

This is a repository copy of *Impact Thresholds of Parameters of Binaural Room Impulse Responses (BRIRs) on Perceptual Reverberation*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/184429/>

Version: Published Version

Article:

Mi, Huan, Kearney, Gavin orcid.org/0000-0002-0692-236X and Daffern, Helena orcid.org/0000-0001-5838-0120 (2022) Impact Thresholds of Parameters of Binaural Room Impulse Responses (BRIRs) on Perceptual Reverberation. Applied Sciences. 2823. ISSN 2076-3417

<https://doi.org/10.3390/app12062823>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:


<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Article

Impact Thresholds of Parameters of Binaural Room Impulse Responses (BRIRs) on Perceptual Reverberation

Huan Mi *, Gavin Kearney  and Helena Daffern 

AudioLab, Department of Electronic Engineering, University of York, York YO10 5DD, UK;
gavin.kearney@york.ac.uk (G.K.); helena.daffern@york.ac.uk (H.D.)

* Correspondence: hm1339@york.ac.uk

Abstract: This paper presents a study on the perceived importance of different acoustic parameters of Binaural Room Impulse Response (BRIR) rendering. A headphone-based listening test was conducted with twenty expert participants. Three BRIRs generated from simulations of three different rooms were convolved with a dry speech signal and used as reference audio samples. Four BRIR parameters, Initial Time Delay Gap (ITDG), Forward Early Reflections (FER), Reverse Early Reflections (RER) and Late Reverberation (LR) were systematically altered and convolved with a speech signal to generate the test conditions. A staircase method was used to obtain the threshold at which each BRIR parameter was perceived as different from the reference audio sample. The average perceived impact threshold of each parameter was then calculated across the twenty participants. Results show that RER removal and ITDG extension have a clear impact on the perceptual reverberation of speech audio. Subjects were less sensitive to FER removal. The effect of LR removal on perceptual reverberation is hard to distinguish. Therefore, RER and ITDG are of particular importance when designing artificial reverberation algorithms, whilst more research is needed to understand the perceptual contribution of LR. Minor changes in FER and LR are less significant.

Keywords: BRIR; ITDG; RER; FER; LR; staircase method; threshold



Citation: Mi, H.; Kearney, G.; Daffern, H. Impact Thresholds of Parameters of Binaural Room Impulse Responses (BRIRs) on Perceptual Reverberation. *Appl. Sci.* **2022**, *12*, 2823. <https://doi.org/10.3390/app12062823>

Academic Editor: Hyunkook Lee

Received: 30 October 2021

Accepted: 1 March 2022

Published: 9 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the increasing popularity of Augmented Reality (AR) technologies, research into the plausible reproduction of virtual acoustic scenes that match the real world has gained importance [1]. AR experiences typically involve the superposition of virtual visual and aural elements onto visual and auditory displays that also show the real world (such as the screen on a mobile phone passing through the video from the camera and AR headphones that pass through real world audio via external microphones). Consequently, the rendering of the virtual audio should seamlessly match the real world acoustic for a convincing experience. Head tracking, binaural filtering using Head Related Transfer Functions (HRTFs) and artificial reverberation are three factors that contribute to producing a realistic spatial hearing experience with accurate localisation and timbre [2]. For reverberation in particular, the development of computationally efficient dynamic reverberation algorithms is essential to provide plausible virtual acoustic rendering on low-cost mobile devices.

Reverberation is usually described by Room Impulse Responses (RIRs) in acoustics research. A RIR is the resultant pressure fluctuation measured at a receiving point due to an impulsive sound source at an arbitrary location in a room [3]. Room auralisation is achieved through convolution of source audio with a computationally derived or physically measured RIR [4]. Acoustic parameters that make up RIRs, including Early Reflections (ERs), Initial Time Delay Gap (ITDG) and Late Reverberation (LR), influence the resultant perceived reverberation.

It is traditionally thought that in order to match reverberation in AR environments to real world reverberation, these parameters need to be as close as possible to corresponding

measured RIRs. Such reverberators should be dynamic, in that the acoustic parameters should change with the six degrees of freedom movement of the listener. They should also be rendered using binaural audio techniques for effective spatialisation. The reverberation proposed in this paper therefore refers to binaural reverberation and the corresponding RIRs are Binaural Room Impulse Responses (BRIRs); however, rendering with real-world BRIRs for AR experiences requires prior knowledge of the location that the AR experience will be consumed. Computational derivation of an unknown room's acoustic for an AR experience, for example through visual depth mapping [5] or on the fly acoustic excitation [6] is challenging and requires a computationally based real-time reverberation model, even if some room acoustics data are physically extracted through measurement.

Alternatives to direct convolution of source material with BRIRs include the use of reverberators entirely based on geometric or wave-based acoustic models or feedback delay networks (FDNs) [7]. Hybrid models that, for example, utilise both FDN reverberation alongside components of the RIR (such as the late reverberation) can also be conceived.

In reality, AR enabled systems should use portable, low-power mobile computing devices and reverberation algorithms with low computational complexity [3]; however, in an effort to simplify artificial reverberation algorithms, it is necessary to understand the influence of the acoustic parameters on the overall perceived reverberation and the thresholds of perception for each of these parameters. Parameters with the most perceptual relevance could therefore be emphasised in the reverberator design and other parameters given less importance.

It is therefore the purpose of this paper to explore the impact thresholds of parameters of BRIRs on perceptual reverberation, providing the foundational research for new artificial reverberation algorithms.

1.1. Composition of Room Impulse Responses

In an enclosed space a proportion of the radiated sound source will be reflected off the room boundaries and eventually decay due to absorption by the room surfaces or air. Any impulsive stimulus will rapidly change the nature of the soundfield from being coherent to partially coherent to non-coherent. Thus, this linear time invariant system is characterised by a Room Impulse Response, consisting of the following components (shown in Figure 1):

- The Direct Sound (DS): The DS reaches a listeners' ears directly from the source before being reflected from the boundaries of the enclosure [8]. Its amplitude is large with less energy loss relative to the reflections because of the shorter propagation path. Its function is to transmit sound information and provide the direction of source.
- The Early Reflections (ERs): These are the sound waves that arrive in a temporal order after being reflected from at least one boundary of the enclosure [8]. They arrive within typically 10 to 80 ms after the direct sound, and typically constitute up to fourth order reflections before the soundfield becomes stochastic. Their energy is reduced by absorption or scattering. The early reflections can increase perceived overall sound pressure level and sound clarity.
- The Late Reverberation (LR): is a chaotic sound field that consists of diffuse reflections [8]. It is an exponentially attenuated dense collection of echoes diffusing in all directions. The echo density is proportional to the square of time. An appropriate amount of late reverberation can contribute to a sense of spatialisation and fullness, although too much can destroy the clarity of the sound.
- The Initial Time Delay Gap (ITDG): This is the time period between the direct sound and the first arriving reflection. ITDG is the main contributor towards the perception of 'presence' [9], an attribute that is recognised as the perceptual sense of feeling boundaries of an enclosed space [10]. It is the hearing-equivalent of 'seeing' the walls of a room [11].

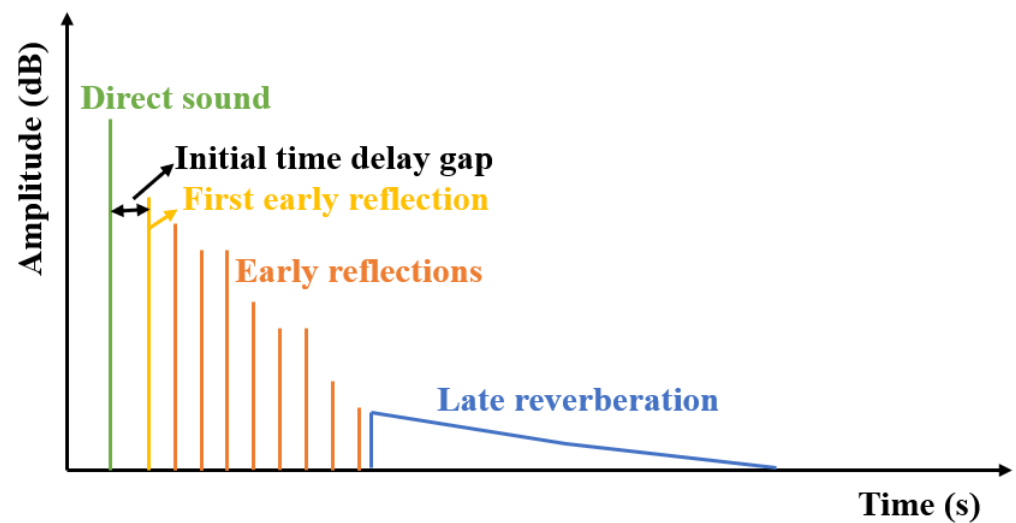


Figure 1. Representation of a room impulse response.

If a binaural dummy head is used to record the impulse, then the resultant IR is known as a Binaural Room Impulse Response (BRIR) and contains both the room characteristics and spectral and temporal binaural cues. Binaural auralisation therefore uses two channels to simulate the binaural listening experience typically over a pair of headphones but also over loudspeakers using cross-talk cancellation methods [12].

1.2. Perception of Different Binaural Room Impulse Response Parameters

Many experts have conducted in-depth research on the perceptual properties of room acoustics and their relation to binaural parameters. Hartmann studied whether the ability to localise sound in a room depends on room acoustics or on the nature of the source signal. His research showed that the localisation of sources with strong attack transients is independent of room reverberation time, although it may depend on the geometry of the room. For sources without strong attack transients, the localisation increases monotonically with the spectral density of the source [13]. This auditory localisation was conducted in a room with a single acoustic reflecting surface whose position changed to simulate floors, ceilings and left or right walls. He measured the steady-state interaural time difference (ITD) and interaural intensity difference (IID) cues available to subjects in different room structures, and compared these data with perceptual judgement [14]. He found that the precedence effect can help the localisation of sound in rooms, but it cannot eliminate all influences of room reflections. Further, the influence of reflections may cause large interaural intensity differences in a room, which have a considerable impact on localisation [15].

Hyde [16] observed that a short ITDG generally indicates an important contributing factor to acoustical quality in a hall through discussing its relation to acoustical intimacy. Beranek [17] reported that the listener's impression of the size of a hall is determined by the time delay of the first major reflection after the direct sound. He also observed that halls that have intimate acoustics had ITDG values at or shorter than 20 ms, and that the shorter the ITDG, the more intimate the experience [17]. He also stated that with a short ITDG, more reflections can occur in the first 80 ms after the arrival of the direct sound, and more early reflections contribute to a greater feeling of intimacy [18].

Research has also shown the relevance of early and late reflections for speech perception in reverberant rooms [19]. The reflections modify the perception of the sound, changing the loudness, timbre and most importantly, the spatial characteristics of the sound [20]. Early reflections are important for localisation of the sound source and to give a listener an impression about the size and the shape of the room, as well as about the place and the orientation of the listener inside the room [21]. Early reflections arriving with the

first 50 ms after the direct sound are integrated for directional cues rather than perceived separately [19]. If late reflections are too weak, a RIR can sound dry. If the late reflections are too strong, the sound is confusing and unintelligible. When a RIR has appropriate strong late reflections, under certain conditions, early reflections have a conducive effect on recognition accuracy and a sense of space [19]. Golzer and Kleinschmidt [19] investigated the importance of certain portions of the impulse response in different contexts. They evaluated the importance of early and late reflections for the accuracy of automatic speech recognition and determined the effective time cutoff between conducive and detrimental portions of the impulse response. They found that when a harmful late portion is removed, early reflections up to a certain critical delay time can carry useful information and contribute to the automatic speech recognition accuracy, and for different room impulse responses, the cutoff time is in the range of 25 to 50 ms [19].

Lindau [22] et al. evaluated physical predictors of the mixing time in binaural room impulse responses. The certain transition time from early reflections to late reverberation tail is called mixing time [23]. By adaptively changing the mixing time in real time, the audible transition time into a homogeneous late reverberation tail can be determined to reduce the length of binaural impulse response [22].

The above research indicates that ITDG, early and late reflections can affect auditory perception in a reverberant environment, but the contribution of each component of the BRIR and the thresholds of perceptibility under different reverberant conditions still remains uncertain.

The paper is organised as follows. Section 2 presents experimental materials and methods. The experimental results are presented in Section 3. Section 4 discusses the results and Section 5 gives the conclusions of this experiment.

2. Materials and Methods

The perception of reverberation relates the actual intensity of a sound source in a reverberant room to the perceived intensity. When a parameter of a BRIR changes, the reverberation effect of the BRIR will change, but this change may not be detected by the listener. Each of the aforementioned parameters of an impulse response needs to be tested for perceptual thresholds so that they can be regarded as important or not in the development of artificial reverberation algorithms.

2.1. Experimental Stimuli

A listening test was designed to measure the perceptual thresholds of four BRIR parameters, Reverse Early Reflection (RER) removal, Forward Early Reflection (FER) removal, Late Reverberation (LR) removal and Initial Time Delay Gap (ITDG) extension of three different rooms. Since in the study of artificial reverberation the RIR is usually divided into two parts, one including direct sound and early reflections and the other including late reverberation [24], DS was not used as a parameter for the measured threshold in this study. Forward Early Reflection (FER) removal has been implemented through the removal of the initial reflections with subsequent tests removing further reflections forward from the initial reflections of the BRIRs. The opposite scenario, known as Reverse Early Reflection (RER) removal is achieved when the latest early reflections (those just before the LR) are removed first, and at each subsequent test render, we traverse further backwards towards the direct sound, removing the earlier reflections. LR cut off is called LR removal and ITDG extended is labelled ITDG extension. A half of a 64-point Hanning window, equivalent to a 0.726 ms transition at 44.1 kHz sample rate, is applied to the transitions between silence and the impulse response to smooth the truncations in RER removal, FER removal and LR removal. The schematic figure to explain these parameters is indicated in Figure 2, where a BRIR with 0.31 s reverberation time is given as an example. Figure 2a presents the BRIR with an extended ITDG of 50 ms. Figure 2b is the same BRIR with ERs reversely cut off by 50 ms. Figure 2c presents the BRIR with ERs forward cut off by 50 ms. Figure 2d shows the BRIR with LR cut off by 465 ms.

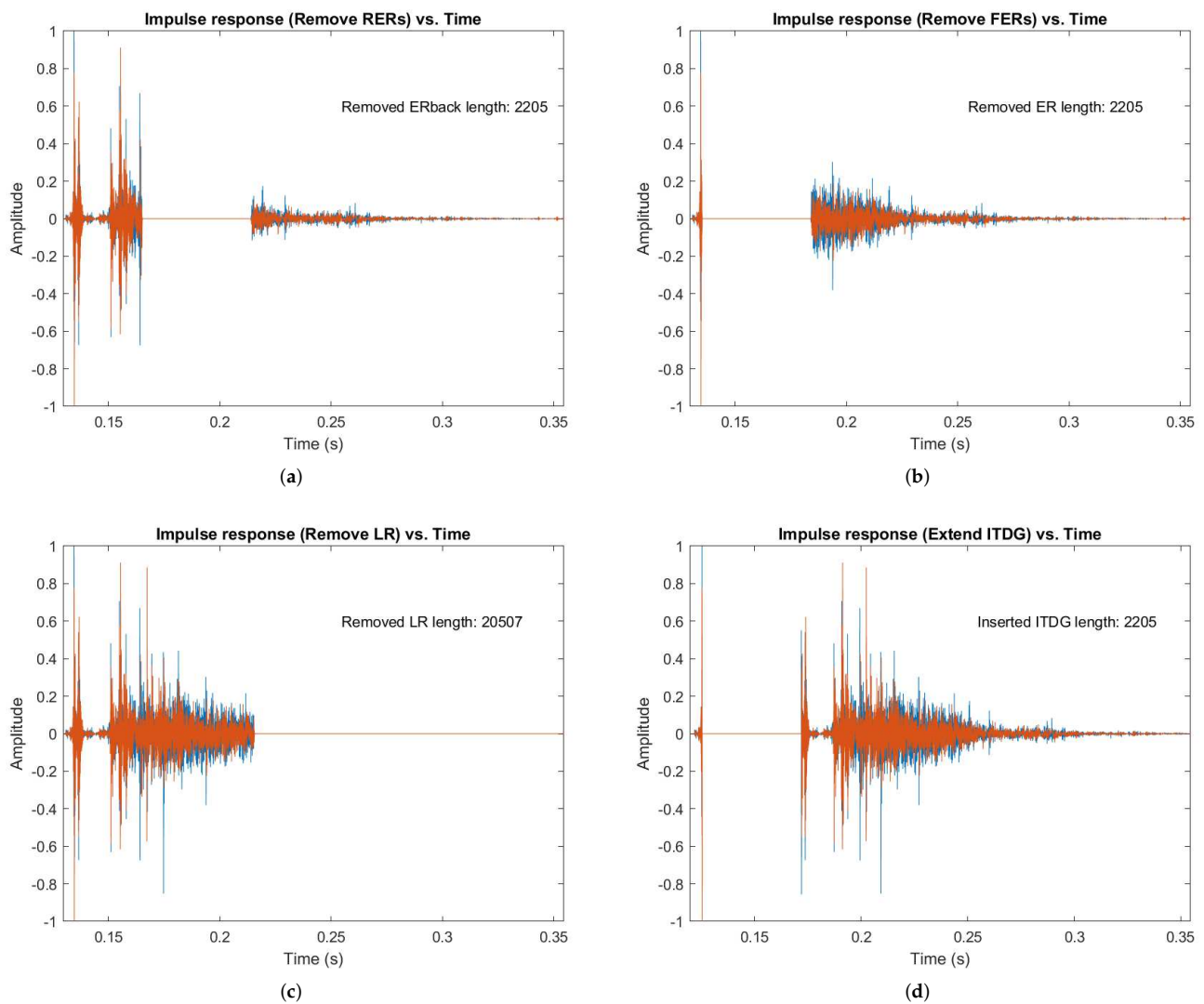


Figure 2. The explanation of changing different BRIR parameters, including RER removal, FER removal, LR removal and ITDG extension. (a) Cut off ERs reversely by 50 ms on the original BRIR. (b) Cut off ERs forward by 50 ms on the original BRIR. (c) Cut off LR by 465 ms on the original BRIR. (d) Extend ITDG by 50 ms on the original BRIR.

Three BRIRs (with 0.31 s, 0.91 s and 1.51 s reverberation time) generated in ODEON [25] were used as the reference impulse signals as shown in Figure 3a–c, and their time frequency spectra are shown in Figure 4a–c, respectively. ODEON is regarded as a useful tool for research in objective and subjective room acoustics [25]. Further, 10 s, 44.1 kHz exponential sine-tone sweeps [26] were used to generate the BRIRs and the HRTF implemented in ODEON was Subject 21 from the CIPIC database [27]. The speaker was placed 4 m away from the front wall of the reverberant room, with the listening position at 0 degrees, 13.5 m from the source. The absorption coefficient was varied by changing the material of the room surface, thus the BRIRs for different reverberation times were measured in the same room model. The materials used in the room surface for different reverberation times are described in Table A1 in Appendix A. The room model used was a cuboid with a length, width and height of approximately 22 m, 16 m and 10 m, respectively. A brief segment of anechoic male speech audio was used as the test signal as shown in Figure 3d. The segment is two-channel anechoic audio of 2.6 s length. Its sample rate is 44.1 kHz and bit depth is 24 bit. The listening test reference audio samples were generated by convolving these three BRIRs with the anechoic male speech audio.

Contrast BRIRs were generated by changing one of the identified acoustic parameters of the reference BRIRs from above four variable acoustic parameters (RER removal, FER removal, LR removal and ITDG extension). These altered BRIRs were convolved with the same anechoic male speech audio to generate the contrast stimuli.

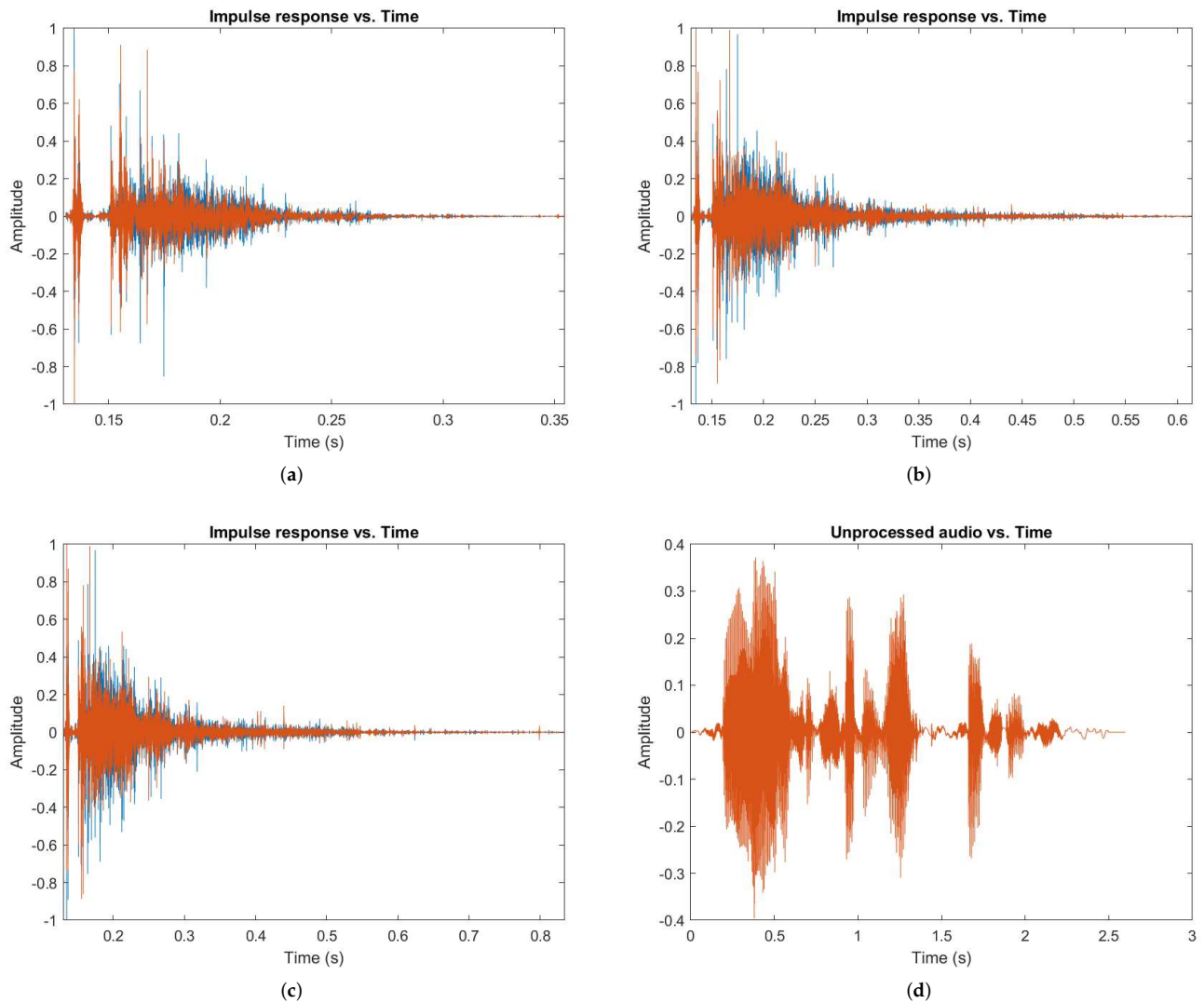


Figure 3. The BRIRs and the test signal used in the listening test. (a) The BRIR with 0.31 s reverberation time. (b) The BRIR with 0.91 s reverberation time. (c) The BRIR with 1.51 s reverberation time. (d) The dry male speech audio signal.

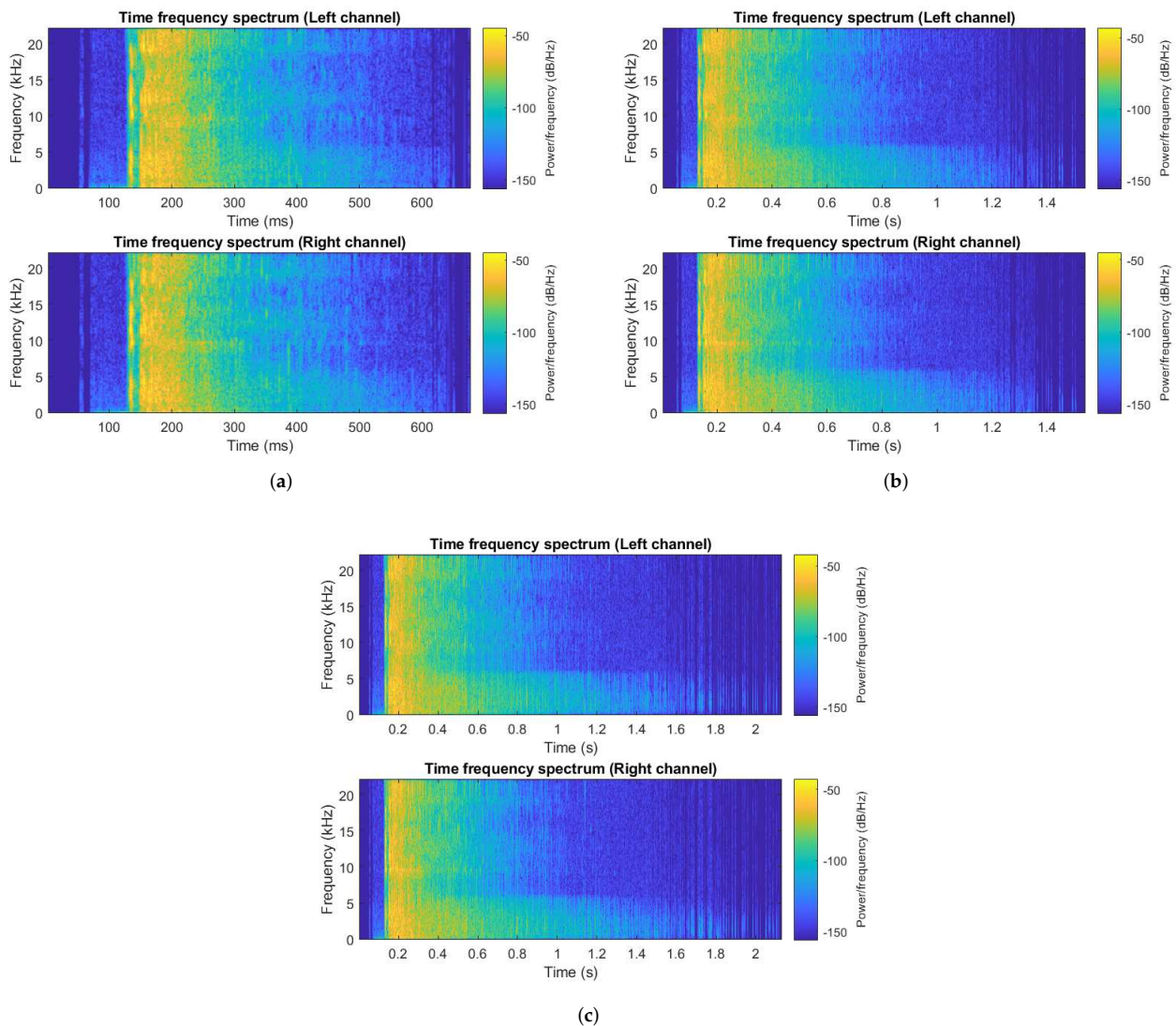


Figure 4. The time frequency spectrum of the BRIRs used in the listening test. (a) The time frequency spectrum of the BRIR with 0.31 s reverberation time. (b) The time frequency spectrum of the BRIR with 0.91 s reverberation time. (c) The time frequency spectrum of the BRIR with 1.51 s reverberation time.

2.2. Experimental Design

Three different reverberation times and four acoustic parameters were employed to evaluate changes in perceptual thresholds of reverberation, so the whole experiment is divided into three groups and each group includes four parts as shown in Table 1. All 12 parts were presented in a random order and a 30 s rest time is set between each part to reduce fatigue and experimental error caused by sustained concentrated listening.

The measurement of perceptual thresholds for each parameter was established using an AB blind test [28] and the staircase method [29]. An AB blind test is a method that can compare two sensory stimuli to judge detectable differences between them [28]. Sample ‘A’ as a reference audio sample and sample ‘B’ as a test audio sample are provided to a subject. The subject compares ‘A’ and ‘B’ and then judges them as the same or not, providing the responses ‘Yes’ or ‘No’. Staircases usually begin with a detectable difference between a reference and test stimulus. This difference is then reduced with predetermined repeat intervals until the participant provides a negative response. At this point, the staircases reverse and the difference increases with predetermined repeat intervals until the participant makes a different response again, triggering another reversal. Predetermined repeat intervals can be set to be the same or different, and this process can be repeated

as needed, until the stimuli reach an asymptotic level. Then, they hover around the plateau as long as the conditions remain unchanged [29]. The staircase method has three predetermined conditions, the start point, the step sizes and the stop point. In this listening test, participants were asked to listen to and compare reference audio samples to audio samples rendered with a single variable acoustic parameter. They responded ‘Yes’ or ‘No’ to the question ‘Are audio samples A and B the same?’.

This experiment involves threshold detection for four BRIR parameters, RER, ITDG, FER and LR. The extended start points of ITDGs were empirically set to 40 ms as it was quite apparent to the authors that the effect could clearly be heard at this large interval for all rooms. Similar to ITDG, the reverse removal start points of ERs are all set as 50 ms, and the forward removal start points of ERs are all set as 35 ms. Reverberation time is closely related to LR, so there are great differences in removal start points of LRs for BRIRs with different reverberation times. Through listening and confirmation of pilot tests, thresholds that produced clearly distinguishable differences from the reference reverberation were set as LR was removed by 465 ms for the BRIR with short reverberation time (0.31 s), 780 ms for the BRIR with medium reverberation time (0.91 s) and 1250 ms for the BRIR with long reverberation time (1.51 s), respectively. By averaging these results, the removal start points of LRs were set as 465 ms for the BRIR with short reverberation time (0.31 s), 780 ms for the BRIR with medium reverberation time (0.91 s) and 1250 ms for the BRIR with long reverberation time (1.51 s), respectively.

The step sizes are not fixed. They are adjusted according to the experimental results. The initial step sizes are 5 ms in ITDG, RER and FER tests. After three ‘Yes’ responses appear, the step sizes are adjusted to 3 ms, and after five ‘Yes’ responses appear, the step sizes are adjusted to 1 ms. For LR tests, the test start point is large with initial step sizes set to 10 ms. After three ‘Yes’ responses appear, the step sizes are adjusted to 5 ms, and after five ‘Yes’ responses appear, the step sizes are adjusted to 3 ms.

There are two kinds of stop conditions. One is to decide the end point at several trials after a predetermined number of ‘Yes’ responses, as this experiment presumes an initial response of ‘No’. The other is to decide a fixed end point through a predetermined number of trials. Obviously, the greater the number of trials, the more reliable the results will be, but also the more time will be consumed to conduct the test. In consideration of reliability and economy of time, these two kinds of stop conditions are combined for each part. In this experiment, after the participants give five ‘Yes’ answers, one stop condition is triggered if they continue to carry out ten trials, or the other stop condition is triggered after thirty fixed trials. If the number of trials in one part does not reach thirty, the first stop condition will be enabled. After the stop condition is triggered, the last ten values of reversals are averaged to obtain the resolution threshold. Figure 5 uses an example to illustrate the start point, step size, stop point and threshold calculation, and Table 1 describes the predetermined conditions of each test part. Since the stopping conditions are the same for all 12 parts (5 ‘Yes’ add 10 trials or 30 trials), they are not presented and repeated in the table.

Table 1. The experimental design and corresponding predetermined conditions of the staircase method (Same stop condition is 5 ‘Yes’ add 10 trials or 30 trials).

Parts		1	2	3	4
Groups		(RER Removal)	(ITDG Extension)	(FER Removal)	(LR Removal)
1 (0.31 s reverb time)	start point:	50 ms	40 ms	35 ms	465 ms
	step size:	5 ms to 3 ms to 1 ms	5 ms to 3 ms to 1 ms	5 ms to 3 ms to 1 ms	10 ms to 5 ms to 3 ms
2 (0.91 s reverb time)	start point:	50 ms	40 ms	35 ms	780 ms
	step size:	5 ms to 3 ms to 1 ms	5 ms to 3 ms to 1 ms	5 ms to 3 ms to 1 ms	10 ms to 5 ms to 3 ms
3 (1.51 s reverb time)	start point:	50 ms	40 ms	35 ms	1250 ms
	step size:	5 ms to 3 ms to 1 ms	5 ms to 3 ms to 1 ms	5 ms to 3 ms to 1 ms	10 ms to 5 ms to 3 ms

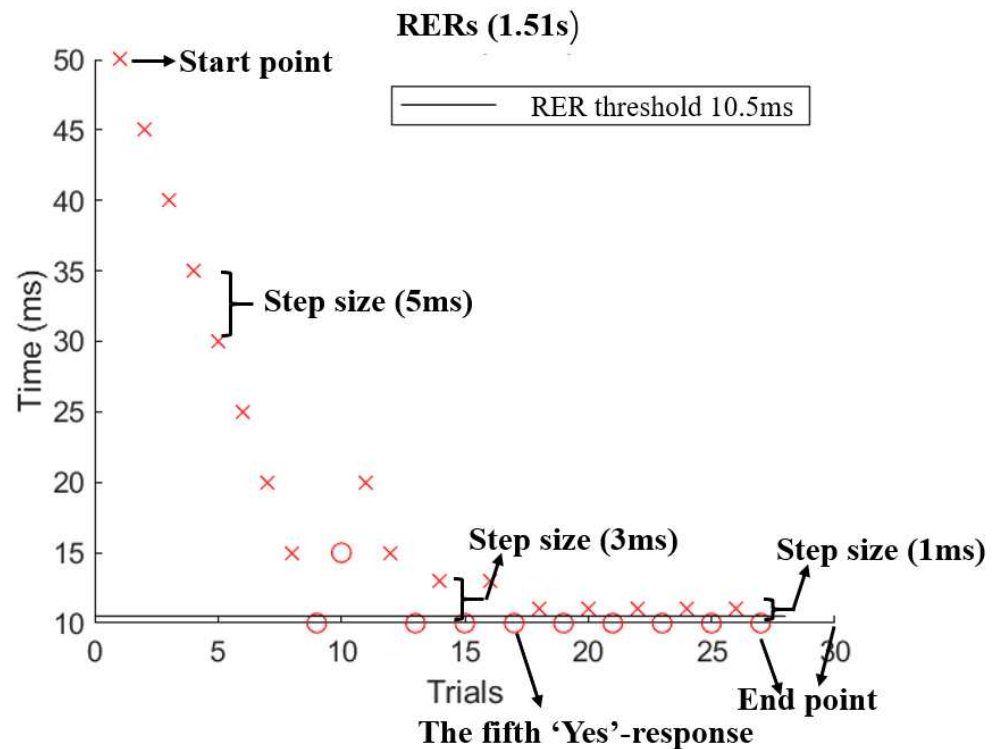


Figure 5. The illustration of predetermined conditions of staircase method.

2.3. Experimental Setup

The listening test was conducted online. The hardware used was the test subjects' own personal computer or laptop and a pair of 'Beyerdynamic DT990 Pro' headphones. Subjects were instructed to conduct the experiment in a quiet listening environment.

The experimental software used was a custom listening test application generated by Appdesigner (a MATLAB environment for application development [30]) by the author. The application's operation interface is shown in Figure 6. Participants can click the 'Play A' or 'Play B' button to play reference audio samples or contrast audio samples, and then click the 'Yes' or 'No' button to make a response. The 'Previous' or 'Next' button can be used to come back to the previous trial or go to the next trial.

2.4. Subjects

Twenty participants were recruited, all classified as expert listeners according to the ITU-R BS.1543-3 recommendation [31]. Each participant was paid to take part in the test which lasted about 1 h. All participants were over 18 years old. Participants were not asked to report their gender. All of these participants come from the AudioLab or music related majors at the University of York and Beijing Contemporary Music Academy. The participants were informed of the purpose of the experiment and the protocol of the experiment prior to conducting their trial.

This study was approved by the University of York Physical Sciences Ethics Committee (Number: Mi111120. Date of approval: 26 November 2020) and participants provided informed consent before taking part.

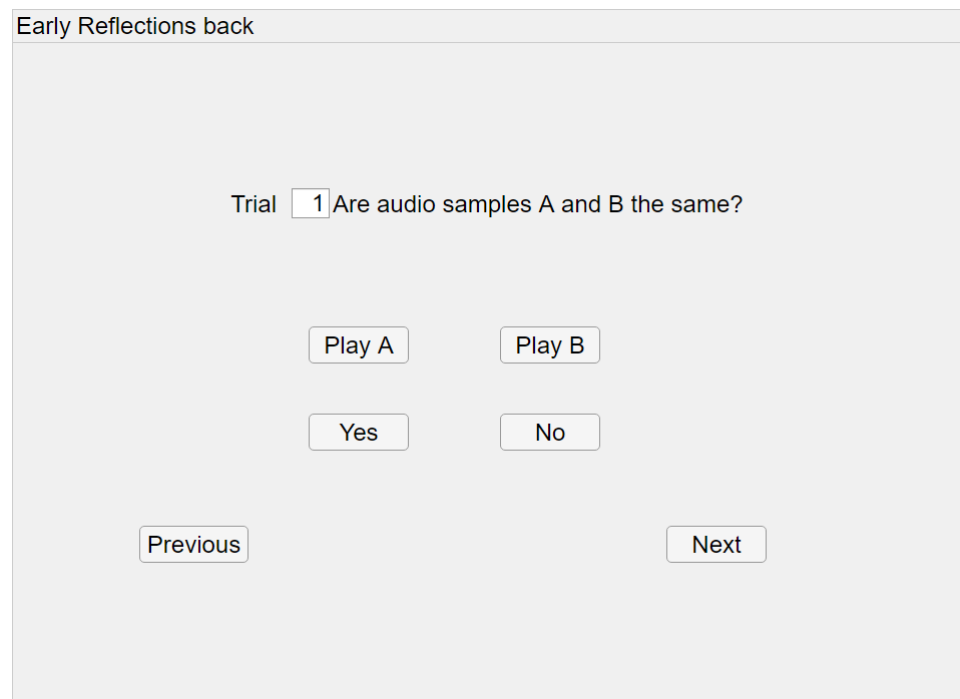


Figure 6. The operation interface of the listening test software.

3. Results

The test data are shown in Tables A2 and A3 in Appendix A. Those data marked with the red colour are the maximum values that the parameters can be changed in this experiment, which means that the participant cannot distinguish any differences between the reference audio samples and the contrast audio samples when the corresponding parameters change, so the maximum values are regarded as the corresponding thresholds. Conversely, those data marked with the blue colour are the minimum values that the parameters can be changed in this experiment, which means that the participants can distinguish the nuances between the reference audio samples and the contrast audio samples when the corresponding parameters change. The minimum values are therefore regarded as the corresponding thresholds.

3.1. ANOVA Test

3.1.1. Data Presentation and Outlier Removal

An ANOVA test is used to analyse whether the reverberation time has a significant effect on the threshold of perceptual reverberation. Because RER removal data do not conform to a normal distribution, and LR removal data do not conform to normal distribution or homogeneity of variance, the two groups of data use non parametric tests i.e., a Kruskal Wallis ANOVA test. However, the data of FER removal and ITDG extension use an ANOVA parametric test.

The outliers of above raw data are removed first. Figure 7a–d are the box plots of RER removal thresholds, ITDG extension thresholds, FER removal thresholds and LR removal thresholds, respectively, and their corresponding outliers are displayed. The data with outliers removed are listed in ascending order in Tables A4 and A5 in Appendix A. Average values, standard deviation values and standard error values of each parameter type are calculated. Those data marked in green are the average values of each parameter type, marked in yellow are the standard deviation values and marked in blue are the standard error values of each parameter type.

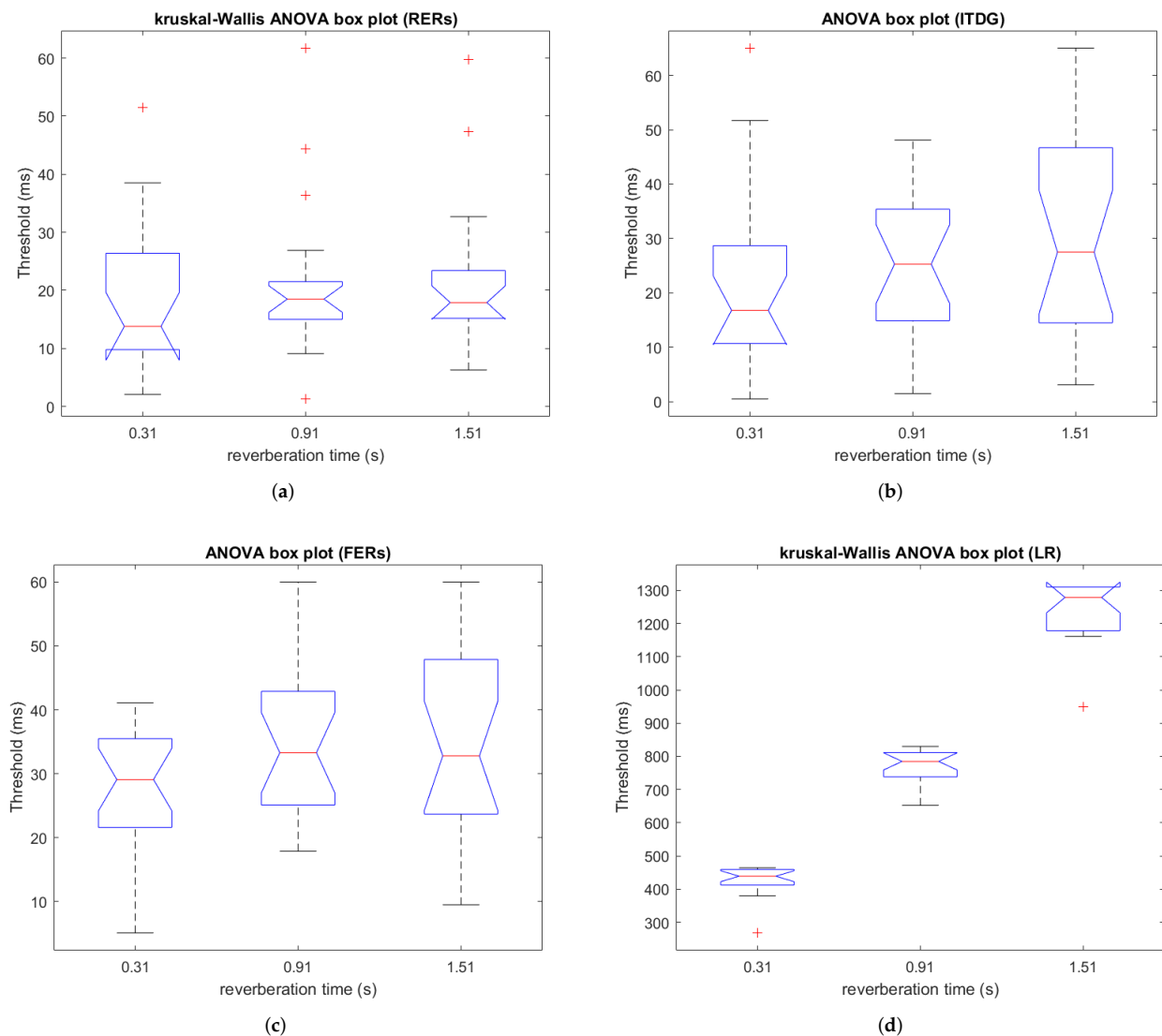


Figure 7. Box plots of the thresholds of RER removal, ITDG extension, FER removal and LR removal. (a) RER removal thresholds. (b) ITDG extension thresholds. (c) FER removal thresholds. (d) LR removal thresholds.

3.1.2. Analysis of Room Differences

Table 2 includes the ANOVA test results of RER removal, ITDG extension, FER removal and LR removal. From these ANOVA test results, the p value of RER removal thresholds is bigger than 0.05, so there is no significant difference in the three groups of RER removal thresholds with 0.31 s, 0.91 s and 1.51 s reverberation time, respectively. This means that reverberation time will not affect the final average threshold of RER removal. Similarly, reverberation time will not affect the final average thresholds of ITDG extension and FER removal. However, the p value of LR removal thresholds is smaller than 0.05, so there is a significant difference in the three groups of LR removal thresholds with 0.31 s, 0.91 s and 1.51 s reverberation time, respectively. It means that reverberation time will affect the final average threshold of LR removal.

Table 2. The ANOVA test results of RER removal, ITDG extension, FER removal and LR removal (Values marked with * indicate significant differences).

DF = 2 Significance Level = 0.05	RER Removal	FER Removal	ITDG Extension	LR Removal
<i>p</i> value (ANOVA)		0.1093	0.1692	
Significant difference (ANOVA)		N	N	
<i>p</i> value (Kruskal-Wallis ANOVA)	0.3901			* 3.84768×10^{-12}
Significant difference (Kruskal-Wallis ANOVA)	N			Y

3.2. Results Analysis for Each Parameter Type

The thresholds with outliers removed in each parameter type are analysed as percentages. Figures 8 and 9 provide the percentage of each impact threshold range of RER and FER removal on perceptual reverberation, respectively. The data in Figure 8 show that for BRIRs with 0.31 s, 0.91 s or 1.51 s reverberation time, most thresholds that influence participants' perceptual reverberation distribute between 0 and 30 ms, and about half of the thresholds are concentrated in 10–20 ms. Further, 26.3% of the thresholds are concentrated between 0–10 ms with 0.31 s reverberation time. These illustrate that when the RER is only removed a little, the reverberation difference can be clearly perceived, so people are sensitive to the RER of BRIRs.

Figure 9 shows that the threshold distribution of FER removal is concentrated in 20–50 ms. These prove that the effect of FER removal on perceptual reverberation is weaker than RER removal. When the change of FER is within 20 ms, the effect of FER removal on perceptual reverberation is negligible.

As shown in Figure 10, even though the thresholds of ITDG extension almost evenly distribute between 0 and 50 ms, when ITDG is extended for a few milliseconds, participants can clearly perceive the reverberation difference. Overall people are easy to distinguish the impact of ITDG extension on perceptual reverberation.

Finally, reverberation time of BRIRs can influence the time of LR removal, so the ranges of parameter change are relatively large. As Figure 11 shows, for short reverberation time, the threshold distribution concentrates in 400–500 ms, at 0.91 s reverberation time, most resolution thresholds distribute 750–850 ms, and most thresholds concentrate in 1250–1350 ms. As shown in Figures A1–A3 in Appendix B, the grey impulse responses are the original impulse responses and the red impulse responses are the impulse responses with LR removed. Through calculation, the time length of the LR is 465 ms for the BRIR with 0.31 s reverberation time, 1329 ms for the BRIR with 0.91 s reverberation time and 1920 ms for the BRIR with 1.51 s reverberation time. A distinction is made here between the length of the LR and the reverb time to avoid confusion as to why the length of the LR is greater than the reverb time. This means that only when a large amount of LR is removed, the perceptual reverberation can be affected clearly, so it is hard for people to distinguish the difference of perceptual reverberation caused by LR removal and the impact of LR removal on perceptual reverberation is not serious.

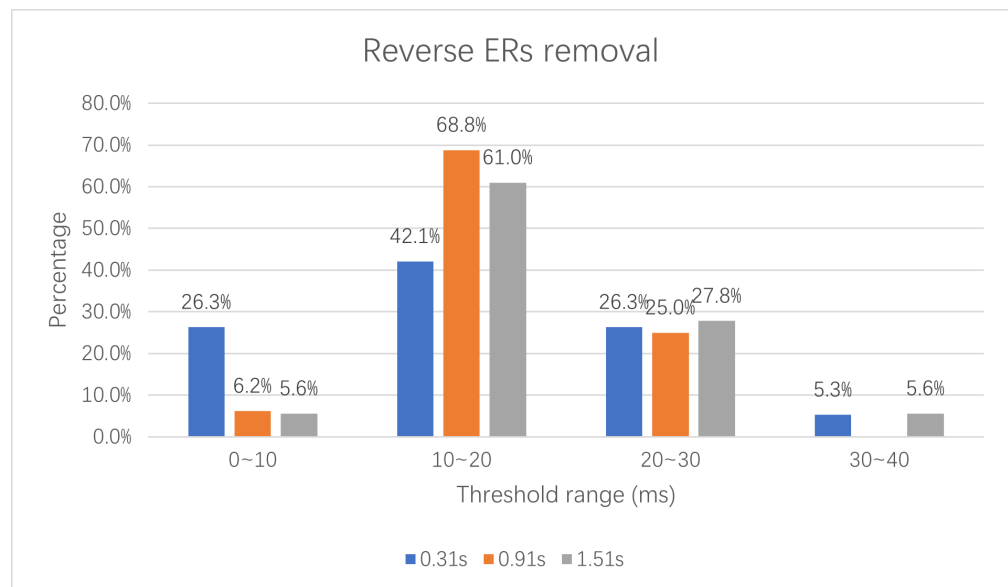


Figure 8. Impact thresholds of RER removal on perceptual reverberation shown as a percentage of threshold ranges for BRIRs with three different reverberation times.

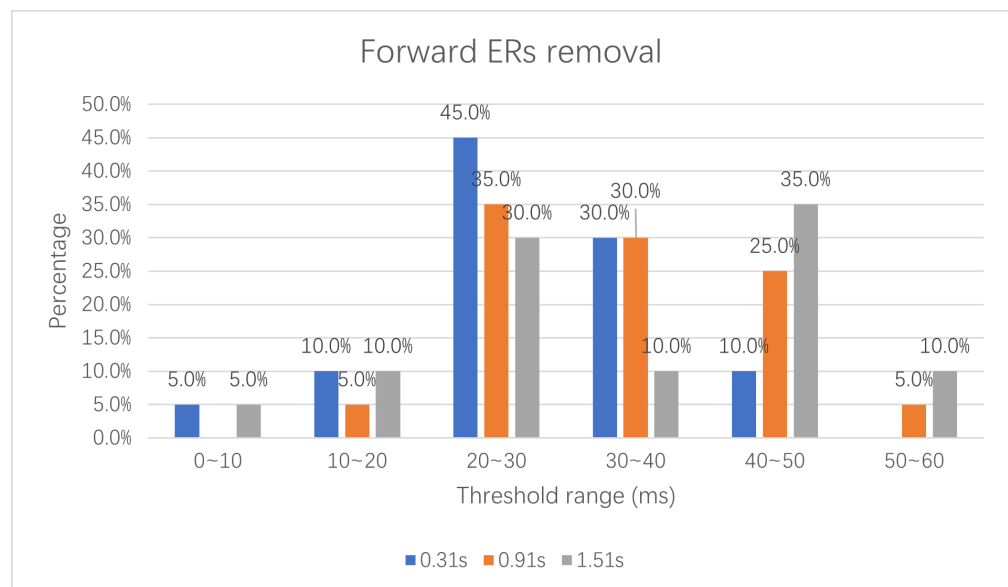


Figure 9. Impact thresholds of FER removal on perceptual reverberation shown as a percentage of threshold ranges for BRIRs with three different reverberation times.

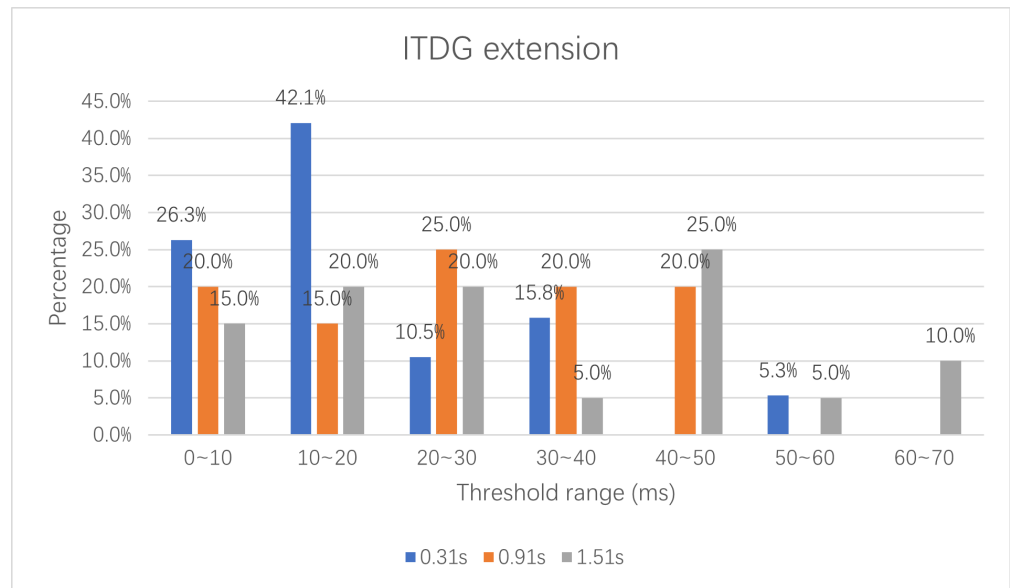


Figure 10. Impact thresholds of ITDG removal on perceptual reverberation shown as a percentage of threshold ranges for BRIRs with three different reverberation times.

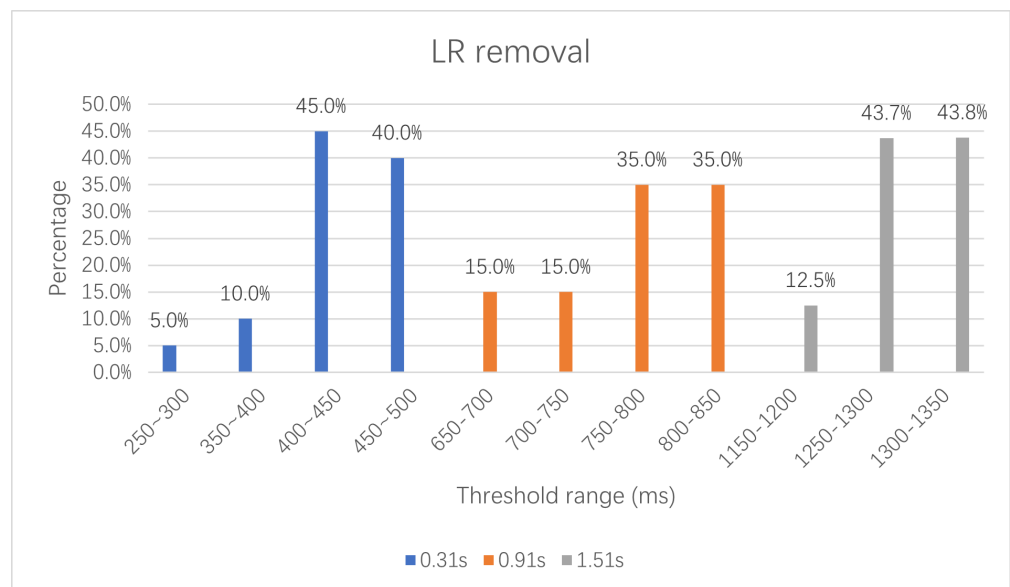


Figure 11. Impact thresholds of LR removal on perceptual reverberation shown as a percentage of threshold ranges for BRIRs with three different reverberation times.

3.3. The Average Threshold Analysis of Each Parameter Type

In order to obtain generic impact thresholds of these parameters on perceptual reverberation, appropriate average thresholds are necessary. The error bars of the average thresholds of each parameter with different reverberation time are presented in Figure 12. The error bars indicate the standard error of test data without outliers of each parameter type. As shown in Figure 12a, the average impact thresholds of RER removal are 15.81 ms, 17.49 ms and 18.18 ms corresponding to 0.31 s, 0.91 s and 1.51 s reverberation time, respectively. For FER removal, as shown in Figure 12b, the corresponding average thresholds for these three reverberation time are 27.68 ms, 34.33 ms and 34.6 ms. Its average thresholds are higher than the average threshold of RER removal, so its effect on perceptual reverberation is less than RER removal. When ITDG is extended, as shown in Figure 12c, the average thresholds are 18.37 ms, 25.21 ms and 30.75 ms, respectively. As shown in Figure 12d,

the average impact thresholds of LR removal on perceptual reverberation are 435.52 ms, 771.16 ms and 1276.9 ms corresponding to 0.31 s, 0.91 s and 1.51 s reverberation time.

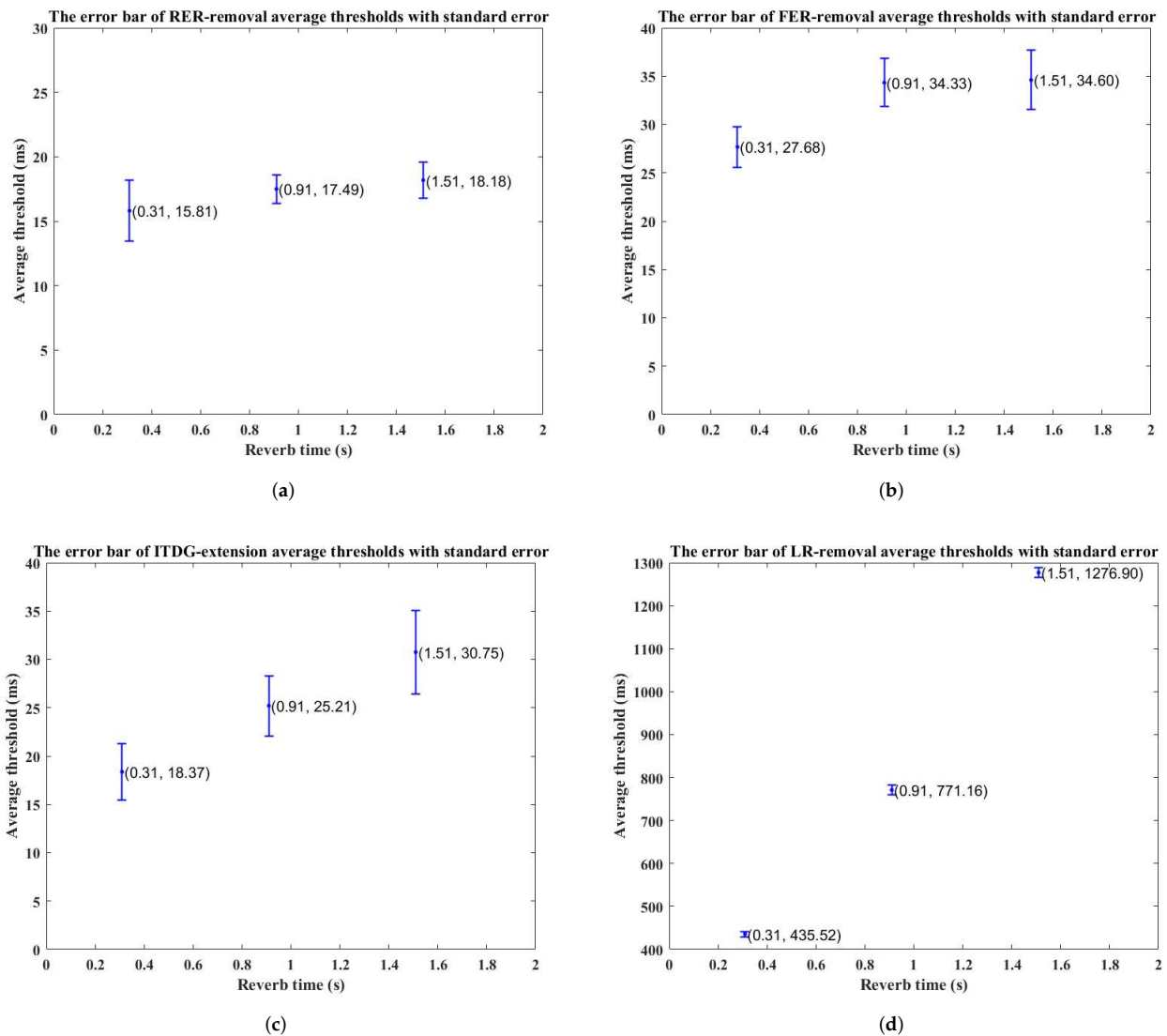


Figure 12. The error bars of the average thresholds distribution with standard errors of each parameter for different reverberation times. (a) The error bar of RER removal. (b) The error bar of FER removal. (c) The error bar of ITDG extension. (d) The error bar of LR removal.

3.4. The Standard Deviation Analysis of Each Parameter Type

As shown in Tables A4 and A5, the standard deviation of RER removal thresholds are 10.38 ms, 4.37 ms and 6.01 ms corresponding to 0.31 s, 0.91 s and 1.51 s reverberation time. The standard deviation of FER removal thresholds are 9.35 ms, 11.14 ms and 13.72 ms, respectively. The standard deviation of ITDG extension thresholds are 12.62 ms, 13.89 ms and 19.39 ms, respectively, and the standard deviation of LR removal thresholds are 27.59 ms, 50.64 ms and 44.10 ms, respectively. Through the above data comparison, RER removal, FER removal and ITDG extension are less discrete compared with LR removal. The results mean that with the exception of LR removal, the distribution of the other three parameters is more concentrated, and for LR removal, some people can distinguish the difference of perceptual reverberation clearly while other people hardly distinguish the difference of perceptual reverberation.

4. Discussion

Through the analysis of the collected data, according to Tables A4 and A5 in Appendix A, the maximum thresholds of LR removal at a long reverberation time appear most frequently. Further, some maximum thresholds of LR removal appear at a medium reverberation time. Although the maximum LR removal threshold does not occur at short reverberation times, the average LR removal threshold for the short reverberation time is close to its LR length. The results suggest that people's perception for LR removal is therefore insensitive, and that only when LR is completely or almost entirely removed are people able to perceive a difference to the reference reverberation.

Through analysis, RER removal and ITDG extension have the most influence on perceptual reverberation of the parameters assessed. ERs improve speech intelligibility by increasing the loudness of the direct sound [32]. Table A6 in Appendix A shows a comparison between the energy corresponding to the reference BRIR versus the BRIR with early reflections removed at different reverberation times. The energy values were computed through integrated Loudness according to EBU R 128 standards [33] (in Loudness Units relative to Full Scale (LUFS)) and the ITU-R BS.1770-4 [34] (still refers to Loudness, K-weighted, relative to full scale 'LKFS', but LKFS and LUFS are equivalent). ITU-R BS.1770-4 [34] also defines that the LKFS unit is equivalent to a decibel in that an increase in the level of a signal by 1 dB will cause the loudness reading to increase by 1 LKFS. It can be seen that the energy of the BRIR is reduced by approximately 0.6 to 3 LUFS when the FER or RER is removed at the corresponding threshold (the one exception is that the energy of the RER removed increases by approximately 1 LUFS at 0.31 s reverberation time). It is shown in [35] that listeners can distinguish a change in sound level of about 1 dB in their most sensitive sound level range (about 35 to 80 dB SPL). So the difference one hears when the early reflections are removed may be due to a difference caused by the level.

For a speech signal, the ERs are limited to around 50 ms. For RER removal under different reverberation times, their average resolution thresholds are from 15 ms to 18 ms. This means that when ERs are reversely removed 30% to 36%, people can perceive the change of the reverberation. For FER removals under different reverberation time, their resolution thresholds are from 27 ms to 34 ms. When ERs are forward removed over 50% to about 68%, people can perceive the change of the reverberation. Compared with FER removal, RER removal should be a greater consideration when designing reverberation algorithms. When ITDG is extended by about 18–30 ms, the difference of reverberation can be perceived clearly. To contextualise, research shows that ITDG is typically in the range of 20 ms to 60 ms in a large concert hall and from about 8 ms to 27 ms in chamber music halls [36]. ITDG extension thresholds of 18–30 ms is enough to affect the perceptual reverberation, so ITDG should be another major consideration in designing artificial reverberation algorithms. For LR, the average thresholds measured at different reverberation times are 435 ms, 771 ms and 1276 ms, respectively. Before a difference was perceived, the LR was more than halved at all three reverberation times, and at 90% removal for the short reverberation time, suggesting that people are less sensitive to the reverberation difference of male speech signals caused by LR changes. Therefore, to a certain extent, in the design of reverberation algorithms, it is not necessary to give priority to some small changes in the later reverberation, if these changes do not have major impact on the perception of reverberation overall. Theoretically, expert listeners and those experienced with acoustic experiments can more accurately distinguish the impact of BRIR parameters on perceptual reverberation. Therefore, with the test data provided by them, the calculated average threshold should be lower than the generic average threshold of the public, so their average threshold should represent the generic threshold.

Overall, comparison of the above four BRIR parameter types reveals that RER and ITDG require most attention in the design of artificial binaural reverberation algorithms. Slight changes in FER and LR need not be over considered.

However, because of the limitation of test time and hearing fatigue, this experiment just used male speech as the test signal. Actually, these parameters may have different effects on

perceptual reverberation of different audio signals. Therefore, further experiments should test a variety of different noise and musical stimuli to further test the influence of BRIR parameters on perceptual reverberation. Furthermore, this experiment is a static binaural reverberation parameter test rather than a dynamic one, so further experiments should also include head tracked conditions.

Furthermore, due to the COVID-19 pandemic environment, some experimental parameters were not controlled across subjects. This listening test was conducted online, so the listening environment, the type of headphones (although advised), and the volume of audio samples were dependent on the preference of the test subject. This should be avoided in the future experiments.

5. Conclusions

The purpose of this experiment was to find out the average perceptual thresholds of four BRIR parameters to represent the generic thresholds on perceptual reverberation, and determine whether these parameters have significant impact on perceptual reverberation. The measurement of these thresholds lays a foundation for the realisation of more idealised artificial reverberation algorithms, which can be applied in AR scenarios to create more plausible reverberation effects. This research makes the following conclusion through the measurement and the analysis for ITDG, FER, RER and LR:

- The average thresholds of RER removal are 15.81 ms, 17.49 ms and 18.18 ms corresponding to 0.31 s, 0.91 s and 1.51 s reverberation time, respectively. The average thresholds of ITDG extension are 18.37 ms, 25.21 ms and 30.75 ms. The reverse removal of ERs and extension of ITDG causes relatively obvious effects on reverberation perception of speech audio, so RER and ITDG should be focused on when designing artificial reverberation algorithms.
- The average thresholds of FER removal are 27.68 ms, 34.33 ms and 34.6 ms, respectively, for each reverberation time. Generally, subjects were less sensitive to FER removal; therefore, FER is less of a concern when designing a reverberation algorithms.
- The average thresholds of LR removal are 435.52 ms, 771.16 ms and 1276.9 ms, respectively. LR removal has a small influence on perceptual reverberation, so when achieving an artificial reverberation algorithm, small changes in LR may not be significant.
- The ANOVA test shows that reverberation time does not affect the thresholds of RER removal, ITDG extension and FER removal on perceptual reverberation, but the thresholds of LR removal on perceptual reverberation are impacted by reverberation time.

Early reflections are achieved by finite impulse response (FIR) delay lines and late reverberation by infinite impulse response (IIR) filters. Whilst focusing on hybrid FIR and IIR filters would certainly theoretically lead to more accuracy, exploiting perceptual sensitivity to reverberation could reduce computational resources for algorithm design. Based on the experimental findings, it appears that LR removal has a smaller effect on perceptual reverberation, and although FER does not affect perceptual reverberation as much as RER, overall ER can have an obvious effect on perceptual reverberation. ITDG can also have a significant effect on perceptual reverberation. In order to balance reverberation accuracy and algorithmic efficiency, perceptually motivated reverberation algorithms should focus on ERs and ITDG. When designing reverberation algorithms, early reflections are achieved by means of FIR delay lines. More finite impulse responses require more computational cost, so the number of finite impulse responses can be minimised, depending on the measured threshold, without affecting reverberation perception. ITDG has a large impact on perceptual reverberation, so controlling the time period between direct sound and early reflections as accurately as possible when designing a reverberation algorithm can improve efficiency in the design.

Author Contributions: Conceptualisation, H.M. and G.K.; methodology, H.M., G.K. and H.D.; software, H.M.; validation, H.M., G.K. and H.D.; formal analysis, H.M., G.K. and H.D.; investigation, H.M.; resources, H.M., G.K. and H.D.; data curation, H.M.; writing—original draft preparation, H.M.; writing—review and editing, G.K. and H.D.; visualisation, H.M.; supervision, G.K. and H.D.; project administration, H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Engineering and Physical Sciences Research Council IAA project “MINERVA: Musical Interactions in Networked Experiences using Real-time Virtual Audio”.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of University of York (protocol code Mi111120 and 26 November 2020 of approval).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: The data presented in this study are available in insert article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BRIR	Binaural room impulse response
ERs	Early reflections
FDN	Feedback delay networks
FER	Forward early reflections
FIR	Finite impulse response
IID	Interaural intensity difference
IIR	Infinite impulse response
ITD	Interaural time difference
ITDG	Initial time delay gap
LR	Late reverberation
RER	Reverse early reflections
RIR	Room impulse response

Appendix A

Table A1. The materials used in the room surface for different reverb times (The material description for each material number can be found in ODEON [25]).

Surface Number	Surface Name	Material Number (0.31 s Reverb Time)	Material Number (0.91 s Reverb Time)	Material Number (1.51 s Reverb Time)	Area (m ²)
1001	Podium floor	70	20	20	78.00
1002	Main audience floor	70	40	20	259.26
2001	End wall behind podium	1004	1004	1004	75.00
−2002	Podium side wall, South + North	1004	1004	1004	58.70
2002	Podium side wall, South + North	1004	1004	1004	58.70
−2003	Side wall, audience area South + North	11,009	11,009	11,009	139.52
−2003	Side wall, audience area South + North	11,009	11,009	11,009	139.52
2004	Rear wall behind audience	1	11,009	11,009	119.04
3001	Podium ceiling	3023	3023	3023	84.50
3002	Ceiling over audience	3023	3023	3023	256.00

Table A2. The test threshold data of each participant (unit: ms).

Parameter Types Participants	0.31 s	0.31 s	0.31 s	0.31 s	0.91 s	0.91 s
	RER Removal	ITDG Extension	FER Removal	LR Removal	RER Removal	ITDG Extension
1	27.5	17.9	20.7	380.1	36.3	24.5
2	2.1	8.9	17.3	415.3	14.7	27.7
3	38.5	34.7	39.3	464.5	26.9	37.1
4	20.7	65	35.7	450.5	21.7	33.7
5	13.9	17.1	23.9	435.1	15.9	23.3
6	12.3	9.3	22.5	268.3	18.3	24.7
7	3.3	19.3	20.5	400.5	1.3	7.7
8	7.3	12.1	22.7	380.1	10.5	6.1
9	18.7	14.3	30.1	437.7	18.1	48.1
10	2.7	2.7	5.1	409.5	15.1	12.9
11	16.3	2.5	31.3	451.9	18.9	40.5
12	2.7	14.5	28.7	423.7	15.5	31.3
13	29.9	23.1	40.1	455.3	21.3	25.9
14	13.7	15.1	29.1	464.1	19.1	16.9
15	12.3	16.5	29.1	429.7	9.1	30.5
16	51.5	51.7	36.7	464.5	61.7	45.1
17	12.5	31.5	35.3	440.5	18.7	1.5
18	25.3	27.1	29.9	464.1	21.1	17.9
19	28.1	30.3	41.1	463.2	44.3	43.3
20	12.5	0.5	14.5	444.5	14.9	5.5

Table A3. The test threshold data of each participant (unit: ms).

Parameter Types Participants	0.91 s	0.91 s	1.51 s	1.51 s	1.51 s	1.51 s
	FER Removal	LR Removal	RER Removal	ITDG Extension	FER Removal	LR Removal
1	32.7	731.9	24.7	65	25.9	1310
2	29.7	730.9	17.5	13.3	28.3	950
3	33.7	830	32.7	50.7	49.1	1310
4	45.3	830	22.3	65	46.7	1310
5	22.9	803.3	18.1	13.7	22.7	1195
6	27.3	696.5	17.7	18.9	40.5	950
7	32.9	804.9	6.3	40.7	40.5	1291.1
8	21.5	652.5	15.7	3.7	19.9	950
9	28.3	744.9	15.1	3.1	24.7	1282.1
10	20.3	756.7	17.7	20.5	9.5	1252.1
11	20.7	787.1	22.3	15.3	44.1	1310
12	34.9	820	14.1	46.7	30.7	1310

Table A3. *Cont.*

Parameter Types	0.91 s FER Removal	0.91 s LR Removal	1.51 s RER Removal	1.51 s ITDG Extension	1.51 s FER Removal	1.51 s LR Removal
Participants						
13	60	752.3	15.3	46.7	60	950
14	47.3	782.5	24.5	27.3	34.9	1259.3
15	44.9	830	10.7	31.3	49.1	1310
16	40.9	798.7	59.7	49.3	49.7	1275.1
17	39.3	818.1	19.9	26.1	50.1	1310
18	39.7	759.5	11.9	27.7	25.3	1260.5
19	46.3	796.9	47.3	44.5	22.7	1283.7
20	17.9	696.5	20.7	5.5	17.5	1161.5

Table A4. The test threshold data with outliers removed for each parameter type are arranged in ascending order, and their average values, standard deviation values and standard error values (unit: ms).

Parameter Types	0.31 s RER Removal	0.31 s ITDG Extension	0.31 s FER Removal	0.31 s LR Removal	0.91 s RER Removal	0.91 s ITDG Extension
	2.1	0.5	5.1			1.5
	2.7	2.5	14.5	380.1	9.1	5.5
	2.7	2.7	17.3	380.1	10.5	6.1
	3.3	8.9	20.5	400.5	14.7	7.7
	7.3	9.3	20.7	409.5	14.9	12.9
	12.3	12.1	22.5	415.3	15.1	16.9
	12.3	14.3	22.7	423.7	15.5	17.9
	12.5	14.5	23.9	429.7	15.9	23.3
	12.5	15.1	28.7	435.1	18.1	24.5
	13.7	16.5	29.1	437.7	18.3	24.7
	13.9	17.1	29.1	440.5	18.7	25.9
	16.3	17.9	29.9	444.5	18.9	27.7
	18.7	19.3	30.1	450.5	19.1	30.5
	20.7	23.1	31.3	451.9	21.1	31.3
	25.3	27.1	35.3	455.3	21.3	33.7
	27.5	30.3	35.7	463.2	21.7	37.1
	28.1	31.5	36.7	464.1	26.9	40.5
	29.9	34.7	39.3	464.1		43.3
	38.5	51.7	40.1	464.5		45.1
			41.1	464.5		48.1
Average	15.81	18.37	27.68	435.52	17.49	25.21
Standard Deviation	10.38	12.62	9.35	27.59	4.37	13.89
Standard Error	2.38	2.90	2.09	6.33	1.09	3.11

Table A5. The test threshold data with outliers removed for each parameter type are arranged in ascending order, and their average values, standard deviation values and standard error values (unit: ms).

Parameter Types	0.91 s FER Removal	0.91 s LR Removal	1.51 s RER Removal	1.51 s ITDG Extension	1.51 s FER Removal	1.51 s LR Removal
	17.9	652.5	6.3	3.1	9.5	
	20.3	696.5	10.7	3.7	17.5	
	20.7	696.5	11.9	5.5	19.9	
	21.5	730.9	14.1	13.3	22.7	
	22.9	731.9	15.1	13.7	22.7	1161.5
	27.3	744.9	15.3	15.3	24.7	1195
	28.3	752.3	15.7	18.9	25.3	1252.1
	29.7	756.7	17.5	20.5	25.9	1259.3
	32.7	759.5	17.7	26.1	28.3	1260.5
	32.9	782.5	17.7	27.3	30.7	1275.1
	33.7	787.1	18.1	27.7	34.9	1282.1
	34.9	796.9	19.9	31.3	40.5	1283.7
	39.3	798.7	20.7	40.7	40.5	1291.1
	39.7	803.3	22.3	44.5	44.1	1310
	40.9	804.9	22.3	46.7	46.7	1310
	44.9	818.1	24.5	46.7	49.1	1310
	45.3	820	24.7	49.3	49.1	1310
	46.3	830	32.7	50.7	49.7	1310
	47.3	830		65	50.1	1310
	60	830		65	60	1310
Average	34.33	771.16	18.18	30.75	34.60	1276.90
Standard Deviation	11.14	50.64	6.01	19.39	13.72	44.10
Standard Error	2.49	11.32	1.42	4.34	3.07	11.02

Table A6. Comparison between the energy corresponding to the reference BRIR versus the BRIR with early reflections removed at different reverberation times (the amount of early reflection removal corresponds to the measured threshold in this paper).

Reverb Time	0.31 s	0.91 s	1.51 s
	Energy (LUFS)		
Reference BRIR	−22.4044	−19.436	−19.3806
BRIR with FER removal	−24.4902	−22.4524	−20.9696
reference-FER removal	2.0858	3.0164	1.589
BRIR with RER removal	−21.0237	−20.0485	−20.0019
reference-RER removal	−1.3807	0.6125	0.6213

Appendix B

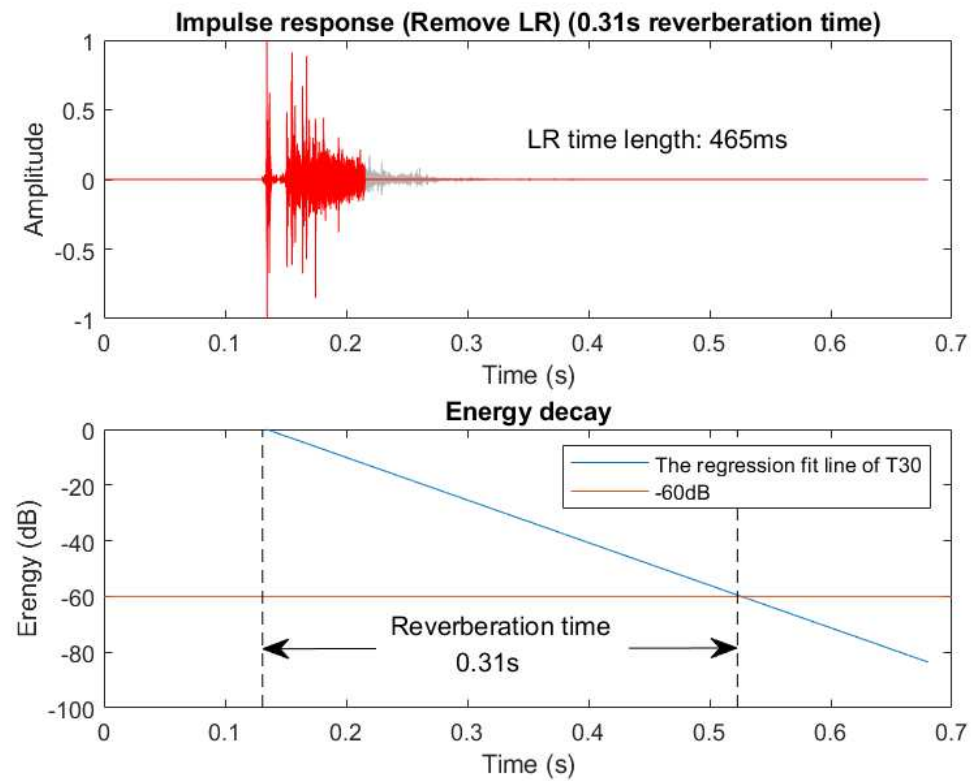


Figure A1. The time length of the LR for the BRIR with 0.31 s reverberation time.

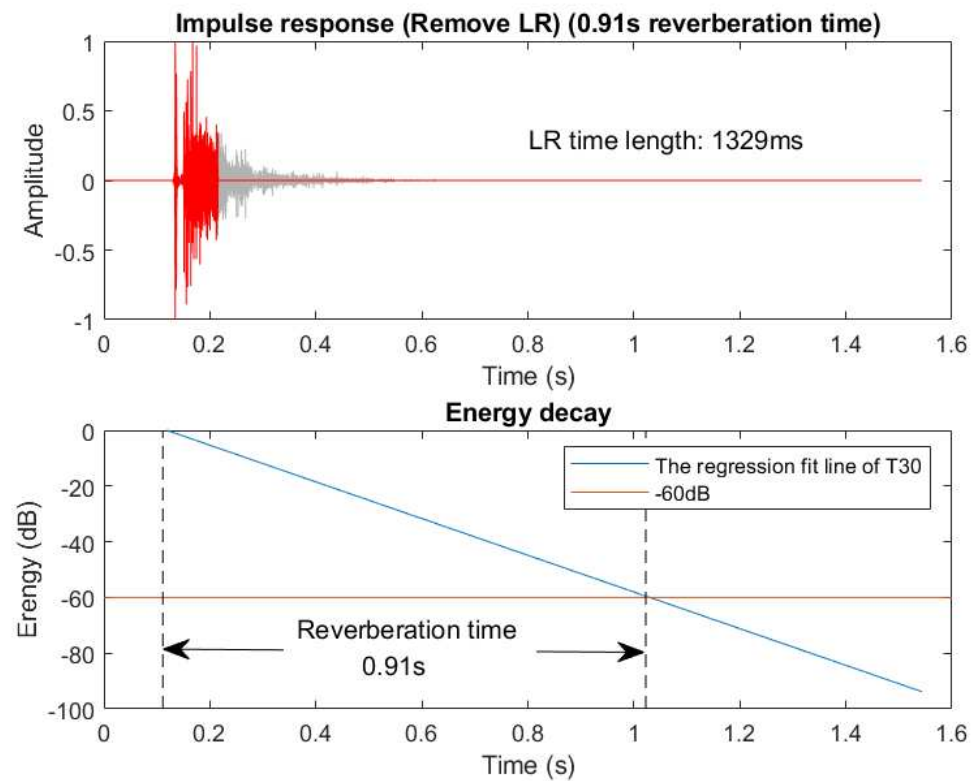


Figure A2. The time length of the LR for the BRIR with 0.91 s reverberation time.

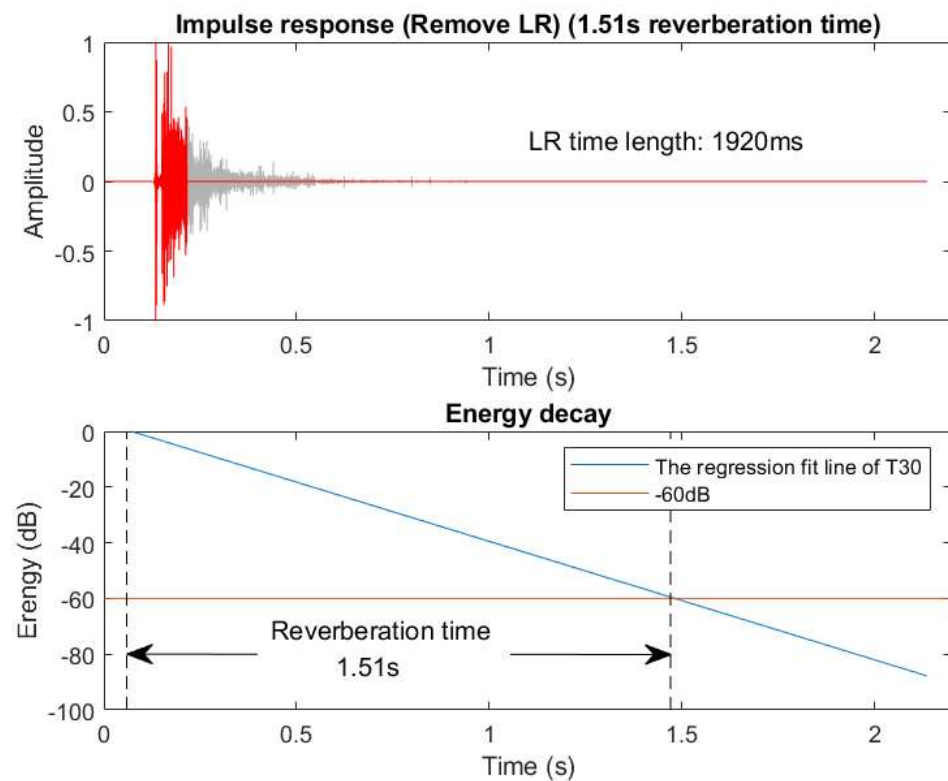


Figure A3. The time length of the LR for the BRIR with 1.51 s reverberation time.

References

1. Klein, F.; Werner, S. *The Relevance of Auditory Adaptation Effects for the Listening Experience in Virtual Acoustic Environments*; Audio Engineering Society Convention 144; Audio Engineering Society: New York, NY, USA, 2018.
2. Begault, D.R.; Wenzel, E.M.; Anderson, M.R. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *J. Audio Eng. Soc.* **2001**, *49*, 904–916. [[PubMed](#)]
3. Hacıhabiboğlu, H.; Murtagh, F. Perceptual simplification for model-based binaural room auralisation. *Appl. Acoust.* **2008**, *69*, 715–727. [[CrossRef](#)]
4. Kleiner, M.; Dalenbäck, B.I.; Svensson, P. Auralization—an overview. *J. Audio Eng. Soc.* **1993**, *41*, 861–875.
5. Scherer, S.A.; Dube, D.; Zell, A. Using depth in visual simultaneous localisation and mapping. In Proceedings of the 2012 IEEE International Conference on Robotics and Automation, St. Paul, Minnesota, USA, 14–18 May 2012; pp. 5216–5221.
6. Dokmanić, I.; Parhizkar, R.; Walther, A.; Lu, Y.M.; Vetterli, M. Acoustic echoes reveal room shape. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12186–12191. [[CrossRef](#)] [[PubMed](#)]
7. Valimaki, V.; Parker, J.D.; Savioja, L.; Smith, J.O.; Abel, J.S. Fifty years of artificial reverberation. *IEEE Trans. Audio, Speech, Lang. Process.* **2012**, *20*, 1421–1448. [[CrossRef](#)]
8. Howard, D.; Angus, J. *Acoustics and Psychoacoustics*; Routledge: Abingdon-on-Thames, UK, 2013.
9. Kaplanis, N.; Bech, S.; Jensen, S.H.; van Waterschoot, T. Perception of reverberation in small rooms: A literature study. In Proceedings of the Audio Engineering Society Conference: 55th International Conference: Spatial Audio, Helsinki, Finland, 27–29 August 2014; Audio Engineering Society: New York, NY, USA, 2014.
10. Rumsey, F. Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. *J. Audio Eng. Soc.* **2002**, *50*, 651–666.
11. Beranek, L.L. Concert Halls and Opera Houses: Music, Acoustics, and Architecture. *J. Acoust. Soc. Am.* **2005**, *117*, 987. [[CrossRef](#)]
12. Jot, J.M.; Larcher, V.; Warusfel, O. *Digital Signal Processing Issues in the Context of Binaural and Transaural Stereophony*; Audio Engineering Society Convention 98; Audio Engineering Society: New York, NY, USA, 1995.
13. Hartmann, W.M. Localization of sound in rooms. *J. Acoust. Soc. Am.* **1983**, *74*, 1380–1391. [[CrossRef](#)] [[PubMed](#)]
14. Rakerd, B.; Hartmann, W. Localization of sound in rooms, II: The effects of a single reflecting surface. *J. Acoust. Soc. Am.* **1985**, *78*, 524–533. [[CrossRef](#)] [[PubMed](#)]
15. Rakerd, B.; Hartmann, W.M. Localization of sound in rooms. V. Binaural coherence and human sensitivity to interaural time differences in noise. *J. Acoust. Soc. Am.* **2010**, *128*, 3052–3063. [[CrossRef](#)] [[PubMed](#)]
16. Hyde, J.R. Discussion of the Relation between Initial Time Delay Gap (ITDG) and Acoustical Intimacy: Leo Beranek's Final Thoughts on the Subject, Documented. *Acoustics* **2019**, *1*, 561–569. [[CrossRef](#)]
17. Beranek, L.L. Concert hall acoustics—1992. *J. Acoust. Soc. Am.* **1992**, *92*, 1–39. [[CrossRef](#)]

18. Beranek, L. *Concert Halls and Opera houses: Music, Acoustics, and Architecture*; Springer: Berlin/Heidelberg, Germany, 2004.
19. Gölzer, H.; Kleinschmidt, M. Importance of early and late reflections for automatic speech recognition in reverberant environments. In Proceedings of the Elektronische Sprachsignalverarbeitung (ESSV), Karlsruhe, Germany, 3 March 2003. Available: http://medi.uni-oldenburg.de/members/michael/papers/Goelzer_Kleinschmidt_ESSV2003.pdf. (Accessed on 3 March 2022).
20. Gardner, W.G. Reverberation algorithms. In *Applications of Digital Signal Processing to Audio and Acoustics*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 85–131.
21. Kuttruff, K.H. Auralization of impulse responses modeled on the basis of ray-tracing results. *J. Audio Eng. Soc.* **1993**, *41*, 876–880.
22. Lindau, A.; Kosanke, L.; Weinzierl, S. *Perceptual Evaluation of Physical Predictors of the Mixing Time in Binaural Room Impulse Responses*; Audio Engineering Society Convention 128; Audio Engineering Society: New York, NY, USA, 2010.
23. Defrance, G.; Polack, J.D. Measuring the mixing time in auditoria. *J. Acoust. Soc. Am.* **2008**, *123*, 3499. [[CrossRef](#)]
24. Väänänen, R. Efficient Modeling and Simulation of Room Reverberation. Master's Thesis, Helsinki University of Technology, Helsinki, Finland, 1997.
25. Naylor, G.M. ODEON—Another hybrid room acoustical model. *Appl. Acoust.* **1993**, *38*, 131–143. [[CrossRef](#)]
26. Farina, A. *Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique*; Audio Engineering Society Convention 108; Audio Engineering Society: New York, NY, USA, 2000.
27. Algazi, V.R.; Duda, R.O.; Thompson, D.M.; Avendano, C. The cipic hrtf database. In Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575), New Platz, NY, USA, 24 October 2001; pp. 99–102.
28. Boley, J.; Lester, M. *Statistical Analysis of ABX Results Using Signal Detection Theory*; Audio Engineering Society Convention 127; Audio Engineering Society: New York, NY, USA, 2009.
29. Cornsweet, T.N. The staircase-method in psychophysics. *Am. J. Psychol.* **1962**, *75*, 485–491. [[CrossRef](#)] [[PubMed](#)]
30. Brodén, D.A.; Paridari, K.; Nordström, L. MATLAB applications to generate synthetic electricity load profiles of office buildings and detached houses. In Proceedings of the 2017 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia), Auckland, New Zealand, 4–7 December 2017; pp. 1–6.
31. ITU. *Method for the Subjective Assessment of Intermediate Quality Level of Audio Systems*; BS Series; ITU: Geneva, Switzerland, 2014.
32. Srinivasan, N.K.; Stansell, M.; Gallun, F.J. The role of early and late reflections on spatial release from masking: Effects of age and hearing loss. *J. Acoust. Soc. Am.* **2017**, *141*, EL185–EL191. [[CrossRef](#)] [[PubMed](#)]
33. EBU-Recommendation. *Loudness Normalisation and Permitted Maximum level of Audio Signals*; European Broadcasting Union: Geneva, Switzerland, 2011.
34. ITU. *Algorithms to Measure Audio Programme Loudness and True-Peak Audio Level*; BS Series; ITU: Geneva, Switzerland, 2011.
35. National Research Council. *Hearing Loss: Determining Eligibility for Social Security Benefits*; National Academies Press: Washington, DC, USA, 2004.
36. Ermann, M. *Architectural Acoustics Illustrated*; John Wiley & Sons: Hoboken, NJ, USA, 2015.