

Untersuchung zur Sprachverständlichkeit bei Schulkindern unter Variation der tieffrequenten Nachhallzeit

Dustin Selbach¹, Jan Selzer², André Fiebig³

¹ *Fachgebiet Audiokommunikation, TU Berlin, 10587 Berlin, Deutschland, Email: d.selbach@campus.tu-berlin.de*

² *Institut für Arbeitsschutz der DGUV, 53757 Sankt Augustin, Deutschland, Email: jan.selzer@dguv.de*

³ *Fachgebiet Technische Akustik, TU Berlin, 10587 Berlin, Deutschland, Email: andre.fiebig@tu-berlin.de*

Einleitung

An Unterrichtsräume werden hohe Anforderungen an die Nachhallzeit (NHZ) gestellt [1]. Studien zeigen, dass Sprachverstehen von Kindern und die NHZ im Raum in einem direkten Zusammenhang stehen [2, 3]. In den genannten Studien werden mittlere NHZ (z. B. Mittelwert der NHZ aus den Oktaven zwischen 250 Hz - 2 kHz [3]) herangezogen. Studien zum Einfluss von tieffrequentem Nachhall auf die Sprachverständlichkeit (SV) unter Anwendung von Hörversuchen sind wenige bekannt [4, 5]. Keine der genannten Studien bezieht Kinder als maßgebendes Auditorium in Bildungsstätten, und im Vergleich mit Erwachsenen stärker beeinflussten Gruppe [2], in die Hörversuche ein. Um auf diesem Gebiet weitere Erkenntnisse zu sammeln, wurde ein Hörversuch mit Kindern realisiert, in dem systematisch die NHZ in den Oktavbändern 125 Hz und 250 Hz variiert wurde.

Stand der Technik und Forschung

Zum Stand der Technik sind maßgeblich die Regelwerke ASR A3.7 [6] sowie die DIN 18041 [1] zu nennen. Anzu merken ist, dass für Unterrichtsräume beide Regelwerke identische Anforderungen an die mittlere Soll-NHZ stellen. Verschieden sind die Anforderungen im zu betrachtenden Frequenzbereich. Während die ASR A3.7 sich auf die Oktavbänder zwischen 250 Hz - 2000 Hz beschränkt, erweitert die DIN 18041 diese Anforderungen für den Oktavbereich von 125 Hz - 4000 Hz. In den durch die ASR A3.7 abgedeckten Oktaven sind Toleranzen von $\pm 20\%$ in beiden Regelwerken vorgesehen, in den übrigen Oktaven sind in DIN 18041 auch größere Abweichungen zulässig (vgl. hierzu Abb. 2).

Beide Regelwerke weisen darauf hin, dass im Kontext einer inklusiven Pädagogik verkürzte NHZ erforderlich seien. Im Gegensatz zur ASR A3.7 formuliert die DIN 18041 konkrete Anforderungen für Inklusion.

Um die Vergleichbarkeit mit dem staatlichen Regelwerk beizubehalten, wird Gleichung 1 zur Berechnung von Soll-NHZ für Unterrichtsräume herangezogen [1, 6].

$$T_{\text{Soll}} = \left(0.32 \lg \frac{V}{\text{m}^3} - 0.17 \right) \text{ s} \quad (1)$$

Yang et al. zeigen, dass Kinder im Grundschulalter geringere SV erreichen als Erwachsene [2]. Weiterhin wurde dargelegt, dass der Einfluss der NHZ auf die SV in Räumen mit zunehmenden Alter sinkt [2].

Demnach wird unterstellt, dass innerhalb der Zielgruppe der jüngeren Klassenstufen größere Beeinträchtigungen durch Änderungen im spektralen Verlauf der NHZ zu er-

warten sind.

Studien, die explizit tieffrequenten Nachhall auf die SV untersuchen, zeigen, dass das Oktavband um 250 Hz das Sprachverstehen beeinflussen kann [4, 5, 7]. Wu et al. ziehen zur Beschreibung des tieffrequenten Nachhalls das Bassverhältnis (BR):

$$BR = \frac{T_{125 \text{ Hz}} + T_{250 \text{ Hz}}}{T_{500 \text{ Hz}} + T_{1000 \text{ Hz}}} \quad (2)$$

heran und zeigen einen signifikanten Zusammenhang zwischen SV und BR [5]. Reents und Ala kommen in ihren Untersuchungen zu dem Schluss, dass die NHZ im Oktavband um 250 Hz Effekte auf die SV ausüben kann. Einen Einfluss durch der NHZ in den Oktaven < 250 Hz auf die SV kann in den Studien nicht signifikant nachgewiesen werden [4, 7].

Methoden

Der Großteil der vorgestellten Studien bedient sich raumakustischer Simulationen [2, 4, 5, 7], welche den Vorteil haben kostengünstig beliebige Zielkonditionen herzustellen. Aus dem genannten Grund wird ebenfalls von dieser Möglichkeit Gebrauch gemacht. Die SV wird unter Verwendung des Oldenburger Kinderreimtests (OLKI) [8] in einer Stichprobe der fünften Klasse ($n = 23$, $\bar{x} 10.2$ Jahre $\sigma = 0.4$) gemessen.

Raumakustische Simulationen

Klatte et al. berichten in ihrer Untersuchungen ein durchschnittliches Raumvolumen von 231 m^3 , wohingegen Yang et al. sich auf ein Raumvolumen von 198 m^3 berufen [2, 3]. In Verbindung mit den Empfehlungen der DGUV Information 202-090 [9] wurden die Raumdimensionen zu $9.6 \text{ m} \times 7.5 \text{ m} \times 3.0 \text{ m} = 216 \text{ m}^3$ festgesetzt. In Abbildung 1 ist der modellierte Klassenraum dargestellt. Zur Berechnung der Impulsantwort wird die Schallquelle im Abstand von 3.75 m zur Fensterbandwand, 1.24 m vor der Tafelwand und in einer Höhe von 1.60 m positioniert. Zur Berücksichtigung einer Richtcharakteristik der Quelle wurde eine Sopranistin angenommen [10], da für Sprecher keine Daten verfügbar waren. Als Empfängerposition ist der freie Platz in der zweiten Tischreihe am Fenster gewählt worden. Dieser liegt im Abstand von 1.625 m von der Fensterbandwand, und 5.71 m von der Tafelwand entfernt. Die Höhe wurde mit 0.80 m angesetzt.

Die Außenohrübertragungsfunktion (HRTF) von Kindern weicht deutlich von denen der Erwachsenen ab [11]. Im Kontext tiefer Frequenzen treten interaurale Laufzeitdifferenzen (ITD) in den Fokus. Um Wiedergabe-

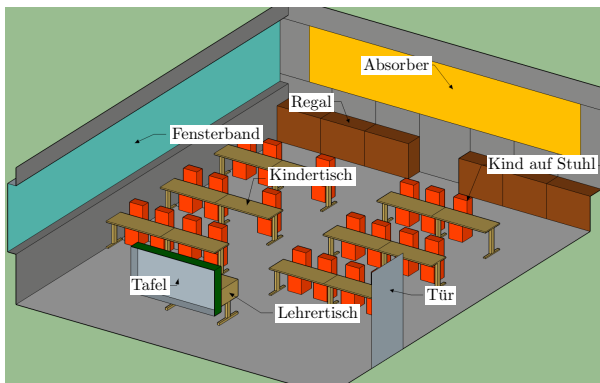


Abbildung 1: Einsicht in den modellierten Klassenraum

fehler zu reduzieren wurden die ITD anhand individueller, anthropometrischer Daten der Probanden angepasst. Grundlage der Anpassung ist das in [12] beschriebene Modell, welches bereits in anderen Untersuchungen erfolgreich eingesetzt wurde [13]. Die Basis der Anpassung ist der vollständige HRTF-Datensatz des durch Schmitz entwickelten Kunstkopfes [14].

Zur Berechnung der binauralen Raumimpulsantwort (BRIR) wurde RAVEN verwendet, welches Spiegelschallquellen und RayTracing verbindet [15]. In dieser Arbeit wurden Spiegelschallquellen 2. Ordnung sowie 50000 Partikel zur Berechnung der Impulsantwort verwendet. Die Raumtemperatur wurde mit 21° C berücksichtigt. Als initiale Absorptions- und Streugrade wurden Kennwerte in Anlehnung an die DIN 18041 [1] herangezogen.

Um den Einfluss der Nachhallzeit der zu untersuchenden Oktavbänder durch einen Hörversuch zu ermitteln, sind Variationen erforderlich. In Anlehnung an Wu et al. wird das BR zur Beschreibung dieser Variationen herangezogen. Für die mittleren Oktaven werden die Anforderungen der Raumgruppe A3 festgesetzt: $T_{500\text{ Hz}} = T_{1000\text{ Hz}} = T_{\text{Soll,A3}}$. In den Oktaven um 125 Hz und 250 Hz wird systematisch variiert. Ziel ist eine konstante Schrittweite von $\Delta\text{BR} = 0.2$ um den ausgeglichen Verlauf von $\text{BR} = 1.0$. Abbildung 2 dokumentiert die resultierenden Zielwerte. Die BRIR wurden mithilfe des Filters der FABIAN Datenbank [16] für den verwendeten Kopfhörer Audio-Technica ATH M50x entzerrt.

Hörversuch

Zur Erfassung der SV mittels Hörversuch kommt der evaluierte OLKI [8] zum Einsatz, welcher bereits in früheren Untersuchungen verwendet wurde [3]. Im Kontext der zum Zeitpunkt der Datenaufnahme vorherrschenden pandemischen Lage ist der Hörversuch in der Schule durchgeführt worden, um notwendige Bewegungen auf ein Mindestmaß zu reduzieren. Die Messung der SV erfolgte im Störgeräusch, mit einem Signal-Rausch-Abstand (SNR) von -2 dB. In Vorversuchen wurde festgestellt, dass hierdurch die SV im überschwelligen Bereich zu erwarten ist und gleichzeitig die SV ausreichend beeinträchtigt wird, sodass Änderungen durch die Raumakustik sichtbar in Erscheinung treten können. Die SV wird, entsprechend der Empfehlungen des OLKI, aus Testlisten mit 12 Antworten ermittelt [17, S.35]. Die Umsetzung erfolgt unter Verwendung der Oldenburger Messprogram-

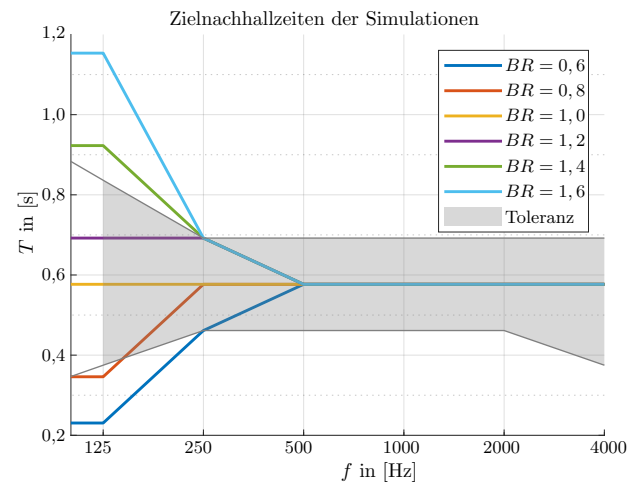


Abbildung 2: Zielwerte des spektralen Verlaufs der Nachhallzeit für die Simulationen mit Darstellung der zulässigen Toleranzen nach DIN 18041:2016 für die Raumgruppe A3 nach [1].

me R&D (OMA). Sprach- und Störsignal werden hierbei mit der BRIR gefaltet. Die Kalibrierung erfolgte innerhalb der Messumgebung unter Verwendung eines Kunstkopfes (HEAD acoustics HSU III.2) für den linken Kanal der Kondition $\text{BR} = 0.6$ und ist anschließend auf den rechten Kanal sowie die übrigen Situationen übertragen und kontrolliert worden. Zur Kalibrierung wurde ein Rauschsignal mit 80 dB eingesetzt. Die Kalibrierung der Zielpegel erfolgte in Vorbereitung zum Versuchsaufbau und wurde vor Ort stichprobenartig überprüft. Als Soundkarte kam die durch die Entwickler der Software empfohlene ESI MAYA22 USB zum Einsatz [17].

Vor Versuchsdurchführung sind sämtliche vorkommende Begriffe im Klassenverband spielerisch erarbeitet worden. Während der Durchführung des Hörversuchs wurde den Kindern ausschließlich das Antwortfenster des OLKI auf einem separaten Bildschirm gezeigt. Ungenutzte Monitorbereiche werden mit schwarzem Hintergrund präsentiert. Eine Computermouse kommt als Eingabewerkzeug zum Einsatz, während die Stimuli über Kopfhörer dargeboten werden. Das Studiendesign wurde durch die Ethikkommission der TU Berlin als ethisch unbedenklich eingestuft.

In Abbildung 3 ist der Aufbau im Klassenraum während der Überprüfung dargestellt. Weiterhin konnte in Vorversuchen festgestellt werden, dass vereinzelte VP Auffälligkeiten in der Bearbeitung aufwiesen und sich Trainingseffekte einstellen. Um die VP an den Versuch zu gewöhnen wurde die Testliste 1 mit der Kondition $\text{BR} = 1.0$ belegt. Über die folgenden 6 Listen (2-7) wurden randomisiert die akustischen Bedingungen dargeboten. Zur Erfassung großer Auffälligkeiten, welche zum Ausschluss der Auswertung führen, wurden die Kriterien nach Gleichung 3 angesetzt:

$$\text{Ausreißer} = \begin{cases} SV < 33,3\% - \sigma \\ SV < 33,3\% + \sigma \\ \overline{SV} < 33,3\% \end{cases} \quad (3)$$



Abbildung 3: Hörversuchsaufbau in der Schule während stichprobenartiger Überprüfung der Kalibrierung

Der gemessene Wert der SV in jeder Liste wird mit der Ratewahrscheinlichkeit ($Rw = 33.3\%$) abzüglich der Standardabweichung (σ) verglichen, um sicherzustellen, dass jede Liste ausreichend konzentriert bearbeitet wurde. Erreicht eine Versuchsperson (VP) in keiner Liste die Rw , korrigiert um σ , wird unterstellt, dass grundsätzlich geraten werden musste. Ebenso wird gefordert, dass die mittlere SV über alle Testlisten oberhalb der Rw liegt, um sicherzustellen, dass überschwellige Reize vorliegen. Die Auswertung der Daten erfolgte in RStudio.

Ergebnisse

Im Boxplot in Abbildung 4 ist das BR, ermittelt aus der monauralen RIR, über die Kategorien der Zielwerte aufgetragen. Es sind deutlich abgrenzbare Kategorien erkennbar. Hinsichtlich potentieller Trainingseffekte konnte

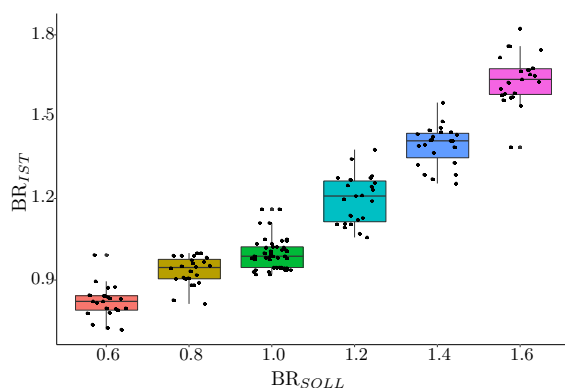


Abbildung 4: Abhängigkeit des BR, ausgewertet aus der monauralen RIR von den zuvor gesetzten Zielkategorien am Ort des Hörenden. Durch • ist die einzelne RIR dargestellt. Die X-Achsenverschiebung der Punkte dient lediglich der übersichtlicheren Darstellung.

unter Verwendung des Friedman-Tests kein signifikanter Unterschied zwischen den Testlisten festgestellt werden ($\chi^2(6) = 10.6, p = 0.10, n = 23$). Durch die Kriterien für Ausreißer wird eine VP weniger erkannt, wenn die Testliste 1 aus der Analyse entfernt wird. Demnach sind für eine VP Trainingseffekte zu unterstellen, sodass in der weiteren Datenauswertung lediglich die Testlisten 2-7 berücksichtigt sind. Insgesamt sind drei VP aus der Da-

tenauswertung entfernt worden. Zwei davon wurden als Ausreißer erfasst. Für eine VP lagen nicht für sämtliche Konditionen Antworten vor.

In Tabelle 1 sind die Ergebnisse der Post-hoc Analyse unter Anwendung eines gepaarten Wilcoxon-Tests mit Bonferroni-Korrektur angegeben. Hierbei konnte kein signifikanter Unterschied zwischen den Gruppen festgestellt werden.

Tabelle 1: Ergebnisse der Post-hoc Analyse (gepaarter Wilcoxon-Test mit Bonferroni-Korrektur) der gemessenen Sprachverständlichkeit in Zuordnung der Ziel-Kategorien ($n = 20$)

Gruppe 1	Gruppe 2	W	p	p _{korrr}
BR = 0.6	BR = 0.8	93.0	0.447	1
BR = 0.6	BR = 1.0	60.0	1	1
BR = 0.6	BR = 1.2	86.0	0.731	1
BR = 0.6	BR = 1.4	63.0	0.814	1
BR = 0.6	BR = 1.6	39.5	0.430	1
BR = 0.8	BR = 1.0	76.5	0.466	1
BR = 0.8	BR = 1.2	42.5	0.192	1
BR = 0.8	BR = 1.4	51.5	0.244	1
BR = 0.8	BR = 1.6	46.5	0.162	1
BR = 1.0	BR = 1.2	91.0	0.613	1
BR = 1.0	BR = 1.4	54.5	0.498	1
BR = 1.0	BR = 1.6	46.5	0.729	1
BR = 1.2	BR = 1.4	69.0	0.739	1
BR = 1.2	BR = 1.6	62.5	0.795	1
BR = 1.4	BR = 1.6	71.0	0.896	1

Weiterhin wurde unter Einbezug des BR aus den monauralen RIR ein Random-Intercept-Model mithilfe des lmer-Pakets berechnet [18]. Das Modell wird durch Gleichung 4 beschrieben:

$$Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + e_{ij} \quad (4)$$

mit:

- Y_{ij} : durch das Modell prognostizierte SV
- $\beta_{0j} = \beta_0 + u_{0j}$: Y-Achsenabschnitt
- $u_{0j} \sim N(0, \sigma_{u0}^2)$: zufälliger Schnittpunkt
- β_1 : Steigung
- X_{ij} : fixed Effekt; hier BR
- j : Clustervariable; hier VP
- i : Anzahl der Beobachtungen für jede VP

In Abbildung 5 sind die erfassten SV in Abhängigkeit des BR aufgetragen. Allgemein liegen die ermittelten Sprachverständlichkeiten im überschwelligen Bereich. Weiterhin ist die Regressionsgerade eingezeichnet. Die Analyse zeigt, dass durch das BR im errechneten Modell eine statistisch nicht signifikante ($p = 0.19$) Varianzaufklärung von $R_m^2 = 1.2\%$ ermöglicht wird und bestätigt damit die Ergebnisse der Post-hoc Analyse. Weiterhin wurde ein $R_k^2 = 18.9\%$ ermittelt, welches sich aufgrund der Unterschiede zwischen den Teilnehmenden erklärt.

Diskussion und Ausblick

Wie Abbildung 4 verdeutlicht, ist ein breiter Bereich im BR durch das Versuchsdesign abgedeckt worden. Die durch den OLKI erfasste SV zeigte in dieser Untersuchung keinen signifikanten Zusammenhang zwischen BR und SV. Die Ergebnisse sind damit konsistent mit den Er-

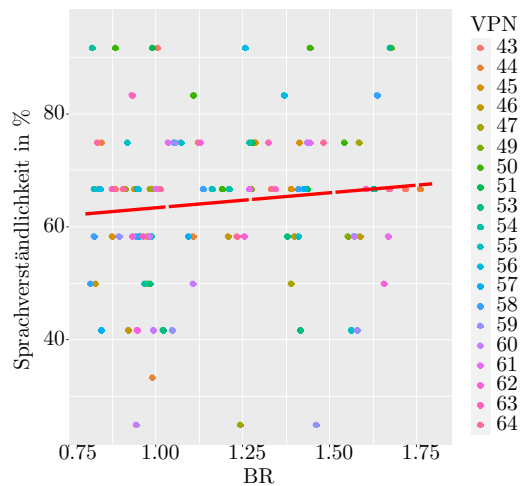


Abbildung 5: Ermittelte Sprachverständlichkeiten in Abhängigkeit des BR. In Rot dargestellt das Ergebnis des Linear-Mixed-Models.

gebnissen des Reimtests in der Untersuchung von Reents [4], obwohl die besonders vulnerable Gruppe der Kinder, die als sensibler für raumakustische Änderungen angenommen werden, für den Hörversuch herangezogen wurden.

Während im hier vorgestellten Verfahren 12 Abfragen je Kondition durchgeführt wurden, berichtet Wu von 25 Abfragen [5] und findet systematische Effekte bei Probanden in der Altersgruppe der 19-21 Jährigen. Im Kontrast dazu werden durch Reents 10 Abfragen für jede Kondition berichtet und mittels Paarvergleich systematische Effekte im Oktavband um 250 Hz gefunden. Demnach wird angenommen, dass die Nutzung der 24 Abfragen je Testliste im OLKI das Potential hat entsprechende Effekte genauer aufzuklären.

Folgeuntersuchungen erscheinen notwendig, um altersabhängige Aussagen hinsichtlich des Einflusses von tieffrequentem Nachhall treffen zu können sowie den Einfluss der absoluten Nachhallzeit mit variierendem BR zu untersuchen.

Literatur

- [1] DIN 18041: Hörsamkeit in Räumen - Anforderungen, Empfehlungen und Hinweise für die Planung. Beuth (2016), Berlin
- [2] Yang, W., Bradley, J.S.: Effects of room acoustics on the intelligibility of speech in classrooms for young children. The Journal of the Acoustical Society of America 125(2) (2009), S. 922-933
- [3] Klatte, M., Hellbrück, J., Seidel, J., Leistner, P.: Effects of Classroom Acoustics on Performance and Well-Being in Elementary School Children: A Field Study. Environment and Behavior 42(5) (2010), S. 659-692
- [4] Reents, P.: Einfluss der Nachhallzeit von tiefen Frequenzen auf die Sprachqualität. Diplomarbeit (2007), Institut für Breitbandkommunikation der Universität Graz, Graz, https://www2.spsc.tugraz.at/www-archive/downloads/DA_Reents.pdf
- [5] Wu, S., Peng, J., Bi, Z.: Chinese speech intelligibility in low frequency reverberation and noise in a simulated classroom. Acta Acustica united with Acustica 100(6) (2018), S. 1067-1072
- [6] ASR A 3.7: „Lärm“, Technische Regeln für Arbeitsstätten. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (2021), Dortmund
- [7] Ala, A.: Speech Intelligibility and Reverberation at Low Frequencies (2020), Fachgebiet Audiokommunikation der Technischen Universität Berlin, Berlin, https://www2.ak.tu-berlin.de/~akgroup/ak_pub/abschlussarbeiten/2020/Ala_MasA.pdf
- [8] Wagener, K.C., Brand, T., Kollmeier, B.: Evaluation des Oldenburger Kinder-Reimtests in Ruhe und im Störgeräusch. HNO, 54(3) (2006) S. 171-178
- [9] DGUV Information 202-090: Klasse(n) - Räume für Schulen - Empfehlungen für gesundheits- und lernfördernde Klassenzimmer. Deutsche Gesetzliche Unfallversicherung e. V. (2018), Berlin
- [10] Weinzierl, S. et al.: A Database of Anechoic Microphone Array Measurements of Musical Instruments. (2017) <http://dx.doi.org/10.14279/depositonce-5861.2>.
- [11] Fels, J., Buthmann P., Vorländer, M.: Head-related transfer functions of children. Acta Acustica united with Acustica, 90(5) (2004) S. 918-927
- [12] Bomhardt, R., Fels, J.: Analytical interaural time difference model for the individualization of arbitrary Head-Related Impulse Responses. Audio Engineering Society Convention 137. Audio Engineering Society (2014)
- [13] Loh, K., Fintor, E., Nolden, S. Fels, J.: Children's intentional switching of auditory selective attention in spatial and noisy acoustic environments in comparison to adults. Developmental Psychology (2021)
- [14] Schmitz, A.: Ein neues digitales Kopfhörersystem. Acta acustica united with acustica, 81(4) (1995) S. 416-420
- [15] Schröder, D., Vorländer, M.: RAVEN: A real-time framework for the auralization of interactive virtual environments. Forum acusticum. Aalborg Denmark (2011), S. 1541-1546
- [16] Brinkmann, F.: „The FABIAN head-related transfer function data base.“ (2017) <https://www.depositonce.tu-berlin.de/bitstream/11303/6153/8/Documentation.pdf>
- [17] gGmbH, HörTech: Bedienungsanleitung „Oldenburger Kinder-Reimtest“ für „Oldenburger Messprogramme“ ab Release 2.2.0.0. (2019)
- [18] Kuznetsova A., Brockhoff P. B., Christensen R. H. B.: lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software, 82(13) (2017) 1-27