulse response at the receiver due to each element j is reconstructed by

$$p_j(t) = \sum_{k=1}^{N_k} \left[\sqrt{\rho\, c^2\, w_{k,j}(t)}\, q_j(t) * h_k(t) \right], \tag{3}$$

where N_k is the number of octave bands, $h_k(t)$ is the impulse response of the octave band filter k, ρ is the density of air, and $q_j(t)$ is a Gaussian noise signal for element j with zero mean and unit variance. Uncorrelated noise signals for each element are used. Due to the randomness in $q_j(t)$, repeating the realisation of Eq. (3) will give slightly different results each time. The impulse response from AR of loudspeaker i is found by

$$S_{i,AR}(t) = \sum_{j=1}^{N_{elem}} g_{j,i}\; p_j(t), \tag{4}$$

where N_{elem} is the number of surface elements, $g_{j,i}$ is the gain factor for element j and loudspeaker channel i. The gain factors are normalised such that $\sum_i g_{j,i} = 1$.

III. REPRODUCTION TECHNIQUES

A. Higher order Ambisonics

The procedure described here is heavily inspired by the method of the LoRA toolbox[2] for auralisations from the room acoustical simulation tool ODEON[1] using higher order Ambisonics (HOA). Ambisonic auralisation is based on decomposing the sound field into spherical harmonic components. The decomposed sound field is then encoded to obtain ambisonic signals for the loudspeakers in the array. The encoding and decoding is done in a single step, and the gain factors in Eqs. (2) and (4) are defined from the encoding and decoding.[2]

A known issue for HOA is that the reconstruction of the sound field is theoretically only valid for the centre of the array.[2] Positions outside of the centre will have errors that depend on the distance from the centre and on the frequency. The error of not listening exactly in the centre increases with frequency. A frequency limit at which the reproduction error on the pressure is around 4 %[13] can be found as[2]

$$f_{lim} = (Mc)/(2\pi r), \tag{5}$$

where M is the ambisonic order, c is the speed of sound in air and r is the radius from the centre within which the receiver is assumed to be located. Therefore, another decoding called *max* r_E has been suggested, which is a method that attempts to preserve the total energy within the radius r.[2] Two version of HOA are implemented here. One where basic decoding is used for both ISM and AR in the entire frequency range, and one where $maxr_E$-decoding is applied to the ISM part above the defined frequency limit. Only basic decoding is used for AR due to the diffuseness of this part of the sound field. This is similar to the procedure in the LoRA toolbox,[2] where basic encoding is used for the late reflections.

For "max r_E"-decoding the frequency response for loudspeaker i must thus be determined in two parts that are high and low pass filtered respectively. The total ISM frequency response is thus found as

$$P_{i,ISM} = \sum_{q=1}^{N_q} P_q(f) \left[g_{q,i,l}\, H_l(f) + g_{q,i,h}\, H_h(f) \right], \tag{6}$$

where H_l is a low pass filter with a cut-off frequency corresponding to the limiting frequency, H_h is a high pass filter with the same cut-off frequency, $g_{q,i,l}$ is the gain factor corresponding to basic decoding and $g_{q,i,h}$ is the gain factor corresponding to "max r_E"-decoding. In the rest of this note basic HOA refers to higher order Ambisonics with basic decoding for the full frequency range and HOA "max r_E" refers to higher order Ambisonics with "max r_E"-decoding for high frequency ISM.

B. Vector-base amplitude panning

Pulkki[14] has developed a method of panning the incoming sound to the nearest loudspeakers by the method he calls vector-base amplitude panning. When this method is applied in 3D, the loudspeaker array is divided into bases consisting of 3 loudspeaker. For each reflection, the base nearest to the direction of the reflection is found and is referred to as the active triangle. In PARISM, a reflection can be either the contribution from an image source or the contribution from an AR surface element. The reflection is panned to the 3 loudspeakers in the active triangle with according gain factors. Therefore, the gain factors in Eqs. (2) and (4) have three non-zero values.

C. Nearest loudspeaker

The simplest method of panning the impulse response to the loudspeaker array is to choose the loudspeakers nearest to the directions of the reflections in the impulse response. For the image source method part of PARISM this means that a loudspeaker is picked for each image source, and for AR a loudspeaker is picked for each surface element. The gain factors in Eqs. (2) and (4) are thus 1 for the active loudspeaker and zero for all others.

IV. METHOD OF VALIDATION

The test case is a rectangular room. It is based on an existing room at the Technical University of Denmark (DTU), but no measurements from the room are included in this work. Two walls of the room are absorbing and two diffusers are placed on the floor. Note that the somewhat unusual surface arrangement of the room is due to it being designed for a listening experiment, not reported here. Fig. 1 shows a sketch of the room, where the two grey walls are those that are absorbing and the two grey areas on the floor are where the diffusers are placed. The volume of the room is 42.8 m^3 and the Schroeder frequency of the room is around 200 Hz.

The absorbing characteristics of the floor, the ceiling, the hard walls and the diffusers are described by absorption coefficients, whereas complex reflection factors are used for the two absorbing walls. The material characteristic impedance of the absorbers is calculated using Miki's model[15] for a porous absorber with a flow resistivity of 47.7 $\frac{kPa\ s}{m^2}$. The reflection factor is then found by modelling the reflection from a porous absorber with a thickness of 4 cm on rigid backing.[16] Scattering coefficients describe the scattering characteristics of the surfaces. The values for the absorption and scattering coefficients are chosen based on realistic values for the materials in the physical room at DTU, and they can be found in Table I.

The sampling frequency of the simulations is 23 kHz in order to include frequencies up to the octave band with 8 kHz as centre frequency.

The impulse response of the room is simulated with PARISM for six source-receiver combinations, see Table II. The spatial mean and standard deviation of the monaural room acoustical parameters reverberation time T_{20}, early decay time EDT, deutlichkeit D_{50} and sound strength G, as defined in ISO 3382-1,[17] are shown in Fig. 2. The speech transmission index[18] (STI) was also calculated. It is a single-number rating of the speech intelligibility, calculated from modulation transfer functions (MTFs) with different modulation and carrier frequencies. The carrier signals used here are octave bandpass filtered noise signals, and the weights applied to the octave bands are those defined by Houtgast and Steeneken[19] that are not gender specific. The mean STI is 0.728 with a standard deviation of 0.0096.

The binaural parameter interaural cross correlation[17] (IACC) is used to evaluate the spatial characteristics of the reproduced sound field. IACC calculated from the first 80 ms of an impulse response is believed to be related to the appar-

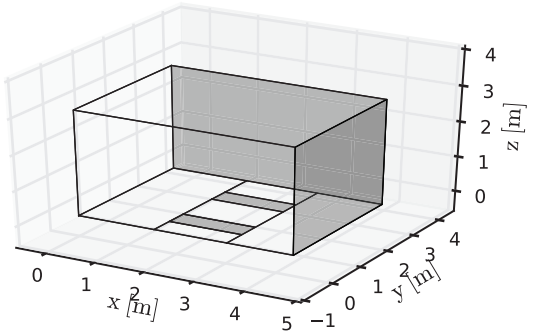

FIG. 1. The room (dimensions: 4.38 x 3.29 x 2.97 m) used for the simulations. The two grey walls are the absorbing walls and the grey areas of the floor are where the scattering is increased to represent diffusers.

TABLE I. The input data for the simulations. α_{fw} is the absorption coefficient of the non-absorbing walls and the floor, α_{ceil} is the absorption coefficient of the ceiling, α_{dif} absorption coefficient of the diffusers, s_{fwc} is the scattering coefficient of the walls, s_{w2} is the scattering coefficient of the left and right walls on Fig. 1, and s_{dif} is the scattering coefficient of the diffusers.

f [Hz]	63	125	250	500	1 k	2 k	4 k	8 k
α_{fw}	0.021	0.025	0.034	0.035	0.035	0.057	0.068	0.070
α_{ceil}	0.23	0.18	0.11	0.071	0.070	0.070	0.070	0.070
α_{dif}	0.090	0.17	0.21	0.64	0.81	0.70	0.46	0.44
s_{fwc}	0.021	0.030	0.039	0.040	0.050	0.077	0.080	0.080
s_{w2}	0.031	0.041	0.040	0.040	0.060	0.080	0.090	0.110
s_{dif}	0.025	0.089	0.31	0.78	0.99	0.94	0.76	0.74

TABLE II. The source and receiver positions

	Source			Receiver		
	x [m]	y [m]	z [m]	x [m]	y [m]	z [m]
Pair 1	2.8	1.5	1.5	1.8	1.5	1.2
Pair 2	3.4	1.3	1.5	1.4	2.4	1.2
Pair 3	2.6	1.1	1.5	3.3	2.3	1.2
Pair 4	1.5	2.1	1.5	2.7	1.1	1.2
Pair 5	1.5	2.1	1.5	1.1	1.0	1.2
Pair 6	2.6	1.1	1.5	1.2	1.1	1.2

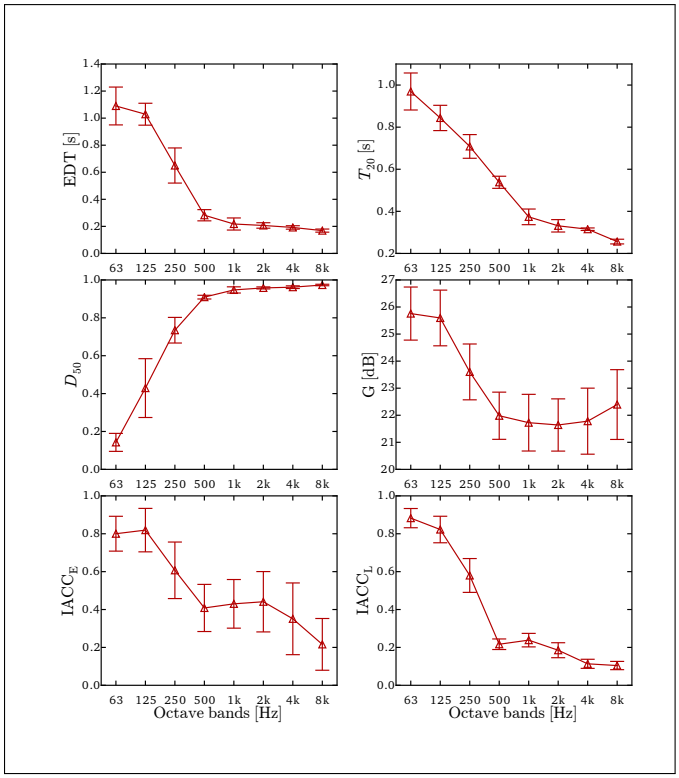

FIG. 2. The mean and standard deviation of T_{20}, EDT, D_{50}, G, IACC_E and IACC_L of the single-point PARISM simulations in the six source-receiver combinations.

ent source width and IACC calculated from the part of the impulse response after 80 ms is believed to be related to the listener envelopment.[20] They will hereafter be referred to as early IACC (IACC_E) and late IACC (IACC_L), respectively. The head related impulse responses (HRIRs) used in the IACC calculations are taken from the CIPIC HRTF database.[21]

Via three reproduction methods, the impulse responses at the center of the array are simulated by summing the 64 channel loudspeaker impulse responses, and the same acoustic parameters are calculated. Taking the sum of the loudspeaker impulse responses is an idealised representation of a measurement with an omnidirectional microphone placed exactly in the centre of the array, assuming that the array is placed in an anechoic room, and that the loudspeakers and the microphone have flat frequency responses. Two further receiver positions in the loudspeaker array are considered, and all three are illustrated in Fig. 3, where it is seen that position 2 roughly corresponds to the location of the right ear of a listener with their head in the centre of the array. To obtain the impulse responses that correspond to measuring in positions 2 and 3, delays and attenuation factors are calculated with the loudspeakers modelled as point sources to take into account that the distances to the receiver from the loudspeakers are no longer the same as they are for the center of the loudspeaker array. These delays are applied to their corresponding loudspeaker impulse responses before the summation.

The regarded loudspeaker array corresponds to a physical setup at the Technical University of Denmark and is known as AVIL (Audio Visual Immersion Lab), see Fig. 4. It consists of 64 loudspeakers placed on a sphere with a radius of 2.4

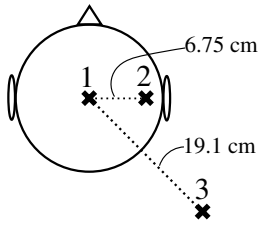

FIG. 3. Illustration of the simulated receiver positions, relative to a head placed in the center of the loudspeaker array. Position 1 is thus the centre of the array.

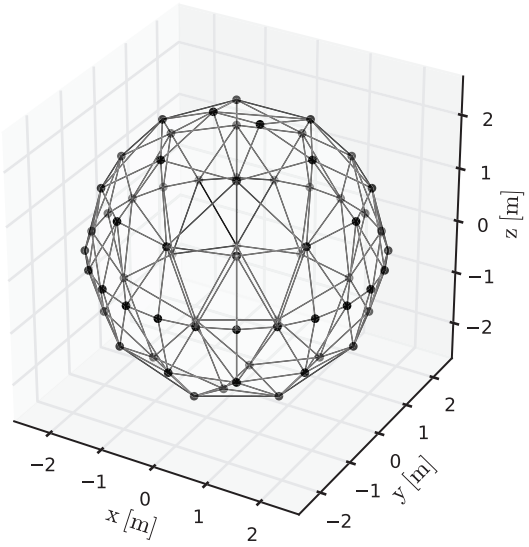

FIG. 4. The array of 64 loudspeakers

m. It has been chosen not to use mixed order for the present investigation, and the highest single Ambisonics order for the setup is five. This is used in Eq. (6) along with a radius of r = 0.1 m, which roughly corresponds to the radius of a human head. The cut-off frequency between basic HOA decoding and $max - r_E$-decoding is then chosen to be 2828 Hz, which corresponds to the upper cut-off frequency of the 2 kHz octave band.[2]

V. RESULTS

The errors of EDT, T_{20}, D_{50} and G of the processed loudspeaker array responses with reference to the single-point PARISM response are calculated for the six simulated source-receiver combinations, and the spatial mean and standard deviations are shown in Figs. 5 to 7 for the four different reproduction methods; HOA max r_E, basic HOA, VBAP and NLS. For EDT and T_{20} the relative errors are given, whereas the absolute errors are given for D_{50} and G due to how their just noticeable differences (JNDs) are defined. An absolute error

is defined as

$$\Delta X = X_{LS} - X_{PARIS\,M}, \tag{7}$$

where X_{LS} refers to a parameter calculated from the loudspeaker array response and $X_{PARIS\,M}$ refers to a parameter calculated from the single-point PARISM response. A relative error is defined as

$$\delta X = \frac{X_{LS} - X_{PARIS\,M}}{X_{PARIS\,M}}\ 100\%, \tag{8}$$

where the definitions are the same as for Eq. (7). The grey areas indicate the JNDs as defined in ISO 3382-1.[17]

HOA "max r_E" has large errors at high frequencies in the centre position of the array, see Fig. 5. G is overestimated and T_{20} and EDT are underestimated. ΔG is similar to what was seen for the validation of LoRA toolbox[2] and is due to the "max r_E"-decoding. The large errors in the reverberation time are possibly due to the fact that the "max r_E"-decoding is only applied to the ISM part of the impulse response, which is dominant in the early part of the response. The "max r_E"-decoding increases the high frequency energy, but in the present implementation it does so only for the early part of the impulse response where ISM is dominant. This will increase the slope of the total decay, thus decreasing the reverberation time. The errors for the basic HOA, VBAP and NLS are close to zero in the centre position. In positions 2 and 3, see Figs. 6 and 7, there are negative errors on G at high frequencies for basic HOA and VBAP. In position 2, there are errors on the "max r_E" HOA parameters, so for this position the error introduced by the "max r_E"-decoding is larger than the error of HOA. For position 3, the error on G is smaller for HOA "max r_E" than basic HOA, and "max r_E"-decoding must be said to improve the reproduction for this off-centre position in the array. Overall, the errors on these monaural parameters are smallest for the NL reproduction.

The means and standard deviations of the errors of the STI of the reproductions are shown in Fig. 8, where all errors are within one JND. Generally, the HOA reproductions overestimate the STI slightly for positions 1 and 2.

STI is a single-number parameter that is derived from modulation transfer functions (MTFs). For speech, the MTFs for the 500 Hz and 2 kHz octave bands are of particular importance, so these are regarded. The means and standard deviations of the errors of the MFTs as a functions of the modulation frequencies 0.63, 1, 1.16, 2.5, 4, 6.3 and 10 Hz are shown in Figs. 9 and 10. A positive error means that the modulation of the reproduction is higher than that of the PARISM response. A higher modulation occurs when the energy ratio of the early to the late part of the impulse response is higher. It is seen that the errors using the HOA "max r_E" are largest at 2 kHz, which is probably ascribed to the fact that the "max r_E"-decoding is applied only to the ISM. This again indicates that applying the "max r_E"-decoding only on ISM changes the ratio between early and late energy. The errors for HOA, VBAP, NL are similar.

Lastly, the errors of early IACC and late IACC are compared. Position 2 in the array is not used because the HRIRs in themselves introduce a shifts from the centre that will roughly correspond to the distance to this positions. Position 3 is included to see the potential influence on IACC for a listener

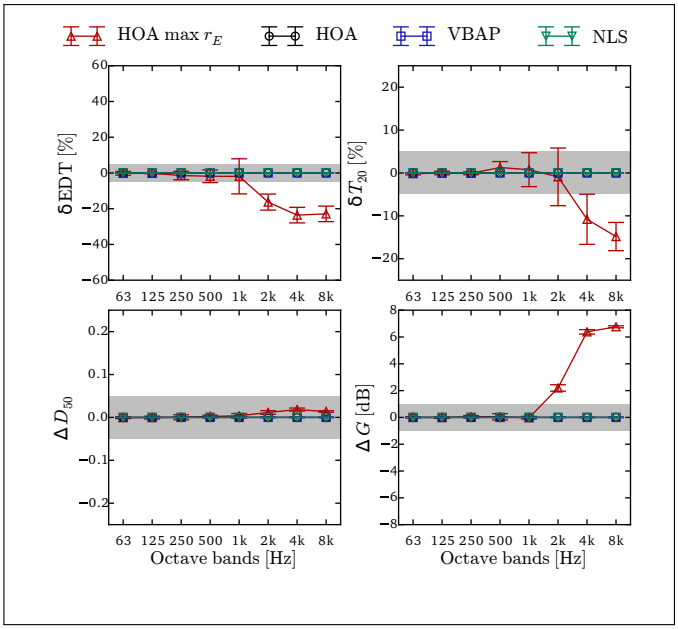

FIG. 5. The errors of the room acoustical parameters calculated for receiver **position 1**. The grey area indicates the JNDs of the parameters.

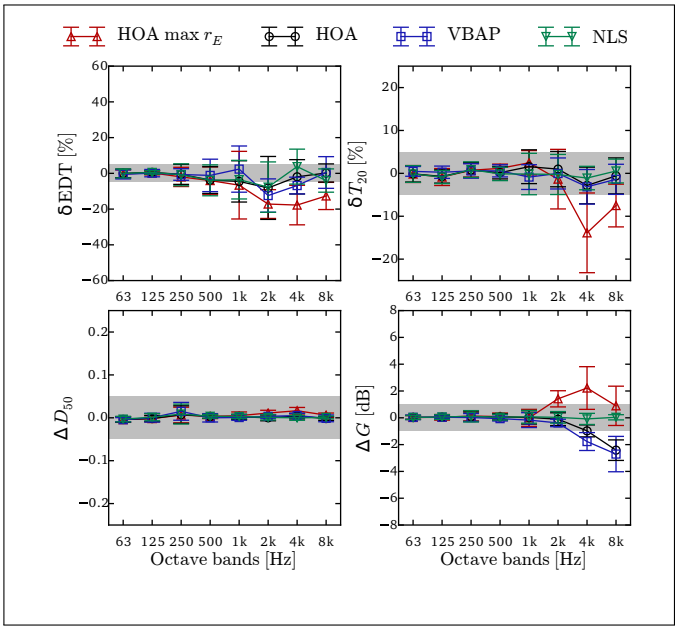

FIG. 6. The errors of the room acoustical parameters calculated for receiver **position 2**. The grey area indicates the JNDs of the parameters.

that is not positioned in the centre of the array. In Fig. 11, it is seen that the reproduction errors on the early IACC are nearly within one JND if the listener is positioned in the centre of the array. But for the late IACC, there are large reproduction errors with HOA "max r_E", HOA and VBAP at high frequencies. Positive errors mean that the reproduced sound field is less diffuse than the original simulation. The NL reproductions produced the smallest errors for the monaural parameters, and Fig. 12 shows that this method is also robust when

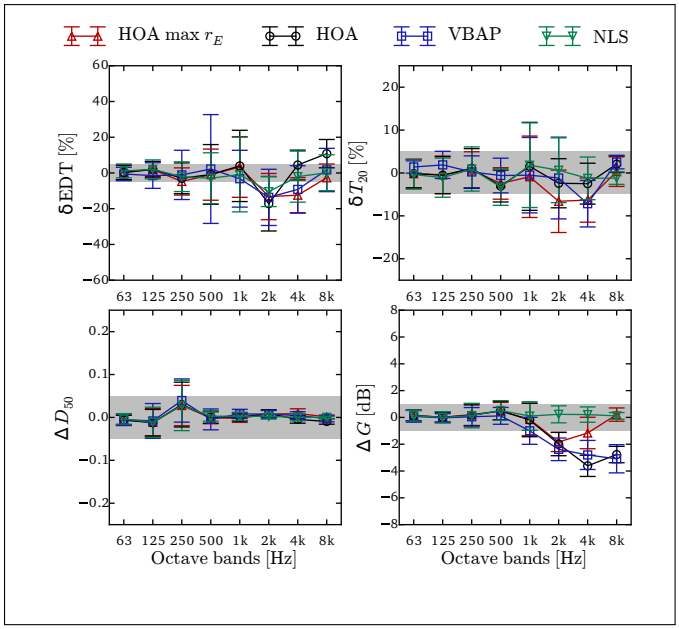

FIG. 7. The errors of the room acoustical parameters calculated for receiver **position 3**. The grey area indicates the JNDs of the parameters.

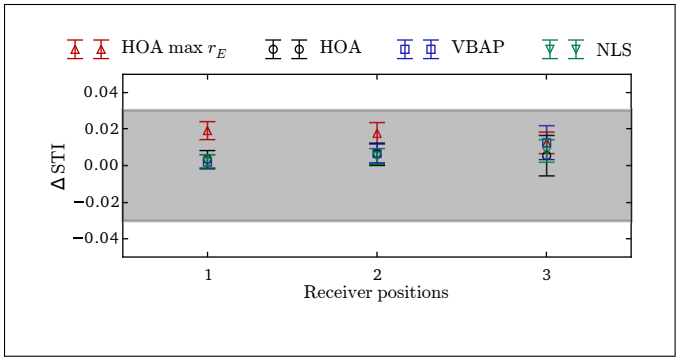

FIG. 8. The errors of the STI of the reproductions. The grey area indicates the JND.

regarding IACC in positions outside of the centre of the array.

VI. CONCLUDING REMARKS

Three reproduction methods have been applied to be able to create loudspeaker auralisations from the phased geometrical room acoustical simulation method PARISM. With respect to the objective room acoustical parameters regarded here, it was shown that by choosing the best reproduction technique most reproduction errors are below one JND. Panning with the nearest loudspeaker method seems to be most robust for the regarded parameters and receiver positions. It should be noted that the loudspeaker array used for the present investigation is a relatively dense array. An array with fewer loudspeakers might behave differently. This would also mean that the order of the HOA auralisations must be reduced because the maximum order depends on the number of loudspeakers.

To complete the validation of the PARISM auralisations,

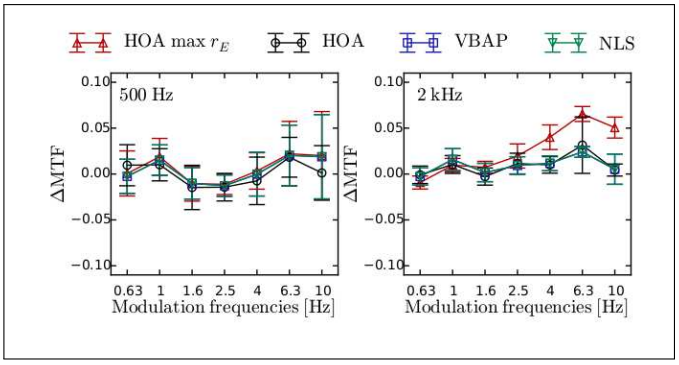

FIG. 9. The errors of the MTFs calculated for receiver **position 1**.

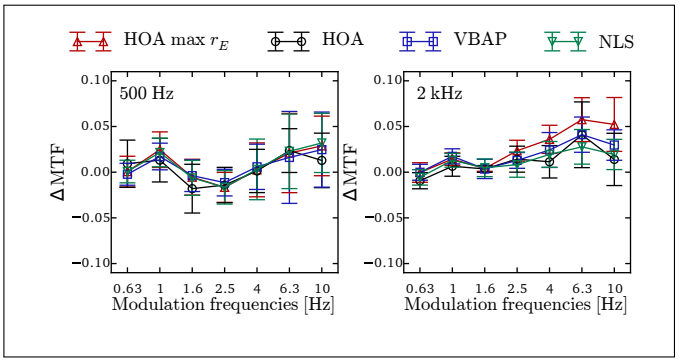

FIG. 10. The errors of the MTFs calculated for receiver **position 2**.

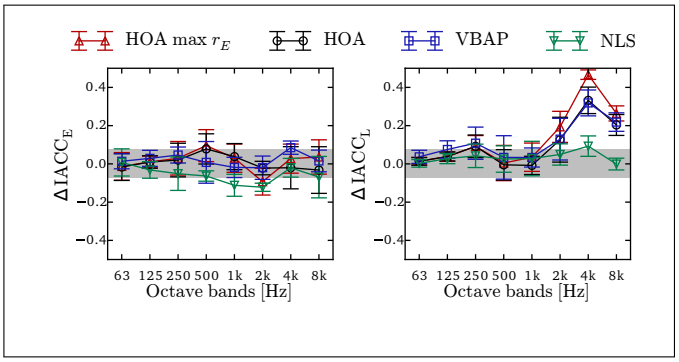

FIG. 11. The error of the IACC for **position 1** in the array. *Left:* early (first 80 ms). *Right:* late (after the first 80 ms).

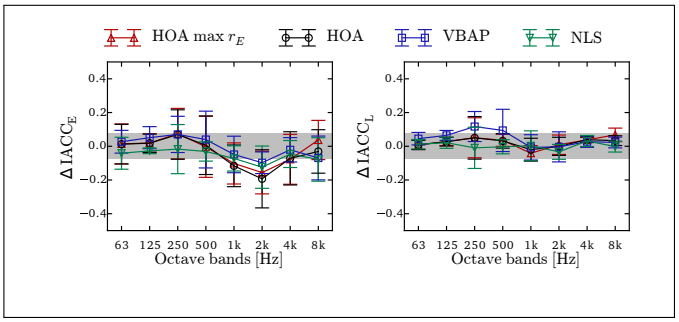

FIG. 12. The error of the IACC for **position 3** in the array. *Left:* early (first 80 ms). *Right:* late (after the first 80 ms).

perceptual studies should be carried out. It is of interest to determine whether the colouration that seems to occur for HOA with "max r_E"-decoding will be perceived, and whether this is more perceivable than the colouration that can occur for basic HOA by listening off-centre positions in the array. Since NL seemed to be the most robust method, it is furthermore relevant to investigate whether there are errors in the spatial perception that are not captured by this objective study. This could for instance be localisation errors for sources that do not coincide with a loudspeaker position or the perception of a moving source.

[1] C. Lynge, G. Koutsouris, and J. Gil, "ODEON manual version 13", ODEON, Kgs. Lyngby, `https://odeon.dk/wp-content/uploads/2017/09/ODEON_Manual.pdf` (2016).

[2] S. Favrot and J. M. Buchholz, "LoRA: A loudspeaker-based room auralization system", Acta Acust. united Ac. **96**, 364–375 (2010).

[3] A. Wabnitz, N. Epain, C. Jin, and A. van Schaik, "Room acoustics simulation for multichannel microphone arrays", in *Proc. International Symposium on Room Acoustics* (European Acoustics Association, Melbourne) (2010).

[4] C.-H. Jeong, J.-G. Ih, and J. H. Rindel, "An approximate treatment of reflection coefficient in the phased beam tracing method for the simulation of enclosed sound fields at medium frequencies", Appl. Acoust. **69**, 601–613 (2008).

[5] A. Wareing and M. Hodgson, "Beam-tracing model for predicting sound fields in rooms with multilayer bounding surfaces", J. Acoust. Soc. Am. **118**, 2321–2331 (2005).

[6] M. Aretz, P. Dietrich, and M. Vorländer, "Application of the mirror source method for low frequency sound prediction in rectangular rooms", Acta Acust. united Ac. **100**, 306–319 (2014).

[7] J. S. Suh and P. A. Nelson, "Measurement of transient response of rooms and comparison with geometrical acoustic models", J. Acoust. Soc. Am. **105**, 2304–2317 (1999).

[8] G. Marbjerg, J. Brunskog, C.-H. Jeong, and E. Nilsson, "Description and validation of a combined phased acoustical radiosity and image source model for predicting sound fields in rooms", J. Acoust. Soc. Am. **138**, 1457–1468 (2015).

[9] G. Marbjerg, J. Brunskog, and C.-H. Jeong, "The difficulties of simulating the acoustics of an empty rectangular room with an absorbing ceiling", Appl. Acoust. **141**, 35–45 (2018).

[10] G. I. Koutsouris, J. Brunskog, C.-H. Jeong, and F. Jacobsen, "Combination of acoustical radiosity and the image source method", J. Acoust. Soc. Am. **133**, 3963–3974 (2013).

[11] F. P. Mechel, "Improved mirror source method in room acoustics", J. Sound Vib. **256**, 873–940 (2002).

[12] E.-M. Nosal, M. Hodgson, and I. Ashdown, "Improved algoritms and methods for room sound-field prediction by acoustical radiosity in arbitrary polyhedral rooms", J. Acoust. Soc. Am. **116**, 970–980 (2004).

[13] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers", IEEE Trans. Speech and Audio Process. **9**, 697–707 (2001).

[14] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning", J. Audio Eng. Soc. **45**, 456–466 (1997).

[15] Y. Miki, "Acoustical properties of porous materials - modifications of delany-bazley", J. Acoust. Soc. Jpn. **11**, 19–24 (1990).

[16] J.-F. Allard, *Propagation of Sound in Porous Media: Modelling Sound Absorbing Materials, chap. 3* (Elsevier Applied Science, London) (2009).

[17] "ISO standard 3382-1 - measurement of room acoustic parameters - part 1: Performance spaces", International Organization for Standardization, Geneva, Switzerland (2009).

[18] "IEC 60268-16 - objective rating of speech intelligibility by speech transmission index", European Committee for Electrotechnical Standardization, Brussels, Belgium (2011).

[19] T. Houtgast and H. J. M. Steeneken, "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria", J. Acoust. Soc. Am. **77**, 1069–1077 (1985).

[20] T. Okano, L. L. Beranek, and T. Hidaka, "Relations among interaural cross-correlation coefficient ($IACC_E$), lateral fraction (LF_E), and apparent source width (ASW) in concert halls", J. Acoust. Soc. Am. **104**, 255–265 (1998).

[21] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The cipic hrtf database", in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustic* (IEEE, New Paltz, NY) (2001).

APPENDIX B

Effect of head-movement on sound-field auditory steady state response measurement

PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Effect of Head-Movement on Sound-Field Auditory Steady State Response Measurement

Sreeram KAITHALI NARAYANAN[1]; Søren LAUGESEN[2]; Valentina ZAPATA-RODRIGUEZ[2,3]; Jonas BRUNSKOG[3]; Cheol-Ho JEONG[3]

[1] Department of Electronic Systems, Aalborg University*

[2]Interacoustics Research Unit

[3]Acoustic Technology Group, Department of Electrical Engineering, Technical University of Denmark

ABSTRACT

Sound-field auditory steady-state response (sound-field ASSR) measurement is an objective alternative for hearing aid fitting validation in infants. For including the hearing aid in the signal path, the stimulus is presented via a loudspeaker. In this case, ASSR can be affected by the head orientation of the participants and the measurement room resulting from the change in degree of modulation of the stimulus in the measurement room. Eleven normal hearing participants were tested for three static head-orientations relative to the loudspeaker in two different rooms. A speech modified NB CE-Chirps stimulus was used to eventually force the hearing aid under test to provide the correct gain. The rooms chosen for measurement were an IEC listening room (T30 of about 0.5 s) and an anechoic chamber (reference condition). A dynamic head-orientation condition comparable to a real head-movement was simulated by randomly combining the responses from the three static head-orientation measurements. The results show a limited influence of head orientation on ASSR level. However, at 4 kHz, a significant reduction in ASSR level was observed when the test ear was oriented away from the loudspeaker. The overall mean ASSR level in the IEC room was reduced by 2.5 dB with reference to the anechoic condition.

Keywords: auditory steady-state response (ASSR), hearing-aid validation, head-movement

1. INTRODUCTION

The improvement in newborn hearing screening has resulted in infants being prescribed with hearing aids at a tender age down to two months. The standard validation tools like aided audiometry or questionnaires are ineffective in these circumstances. Hence, an objective electrophysiological method called steady-state auditory response (ASSR) measurement is considered (1). In sound-field ASSR, the stimulus is presented through a loudspeaker to include the hearing aid in the signal path. This approach comes with several challenges. The reverberation time and background noise of the measurement room can affect the modulation depth of the signal resulting in a reduced ASSR level (2). The head-orientation of the infant during the recording may also affect the ASSR level. These two challenges associated with the sound field ASSR were investigated in this study.

The effect of the room on the ASSR was first examined using simulated room environments (2). This study instead measured sound-field ASSR in two real rooms to investigate the effect of the room. The two rooms were an anechoic chamber and an IEC standard listening room (T_{30}~0.5 s) The ASSR was recorded at three pre-defined static head-orientations to investigate the effect of head-movement. The static head-orientations were considered as the analysis blocks (“epochs”) will be rejected during the actual head-movement due to increase in EEG noise floor resulting from the muscle activity and strain during the movement. Hence, the resultant ASSR in this condition would be an average across the different static head orientations over the whole measurement period. A dynamic head-orientation

[1] skn@es.aau.dk

* The research pertaining to this paper was undertaken during the first author’s tenure at Acoustic Technology Group, Department of Electrical Engineering, Technical University of Denmark (DTU).

condition was simulated from the static head-orientation responses for mimicking the response of natural head-movement during the measurement.

The recorded responses were post-processed and statistically analyzed for comparing the ASSR level, detection time, EEG-noise levels, and detection rate for the defined measurement conditions. For choosing a favorable position for the ASSR measurement in the room, modulation power analysis, a useful tool in characterizing the efficacy of the stimuli for ASSR measurement (3), was performed.

2. METHODS AND MATERIAL

The measurements were carried out on 11 normal hearing test subjects (6 males and 5 females) with a mean age of 25 years. All the participants gave written informed consent and were compensated for their participation. The study was conducted under the approval of the Science-Ethics Committee for the capital region of Denmark.

The stimulus used for the measurement was an ISTS-modified NB CE-Chirp (3). This stimulus has speech like properties, which is critical in the validation of hearing aids in their normal mode of operation, to ensure that correct gain and signal processing features are activated. Each one-octave band CE-chirp (4) was modified by applying the frequency-specific envelope of the International Speech Test Signal (5). The response evoked by each octave-band is identified by the repetition rate at which it was presented. Four-octave band CE-chirps were used, centered at 0.5, 1, 2, and 4 kHz with repetition rates of 90.8, 94.7, 102.5 and 96.7 Hz, respectively. The basic NB-CE chirp was created with a 32-kHz sampling frequency and 65536 samples per epoch, corresponding to a period of 2.048 seconds.

The participants were lying down comfortably on a bed, and the room was darkened during the measurement. Only one ear was stimulated at a time, while the other ear was blocked using an earplug. Standard 4-electrode montage (high forehead reference, ipsi- and contra-lateral mastoids active, and cheek ground) were used for the ASSR recording. A custom-made MATLAB software loaded to a laptop controlled the playback and recording. The stimulus was routed to an in-house built two-way coaxial loudspeaker, placed 1 meter above the test subject's head, through an audio chain comprising of an RME Fireface UC soundcard, a graphic equalizer, a custom-made crossover filter, two attenuators, and an audio power amplifier. The Interacoustics Eclipse unit was used as the front-end to deliver the line-level EEG signal to the RME Fireface UC soundcard.

To evaluate the effect of head-movement, three static head-orientation conditions were defined, namely, No Head-Movement (NHM), Towards the Speaker (TS) and Away from the Speaker (AS). These head-orientation conditions were characterized concerning the position of the test ear in relation to the loudspeaker. In the NHM condition, the participants were instructed to not to move their head and look straight at the speaker for the entire duration of the measurement. Then, the participants were asked to move the test ear towards the speaker and hold that position for the TS condition. For AS, the participants moved the test ear away from the speaker. This is illustrated in Figure 1. The head-orientation of the participants for the TS and AS conditions were not constrained to a specific angle but were left to the test subject's comfort and convenience, to keep muscle artifacts in the EEG signals to a minimum.

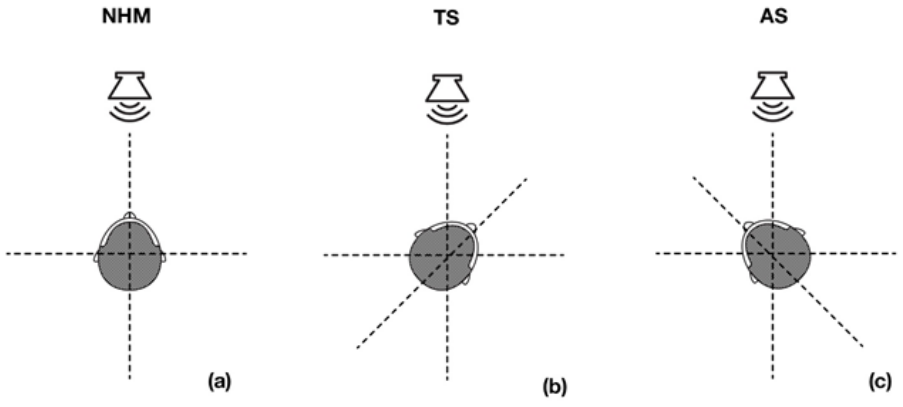

Figure 1: Static head-orientation conditions with reference to left ear (stimulated ear): (a) No Head-Movement (NHM), (b) Towards the Speaker (TS), and (c) Away from the speaker (AS)

Two room conditions were identified to understand the effect of the room on the ASSR measurement. The first one was an anechoic chamber of dimension 4.8m x 4.1 m x 2.9 m (L x W x H). This chamber is considered anechoic from 100 Hz to 10 kHz. The second one was IEC 60268-13 standard listening room henceforth referred to as the IEC room. The IEC room (7.52 m x 4.74 m x 2.76 m) with a T_{30} of about 0.5 seconds is comparable to a clinical room, albeit somewhat larger than a typical clinic. The anechoic chamber was considered the reference condition. The results from the IEC room were compared with those from the anechoic chamber to understand the effect of the room.

Before conducting the actual ASSR measurement, an un-normalized modulation power analysis of the stimulus was performed. This analysis intends to predict the ASSR efficacy of the stimulus (3) in a given room and head-orientation condition. The stimulus was convolved with the impulse responses recorded using a HATS and a DIRAC system for every defined room and head-movement condition. This convolved stimulus was then passed through a gammatone filter bank having 24 filters $1/24^{th}$ octave spaced apart representing the auditory filter with respect to the octave bands of interest (0.5 through 4 kHz). Edge effects were removed for the aperiodic speech-modified stimulus by multiplying the sample epochs of each Hilbert envelope by a raised cosine window. Then a discrete Fourier transform was applied on the gated envelopes, and the resulting modulation power was averaged across the epochs and the 24 gammtone filter bands. The modulation power at each octave frequency band was evaluated at the fundamental corresponding repetition rate (ignoring the higher harmonics). The modulation power analysis was used to screen out un-favorable measurement positions.

The ASSR was recorded for all the participants at first in the anechoic chamber and then in the IEC room for the three defined static head-orientation conditions and for both ears. Special care was taken to measure with low EEG noise levels (henceforth referred to as noise level), meeting the rejection ratio of 40 μV. A simple F-detector (6) comparing the power at the repetition rate to the average power across 10 frequency bins on either side of the repetition rate was used to evaluate the detection of the first harmonic at the repetition rate in question. A Bonferroni corrected 1% error rate was used for this test. The measurement was concluded once the F-statistics for all four octave-bands were above the F-critical value, or after a maximum of 20 minutes of recording time. Continuous averaging of response and noise frequency bins improved the signal to noise ratio.

A simulated dynamic head-orientation condition comparable to real head-movement was constructed by combining one-third of each static head-orientation recording for each participant. Initially, the responses were truncated to the length of the recording with the least number of epochs, and these responses were divided into three parts. Then, one of the three parts of each static head-orientation condition was chosen randomly and combined in a random order to simulate the dynamic head-orientation condition. This is illustrated in Figure 2.

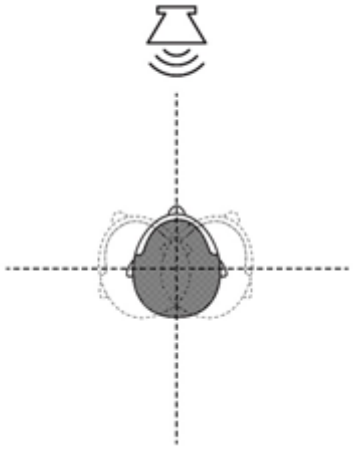

Figure 2: Illustration of simulated dynamic head-orientation condition.

Subsequently, the recorded responses were post-processed along with the simulated dynamic head-movement response. Three primary outcomes: noise-corrected ASSR level, detection time, and noise level were considered. In addition, detection rates were calculated. The noise corrected ASSR level (referred as ASSR level hereafter) was calculated as the difference between the estimated power at the response harmonic and the noise power (6) converted to dB for a better assessment of relative

changes across conditions. The detection time was determined as the first time a response was detected with an F-detector using a 5% error rate in the successive weighted average. Any corrections for the repeated measure was thus ignored. A $\log_{10}$ transformation (relative to 1 s) of the determined detection time was performed, and this log-transformed detection time is termed T_{log}. The log transformation was found to yield a more even distribution of residuals in the statistical analysis (3). The noise level was estimated by averaging the noise power in the 10 bands on either side of the response harmonic with respect to a specific repetition rate. The frequency bins close to 50 Hz line-noise, other know interferences, and any other repetition rate harmonics were excluded. The noise level thus determined was later adjusted to the level that would have been found at a testing time of 100 s. This was achieved by adding 10 x $\log_{10}$ (t/100) to the noise level (in dB) of the whole recording, where t is the full recording time. This level is termed L_{100}. This adjustment was necessary as the noise is expected to be lower with longer measurement time. As explained above, some measurements were concluded after reaching the F-critical value, which means that the recording time of each measurement need not be 20 minutes.

The three outcomes of the post-processing: ASSR level, T_{log}, and L_{100} were analyzed using a linear mixed model, whereas the detection rate was analyzed using non-parametric statistical tests. In the mixed effects model, the participants (TP: 1, 2, 3..., 11) were considered a random effect. The fixed effects considered were: test ear (TE: Left and Right), room conditions (RC: Anechoic and IEC), head-orientation condition (HOC: NHM, TS, AS, DHM), and stimulus frequency (FREQ: 0.5, 1, 2, and 4 kHz). Estimated marginal means (7) were determined for the best-fit model, and a post-hoc analysis comparing the means was performed using the Tukey method (8).

3. Results

The mixed model ANOVA results for the three primary outcomes are shown in *Table 1*. For the ASSR level and L_{100} models, the test subject was found to be significant ($p<0.001$***).

Table 1: Summarized mixed model ANOVA results. Only the significant main effects and interactions are presented. ***p<0.001, **p<0.01, *p<0.05

	ASSR level		Detection Time (T_{log})		Noise Level (L_{100})	
Factor	F statistics	Pr(>F)	F statistics	Pr(>F)	F statistics	Pr(>F)
FREQ	F(3,612)=35	<0.001***	F(3,662)=6	<0.001***		
RC	F(1,612)=83	<0.001***	F(1,662)=4	0.011**	F(1,678)=41	<0.001***
HOC	F(3,612)=11	<0.001***				
HOC:FREQ	F(9,612)=6	<0.001***				
HOC:RC					F(3,678)=9	<0.001***
HOC:TE					F(3,678)=4	0.006**
RC:HOC:TE					F(3,678)=5	0.002**

3.1 ASSR Level

The plot concerning the effect of stimulus frequency, head-orientation, and room conditions is shown in Figure 3. Even though the third order interaction is not significant, it helps to visualize all the significant effects in a single plot, in combination with post-hoc testing (Tukey HSD). Regarding the main effect of FREQ, the post-hoc analysis showed that levels at 0.5 and 1 kHz were significantly smaller than those at 2 and 4 kHz (all $p<0.001$). The significant main effect of room (RC) indicates lower levels in the IEC room. The overall mean reduction in ASSR level in the IEC room was approximately 2.5dB. For the main effect of head orientation (HOC), the post-hoc analysis shows that only the AS condition is significantly different from the other three (all $p<0.05$). Finally, regarding the significant HOC:FREQ interaction, the post-hoc results show a significant difference only for the AS condition at 4 kHz ($p<0.001$).

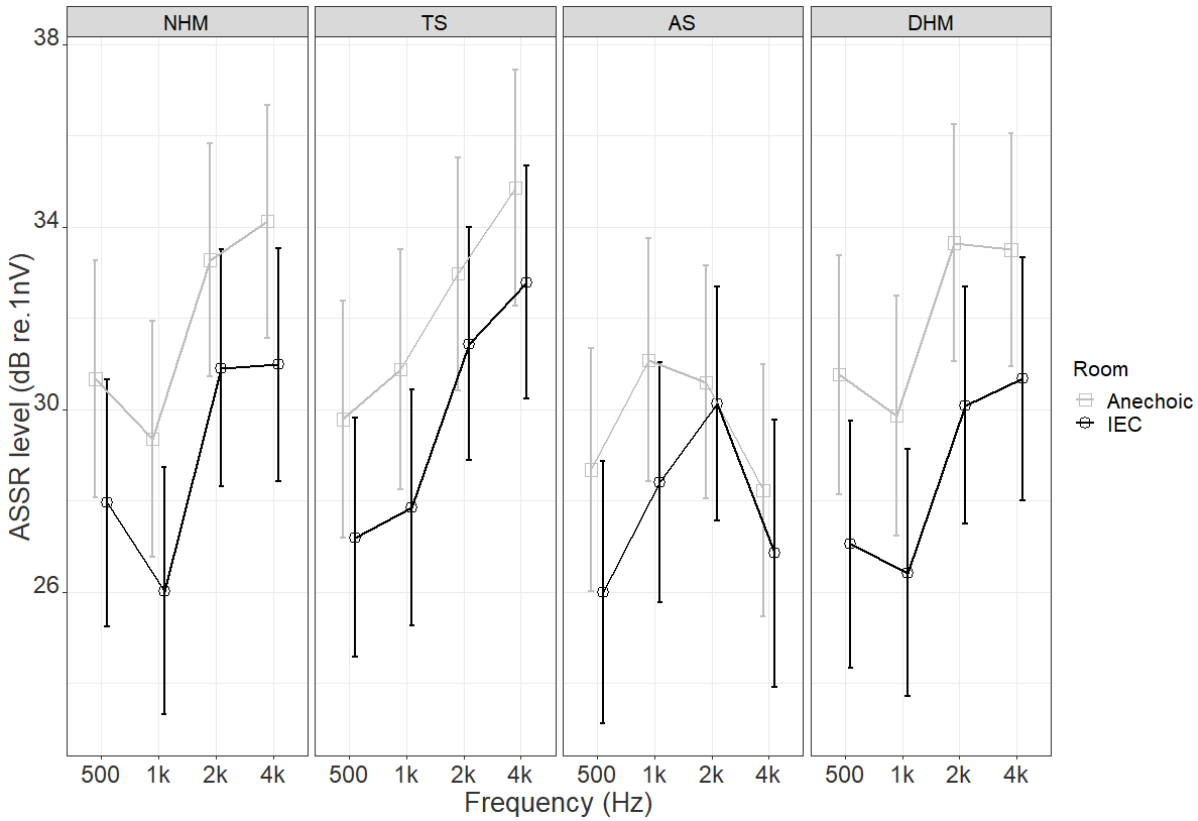

Figure 3: Estimated marginal mean ASSR level plotted across stimulus frequency, room condition, and head-orientation. Error bar indicates 95% confidence interval.

3.2 Detection Time

Figure 4 shows that the higher frequencies (2 and 4 kHz) had lower detection times than the lower frequencies (0.5 and 1 kHz). Also, from Figure 5, it can be seen that detection times in the IEC room were significantly higher than those from the anechoic chamber.

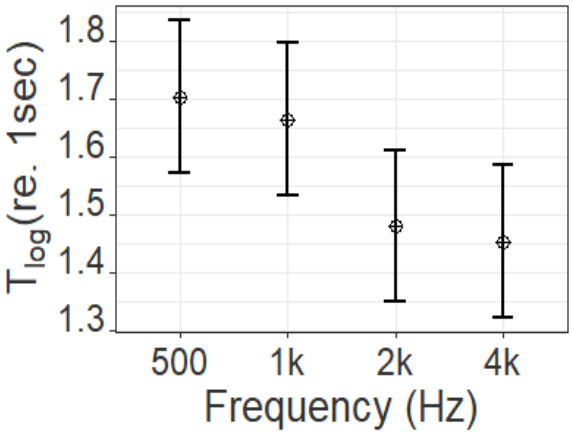

Figure 4: Estimated marginal mean detection time corresponding to stimulus frequency. Error bar indicates 95% confidence interval

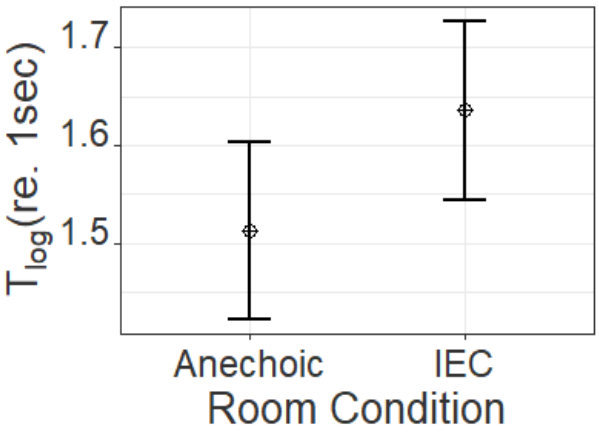

Figure 5: Estimated marginal mean detection time corresponding to room conditions. Error bar indicates 95% confidence interval

3.3 Noise Level

Figure 6 shows the plot concerning the effect of head-orientation, test ear, and room condition on L_{100}. The most interesting result is the significant reduction in noise level in the IEC room. The post-hoc results for the interaction RC:HOC showed that the major contributor to the overall decrease in the noise level in the IEC room was the TS condition. For the significant HOC:TE interaction, the post-hoc result showed a significant difference between the NHM and TS condition for the left ear. The post-hoc analysis on the third order RC:HOC:TE interaction revealed that the difference in noise level at IEC and anechoic chamber were most significant for TS condition for both left and right ears (both $p<0.001$).

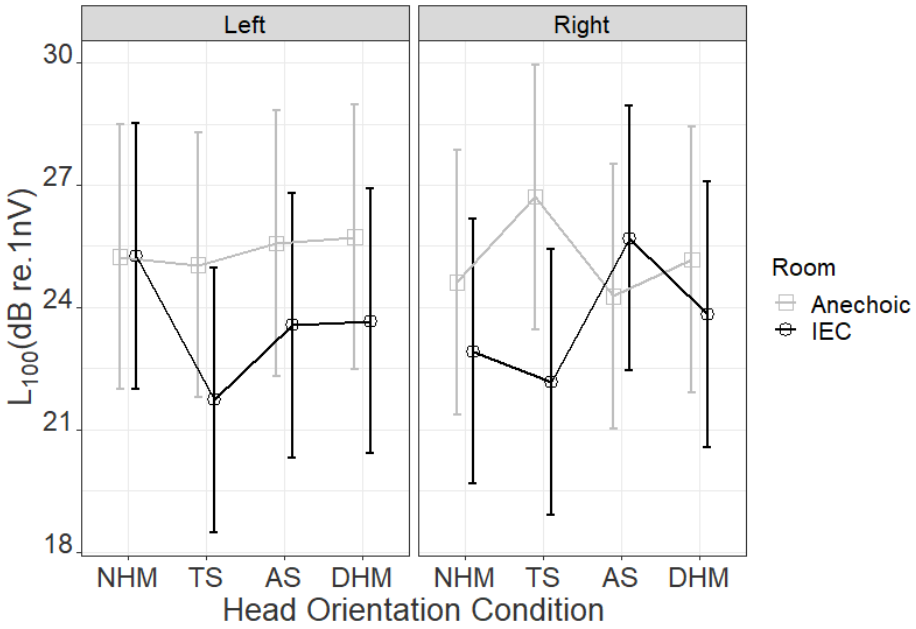

Figure 6: Estimated marginal mean noise level L_{100} (dB re. 1nV) plotted across head-orientation, test ear, and room condition. Error bar indicates 95% confidence interval.

3.4 Detection Rate

An overall detection rate of 91% was obtained (63 non-detections out of 703 data points). The results presented in Table 2 show higher detection at higher frequencies (2 and 4 kHz). The AS condition had the lowest detection rates in both anechoic and IEC room. A Fisher exact test revealed that the reduction in aggregate detection rate at IEC room (89%) compared to the anechoic chamber (93%) was significant ($p=0.017$).

Table 2: Detection rate in percentage across stimulus frequency, head-orientation conditions, and room conditions.

FREQ	Anechoic Chamber				IEC Room				FREQ
	NHM	TS	AS	DHM	NHM	TS	AS	DHM	(Aggregate %)
0.5 kHz	95	95	86	91	82	91	68	82	86
1 kHz	95	91	86	91	82	95	91	82	89
2 kHz	100	100	100	95	95	100	100	95	98
4 kHz	100	95	77	100	100	100	64	86	90
HOC (Aggregate %)	98	95	87	94	90	97	81	86	
RC (Aggregate %)	93				89				

4. DISCUSSION

The general trend in the ASSR levels concerning the stimulus frequency is in good agreement with the previous study (3), with higher ASSR levels towards higher frequencies (except for the AS condition at 4 kHz). It is concluded that the effect of head-movement on ASSR level was limited because the ASSR level for DHM condition corresponding to real head-movement was not significantly different from NHM condition. However, if the head is oriented as in AS condition for the entire time of recording, this could result in a significant reduction of ASSR level at 4 kHz. The observed decrease in ASSR level at 4 kHz for AS condition can be attributed to the head-shadow effect. The mean reduction of ASSR level of 2.5 dB in the IEC room compared to the anechoic chamber is associated with the change in acoustics. The reflections from the surfaces and edges in the IEC room can alter the degree of modulation of the stimulus resulting in a reduced ASSR level.

The detection times generally showed an inverse pattern relative to the ASSR level, in agreement with the previous study (3). It is expected that an increase in ASSR level is accompanied by a reduced detection time, as the criterion signal-to-noise ratio is reached faster with a strong signal. The detection rates also showed close relation to the ASSR level, with the higher detection rates at higher frequencies and a significantly reduced detection rate in the IEC room.

The noise level results were peculiar; the reduction in noise level in the IEC room was unexpected. It could be speculated that the participants were more relaxed and comfortable in the IEC room. During the informal discussion after the measurement, some participant commented that the sound from the speaker in IEC room was more diffused and appealing than the sharp and uncomfortable sound in the anechoic chamber. However, a couple of participants felt the anechoic chamber to be more relaxing. There can also be an order effect as the measurements were first performed in the anechoic chamber. Also, most of the participants did not have prior experience with participating in electrophysiological experiments. All these could have resulted in an increased noise level in the anechoic chamber.

5. CONCLUSIONS

The aim of the study was two-fold: primarily to understand the effect of head-movement on sound-field ASSR and secondarily to understand the influence of the measurement room. The measurements were carried out in an IEC standard listening room and anechoic chamber with three predefined static head-orientation conditions.

The main findings from the study are:

- The comparison of results for NHM and DHM head-orientation condition indicates a minimal effect of head-movement on ASSR. DHM is considered comparable to a real head-movement. However, the AS condition at 4 kHz can influence the overall ASSR level and can have serious implication in a clinical setting.

- The change in acoustics in the IEC room resulted in an overall mean reduction of 2.5 dB in the ASSR level compared to the anechoic chamber. This establishes the influence of the measurement room on ASSR measurements.

ACKNOWLEDGEMENTS

The study was carried out as part of the master thesis project and further research assistantship of the first author at Acoustic Technology Group, Department of Electrical Engineering, Technical University of Denmark. The study was co-funded by Technical University of Denmark and William Demant Foundation.

REFERENCES

1. Picton TW, Durieux-Smith A, Champagne SC, Moran LM, Giguere C, Beauregard Y. Objective Evaluation of Aided Thresholds Using Auditory Steady-State Responses. J Am Acad Audiol. 1998;9(5):315–31.
2. Zapata Rodriguez, Valentina; Laugesen, Søren; Jeong, Cheol-Ho; Brunskog, Jonas; Harte JM. Towards Predicting Room Acoustical Effects on Sound-Field ASSR from Stimulus Modulation Power. A41st MidWinter Meet Assoc Res Otolaryngol San Diego, United States. 2018;
3. Laugesen S, Rieck JE, Elberling C, Dau T, Harte JM. On the Cost of Introducing Speech-Like Properties to a Stimulus for Auditory Steady-State Response Measurements. Trends Hear. 2018;22:1–11.
4. Elberling C, Don M. A direct approach for the design of chirp stimuli used for the recording of auditory brainstem responses. J Acoust Soc Am. 2010;128(5):2955–64.
5. Holube I, Fredelake S, Vlaming M, Kollmeier B. Development and analysis of an international speech test signal (ISTS). Int J Audiol. 2010;49(12):891–903.
6. Dobie RA, Wilson MJ. A comparison of t test, F test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids . J Acoust Soc Am. 2005;100(4):2236–46.
7. Searle SR, Speed FM, Milliken GA. Population Marginal Means in the Linear Model : An Alternative to Least Squares Means Author (s): S . R . Searle , F . M . Speed and G . A . Milliken Published by : Taylor & Francis , Ltd . on behalf of the American Statistical Association Stable URL : . 2019;34(4):216–21.
8. Jean Dunn O. Multiple Comparisons Among Means. J Am Stat Assoc. 1961;56(293):52–64.

APPENDIX C

The influence of overlapping band filters on octave band decay curves

Acta Acustica united with Acustica
Vol. 104 (2018) 943 – 946

Fast Track Articles
DOI 10.3813/AAA.919259

The Influence of Overlapping Band Filters on Octave Band Decay Curves

Gerd Marbjerg[1)], Jonas Brunskog[1)], Cheol-Ho Jeong[1)], Valentina Zapata-Rodriguez[2,1)]
[1)] Acoustic Technology, Department of Electrical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark. jbr@elektro.dtu.dk
[2)] Interacoustics Research Unit, 2800 Kgs. Lyngby, Denmark

Summary
This study showed that the overlap of practically-used bandpass filters can influence the octave band decay curves, especially if the decays are calculated from a filtered impulse response that has been created from octave band energy responses. Energy from a frequency band with a long reverberation time can leak into a neighbouring band with a shorter reverberation time. This also means that neither octave band decays from a measured response are independent, nor are measured octave band reverberation times.

1. Introduction

When using energy-based geometrical room acoustic modelling techniques, room acoustical parameters are normally calculated at the centre frequency of octave bands. This assumes that the energy response in each band only depends on the material properties of the very band. If these results are to be used for auralisations, it is necessary to create a total full bandwidth pressure impulse response from the octave band energy impulse responses [1, 2, 3, 4]. Here, full bandwidth response refers to a response that covers the entire frequency range of interest, typically the audible range. The full bandwidth pressure response can be obtained by first creating octave band pressure impulse responses and then summing these.

If the full bandwidth pressure impulse response is re-filtered into octave band impulse responses, the decays of these are unlikely to be the same as the decays of the responses from before the summation, because of overlaps between adjacent bands that cause energy leakage. Reverberation times are often calculated from the octave band responses, expecting these to be valid also for the full bandwidth impulse response. This letter demonstrates that they are not necessarily so. When measured impulse responses are processed with bandpass filters, the decays in the bands are not independent, and simulations assuming independent bands therefore do not correspond to measurements.

This study focuses on decay curves and reverberation times, because the effect of the frequency leakage is much larger when considering the energy decay than the total energy or the steady state response.

The phenomenon is not limited to energy-based models, because one might use a pressure-based model to calculate an impulse response within a octave band and determine the decay from this. In this case, it is also possible that the obtained decay will not correspond to one that would be obtained if the impulse response of a wider frequency range had been calculated and then filtered to the octave band.

To the best knowledge of the authors, this issue of overlapping bands has not yet been sufficiently discussed in this application field. A related issue for measurements of narrow band decays, is the influence of the time responses of the filters, which has been studied [5, 6]. The present study illustrates how the overlapping bands influences the decays from energy-based models, and investigates this through simple examples.

Received 23 March 2018,
accepted 1 September 2018,
published online 10 October 2018.

2. Full bandwidth impulse response

A full bandwidth impulse response from energy-based methods is often found by first determining octave band impulse responses and then taking the sum of those as

$$p(t) = \sum_b p_b(t), \tag{1}$$

where $p_b(t)$ is the impulse response of the octave band b. $p_b(t)$ can for instance be found with an octave band noise signal that is used to fill an energy impulse response. The octave band impulse response will in that case be given by

$$p_{b,noise}(t) = \big(n(t) * h_b(t)\big)\sqrt{w_b(t)}, \tag{2}$$

where $w_b(t)$ is the energy impulse response of the band b, $n(t)$ is a Gaussian noise signal, and $h_b(t)$ is the impulse response of the filter of band b. $n(t) * h_b(t)$ is thus a octave band noise signal of which the content will be mainly within the cutoff frequencies of the band b, but there will be some content outside depending on the sharpness of the filter. The method of Equation (2) is, for example, used to obtain a pressure impulse response from acoustical radiosity in the simulation tool PARISM [4]. A Poisson process with random sign can also be used rather than the Gaussian noise signal [3].

To compare with Equation (2), a method that does not employ bandpass filters is tested. The rational behind this approach is to limit the overlaps of the octave band responses. For this, sine functions of random phases are used, and the impulse response within a single band is then given by

$$p_{b,sin}(t) = \sum_{l=1}^{L} \sin\left(2\,\pi\; f_l\, t + \varphi_l\right) \sqrt{w_b(t)}, \qquad (3)$$

where L is the number of included sines within band b and φ_l is a random phase. f_l refers to frequencies between the lower and upper cutoff frequencies of band b. With this formulation, the overlap of the bands in the creation of the full bandwidth response only comes from the attenuation of the sines due to the decay factor $\sqrt{w_b(t)}$.

Regardless of whether $p_{b,noise}$ or $p_{b,sin}$ is used, there will be an overlap of the bands if the full bandwidth response is refiltered with non-ideal filters. By comparing $p_{b,noise}$ and $p_{b,sin}$, it can be determined how much of the total effect is due to the fact the filters in $p_{b,noise}$ overlap, and how much is due to the overlap of the filters for refiltering. The refiltered octave band response is found as

$$p_{b,RF}(t) = p(t) * h_b(t), \qquad (4)$$

where the subscript RF denotes that it is the refiltered response, and $p(t)$ is found with Equation (1). In the following $p_{b,noise,RF}$ refers to a $p_{b,RF}$ using $p_{b,noise}$ in Equation (1), and $p_{b,sin,RF}$ refers to a $p_{b,RF}$ using $p_{b,sin}$ in Equation (1).

3. Example with geometrical room acoustics

Firstly, the influence of overlapping bands on decay curves is illustrated with an example using the room acoustical simulation tool CARISM [7] (Combined Acoustical Radiosity – Image Source Method). CARISM is an energy-based combination of acoustical radiosity (AR) and the image source method (ISM), and the results from CARISM are octave band energy impulse responses. The method of Equation (2) is applied to obtain a pressure impulse response, and the filters used there and in the refiltering are octave bandpass filters constructed from the 7th order high- and low-pass Butterworth filters. The bandpass filters are constructed such that the sum of their frequency responses is flat, and they meet the requirements of IEC 61260-1 [8].

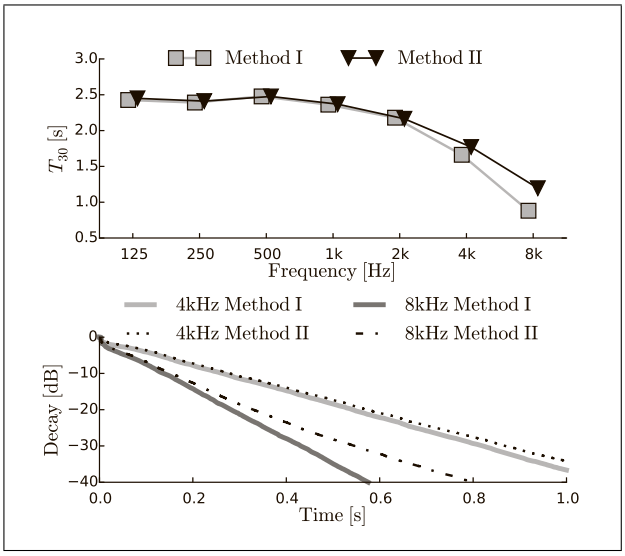

Figure 1. Reverberation times T_{30} (above) and decay curves (below) from a CARISM simulation.

Octave band decay curves and reverberation times from CARISM can then be obtained with two methods. **Method I:** Directly from the octave band energy impulse responses ($w_b(t)$), where the octave band results are independent of each other. This is the standard method in CARISM. **Method II:** By filtering the full bandwidth impulse response that is constructed, Equation (4).

The chosen test room is based on an existing room at the laboratories of the Technical University of Denmark and has dimensions [4.38 × 3.29 × 2.7] m. The calculations are done in the octave bands from 125 Hz to 8 kHz, and the sampling frequency for the pressure impulse response is 24 kHz. The scattering coefficients of all surfaces are set to [0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1] for the eight octave bands, respectively. The absorption coefficient is 0.05 for all surfaces and frequencies. The air absorption is determined according to ISO 9631-1 [9], and since the surface absorption is frequency-independent most of the differences in reverberation times over frequency will be due to the air absorption. The reverberation times T_{30}, calculated with both methods I and II, are plotted in Figure 1. Differences are seen between the two methods at 4 and 8 kHz, and that values obtained with method II are higher than those of method I to II.

The decay curves for the 4 kHz and 8 kHz bands from methods I and II are plotted in the lower part of Figure 1. The 8 kHz curve of method II follows the 8 kHz method I curve in the very early part, and then the slope changes to be more similar to that of the 4 kHz method I curve. This indicates that energy from the 4 kHz band is influencing the 8 kHz band in the part of the decay where the energy is low in the 8 kHz band. It is also observed that the curve of 8 kHz band of method II is tending more towards being double-sloped than that of method I. The 4 kHz curves of the two methods are more similar.

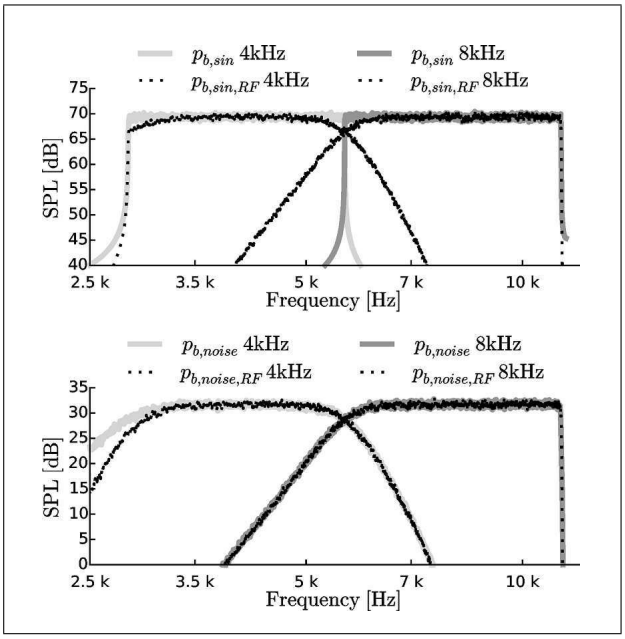

Figure 2. Frequency responses.

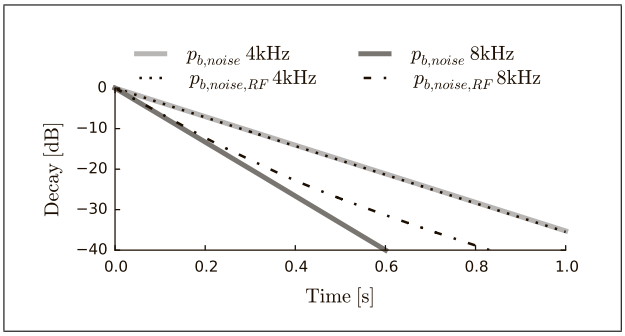

Figure 3. Decay curves.

Table I. Percentage differences in T_{30} and EDT from p_b to $p_{b,RF}$. Note that the JND is 5%.

	7th order filters		9th order filters	
	4 kHz	8 kHz	4 kHz	8 kHz
$\Delta T_{30,p_{b,noise}}$	-0.19%	38%	-0.15%	32%
$\Delta T_{30,p_{b,sin}}$	-0.13%	29%	-0.10%	23%
$\Delta EDT_{p_{b,noise}}$	-2.00%	7.3%	-1.50%	5.6%
$\Delta EDT_{p_{b,sin}}$	-1.31%	3.0%	-1.01%	3.9%

4. Examples with exponential decays

The energy impulse responses of Eqs. (2) and (3) are then chosen to decay exponentially. We thus set $\sqrt{w_b(t)} := \mathrm{e}^{-\delta_b t}$, rather than determining $w_b(t)$ through a simulation as in Section 3. δ_b is the exponential decay constant, from which the reverberation time can be found by $T = 3\ \ln(10)/\delta$. The initial filters used here are the same as in the example of Figure 1.

As in the previous example, the 4 kHz and 8 kHz octave bands are used. Firstly, the frequency responses of the two bands are regarded with both reverberation times set to 0.9 s. The spacing of the frequencies f_l in Equation (3) is 1 Hz. The realisations of $p_{b,noise}$ and $p_{b,sin}$ were repeated 200 times, because each realisation will be slightly different due to the random noise in $p_{b,noise}$ and the random phases in $p_{b,sin}$. The frequency responses are calculated by the Fourier transform of the impulse responses and the means of the squared magnitudes of the 200 realisations of frequency responses are plotted in Figure 2. The overlap of the $p_{b,sin}$ frequency responses is very small (0.085% of the total energy), which makes good sense. The overlap of the frequency responses of $p_{b,sin,RF}$ is then much larger (4.3% of the total energy). The frequency responses of $p_{b,noise}$ overlap much already, so it is barely increased for $p_{b,noise,RF}$ (from 4.3% to 4.4% of the total energy).

A difference in the reverberation times between the 4 kHz and 8 kHz octave bands is then introduced. They are set to 1.7 and 0.9 s, respectively, which are taken from the example of Figure 1.

The mean decay curves from 200 realisations of $p_{b,noise}$ and $p_{b,noise,RF}$ are shown in Figure 3. $p_{b,sin}$ and $p_{b,sin,RF}$ are left out of this figure because they are very similar. The decay curve of $p_{b,noise,RF}$ in the 8 kHz band is very much influenced by the 4 kHz band. It is not single-sloped and follows the one of 8 kHz p_b for the very first part of the decay, but in the later part the slope approaches that of the 4 kHz $p_{b,noise}$ decay. The $p_{b,noise}$ and $p_{b,noise,RF}$ decays for the 4 kHz band coincide, which confirms that the leakage between bands mostly influences the band with the shorter reverberation time.

The mean reverberation time (T_{30}) and mean early decay time (EDT) of the 200 realisations were also calculated, and the relative differences between those from $p_{b,noise,RF}$ and $p_{b,sin,RF}$, and those from $p_{b,sin}$ and $p_{b,noise}$ are calculated as

$$\begin{aligned} \Delta T_{30,p_b} &= \left(T_{30,p_{b,RF}} - T_{30,p_b}\right)/T_{30,p_b} \cdot 100\% \qquad (5) \\ \Delta \mathrm{EDT}_{p_b} &= \left(\mathrm{EDT}_{pb,RF} - \mathrm{EDT}_{p_b}\right)/\mathrm{EDT}_{p_b} \cdot 100\%, \end{aligned}$$

where the T_{30} and EDT values are means of the 200 realisations. In Table I, it is seen that the difference is obviously largest for the 8 kHz band, and that it is the reverberation time that is most influenced. For the 8 kHz bands, the changes in reverberation times are above the just noticeable difference, which is stated as 5% in ISO 3382-1 [10]. For the 8 kHz early decay times, it is only $\Delta EDT_{p_{b,sin}}$ that is below the just noticeable difference. For the 4 kHz values there are also small changes, all below the just noticeable difference. But since the differences are consistently reductions, they cannot be random and must stem from the leakage. The changes are generally smaller for $p_{b,sin,RF}$, but still large enough to show that the overlap of the refiltering filters can create a noticeable difference.

In order to test the influence of the filter design, higher order Butterworth filters are tested. The bandpass filters are then created from the 9th order filters rather than the 7th order. The differences in the EDT and T_{30} with these filters are shown on the right side of Table I. The differences caused by the leakage are smaller when the filters

are sharper, but it is still only the 4 kHz $\Delta EDT_{p_{b,sin}}$ that is below the just noticeable difference.

When choosing and designing filters, their computational cost and stability should be considered. For the present Butterworth filters, the highest possible order for stable filters is 7, if the 125 Hz octave band should be included. Moreover, when filtering to obtain decays in octave bands possible, ringing of the filters in the time domain should also be considered, because it can influence the decays [6]. Ringing in the time domain tends to increase when the filter are sharper in the frequency domain.

5. Concluding remarks

When creating full bandwidth pressure impulse responses from octave band energy responses, the overlaps of the applied bandpass filters influence the decays of the octave bands. The effect can be important when looking at decays, even when the leakage in energy is marginal. If an octave band has a neighbouring band with a slower decay than itself, leakage from the slowly decaying band will make its decay slower when calculated from the full bandwidth impulse response. The shape of the decay curve will furthermore tend to be double-sloped. Even if the construction of the full bandwidth response is done such that there is hardly any overlap between the octave band responses, the overlaps of the filters used for refiltering the full bandwidth response are big enough for spillover between the bands to influence the reverberation times. This indicates that the same will be true when filtering and processing a measured impulse response, which means that the assumption of independent bands in simulations is an approximation and may lead to noticeable errors. Finally, the design of the bandpass filters has an influence on leakage, and sharper filters naturally reduce the effect. But even with sharper filters than required in IEC 61260-1 [8], the influence on the reverberation time of non-ideal filtering is found to be higher than the JND.

Acknowledgements

The authors are grateful for discussions with James Harte and Søren Laugesen that lead to the findings in this paper.

References

[1] H. Kuttruff: Auralization of impulse responses modeled on the basis of ray-tracing results. J. Audio Eng. Soc. **41** (1993) 876–880.

[2] C. Lynge, G. Koutsouris, J. Gil: ODEON manual version 13. ODEON, Kgs. Lyngby, 2016.

[3] D. Schröder: Physically based real-time auralization of interactive virtual environments. Ph.D. thesis, RWTH Aachen, 2011.

[4] G. Marbjerg, J. Brunskog, C.-H. Jeong, E. Nilsson: Description and validation of a combined phased acoustical radiosity and image source model for predicting sound fields in rooms. J. Acoust. Soc. Am. **138** (2015) 1457–1468.

[5] M. Kob, M. Vorländer: Band filters and short reverberation times. Acta Acust. united Ac. **86**, (2000) 350–357.

[6] F. Jacobsen: A note on acoustic decay measurements. J. Sound Vib. **115** (1987) 163–170.

[7] G. I. Koutsouris, J. Brunskog, C.-H. Jeong, F. Jacobsen: Combination of acoustical radiosity and the image source method. J. Acoust. Soc. Am. **133** (2013) 3963–3974.

[8] IEC 61260-1: Octave-band and fractional-octave-band filters - Part 1: Specifications. European Committee for Electrotechnical Standardization, Brussels, Belgium, 2014.

[9] ISO standard 9631-1: Attenuation of sound during propagation outdoors - Part 1: Calculation of the absorption of sound by the atmosphere. International Organization for Standardization, Geneva, Switzerland, 1993.

[10] ISO standard 3382-1: Measurement of room acoustic parameters - Part 1: Performance spaces. International Organization for Standardization, Geneva, Switzerland, 2009.

Bibliography

Abouchacra, K. S., Koehnke, J., Besing, J., & Letowski, T. (2011). Sentence recognition in the presence of competing speech messages presented in audiometric booths with reverberation times of 0.4 and 0.6 seconds. *Archives of Acoustics*, *36*(1), 3–14. https://doi.org/10.2478/v10168-011-0001-4

Ahrens, A., Marschall, M., & Dau, T. (2019). Measuring and modeling speech intelligibility in real and loudspeaker-based virtual sound environments. *Hearing research*, *377*, 307–317.

Ahrens, J., & Spors, S. (2008). An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions. *Acta Acustica United With Acustica*, *94*(6), 988–999. https://doi.org/10.3813/aaa.918115

Akaike, H. (1974). A new look at the statistical model identification. *Ieee Transactions on Automatic Control*, *AC-19*(6), 716–23, 716–723. https://doi.org/10.1109/TAC.1974.1100705

ANSIS3.5. (1997). Methods for the calculation of the speech intelligibility index. new york: American national standards institute.

Arweiler, I., Buchholz, J., & Dau, T. (2009). Speech intelligibility enhancement by early reflections.

ASHA. (1991). American speech-language-hearing association: Sound field measurement tutorial.

AS/NZS2107. (2000). Acoustics – recommended design sound levels and reverberation times for building interiors. australian/new zealand standards: Sydney/wellington.

Bamford, J., & McSporran, E. (1993). Visual reinforcement audiometry. *Paediatric audiology, 0–5 years*, 124–154.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (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

Bistafa, S. R., & Bradley, J. S. (2000). Reverberation time and maximum background-noise level for classrooms from a comparative study of speech intelligibility metrics. *Journal of the Acoustical Society of America*, *107*(2), 861–875. https://doi.org/10.1121/1.428268

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., & Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, *44*(2), 264–285. https://doi.org/10.1044/1092-4388(2001/022)

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

Boothroy, A. (1984). Auditory perception of speech contrasts by subjects with sensori neural hearing loss. *Journal of Speech and Hearing Research*, *27*(1), 134–144, 134–144.

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. *Journal of the Acoustical Society of America*, *106*(4 I), 1820–1828. https://doi.org/10.1121/1.427932

BSA. (2018). *British society of audiology. practice guidance: Guidance on the verification of hearing devices using probe microphone measurements*. Retrieved 2018, from https://www.thebsa.org.uk/wp-content/uploads/2018/05/REMS-2018.pdf

BSA. (2019a). *British society of audiology. assessment of speech understanding in noise in adults with hearing difficulties*. Retrieved 2019, from https://www.thebsa.org.uk/wp-content/uploads/2019/04/OD104-80-BSA-Practice-Guidance-Speech-in-Noise-FINAL.Feb-2019.pdf

BSA. (2019b). *British society of audiology. practice guidance: The acoustics of sound field audiometry in clinical audiological applications*. Retrieved 2019, from 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. (2015). Automated auditory response detection: Further improvement of the statistical test strategy by using progressive test steps of iteration. *International Journal of Audiology*, *54*(8), 568–572. https://doi.org/10.3109/14992027.2015.1017659

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

Clarke, S. (2011). Acoustic design approach for hospitals. *Australian Acoustical Society-conference 2011, Acoustics 2011: Breaking New Ground*, 680–685.

Cubick, J., & Dau, T. (2016). Validation of a virtual sound environment system for testing hearing aids. *Acta Acustica United With Acustica*, *102*(3), 547–557. https://doi.org/10.3813/aaa.918972

Damarla, V. K. S., & 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

Daniel, J. (2003). Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format, In *Audio engineering society conference: 23rd international conference: Signal processing in audio recording and reproduction*. Audio Engineering Society.

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. https://doi.org/10.3109/03005364000000148

DBR. (2018). *Bygningsreglement: Vejledning om lydforhold § 368 - § 376. danish transport, construction and housing authority, 2017. copenhagen, denmark*. Retrieved 2018, from http://bygningsreglementet.dk/Tekniske-bestemmelser/17/Vejledninger

Dillon, H., & Walker, G. (1982). Comparison of stimuli used in sound field audiometric testing. *Journal of the Acoustical Society of America*, *71*(1), 161–172. https://doi.org/10.1121/1.387343

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), 100–11, 100–111. https://doi.org/10.1097/00003446-200104000-00003

DIN18041. (2016). Acoustic quality in rooms - requirements, recommendations and indications for planning. berlin.

Dirks, D., Stream, R., & Wilson, R. (1972). Speech audiometry - earphone and sound field. *Journal of Speech and Hearing Disorders*, *37*(2), 162–176. https://doi.org/10.1044/jshd.3702.176

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. *Journal of the Acoustical Society of America*, *100*(4), 2236–2246. https://doi.org/10.1121/1.417933

Don, M., & Elberling, C. (1994). Evaluating residual background noise in human auditory brain-stem responses. *Journal of the Acoustical Society of America*, *96*(5 PART 1), 2746–2757, 2746–2757.

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. https://doi.org/10.1097/aud.0000000000000199

Elberling, C., & Don, M. (2010). A direct approach for the design of chirp stimuli used for the recording of auditory brainstem responses. *Journal of the Acoustical Society of America*, *128*(5), 2955–2964. https://doi.org/10.1121/1.3489111

Elberling, C., Don, M., Cebulla, M., & Stürzebecher, E. (2007). Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay. *Journal of the Acoustical Society of America*, *122*(5), 2772–2785. https://doi.org/10.1121/1.2783985

Erenberg, A., Lemons, J., Sia, C., Tunkel, D., Ziring, P., Adams, M., Holstrum, J., McPherson, M., Paneth, N., & Strickland, B. (1999). Newborn and infant hearing

loss: Detection and intervention. *Pediatrics*, *103*(2), 527–530. https://doi.org/10.1542/peds.103.2.527

Evans, J., & Himmel, C. (2012). Acoustical standards and criteria documentation of sustainability in hospital design and construction.

Fausti, P., Santoni, A., & Secchi, S. (2019). Noise control in hospitals: Considerations on regulations, design and real situations. *Inter-noise 2019 Madrid - 48th International Congress and Exhibition on Noise Control Engineering*.

Favrot, S., & Buchholz, J. (2010). Lora: A loudspeaker-based room auralization system. *Acustica United With Acta Acustica*, *96*(2), 364–375. https://doi.org/10.3813/aaa.918285

Favrot, S., & Buchholz, J. M. (2009). Validation of a loudspeaker-based room auralization system using speech intelligibility measures. *126th Audio Engineering Society Convention 2009*, *2*, 579–587.

Grimm, G., Ewert, S., & Hohmann, V. (2015). Evaluation of spatial audio reproduction schemes for application in hearing aid research. *Acta Acustica United With Acustica*, *101*(4), 842–854. https://doi.org/10.3813/aaa.918878

Hernández-Pérez, 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

Hirschorn, M., & Singer, E. (1989). Reverberation time descriptors for audiometric testing environments. *Sound and Vibration*, *23*(12), 24–6, 24–26.

Hodgson, M., & Nosal, E. M. (2002). Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms. *Journal of the Acoustical Society of America*, *111*(2), 931–939. https://doi.org/10.1121/1.1428264

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. https://doi.org/10.3109/14992027.2010.506889

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

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

HTM08-01. (2013). Health technical memorandum htm 08-01: Acoustics. uk department of health.

HTM2045. (1996). Health technical memorandum 2045 acoustics: Design considerations. HMSO, London, UK.

Ickes, M. A., Hawkins, D. B., & Cooper, W. A. (1991). Effect of reference microphone location and loudspeaker azimuth on probe tube microphone measurements. *Journal of the American Academy of Audiology*, *2*(3), 156–63, 156–163.

IEC268-13. (1985). 268-13,“sound system equipment-part 13: Listening tests on loudspeakers”.

Iglehart, F. (2019). Speech perception in classroom acoustics by children with hearing loss and wearing hearing aids. *American Journal of Audiology*, 1–12, 1–12. https://doi.org/10.1044/2019_aja-19-0010

ISO3382-1. (2009). Acoustics – measurement of room acoustic parameters – part 1: Performance spaces.

ISO60268-16. (2011). Sound system equipment – part 16: Objective rating of speech intelligibility by speech transmission index.

ISO8253-2. (2010). Acoustics – audiometric test methods – part 2: Sound field audiometry with pure-tone and narrow-band test signals.

ISO9921. (2003). Ergonomics – assessment of speech communication.

ISO/TR22411. (2008). Ergonomics data and guidelines for the application of iso/iec guide 71 to products and services to address the needs of older persons and persons with disabilities: Données d'ergonomie et lignes directrices pour l'application du guide iso/cei 71 aux produits et services afin de répondre aux besoins des personnes Âgées et de celles ayant des incapacité.

IST45. (2016). Hljóðvist - flokkun íbúðar- og atvinnuhúsnæðis. icelandic standards, iceland. (acoustic conditions in buildings - sound classification of various types of buildings).

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

Jeong, C.-H. (2011). Guideline for adopting the local reaction assumption for porous absorbers in terms of random incidence absorption coefficients. *Acustica United With Acta Acustica*, *97*(5), 779–790. https://doi.org/10.3813/aaa.918458

Johannesma, P. (1972). The pre-response stimulus ensemble of neurons in the cochlear nucleus, In *Symposium on hearing theory, 1972*. IPO.

John, M. S., Dimitrijevic, A., Van Roon, P., & Picton, T. W. (2001). Multiple auditory steady-state responses to am and fm stimuli. *Audiology and Neuro-otology*, *6*(1), 12–27. https://doi.org/10.1159/000046805

Killion, M. C., & Revit, L. J. (1987). Insertion gain repeatability versus loudspeaker location you want me to put my loudspeaker where? *Ear and Hearing*, *8*(5 SUPPL), 68S–73S, 68S–73S.

Kuwada, S., Batra, R., & Maher, V. (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*(ISAAR Special Issue:), 2331216518789302, 2331216518789302. https://doi.org/10.1177/2331216518789302

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

Machimbarrena, M., Rasmussen, B., & Alves Monteiro, C. R. (2019). Regulatory sound insulation requirements in south america-status for housing, schools, hospitals

and office buildings, In *Inter-noise and noise-con congress and conference proceedings*. Institute of Noise Control Engineering.

Marbjerg, G., Brunskog, J., Jeong, C.-H., & Nilsson, E. (2015). Development and validation of a combined phased acoustical radiosity and image source model for predicting sound fields in rooms. *Journal of the Acoustical Society of America*, *138*(3), 1457–1468. https://doi.org/10.1121/1.4928297

Marbjerg, G., Brunskog, J., Jeong, C.-H., & Zapata Rodriguez, V. (2018). The influence of overlapping band filters on octave band decay curves. *Acustica United With Acta Acustica*, *104*(6), 943–946. https://doi.org/10.3813/aaa.919259

Marbjerg, G., Zapata-Rodriguez, V., Brunskog, J., & Jeong, C.-H. (2020). Comparing loudspeaker array reproduction techniques using a phased combination of the image source method and acoustical radiosity. *Submitted to Acta Acustica*.

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

Miki, Y. (1990). Acoustical properties of porous materials—modifications of delany-bazley models. *Journal of the Acoustical Society of Japan (e)*, *11*(1), 19–24, 19–24. https://doi.org/10.1250/ast.11.19

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

Moeller, M. P., Osberger, M. J., & Eccarius, M. (1986). Language and learning skills of hearing-impaired students. receptive language skills. *Asha Monographs*, (23), 41–53, 41–53.

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

Narayanan, S. K., Laugesen, S., Zapata Rodriguez, V., Brunskog, J., & Jeong, C.-H. (2019). Effect of head-movment on sound-field auditory steady state response meansurement. *Proceedings of 23rd International Congress on Acoustics*, 7855–62.

NBE-CA-82. (1982). Real decreto 2115/82, de 12 de agosto. norma básica de la edificación. nbe-ca-82, sobre condiciones acústicas en los edificios (basic building regulation on acoustic conditions in buildings) (in spanish).

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

Neumann, K., Euler, H. A., Knauth, M., & White, K. (2015). *The global status of newborn and infant hearing screening. presented at the 6th annual coalition for global hearing health conference, washington d.c., october 09-10*. Google Machine Translation Team. Retrieved 2020, from http://conference.usu.edu/SYSTEM/Uploads/pdfs/15233_1964KatrinNeumann.pdf

Nielsen, J. B., & Dau, T. (2010). The danish hearing in noise test. *International Journal of Audiology*, *50*(3), 202–208. https://doi.org/10.3109/14992027.2010.524254

Nocke, C. (2016). The new standard din 18041 - acoustic quality in rooms. *Larmbekampfung*, *11*(2), 50–55.

NS8175. (2012). Lydforhold i bygninger - lydklasser for ulike bygningstyper. standards norway. (acoustic conditions in buildings - sound classification of various types of buildings).

Oreinos, C., & Buchholz, J. M. (2015). Objective analysis of ambisonics for hearing aid applications: Effect of listener's head, room reverberation, and directional microphones. *Journal of the Acoustical Society of America*, *137*(6), 3447–3465. https://doi.org/10.1121/1.4919330

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 and Speech Research*, *9*(1), 15–24. https://doi.org/10.21848/audiol.2013.9.1.15

Payton, K. L., & Braida, L. D. (1999). A method to determine the speech transmission index from speech waveforms. *Journal of the Acoustical Society of America*, *106*(6), 3637–3648. https://doi.org/10.1121/1.428216

Picton, 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*(5), 315–331.

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

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

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

Ptok, M. (2011). Früherkennung von schwerhörigkeiten im neugeborenen- und säuglingsalter, early detection of hearing impairment in newborns and infants. *Deutsches Arzteblatt*, *108*(25), 426–431. 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.

Rasmussen, B. (2018a). Acoustic classification of buildings in europe – main characteristics of national schemes for housing, schools, hospitals and office buildings. *Proceedings of Euronoise 2018*.

Rasmussen, B. (2018b). Acoustic classification of buildings in europe – main characteristics of national schemes for housing, schools, hospitals and office buildings. *Proceedings of Euronoise 2018*.

Rees, A., Green, G., & Kay, R. (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. *Journal of the Acoustical Society of America*, *140*(4), 2670–2679. https://doi.org/10.1121/1.4964505

Rochlin, G. D. (1993). Status of sound field audiometry among audiologists in the united states. *Journal of the American Academy of Audiology*, *4*(2), 59–68, 59–68.

Rønne, F. M. (2013). *Modeling auditory evoked potentials to complex stimuli*. Retrieved 2012, from http://orbit.dtu.dk/fedora/objects/orbit:127704/datastreams/file_250c6e0a-10de-405b-bfa2-2ab780406324/content

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. *Journal of the Acoustical Society of America*, *108*(2), 679–691. https://doi.org/10.1121/1.429600

Ruben, R. (1991). Effectiveness and efficacy of early detection of hearing impairment in children. *Acta Oto-laryngologica*, 127–135.

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

Schroeder. (1981). Modulation transfer-functions - definition and measurement. *Acustica*, *49*(3), 179–182.

Schroeder, & Kuttruff. (1962). On frequency response curves in rooms. comparison of experimental, theoretical and monte carlo results for the average frequency spacing between maxima. *Journal of the Acoustical Society of America*, *34*(1), 76–80, 76–80. https://doi.org/10.1121/1.1909022

Searle, S., Speed, F., & Milliken, G. (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

Seeber, B. U., Kerber, S., & Hafter, E. R. (2010). A system to simulate and reproduce audio-visual environments for spatial hearing research. *Hearing Research*, *260*(1-2), 1–10. https://doi.org/10.1016/j.heares.2009.11.004

Selim, M. H., Mourad, M. E., El-Shennawy, A. M., Elfouly, H. S., Et al. (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.

SFS5907. (2004). Rakennusten akustinen luokitus, finland. (english version, 2005: Acoustic classification of spaces in buildings).

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

Shaw, P., & Greenwood, H. (2012). The effect of head movement and head positioning on sound field audiometry. *International Journal of Audiology*, *51*(6), 499–504. https://doi.org/10.3109/14992027.2012.666361

Shaw, P., & Nikolopoulos, T. (2004). The effect of initial stimulus type for visual reinforcement audiometry. *International Journal of Audiology*, *43*(4), 193–197. https://doi.org/10.1080/14992020400050027

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.

Singh, V. (2015). Newborn hearing screening: Present scenario. *Indian Journal of Community Medicine*, *40*(1), 62–65. 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

SS25268. (2007). + t1:2017, byggakustik - ljudklassning av utrymmen i byggnader - vårdlokaler, undervisningslokaler, dag- och fritidshem, kontor och hotell. swedish standards institute, stockholm, sweden. (acoustics - sound classification of spaces in buildings - institutional premises, rooms for education, preschools and leisure-time centres, rooms for office work and hotels).

Steeneken, H. J. M., & Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *Journal of the Acoustical Society of America*, *67*(1), 318–26, 318–326. https://doi.org/10.1121/1.384464

Stream, R., & Dirks, D. (1974). Effect of loudspeaker position on differences between earphone and free-field thresholds (map and maf). *Journal of Speech and Hearing Research*, *17*(4), 549–568. https://doi.org/10.1044/jshr.1704.549

Stroebel, D., Swanepoel, D. W., & Groenewald, E. (2007). Aided auditory steady-state responses in infants. *International Journal of Audiology*, *46*(6), 287–292. https://doi.org/10.1080/14992020701212630

Swan, I. R., & Gatehouse, S. (1995). The value of routine in-the-ear measurement of hearing aid gain. *British Journal of Audiology*, *29*(5), 271–277. https://doi.org/10.3109/03005369509076742

Sykes, D., Tocci, G. C., & Cavanaugh, W. J. (2012). *Sound & vibration 2.0: Design guidelines for health care facilities*. Springer Science & Business Media.

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

Uhler, K. M., Hunter, S. K., Tierney, E., & Gilley, P. M. (2018). The relationship between mismatch response and the acoustic change complex in normal hearing infants. *Clinical Neurophysiology*, *129*(6), 1148–1160. https://doi.org/10.1016/j.clinph.2018.02.132

UNI11532. (2018). Caratteristiche acustiche interne di ambienti confinati – metodi di progettazione e tecniche di valutazione – parte 1: Requisiti generali (indoor acoustic characteristics of confined environments - design methods and evaluation techniques - part 1: General requirements) (in italian).

Van Wyk, K., Horan, D., & Murphy, K. (2014). A summary of the 2014 fgi and sound vibration guidelines for healthcare facilities. *Internoise 2014 - 43rd International Congress on Noise Control Engineering: Improving the World Through Noise Control*.

Walker, G., Dillon, H., & Byrne, D. (1984). Sound field audiometry - recommended stimuli and procedures. *Ear and Hearing*, *5*(1), 13–21. https://doi.org/10.1097/00003446-198401000-00005

Watson, S. D., Laugesen, S., & Epp, B. (2019). Provoking and minimising potentially destructive binaural stimulation effects in auditory steady-state response (assr) measurements, In *Proceedings of the international symposium on auditory and audiological research*.

Westermann, A., & Buchholz, J. M. (2017). The effect of nearby maskers on speech intelligibility in reverberant, multi-talker environmentsa). *Journal of the Acoustical Society of America*, *141*(3), 2214–2223. https://doi.org/10.1121/1.4979000

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. https://doi.org/10.3109/14992027.2014.989548

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

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

Zapata-Rodríguez, V., Jeong, C.-H., Zaar, J., Laugesen, S., Brunskog, J., & Harte, J. M. (2020b). Evaluation of acoustical characteristics of audiological clinic rooms for sound field audiometric testing. *In preparation for the International Journal of Audiology*.

Zapata-Rodríguez, V., Laugesen, S., Cebulla, M., Jeong, C.-H., Brunskog, J., & Harte, J. M. (2020d). Effect of room acoustics on sound-field auditory steady-state sesponses for hearing aid fitting validation. *In preparation for Trends in Hearing*.

Zapata-Rodríguez, V., Laugesen, S., Jeong, C.-H., Brunskog, J., & Harte, J. M. (2020a). Does room acoustics affect the amplitude of sound-field auditory steady-state responses? *Under review in Trends in Hearing*.

Zapata-Rodríguez, V., Laugesen, S., Marbjerg, G., Jeong, C.-H., Brunskog, J., & Harte, J. M. (2020c). Reproduction of nearby sources using loudspeaker-based virtual acoustic environments. *Submitted to The Journal of the Acoustical Society of America Express Letters*.

Zoontjens, L., & Cockings, T. (2014). Review of design approaches to acoustics in australian hospitals. *Internoise 2014 - 43rd International Congress on Noise Control Engineering: Improving the World Through Noise Control*.

DTU Electrical Engineering
Department of Electrical Engineering
Technical University of Denmark

Ørsteds Plads
Building 348
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 38 00

www.elektro.dtu.dk