buildings

MDPI

Article

Monoaural and Binaural Classroom Acoustics for the Assessment of Speech Comprehension in Primary-School Children

Simone Secchi [1,*], Arianna Astolfi [2], Veronica Amodeo [1], Lucia Bigozzi [3], Fabio Brocchi [1], Giuseppina Puglisi [4], Louena Shtrepi [2] and Giulia Vettori [3]

1 Department of Architecture, University of Florence, 50121 Florence, Italy; veronica.amodeo@unifi.it (V.A.); fabio.brocchi@onda-quadra.it (F.B.)
2 Department of Energy, Politecnico di Torino, 10129 Turin, Italy; arianna.astolfi@polito.it (A.A.); louena.shtrepi@polito.it (L.S.)
3 Department of Education, Languages, Interculture, Literature and Psychology, University of Florence, 50121 Florence, Italy; lucia.bigozzi@unifi.it (L.B.); giulia.vettori@unifi.it (G.V.)
4 Building Heritage, Management, Design and Construction and Workplace Safety (PROGES Department), Politecnico di Torino, 10129 Turin, Italy; giuseppina.puglisi@polito.it
* Correspondence: simone.secchi@unifi.it

Abstract

The paper presents the results of monoaural and binaural acoustic measurements and speech-intelligibility tests conducted in two classrooms representative of typical school buildings in Italy. The sample of two classrooms was selected based on a statistical analysis of reverberation time measurements taken in more than one hundred Italian schools. In the two sample classrooms, measurements were taken both when the room was empty and when it was occupied by pupils. The binaural measurements were conducted on two students who were fitted with special headphones equipped with binaural microphones, with sound sources placed in various positions within the classroom, both in the acoustically untreated classroom (with a long reverberation time) and in the acoustically treated classroom (with a short reverberation time). The results of the speech intelligibility tests conducted on 25 elementary school students confirmed that speech intelligibility improves significantly in the acoustically treated classroom.

Keywords: schools; reverberation time; intelligibility; binaural acoustic parameters; room acoustics; spatial release from masking; interaural time and level differences

1. Introduction

The negative effect of reverberation on children's speech intelligibility has been tested in-field in different classroom scenarios accounting for different types of noise. Astolfi et al. (2012) found that traffic noise is the most interfering noise followed by babble noise, but a lowering in reverberation time brought an increase in speech intelligibility only with traffic noise located outdoors and not with babble noise, whose source was in the middle of the classrooms [1]. In a study by Buss et al. (2017), carried out without reverberation in a laboratory, it has been proved that for children aged 5 to 16 years and for adults, a two-talker masker is more detrimental for speech intelligibility than a one-talker masker, due to the increased cognitive resources needed to perceptually isolate the target in the context [2].

In their review, Leibold and Buss (2019) [3] found that adult-like performance for speech-in-speech recognition is not usually observed until adolescence. Difficulties are due

Check for updates

Academic Editor: Kai Zhang

Received: 7 August 2026
Revised: 12 September 2026
Accepted: 20 September 2026
Published: 23 September 2026

in part to the immature allocation of attention to the relevant target speech and extensive listening experience is required before maturing selective attention abilities. Additionally, working memory seems to play an important role in the development of speech-in-noise recognition. Children with high working memory appear to have high performance on masked speech-recognition tests. The central auditory process is responsible for selective attention, which allows isolation of target and masker speech by auditory grouping promoted by different vocal characteristics, temporal onsets, and binaural cues. Children, however, seem to require more acoustic cues than adults to perceptually isolate target and masker speech [3].

The speech-intelligibility benefit from spatial separation of the source from the masker is known as a Spatial Release from Masking (SRM), which is typically computed as the difference in Speech Recognition Threshold (SRT) between conditions in which the target and masker are co-located versus spatially separated. Speech Recognition Threshold (SRT) is the signal-to-noise ratio (SNR) yielding a certain % correct recognition score (i.e., for example, Speech Recognition Threshold 80%, SRT80 in dB SNR, is the signal-to-noise ratio yielding an 80% correct recognition score). SRM is likely associated with at least two binaural processes, which are the "better-ear listening" and the "binaural unmasking" [4]. The former is related to the listener's ability to detect sound from the ear which has a better signal-to-noise ratio, and the latter to the differences in the interaural phase between signal and noise sources, caused by differences in path distance to each ear, exploited by the brain to detect a signal. These processes are related to the Interaural Time and Level Differences (ITDs and ILDs) [5–7]. ITD is the difference in arrival time of a sound wave at the two ears, due to the path length difference from the source to each of the ears. ILD is caused by the attenuation due to head shadowing at high frequencies. ITD and ILD are the most important cues for the auditory system to localise sound sources.

SRM was found for children as young as 4 years in the case of both asymmetrical [8,9] and symmetrical interference configurations [9]. Asymmetrical interference configuration usually considers the target at 0° azimuth and the maskers at +90° or –90°, while in the symmetrical configuration, the maskers are either at +90° or –90. The asymmetrical configuration provides access to both monaural head shadow and interaural differences cues [10], while when maskers are displaced symmetrically, the monaural head shadow is minimized, resulting in a reduced availability of spatial cues for source segregation [11]. However, no significant difference in SRM has been found in the asymmetrical and symmetrical configurations by Misurelli and Litovsky (2012) [9], except when the target and the maskers consist of the same-sex stimuli in the symmetrical interferer condition [12]. In this case, SRM is significantly reduced because not only monoaural cues are reduced, but also spectral difference cues. Furthermore, no significant differences have been found between adults and children in anechoic conditions [10,11]. On the contrary, Peng and Litovsky (2021) [6] observed SRM immaturity in children between 6 and 15 years of age compared to adults. They found that the adult-child difference of $\sim$3 dB SRM was mainly driven by children's poorer ability to use interaural differences rather than head shadow [6].

The above-referenced studies refer to anechoic conditions in the laboratory, while only a recent study, to the best of the Authors' knowledge, has concerned SRM with reverberation in typical classroom conditions. Peng et al. (2021) [13] found that children in the 7–12 years age range have not reached maturity in using binaural cues that are distorted in reverberation, as adults. They demonstrated 3–8 dB SRM on average in reverberant environments, which was much smaller than that observed for adults (10–13 dB) [13].

Reverberation does not affect the sound-source localization ability [14]. Indeed, the "precedent effect in sound localization" (Wallach et al., 1949) claims that a sound source is localized based on the location of the first arriving sound more than on the directions of the

https://doi.org/10.3390/buildings16193788

reflected sounds arriving later and from many directions unrelated to the source [15]. The "echo threshold" is the briefest delay between the first arriving sound, i.e., the direct sound, and the delayed sound, i.e., the first arriving reflected sound, which allows the listener to perceive one single sound instead of two spatially separated sounds. This is named "fusion", and the localization of the fused signal is at or near the location cued by the first of the two signals, i.e., the direct sound. In the literature, a wide range of echo thresholds is reported, which is from 2 to 100 ms, depending on many factors including age, hearing status, and differences in stimuli [16]. It also depends on the signal since it is roughly 1 ms for short impulses, 50 ms for ongoing speech, and 80 ms for classic/romantic music [4]. Children about 5 years old exhibited significantly longer echo thresholds than adults, which leads to a more robust fusion effect on one side, but also to a deficit in localization of the delayed sound when this threshold is within the limit for adults on the other side, which suggests an immature spatial hearing in children [17,18].

Even if the localization of a sound source is not affected by reverberation, increasing reverberation reduces ITD and ILD and this results in poorer performance in speech identification. However, in the case of adult listeners, it has little effect on the benefit of spatial separation of the target and the masker, i.e., it does not affect the SRM in the case of informational maskers, which is preserved compared to anechoic conditions and is higher compared to the case of energetic maskers [19]. This represents the condition more often found in real environments where listeners use the "precedent effect" to derive localization from signal transients and thus perceive the target as distinct from the masker.

Binaural listening appears to have limited effects on reducing the impact of noise on learning in children. Leist et al. (2022) compared a realistic classroom-noise scenario presented either monaurally, with identical signals delivered to both ears, or binaurally, preserving interaural differences and spatial information, while children performed tasks involving auditory and visual information processing [20]. Word identification was more impaired in the monaural than in the binaural condition, whereas listening comprehension and short-term memory were similarly affected in both conditions. These findings suggest that irrelevant speech effects are mainly driven by cognitive interference between automatic speech processing and verbal information maintenance, while attention capture elicited by spatially separated and changing sound sources plays only a minor role.

Some advantages of using child HATSs (Head And Torso Simulators) rather than adult HATSs and/or omnidirectional measurements have been observed when assessing A-weighted equivalent sound-pressure levels in classrooms, due to their anthropomorphic characteristics [21]. For prominent ear measurements, no differences are observed between adult and child HATSs and the omnidirectional microphone up to 2 kHz. However, differences emerge around 4 kHz between HATS measurements and the omnidirectional microphone, while above 8 kHz, differences can be observed between child and adult HATSs, with higher values recorded by child HATSs. Notably, sharpness values also differ across transducers, with child HATSs showing greater high-frequency content compared with adult HATSs.

Despite the extensive literature on speech recognition in noise and on the use of binaural cues for source segregation, most studies have been conducted under anechoic and laboratory conditions. Evidence collected in real classrooms remains limited, although reverberation and "classroom bubble" noise may substantially degrade the linguistic signal and increase the cognitive load on children. Importantly, speech understanding in classrooms is not an isolated outcome, but a prerequisite for higher-level cognitive processes, such as verbal working memory. Compared to attention or reading comprehension, working memory is more sensitive to the acoustic environment [22] and exhibits stronger plasticity during the primary-school stage [23]. As reported in the literature in suboptimal school acoustic environments, characterized by long reverberation times and background babble noise, the extra cognitive effort required to process and restore degraded auditory signals directly consumes phonological and executive control resources, thereby reducing young children's effective verbal working-memory capacity [24]. Because primary-school children are in a critical developmental window marked by ongoing maturation and high neurocognitive plasticity, verbal working memory serves as a foundational cognitive bottleneck for broader learning activities (including reading comprehension and task attention). Focusing verbal working memory thus provides a more sensitive and direct mechanistic measure of how classroom acoustic quality affects young children's cognitive processing.

Systematic reviews demonstrate that classroom acoustic treatments (e.g., sound absorbers and diffusers) generally reduce reverberation and improve speech perception, attention, and student well-being across diverse educational settings. However, excessive absorption can inadvertently lower signal intensity and Speech Transmission Index (STI), thereby increasing listening effort. Furthermore, the overall effectiveness of these interventions varies depending on specific noise conditions, spatial configurations, and individual cognitive or perceptual vulnerabilities [25].

In this framework, global room-acoustic parameters such as Reverberation Time (RT), Clarity (C_{50}), and Speech Transmission Index (STI) are particularly relevant to describe classroom communication quality. Binaural measures such as interaural time and level differences (ITDs and ILDs), interaural cross-correlation (IACC), and direct-to-reverberant ratio (DRR) may provide complementary information to interpret perceptual differences between rooms with different acoustic quality. While monaural parameters describe the room acoustics, the binaural parameters focus on the listeners perspective. The paper aims to investigate which binaural cues are and are not affected by acoustic treatments and therefore identifying the likely perceptual mechanisms behind the observed cognitive benefit.

The objective of the present study is to compare monoaural and binaural acoustic parameters measured in two "twin" primary-school classrooms in Italy, identical in volume and layout but characterized by very different reverberation, with the results of comprehension tests based on verbal working memory, administered in an ecological setting. A classroom representative of typical reverberant conditions in Italian historical school buildings is compared with an acoustically treated classroom, by analyzing classical room-acoustic parameters (RT, C_{50}, STI) and binaural parameters, as well as children's performance on the Listening Span Test under controlled "classroom bubble" noise.

2. Materials and Methods

To investigate the impact of classroom acoustics on primary-school children's listening and learning performance in an ecological environment, an experimental methodology combining acoustic characterization and cognitive assessment was established. The overall workflow, illustrated in Figure 1, outlines the sequential steps of the investigation. First, a statistical analysis of typical Italian primary-school classrooms guided the selection of two twin case-study rooms in Florence, identical in layout but differing in acoustic treatment (untreated classroom C1 vs. acoustically optimized classroom C2). Subsequent room acoustic measurements—incorporating both monaural and binaural parameters under unoccupied and occupied conditions—were performed to thoroughly characterize the sound field. Finally, comprehension tests were administered under controlled noise conditions to a selected sample of primary-school pupils to evaluate cognitive performance across the two acoustic environments. Each stage of this experimental investigation is detailed in the following subsections.

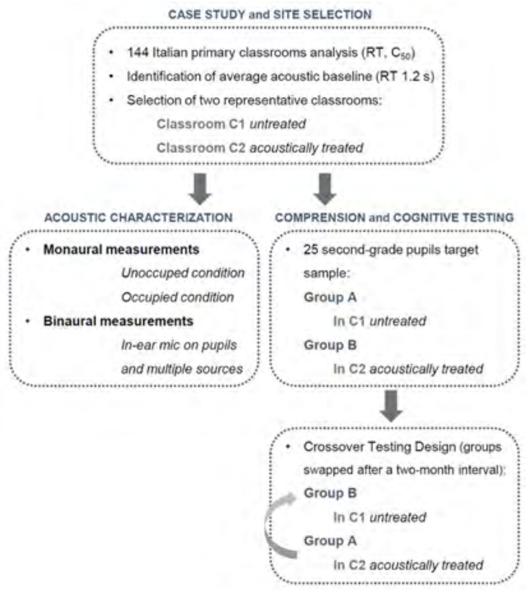

Figure 1. Experimental methodology workflow, including site selection, monaural and binaural acoustic characterization, and the crossover design for comprehension testing in the two sample classrooms (untreated C1 and acoustically treated C2). The arrow indicates the swap between Group A and Group B across the environments (C1 and C2), after two months.

2.1. Description of the Case Study

In order to carry out in a real context the study on the effect of acoustic quality in the comprehension and learning process of primary-school children, it was necessary to identify two sample classrooms. To this end, a statistical analysis was conducted on a sample of one hundred forty-four Italian schools (many of them located in the city of Florence). The results showed that part of this sample was composed of a classroom with no acoustic treatment while another part had some kind of acoustic treatment (usually an absorbing ceiling). The reverberation time and clarity measurements in 1/3 octave frequency bands were carried out in all classrooms (n = 144). All classrooms were not occupied by pupils and normally furnished. The average values of reverberation time between third octave bands of 200–5000 Hz (i.e., between octave bands of 250–4000 Hz) was analyzed based on the requirements of Italian standard UNI 11367 annex C [26].

Figure 2 shows the statistical distribution of average value of reverberation time in the sample of 144 classrooms.

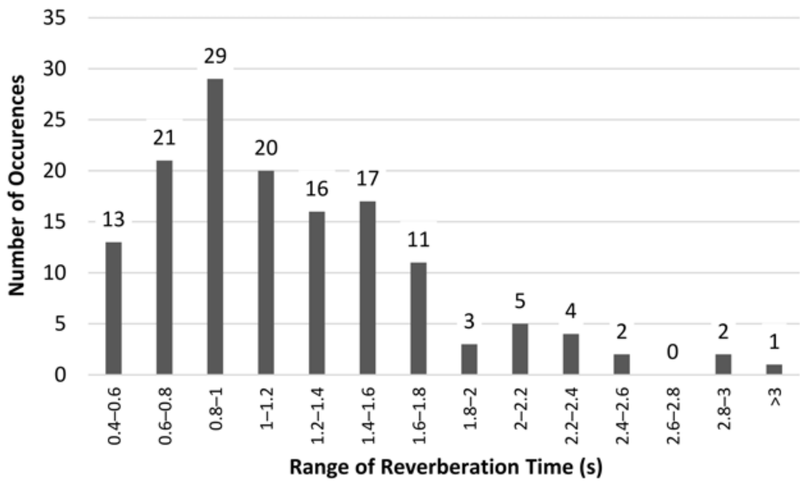

Figure 2. Statistical distribution of results of reverberation time (average 250–4000 Hz) in a group of 144 classrooms.

The average value of the RT in this sample was 1.2 s, the median was 1.1 s, and the 80th percentile was 1.6 s.

To carry out the tests, it was consequently necessary to select a classroom representative of the typical Italian classrooms, i.e., with an average value of RT of about 1.2 s in the octave bands range of 250–4000 Hz.

The sample classroom (C1) of the current study was selected in a primary school of the city of Florence that was built as a convent before 1800; the adaptation to the current scholastic use dates back to the 1950s. This school is representative of many other schools where no acoustic treatment was realized.

The average unoccupied reverberation time of the selected classroom was 1.3 s in the octave band range of 250–4000 Hz; therefore, near to the average value of the sample of 144 Italian classrooms.

To carry out the test it was necessary to select two different classrooms in the same school: the one described above, with high reverberation time (C1), and another with low reverberation time (C2). To this aim, a classroom of the same school building treated with a sound-absorbing ceiling was selected as the second sample classroom.

To better represent the two conditions of optimal and inadequate acoustics, additional sound-absorbing panels were added to the walls of classroom with the ceiling treatment, to obtain in this classroom a value of the reverberation time adequate to the optimum value defined by the Italian standard UNI 11532-2 [27,28] for classrooms accommodating hearing-impaired pupils.

The Italian standard UNI 11532-2 [28] provides the following equation to define the optimum value of reverberation time for classroom housing pupils with hearing defects.

$$RT_{opt} = 0.32lg(V) - 0.17(s)$$

where V is the volume (m^3) of the classroom.

Considering the room (C2) dimensions of 7.2 × 6.2 × 3.8 m^3 (see Figure 3, right), the optimal reverberation time according to UNI 11532-2 is about 0.5 s.

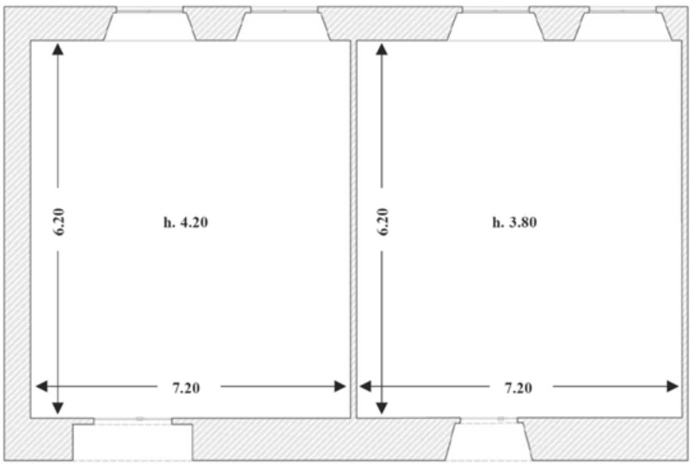

Figure 3. The case study: two twin classrooms with different acoustic quality. The untreated classroom C1 on the left, and the treated classroom C2 on the right.

The average reverberation time obtained in this classroom (C2) after the insertion of additional absorbing panels to the walls was about 0.4 s in the octave bands range of 250–4000 Hz; therefore, about one third of the reverberation time in the other classroom.

In this way, the two test rooms chosen presented the same boundary conditions, except for the sound-absorbing material in one of them (in the ceiling and in part of the walls).

2.2. Room Acoustics Measurements

2.2.1. Monaural Measurements

Different procedures were used for the measurements carried out in the whole sample of 144 classrooms and in the case study of the two classrooms (untreated C1 and treated C2) selected as case study.

Room acoustic measurements carried out in the sample of 144 classrooms were performed with a two-channel class 1 real-time analyzer 01 dB (from 2012 Acoem, Limonest, France) Symphonie with preamplifiers and 1⁄2-inch gross diffuse-field microphones. Before each measurement, the microphones were calibrated.

An impulsive type of sound source (clapper II Bielettro, Castellare di Pescia, PT, Italy) was used for sound-field generation for reverberation and clarity measurements.

In total, 10–16 impulse response measurements were analyzed and averaged for each room, depending on room volume. Every measurement performed was in third-octave frequency bands between 50 and 5000 Hz.

Measurements were performed within rooms in furnished and unoccupied conditions, following ISO 3382-2 [29] and Italian standard UNI 11532-1 [27]. Microphones were placed at a distance of at least 1 m from each surface and at a height of at least 1.2 m to simulate the ear height of a seated child.

From impulse response measurements, spatial average of reverberation time, RT, and Clarity, C_{50}, were obtained for each classroom.

For Speech Transmission Index, STI, measurements, a Talk-Box Bedrock model BT65, and an analyzer Bedrock SM 90 were used (Bedrock, Rijswijk, Netherlands).

With reference to the two case studies C1 and C2, acoustic measurements were performed both in unoccupied and occupied conditions, in the presence of pupils while administering the tests, which are described in a dedicated section below. In the occupied condition, the measurement of the impulse response was performed using the omnidirectional source 01 dB omni, generating a sine sweep signal (frequency range of 20–20.000 Hz), since the clapper produces an impulsive noise which can be harmful to children's hearing. The two sound sources (Clapper and Dodecahedron) are comparable in the frequency range of 0.5–8 kHz [30].

Impulse responses were measured in 18 different positions randomly distributed in each of the two sample classrooms. Background noise was measured in two different points in each classroom.

2.2.2. Binaural Measurements

The two sample classrooms (C1 and C2) were also characterized by binaural measurements in unoccupied conditions. In particular, the following binaural parameters were measured for the two listening positions and two sound sources: Interaural Time Difference (ITD), Interaural Level Difference (ILD), binaural C_{50}, InterAural Cross Correlation (IACC), Direct-to-Reverberant Ratio (DRR) [31].

Interaural Level Difference (ILD) and Interaural Time Difference (ITD) were obtained to investigate the perceived sound localization, while InterAural Cross Correlation (IACC) was used to retrieve information on the sound spatial impression. Direct-to-Reverberant energy Ratio (DRR) was also obtained for the two ears and as the average value between the two ears. Speech Clarity (C_{50}) was evaluated as a frequency-dependent parameter separately for the two ears.

The measurements were performed using a source Talkbox (NTi Audio AG, Liechtenstein) and a dodecahedron (01 dB omni) as sound sources and a binaural microphone set (In-Ear Microphone, Type 4101, Bruel&Kjaer, Danimark). Figure 4 shows the positions of the Talkbox (S1, S2, and S3), dodecahedron (S2 and S3), and listeners (L1 and L2) in the two classrooms C1 and C2. The binaural microphone set was worn by two children (11 years old), who were considered listeners in positions L1 and L2. The measurements were performed by first locating the Talkbox in the three positions and generating a sine sweep of 3 s with a frequency range of 20–20,000 Hz. The same procedure was followed for the dodecahedron measurements in the two positions S2 and S3. The Talkbox S1 position is considered as the target speaker (T), while the Talkbox S2 and S3 are considered as single-talker interferers producing unwanted noise (N). The dodecahedron in S2 and S3 positions is considered a multi-talker interferer producing unwanted noise (N). In this work, the above-mentioned parameters have been analyzed for the target and noise positions.

Figure 4. Experimental setup of binaural measurements: in classroom C1, on the left, and in classroom C2, on the right. Positions of the sources, in light grey (S1, in the target speech position at 3.6 m from the listener, S2, at one meter from the listeners, S3 at one meter from the listener, 120°), and listeners, in dark grey (L1 listener type one, L2 listener type two).

Figures 4–6 show the measurement setup, with the position of the children during the binaural measurements.

Figure 5. Experimental setup of binaural measurements in the classroom C1 (untreated): listener L1 and the two sources (S2, Talkbox at one meter from the listener L1 and dodecahedron S3, at one meter from the listener).

Figure 6. Experimental setup of binaural measurements in the classroom C1: on the (**left**), listener type one (L1); on the (**right**), listener type two (L2).

The broadband Interaural Time Difference ITD was calculated using the threshold method described in [32]. The broadband Interaural Level Difference ILD was calculated as the energy ratio between the left and right ears. Since in human auditory perception, ITD and ILD are used complementarily for the lower and higher frequency range, respectively, the Binaural Room Impulse Responses, BRIRs, were lowpass filtered in the first case and high-pass filtered in the second, using a 10th-order Butterworth filter with a cutoff frequency of 1300 Hz, as in [33]. The broadband Direct-to-Reverberant Energy Ratio, DRR, values for the left and right ears were calculated by exploiting the open-source MATLAB R2023b toolbox in [34], which includes a function that determines the direct sound as the peak of the squared impulse response and returns the DRR value using a time window of 5 ms centered in the peak to select the direct sound [35]. The binaural speech clarity was evaluated in octave bands and provided as the average value from 250 Hz to 2 kHz [28]. Differences have also been obtained between speech clarity from the target source and each noise source, at the left and right ears, respectively, and as a mean value between the ears. These differences have been used to evaluate the gap in speech clarity between the target and the noise [36]. The JND (just noticeable difference) of IACC is assumed to be 0.075 [29]. The ITD JND is 58 μs, while the ILD JND is 1.9 dB, at mid frequencies [37]. The JND for the Direct-to-Reverberant Ratio (DRR) is approximately 2–3 dB for reference DRRs between 0 and +10 dB [38]. For classrooms, the just noticeable difference (JND) for the speech clarity C_{50}, over the frequency range 0.5–2 kHz, is approximately 1.1 dB, while an increase of about 3 dB is generally required to achieve a clearly perceptible improvement in speech intelligibility [39].

2.3. Comprehension Tests

To investigate the effect of noise on verbal working memory in elementary school children, within an ecological school environment, a test consisting of a particular listening test, Listening Span Test [40], was administered within the two case-study classrooms C1 and C2.

Detailed information about the method of analysis is contained in [41], while in the following a summary of boundary conditions and of main parameters analyzed is reported.

The test was carried out in the two selected classrooms, C1 and C2, where a dodecahedral source, placed in a corner position (Figure 7, position M, and Figure 8), reproduced a "classroom babble" recorded in a real classroom and reproduced to give an equivalent SPL in the center of the classroom of 59 dBA. This was considered as a masking sound source (M).

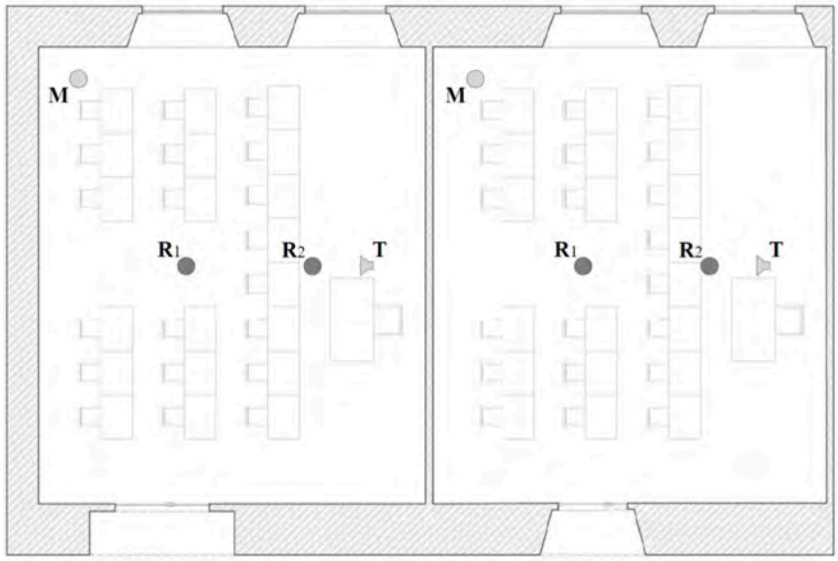

Figure 7. Experimental setup during the comprehension tests: in the classroom C1 on the (**left**), and in the classroom C2 on the (**right**). Positions of the Talkbox (T), microphones (R1 at one meter from the target speech, R2 in the middle), and the masking sound source (dodecahedron) (M).

Figure 8. Experimental setup during the comprehension tests: on the (**left**), positions of the Talkbox (T) and microphones (R2 at one meter from the Talkbox, R1 in the middle) in the classroom C1; on the (**right**), positions of the masking sound source (dodecahedron) (M) in the classroom C2. In the photo the absorbing panels placed near the walls are shown.

Both the dodecahedron reproducing the babble (M) and the Talkbox reproducing the signal (T) were calibrated to give a predetermined level in classroom C1.

The Talkbox was calibrated to give 60 dBA at 1 m of frontal distance, according to ISO 9921:2003 [42], while the omnidirectional source was calibrated to give an average level of about 59 dBA in the center of the classroom. The level of 59 dBA was chosen as representative of bubbling generated by primary-school children [43].

The average SPL generated by the dodecahedron and by the Talkbox in the two classrooms C1 and C2, which were amplified to give the same emission, was different as a consequence of different values of reverberation time as described in Figure 9.

Figure 9 shows the SPLs measured in the center of the classroom and at 1 m in front of the talk box or of the teacher in the classroom C1 and in the classroom C2. Both background noise and noise produced by the loudspeaker in position M (the bubble noise) were almost uniform in the classroom, while Talkbox speech and teacher speech were greater in front of the source. The difference between the position in front of the source and that in the center of the classroom is greater in classroom C2 (green bars in the graph of Figure 9) because of the lower reverberation time.

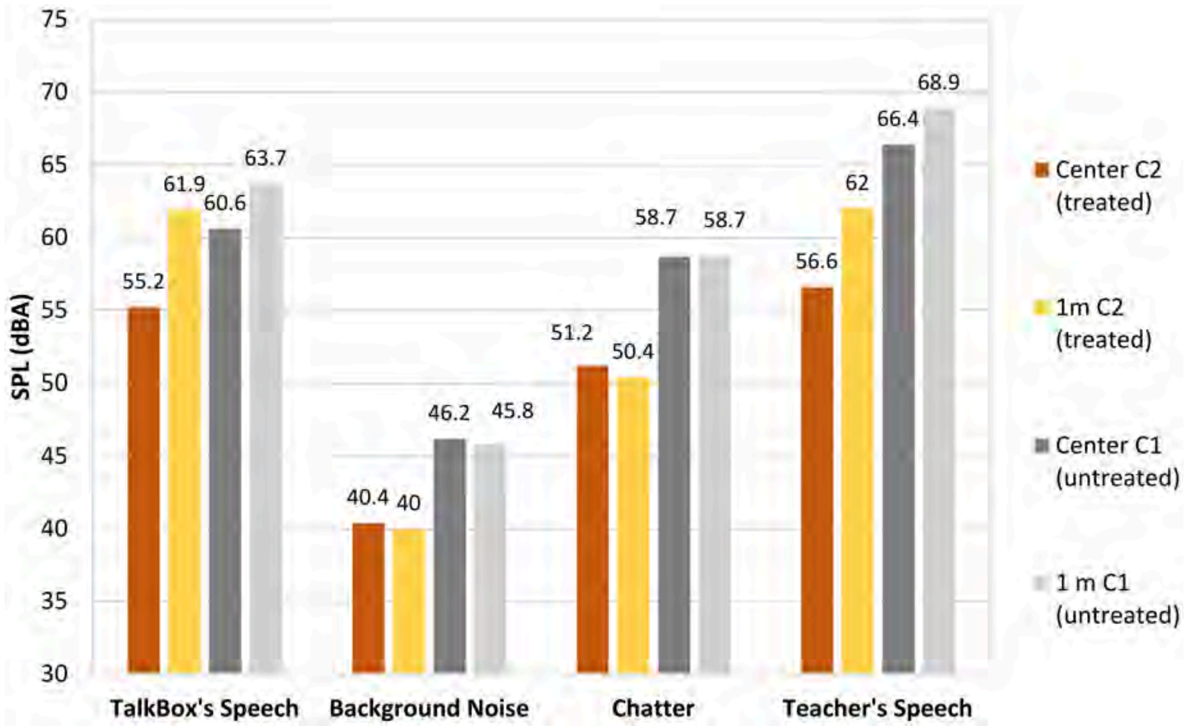

Figure 9. SPLs in classrooms C1 and C2 due to the talk box, to the teacher's speech, to the babble reproduced by the dodecahedron, and to background noise.

Figure 10 shows the spectrum SPL of the chatter noise reproduced by the dodecahedron and measured in the centre of the classroom C1. As shown in Figure 9, the A-weighted SPL corresponding to the spectrum of Figure 10 was 58.7 dBA and the signal reproduced was previously measured in a real classroom and can be considered representative of the noise produced by children babble in a classroom.

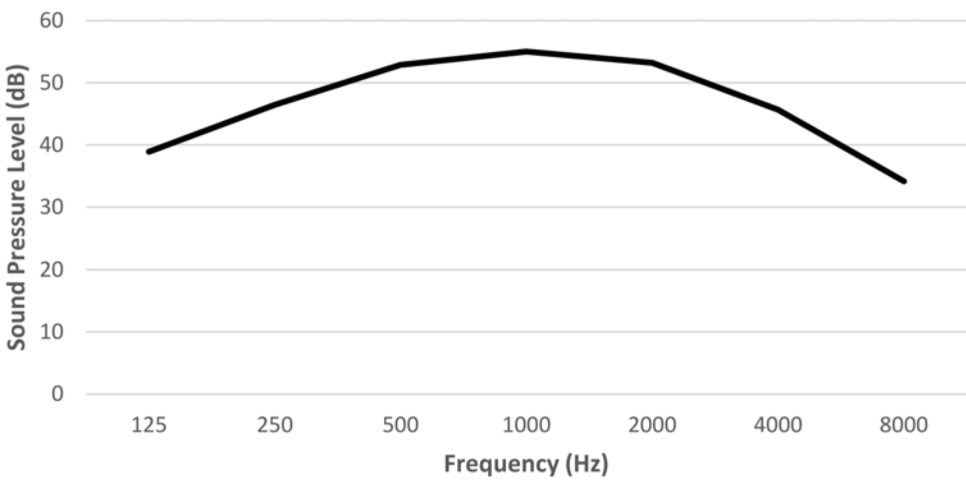

Figure 10. SPL spectrum of the babble noise reproduced by the dodecahedron and measured in the centre of the classrooms C1.

2.4. Participants and Test Design

To carry out the test, a sample of 25 s-grade children was selected and divided into two homogeneous groups (later indicated as Group A and Group B). The children involved had typical cognitive and linguistic development. None of the participants had diagnosed hearing loss or auditory processing deficits, speech, language, or learning disorders, or certified neurodevelopmental or intellectual impairments. Children performed the tests at the same time of day (in the morning from 9 to 12 am) in both classrooms.

At first, the test was administered to group A in the very reverberant classroom (the untreated classroom C1) and to group B in the classroom with acoustic treatment (the treated classroom C2); after 1 month, the test was repeated, inverting the two groups: group A in the classroom C2 with acoustic treatment and group B in the untreated classroom C1. In this way, four combinations of the tests performed were analyzed:

- Group A in the classroom C1 (untreated);
- Group B in the classroom C2 (treated);
- Group A in the classroom C2 (treated);
- Group B in the classroom C1 (untreated).

All the tests were administered by the same researcher/examiner who administered the task (Ph.D. in Psychology), and the same researcher to control the sound pressure level produced by the voice and by the babble noise and the background-noise level. The sound pressure level generated by the voice varied between 59 ÷ 61 dB at 1 m from the mouth and was monitored during all sessions with the Vocal Holter Med dosimeter (Device developed by the Polytechnic University of Turin, the University Otorhinolaryngology 2 Unit of the University of Turin, and PR.O.VOICE Srl, a Polytechnic University of Turin start-up.) [44].

2.5. Statistical Analyses

A statistical comparison of verbal working-memory performance obtained by the same children in the two different acoustic contexts was performed.

To test verbal working memory, a simplified version of Daneman and Carpenter's (1980) Listening Span Test for children from second to fifth grade, described in [41], was used.

The level of performance was determined by coding the children's responses into six indices representing the level of span, evaluation, words recalled, intrusions, inventions and inversions.

In particular, the index "Evaluations" was obtained by assigning one point for each statement heard correctly judged for its truthfulness or falsity by the child; the index "Words recalled" was obtained by assigning one point for each correct word recalled by the child; the index "Span" indicates the sequence of words of maximum length correctly remembered and written down by the child; the index "Intrusions" was obtained by counting each word recalled and written down by the child that was not in the last position in the sentence uttered by the researcher; the index "Inventions" was obtained by counting each word written down by the child with semantic or phonological similarity to the last word but not present in the sentence uttered by the researcher; finally, the index "Inversions" was obtained by counting each word recalled and written down by the child but not in the exact sequence of presentation as in the sentence uttered by the researcher.

The normality of the distribution of the results was determined using the Shapiro–Wilk test.

3. Results and Discussion

In the following, monaural and binaural acoustic-measurements results and comprehension tests results are reported for the two sample classrooms.

3.1. Monaural Acoustic-Measurements Results

Figures 11 and 12 show the average values of reverberation time and clarity C_{50} measured in the two sample classrooms with and without children.

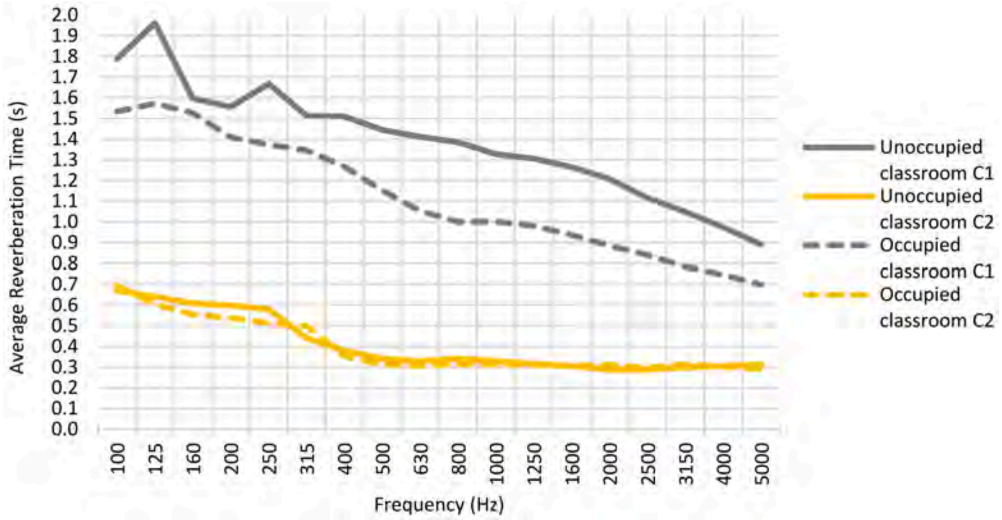

Figure 11. Average reverberation time in the two case-study classrooms C1 and C2 (untreated and treated) with (occupied) and without children (unoccupied).

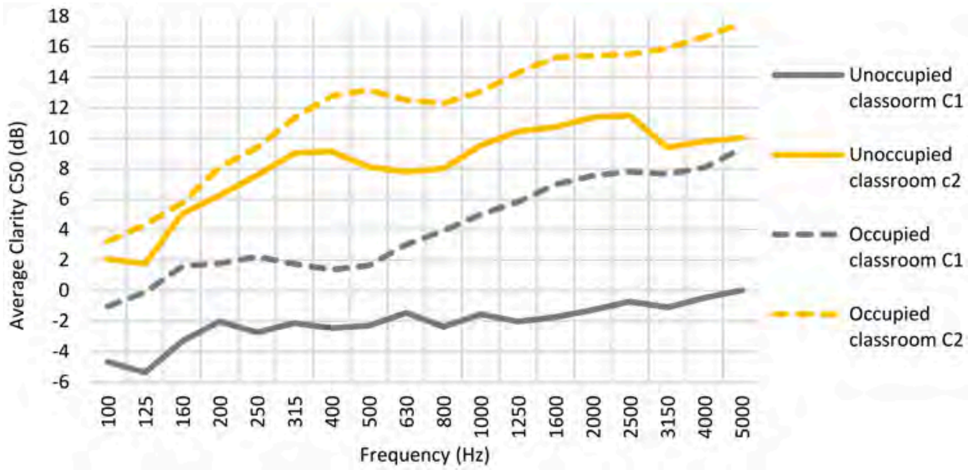

Figure 12. Average clarity C_{50} in the two case-study classrooms, C1 and C2, with (occupied) and without children (unoccupied).

During the measurements in the occupied classrooms there were 25 children, with their schoolbags and coats in the classrooms in addition to the technicians who performed the measurements.

Figure 11 shows that reverberation time in the classroom C1 reduces from 1.3 to 1.0 s from unoccupied to occupied condition, while Figure 12 reports the clarity C_{50} increases from −1.6 to +4.9 dB, as average values in the octave bands frequency range from 250 to 4000 Hz. In the classroom C2, reverberation time is almost the same without and with children, while clarity increases from 9.2 to 13.5 dB in the same frequency range.

Figure 13 shows the Speech Transmission Index, STI, evaluated in the classrooms C1 and C2 with normal background noise and with background noise generated by the dodecahedron.

The calculation of STI was carried out according to IEC 60268-16:2020 [45] with the following input data.

- Reverberation time measured in the occupied condition 1 m from the source and in the center of the classroom.
- Signal (the speech level) was measured in the same positions as the Talkbox with the sound source calibrated to give 60 dBA at 1 m in front of the loudspeaker.

- Background Noise measured both with normal activity in the classroom (Figure 13, bars on the left) and with the chatting noise generated by the omnidirectional source (Figure 13, bars on the right).

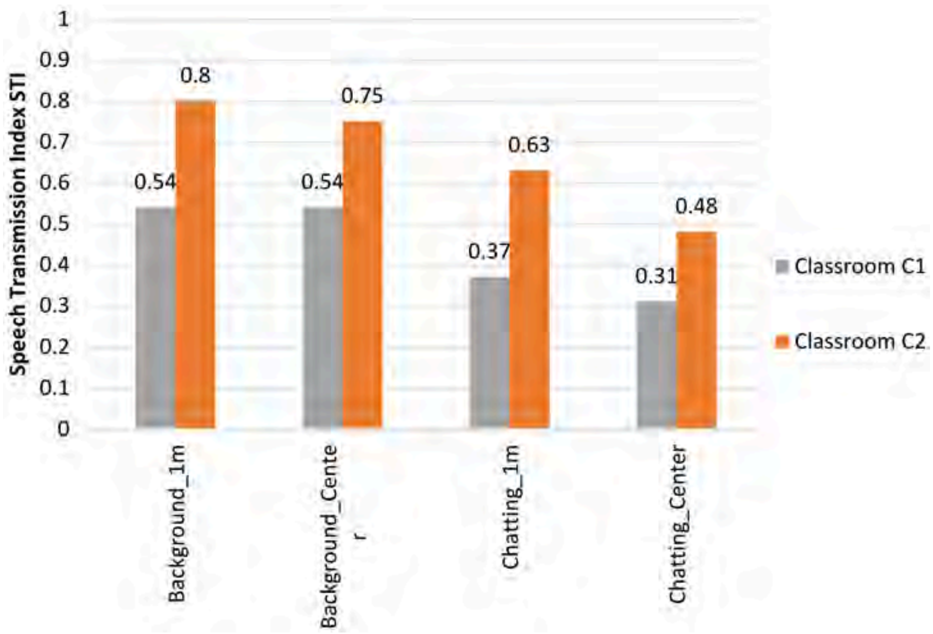

Figure 13. Speech Transmission Index in the two occupied classrooms, C1 and C2, 1 m in front of the Talkbox and in the center of the classroom, with normal background and with background due to the babble reproduced by the dodecahedron.

It is evident that the STI reduces when moving from 1 m in front of the source to the center of the classroom and when introducing the chatting noise, but mainly when considering the reverberant classroom instead of the absorbing classroom.

3.2. Binaural Acoustic Measurements Results

Figure 14 shows the binaural clarity C_{50} evaluated for the two listeners and two sources in three different positions for the treated C2 and untreated C1 classrooms. The following outcomes can be drawn:

- The results obtained from the two children's heads are consistent.
- C_{50} is approximately 2–4 dB higher for the Talkbox source, which has the directivity of a human talker and simulates a single-talker noise source, than for the dodecahedron, an omnidirectional source that simulates multi-talker background noise.
- The left and right ears show similar C_{50} values when the source is positioned at 0° azimuth. As expected, the right ear exhibits higher clarity up to 6–8 dB at the highest frequencies than the left ear when the noise source is located at 120° azimuth.
- C_{50} is higher for all sound sources (target and noise) in the treated classroom across all octave bands, with improvements exceeding 10 dB in most cases. The only exceptions occur at low frequencies and for the dodecahedron source at 1 kHz in S2, where the increase is approximately 4 dB.
- C_{50} generally increases with frequency in both the treated and untreated classrooms and for both ears. The main exceptions are a negative peak at 1 kHz for the source in S2, particularly with the dodecahedron, and the source in S3, where C_{50} remains relatively stable above 250 Hz.
- A local minimum around 1 kHz was observed only in the acoustically treated classroom when the Talkbox and the dodecahedron were positioned at 0° azimuth and 1 m from the listener. This feature was not present in the untreated classroom. A possible

explanation is that the reduction of late reverberation could have made interference effects between the direct sound and residual early reflections more evident. In the untreated classroom, the higher reverberant energy may have masked these spectral irregularities, resulting in a smoother frequency response.

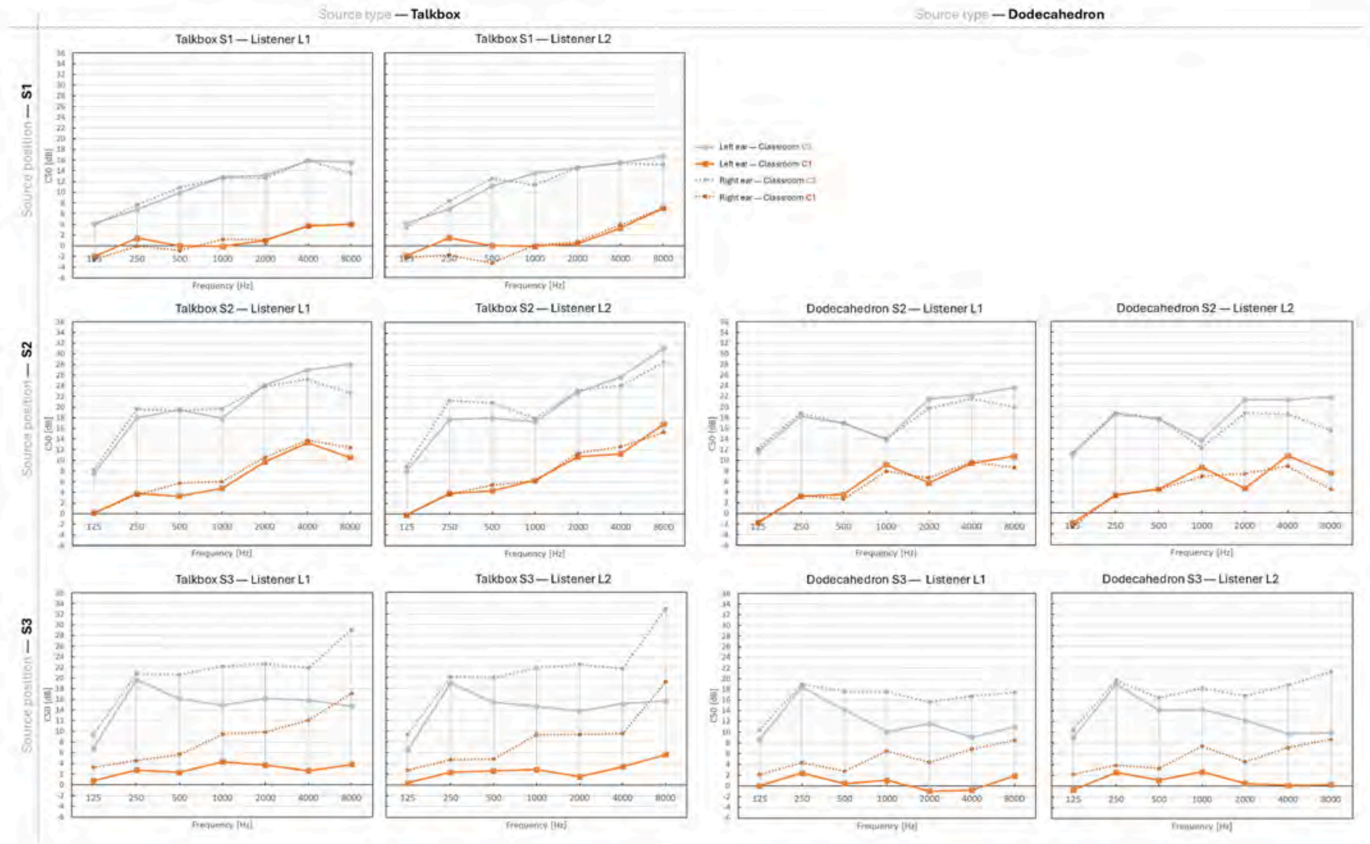

Figure 14. Binaural Clarity in the untreated (C1) and treated (C2) classrooms for two different source types (Talkbox and dodecahedron) located in three positions (S1, S2, and S3).

Table 1 shows the values of ITD, ILD, IACC, C50, and DRR computed from the binaural impulse responses. It can be seen that:

- The results obtained from the two children's heads are consistent across all source positions and acoustic treatments, except for IACCearly and IACCfull. Differences in ILDearly and ILDfull are observed only for the omnidirectional noise source located at 120° azimuth in the acoustically treated classroom/
- C_{50} and DRR (left, right, and mean) are consistently higher in the treated classroom, with differences exceeding their respective JNDs by more than ten and three times, respectively. In contrast, ITD, ILD, and IACC do not exhibit clear trends, as the observed differences are generally within or close to their corresponding JNDs. However, ITD shows a slightly higher increase than JND in the case of the child L1 when the nose source, either the Talkbox or the Dodecahedron, was at 120° azimuth at 1 m from the listener. The head size effect of the child L2 leads to a smaller change in ITD amplitude due to classroom acoustics. The generally higher IACC values observed for both children in the acoustically treated classroom may be related to its lower reverberation, as reverberation tends to decorrelate the signals at the two ears [38].

Table 1. ITD (low-pass filtered with a cut-off frequency of 1.3 kHz), ILD (high-pass filtered with a cut-off frequency of 1.3 kHz), broadband IACC, left ear, right ear, and within-ear mean DRR (broadband) and C_{50} (average from 250 Hz to 2 kHz) values computed from the binaural impulse responses.

	Listener L1										**Listener L2**									
Talker Type	**Single-Talker Noise (Talkbox)**						**Multi-Talker Noise (Dodecahedron)**				**Single-Talker Noise (Talkbox)**						**Multi-Talker Noise (Dodecahedron)**			
Speech/Noise position	T@3.6 m,0°		N@1 m,0°		N@1 m,120°		N@1 m,0°		N@1 m,120°		T@3.6 m,0°		N@1 m,0°		N@1 m,120°		N@1 m,0°		N@1 m,120°	
Classroom	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2	C1	C2
ITDearly (μs)	41.7	62.5	41.7	20.8	541.7	625.0	48.6	27.8	444.4	520.8	55.6	62.5	20.8	20.8	583.3	562.5	55.6	41.7	486.1	465.3
ITDfull (μs)	34.7	62.5	34.7	20.8	541.7	625.0	48.6	27.8	416.7	520.8	55.6	62.5	20.8	20.8	583.3	562.5	55.6	41.7	437.5	465.3
ILDearly (dB)	−0.7	−0.2	−1.2	1.2	−9.9	−10.6	0.8	1.0	−6.2	−6.6	−1.0	−1.7	−0.5	−0.4	−8.4	−10.6	0.0	2.0	−6.5	−9.6
ILDfull (dB)	−0.7	−0.2	−1.1	1.2	−9.0	−10.6	0.7	1.0	−5.3	−6.5	−1.0	−1.7	−0.5	−0.4	−7.6	−10.6	0.0	2.0	−5.4	−9.5
IACCearly (-)	0.5	0.6	0.9	0.8	0.4	0.5	0.6	0.7	0.2	0.3	0.5	0.7	0.8	0.9	0.5	0.4	0.7	0.7	0.3	0.4
IACClate (-)	0.1	0.2	0.2	0.4	0.1	0.3	0.2	0.3	0.3	0.3	0.1	0.2	0.1	0.4	0.1	0.3	0.2	0.4	0.2	0.4
IACCfull (-)	0.4	0.6	0.8	0.8	0.3	0.5	0.5	0.7	0.2	0.3	0.4	0.7	0.8	0.9	0.4	0.4	0.6	0.7	0.2	0.4
C50_L (dB)	−0.1	12.5	5.1	18.6	2.0	13.1	5.6	16.8	0.9	11.3	−0.4	11.6	6.5	19.5	1.2	14.3	5.8	16.8	−0.3	12.7
C50_R (dB)	0.1	12.4	6.7	19.9	6.8	19.7	5.7	16.5	4.3	16.0	−1.2	11.8	7.3	20.0	6.5	20.3	5.4	15.8	4.0	16.3
C50_mean (dB)	0.0	12.5	6.0	19.3	5.1	17.6	5.7	16.7	3.0	14.3	−0.7	11.7	6.9	19.8	4.7	18.3	5.6	16.3	2.4	14.9
DRR_L (dB)	−2.7	5.8	8.5	16.5	−3.6	2.6	4.7	11.2	−7.9	−3.2	−1.8	5.5	8.8	15.8	−3.0	3.8	5.9	9.3	−9.3	−2.3
DRR_R (dB)	−2.4	6.6	8.0	15.0	8.9	14.7	3.8	9.8	0.7	2.8	−2.2	5.9	8.6	15.3	7.7	14.5	5.7	7.1	0.8	5.5
DRR_mean (dB)	−2.6	6.2	8.3	15.8	6.1	11.9	4.3	10.5	−1.7	0.8	−2.0	5.7	8.7	15.5	5.0	11.9	5.8	8.3	−1.8	3.1

https://doi.org/10.3390/buildings16193788

As can be inferred from the values reported in Table 1, and as reported above, the binaural parameters ITD, ILD, C50, and DRR obtained for the two children's heads show consistent results across all source positions and acoustic treatments, with differences within the JND, except for IACC and ILD in one case. The similarity between the two children's head results could represent a promising starting point for evaluating these parameters through measurements performed not with a child HATS, but with in-ear microphones worn by a child. Under anechoic conditions, ITD, ILD, and IACC can be obtained from the head-related transfer function (HRTF), which describes how a given sound-wave input is filtered by the diffraction and reflection properties of the head, pinna, and torso before the sound reaches the eardrum. HRTFs are usually measured using microphones placed at the ears and a moving loudspeaker around the subject. However, the need to remain still makes this method impractical for children. Therefore, Fels and Vörlander (2009) [46] investigated HRTFs and the computed binaural parameters in 95 subjects aged from 6 months to 17.5 years, using the Boundary Element Method (BEM), starting from CAD models of the heads. They found that the ear–shoulder distance, head breadth, and back vertex have the greatest influence, while head height and ear height, despite changing considerably during growth, have only a minor effect. Among pinna dimensions, the depth and breadth of the cavum concha were the most influential [46].

Individual HRTF for sound incidences of +90° and −90° azimuth was measured under anechoic conditions in 23 children aged 7–12 years, revealing variability in ITD (645 ± 44 µs in girls and 666 ± 35 µs, in boys), with inter-individual differences largely outweighing the systematic effect of age [47]. The mean and standard deviation of the SII (Speech Intelligibility Index)-weighted ILDs were 4.5 dB ± 0.7 dB, and no gender difference was found.

These findings highlight substantial variability in children's binaural acoustic characteristics and the challenges to obtain large datasets from this population. However, when tested in-field, the extent to which this inter-individual variability translates into differences in IACC remains to be directly established in primary-school children. The limited differences observed in ITD and ILD, compared with the more pronounced differences in IACC, suggest that the two children's head configurations mainly affect the similarity of the binaural responses rather than their interaural time or level differences. This may reflect differences in the temporal structure of the binaural responses, particularly due to reflections and acoustic conditions.

Table 2 shows the difference between target and noise binaural RIRs, ΔC_{50}, averaged from 250 Hz to 2 kHz. The following findings emerge:

- The results obtained from the two children's heads are consistent.
- The difference in C_{50} between the target and noise sources (ΔC_{50}) is smaller under the more favourable condition of a 120° spatial separation than with co-located target and noise sources, by approximately 2–3 dB.
- ΔC_{50} is about 2 dB smaller with the multi-talker noise source than with the single-talker source.
- When the target and noise sources are co-located in the untreated classroom, ΔC_{50} does not differ between single and multi-talker sources to a great extent. In the treated classroom, instead, the multi-talker noise source results in lower ΔC_{50} values.
- With the noise source positioned at 120° relative to the target source, ΔC_{50} does not differ between the single-talker and multi-talker noise sources in either the untreated or the treated classroom.

https://doi.org/10.3390/buildings16193788

Table 2. Difference between target and noise binaural RIRs C_{50} averaged from 250 Hz to 2 kHz.

			Untreated Classroom (C1)			Treated Classroom (C2)		
			C_{50}_L (dB)	C_{50}_R (dB)	C_{50}_Mean (dB)	C_{50}_L (dB)	C_{50}_R (dB)	C_{50}_Mean (dB)
L2	Single-talker	T@3.6 m, 0° N@1 m, 0°	−6.9	−8.5	−7.6	−6.1	−7.5	−6.8
		T@3.6 m, 0° N@1 m, 120°	−1.6	−7.7	−5.4	−0.5	−7.3	−5.1
	Multi-talker	T@3.6 m, 0° N@1 m, 0°	−6.2	−6.7	−6.3	−4.3	−4.1	−4.2
		T@3.6 m, 0° N@1 m, 120°	−0.1	−5.2	−3.1	1.2	−3.5	−1.8
L1	Single-talker	T@3.6 m, 0° N@1 m, 0°	−5.2	−6.6	−6.0	−7.9	−8.3	−8.1
		T@3.6 m, 0° N@1 m, 120°	−2.2	−6.8	−5.1	−2.7	−8.5	−6.6
	Multi-talker	T@3.6 m, 0° N@1 m, 0°	−5.7	−5.6	−5.7	−5.1	−4.0	−4.7
		T@3.6 m, 0° N@1 m, 120°	−1.0	−4.3	−3.0	−1.1	−4.6	−3.2

3.3. Verbal Working-Memory Task Results

The detailed results obtained with the verbal working memory task are reported in Vettori et al. (2022) [41]. The study investigated whether classroom acoustic quality, specifically reverberation time and background noise, could influence second graders' Verbal Working-Memory (VWM) performance.

The present study integrates the cited work and delves into the aspects inherent to the acoustic response of the two classrooms by analyzing the acoustic response with both monaural and binaural parameters.

VWM is a foundational process for learning, as it enables the simultaneous storage and manipulation of new information. This is critical in numerous school activities, such as when students listen to a teacher while attempting to comprehend and encode new concepts or take notes. It is particularly important to study the efficiency of VWM in early primary-school children, a stage when cognitive processes are still developing. A sample of twenty-five second-graders was administered a validated verbal working-memory task (Listening Span Test by Palladino, 2005) in the two different acoustic conditions given by classroom C1 and C2 with a time frame of 1 month and controlled background noise achieved through soundproofing.

Coding the children's responses yielded significant results for the Span index, for the number of omitted words, and for the words recalled.

No statistically significant differences between the two conditions were observed in the other variables (words invention, inversions, intrusions) taken into account.

As shown in Figure 15, the span index detected, about the maximum sequence of words remembered, was higher in the treated classroom C2 (M = 2.40; $M\ dn$ = 2), compared to the untreated classroom C1 (M = 1.88; $M\ dn$ = 2). A Wilcoxon signed-rank test indicated that this difference was statistically significant T = 122, Z = −2.31, p = 0.021, r = −0.33. The index of words recalled was lower in the classroom C1 (M = 6.60, SD = 3.04) compared to that in classroom C2 (M = 8.53, SD = 3.07). This difference, −1.92, 95% CI [−3.64, −0.20], was statistically significant, $t(24) = -2.30$, $p = 0.030$, Cohen's $d = -0.46$. Finally, the number of words omitted was higher in the classroom C1 (M = 6.68, SD = 3.00) compared to the classroom C2 (M = 4.84, SD = 2.85). This difference, 1.84, 95% CI [0.09, —3.59], was statistically significant, $t(24) = 2.17$, $p = 0.040$, *Cohen's d* = 0.43. A two-way mixed ANOVA revealed a significant main effect of acoustic condition on both Span $F(1,23) = 4.31$, $p = 0.049$, $\eta^2 = 0.16$ and Recalled words $F(1,23) = 5.19$, $p = 0.032$, $\eta^2 = 0.18$, with no significant main or interaction effects observed for testing order ($p > 0.05$).

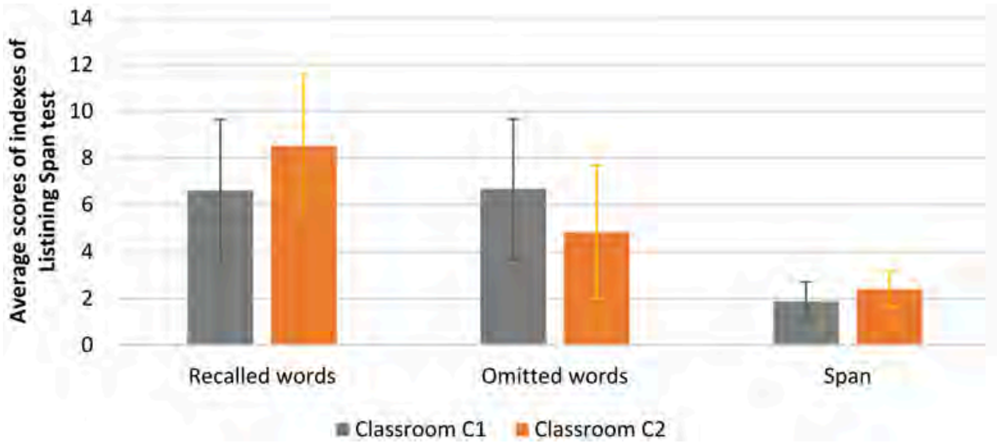

Figure 15. Average scores of recalled words, omitted words and span indexes as a function of the acoustic condition. Error bars indicate standard errors of the means. Note: $p < 0.05$.

The results obtained in the tests confirmed what was expected; the score that the children obtained in the very reverberant classroom C1 was lower than in the classroom with the sound-absorbing treatment C2; "classroom bubble", in the presence of longer reverberation times, affects in a negative way the children's efficiency in successfully pursuing tasks requiring verbal working memory, particularly when a dual-task paradigm is involved. High school-classroom reverberation combined with typical classroom noise can mask or distort speech signals, thereby increasing cognitive load and interfering with semantic processing or task focus. Thus, children who have to perform tests within overly reverberant environments spend some of their cognitive resources understanding and processing phonemes or words masked by reflections, instead of concentrating on keeping the information in mind to answer the task correctly [3].

3.4. Limitations of the Study and Future Developments

The main limitation of this study is the small sample of children who participated in the test, which included 25 pupils divided into two groups. The binaural measurements could be performed only in two representative positions in the two rooms with two different children.

Another limitation of the study is the fact that the sample consists of pupils with a relatively high socio-economic level and does not include pupils with learning problems.

In the future development of the study, it is planned to involve a larger number of pupils from different socio-economic backgrounds.

Furthermore, future developments may include real-time cognitive load analyses, such as measuring the pupil diameter of the children tested.

These developments in the study could further strengthen the results of the analyses conducted.

4. Conclusions

This article presents the results of monoaural and binaural acoustic measurements and comprehension tests conducted in two classrooms representative of the Italian primary-school context. Of the two classrooms, one was acoustically treated in accordance with the Italian standard UNI 11532-2 (classroom C2), while the other was not acoustically treated and represents the average Italian school (classroom C1).

Monoaural measurements were carried out in classrooms C1 and C2 both in occupied and in unoccupied conditions. Results show that average reverberation time in the classroom C1 reduces from 1.3 to 1.0 s from unoccupied to occupied condition, while average clarity C_{50} increases from −1.6 to +4.9 dB. The effect of the students' presence is less pronounced but still noticeable in the acoustically optimized classroom (C2).

In the treated classroom C2, STI reduces when moving from 1 m in front of the source to the center of the classroom and when introducing the chatting noise. This effect is less pronounced in the untreated classroom C1, where the STI remains low regardless of the distance from the speaker.

For binaural measurements, no significant differences are observed for the two children's heads, apart from IACC. Clarity C_{50} is approximately 2–4 dB higher for the source with a human talker directivity than for the omnidirectional source, representing a multi-talker source. In the treated classroom C2, C_{50} exceeds the values measured in the untreated classroom C1 by more than 10 dB in the mid- and high-frequency range. As expected, the right ear exhibits higher clarity than the left ear up to 6–8 dB at the highest frequencies, when the noise source is positioned at an azimuth of 120°.

As expected, the acoustic treatment leads to a clear increase in C_{50} and DRR, with differences exceeding their respective JNDs by more than ten and three times, respectively. Conversely, ITD, ILD, and IACC do not exhibit clear trends, as the observed differences are generally within or close to their corresponding JNDs. Differences between the two children's heads are mainly observed for IACC.

The results also show that ΔC_{50} between the target and noise sources is smaller under the more favorable condition of a 120° spatial azimuth separation than with colocation. Furthermore, the single-talker noise source (Talkbox) produces larger ΔC_{50} than the omnidirectional multi-talker noise source in the treated classroom, indicating a greater detrimental effect. In the untreated room, this effect is not evident.

In conclusion, the benefit appears to originate primarily from improvements in speech clarity and direct-sound dominance rather than from enhanced binaural localization or spatial-segregation cues.

As reported, the aim of the work was to investigate how the binaural measurements contribute to the interpretation of the cognitive results by identifying which auditory cues are affected by the acoustic treatment. It was shown that while ITD, ILD, and IACC remained largely unvaried between classrooms, substantial increases were observed for C_{50} and DRR. These findings suggest that the improved verbal working-memory performance found in the treated classroom is unlikely to be related to enhanced localization cues (ITDs and ILDs). Instead, it may be associated with a more favorable balance between direct and reverberant sound and with increased speech clarity at the listener's ears. It can be suggested that under these conditions, children may require less effort to decode the spoken information, leaving more cognitive resources available for storing and processing in verbal working-memory tasks. Future studies should build upon these results by integrating binaural acoustic metrics alongside multidimensional cognitive profiling of children. Specifically, incorporating non-verbal IQ assessments would allow researchers to explicitly control for baseline individual cognitive variations.

https://doi.org/10.3390/buildings16193788

These findings contribute new knowledge to the field of classroom acoustics. However, further experiments using auralized binaural signals with children are needed to determine whether these parameters are related to speech intelligibility and cognitive performance. For now, the results obtained in the comprehension tests confirmed what was expected; the score of the children in the untreated classroom C1 was lower than that of the children in the treated classroom C2, demonstrating how the bubbling noise in the presence of longer reverberation times negatively affects the children's efficiency in those tests that required the use of verbal working memory.

This is because children in excessively reverberant environments spend part of their cognitive resources on understanding and processing any phonemes or words masked by reflections, instead of concentrating on keeping in mind the information needed to answer the task correctly.

Author Contributions: Conceptualization, S.S., A.A., L.B., L.S. and G.V.; methodology, L.S.; writing—original draft, S.S., A.A., G.P., L.S. and G.V.; writing—review and editing, S.S., A.A., V.A., L.B., F.B., L.S. and G.V.; visualization, L.S.; supervision, S.S. and A.A.; funding acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was founded by the Foundation of the Savings Bank of Florence (Practice number 2017.0809), by the Municipality of Florence (determination n. 02427 of 09.06.2017 of the Municipality of Florence), and by a local manufacturer (Manifattura Maiano spa, contract of 22.06.2018).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of the University of Florence (protocol code 55 and approved on 16 July 2019) for studies involving humans.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in this study are included in the article.

Conflicts of Interest: Authors have received Funding from Manifattura Maiano spa. All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Astolfi, A.; Bottalico, P.; Barbato, G. Subjective and objective speech intelligibility investigations in primary school classrooms. *J. Acoust. Soc. Am.* **2012**, *131*, 247–257. [CrossRef] [PubMed]
2. Buss, E.; Leibold, L.J.; Porter, H.L.; Grose, J.H. Speech recognition in one- and two-talker maskers in school-age children and adults: Development of perceptual masking and glimpsing. *J. Acoust. Soc. Am.* **2017**, *141*, 2650–2660. [CrossRef] [PubMed]
3. Leibold, L.J.; Buss, E. Masked Speech Recognition in School-Age Children. *Front. Psychol.* **2019**, *10*, 1981. [CrossRef] [PubMed]
4. Blauert, J. *The Technology of Binaural Listening*; Springer: Berlin/Heidelberg, Germany, 2012.
5. Bronkhorst, A.W. The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Atten. Percept. Psychophys.* **2015**, *77*, 1465–1487. [CrossRef] [PubMed] [PubMed Central]
6. Peng, E.Z.; Litovsky, R.Y. The role of interaural differences, head shadow, and binaural redundancy in binaural intelligibility benefits among school-aged children. *Trends Hear.* **2021**, *25*, 23312165211045313. [CrossRef] [PubMed]
7. Dieudonné, B.; Francart, T. Redundant Information Is Sometimes More Beneficial Than Spatial Information to Understand Speech in Noise. *Ear Hear.* **2019**, *40*, 545–554. [CrossRef] [PubMed]
8. Ching, T.Y.C.; van Wanrooy, E.; Dillon, H.; Carter, L. Spatial release from masking in normal-hearing children and children who use hearing aids. *J. Acoust. Soc. Am.* **2011**, *129*, 368–375. [CrossRef] [PubMed]
9. Misurelli, S.M.; Litovsky, R.Y. Spatial release from masking in children with normal hearing and with bilateral cochlear implants: Effect of interferer asymmetry. *J. Acoust. Soc. Am.* **2012**, *132*, 380–391. [CrossRef] [PubMed] [PubMed Central]
10. Corbin, N.E.; Buss, E.; Leibold, L.J. Spatial Release from Masking in Children: Effects of Simulated Unilateral Hearing Loss. *Ear Hear.* **2017**, *38*, 223–235. [CrossRef] [PubMed]
11. Griffin, A.M.; Poissant, S.F.; Freyman, R.L. Speech-in-Noise and Quality-of-Life Measures in School-Aged Children with Normal Hearing and with Unilateral Hearing Loss. *Ear Hear.* **2019**, *40*, 887–904. [CrossRef] [PubMed]

12. Misurelli, S.M.; Litovsky, R.Y. Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity. *J. Acoust. Soc. Am.* **2015**, *138*, 319–331. [CrossRef] [PubMed]
13. Peng, Z.E.; Pausch, F.; Fels, J. Spatial release from masking in reverberation for school-age children. *J. Acoust. Soc. Am.* **2021**, *150*, 3263–3274. [CrossRef] [PubMed]
14. Brown, A.D.; Stecker, G.C.; Tollin, D.J. The Precedence Effect in Sound Localization. *J. Assoc. Res. Otolaryngol.* **2015**, *16*, 1–28. [CrossRef] [PubMed]
15. Wallach, H.; Newman, E.B.; Rosenzweig, R. The precedence effect in sound localization. *Am. J. Psychol.* **1949**, *62*, 315–336. [CrossRef]
16. Litovsky, R.Y.; Colburn, H.S.; Yost, W.A.; Guzman, S.J. The precedence effect. *J. Acoust. Soc. Am.* **1999**, *106*, 1633–1654. [CrossRef] [PubMed]
17. Litovsky, R.Y.; Godar, S.P. Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks. *J. Acoust. Soc. Am.* **2010**, *128*, 1979–1991. [CrossRef] [PubMed]
18. Clifton, R.K.; Morrongiello, B.A.; Dowd, J.M. A developmental look at an auditory illusion: The precedence effect. *Dev. Psychol.* **1984**, *17*, 519–536. [CrossRef] [PubMed]
19. Kidd, G., Jr.; Mason, C.R.; Brughera, A.; Hartmann, W.M. The role of reverberation in release from masking due to spatial separation of sources for speech identification. *Acta Acust. United Acust.* **2005**, *91*, 526–536.
20. Leist, L.; Breuer, C.; Yadav, M.; Fremerey, S.; Fels, J.; Raake, A.; Lachmann, T.; Schlittmeier, S.J.; Klatte, M. Differential Effects of Task-Irrelevant Monaural and Binaural Classroom Scenarios on Children's and Adults' Speech Perception, Listening Comprehension, and Visual–Verbal Short-Term Memory. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15998. [CrossRef] [PubMed]
21. Loh, K.; Yadav, M.; Persson Waye, K.; Klatte, M.; Fels, J. Toward Child-Appropriate Acoustic Measurement Methods in Primary Schools and Daycare Centers. *Front. Built Environ.* **2022**, *8*, 688847. [CrossRef]
22. Klatte, M.; Bergström, K.; Lachmann, T. Does noise affect learning? A short review on noise effects on cognitive performance in children. *Front. Psychol.* **2013**, *4*, 578. [CrossRef] [PubMed]
23. Gathercole, S.E.; Pickering, S.J.; Ambridge, B.; Wearing, H. The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* **2004**, *40*, 177–190. [CrossRef] [PubMed]
24. Gheller, F.; Spicciarelli, G.; Scimemi, P.; Arfé, B. The effects of noise on children's cognitive performance: A systematic review. *Environ. Behav.* **2023**, *55*, 698–734. [CrossRef]
25. Mealings, K. The Effect of Classroom Acoustic Treatment on Listening, Learning, and Well-being: A Scoping Review. *Acoust. Aust.* **2023**, *51*, 279–291. [CrossRef]
26. *UNI 11367:2023*; Acustica in Edilizia—Classificazione Acustica Delle Unità Immobiliari—Procedura di Valutazione e Verifica in Opera [Building Acoustics—Acoustic Classification of Building Units—Procedure for On-Site Assessment and Verification]. Ente Italiano di Normazione: Milano, Italy, 2023.
27. *UNI 11532-1:2018*; Caratteristiche Acustiche Interne di Ambienti Confinati—Metodi di Progettazione e Tecniche di Valutazione—Parte 1: Requisiti Generali [Internal Acoustic Characteristics of Confined Spaces—Design Methods and Evaluation Techniques—Part 1: General Requirements]. Ente Italiano di Normazione: Milano, Italy, 2018.
28. *UNI 11532-2*; Caratteristiche Acustiche Interne Di Ambienti Confinati—Metodi Di Progettazione E Tecniche Di Valutazione—Parte 2: Settore Scolastico [Internal Acoustic Characteristics of Confined Spaces—Design Methods and Evaluation Techniques—Part 2: School Sector]. Ente Italiano di Normazione: Milano, Italy, 2015.
29. *ISO 3382-1:2009*; Acoustics—Measurement of Room Acoustic Parameters—Part 1: Performance Spaces. International Organization for Standardization: Geneva, Switzerland, 2009.
30. Astolfi, A.; Minelli, G.; Puglisi, G.E. A basic protocol for the acoustic characterization of small and medium-sized classrooms. *J. Acoust. Soc. Am.* **2022**, *152*, 1646–1659. [CrossRef] [PubMed]
31. Guastamacchia, A.; Rosso, R.G.; Puglisi, G.E.; Riente, F.; Shtrepi, L.; Astolfi, A. Real and Virtual Lecture Rooms: Validation of a Virtual Reality System for the Perceptual Assessment of Room Acoustical Quality. *Acoustics* **2024**, *6*, 933–965. [CrossRef]
32. Katz, B.F.G.; Noisternig, M. A comparative study of interaural time delay estimation methods. *J. Acoust. Soc. Am.* **2014**, *135*, 3530–3540. [CrossRef] [PubMed]
33. Hládek, L.; Ewert, S.D.; Seeber, B.U. Communication conditions in virtual acoustic scenes in an underground station. In *Proceedings of the 2021 Immersive and 3D Audio: From Architecture to Automotive (I3DA), Bologna, Italy, 8–10 September 2021*; IEEE: New York, NY, USA, 2021; pp. 1–8.
34. Hummersone, C. Impulse Response Acoustic Information Calculator, GitHub, 2023. Available online: https://github.com/IoSR-Surrey/MatlabToolbox (accessed on 6 August 2026).
35. Zahorik, P. Direct-to-reverberant energy ratio sensitivity. *J. Acoust. Soc. Am.* **2002**, *112*, 2110–2117. [CrossRef] [PubMed]
36. Lavandier, M.; Jelfs, S.; Culling, J.F.; Watkins, A.J.; Raimond, A.P.; Makin, S.J. Binaural prediction of speech intelligibility in reverberant rooms with multiple noise sources. *J. Acoust. Soc. Am.* **2012**, *131*, 218–231. [CrossRef] [PubMed]

https://doi.org/10.3390/buildings16193788

37. Spencer, N.J.; Hawley, M.L.; Colburn, H.S. Relating interaural difference sensitivities for several parameters measured in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* **2016**, *140*, 1783–1799. [CrossRef] [PubMed]
38. Larsen, E.; Iyer, N.; Lansing, C.R.; Feng, A.S. On the minimum audible difference in direct-to-reverberant energy ratio. *J. Acoust. Soc. Am.* **2008**, *124*, 450–461. [CrossRef] [PubMed]
39. Bradley, J.S.; Reich, R.D.; Norcross, S.G. A just noticeable difference in C_{50} for speech. *Appl. Acoust.* **1999**, *58*, 99–108. [CrossRef]
40. Palladino, P. A measure of verbal working memory for children from second to fifth grade: Norms and validation. *Psicol. Clin. Dello Svilupp.* **2005**, *9*, 129–149.
41. Vettori, G.; Di Leonardo, L.; Secchi, S.; Astolfi, A.; Bigozzi, L. Primary school children's verbal working memory performances in classrooms with different acoustic conditions. *Cogn. Dev.* **2022**, *64*, 101256. [CrossRef]
42. *ISO 9921:2003*; Ergonomics—Assessment of Speech Communication. International Organization for Standardization: Geneva, Switzerland, 2003.
43. Shield, B.; Dockrell, J.E. External and internal noise surveys of London primary schools. *J. Acoust. Soc. Am.* **2004**, *115*, 730–738. [CrossRef] [PubMed]
44. Carullo, A.; Vallan, A.; Astolfi, A.; Pavese, L.; Puglisi, G.E. Validation of calibration procedures and uncertainty estimation of contact-microphone based vocal analyzers. *Measurement* **2015**, *74*, 130–142. [CrossRef]
45. *IEC 60268-16:2020*; Sound System Equipment—Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index. IEC: Geneve, Switzerland, 2020.
46. Fels, J.; Vorländer, M. Anthropometric Parameters Influencing Head-Related Transfer Functions. *Acta Acust. United Acust.* **2009**, *95*, 331–342. [CrossRef]
47. Zenke, K.; Rosen, S. Spatial release of masking in children and adults in non-individualized virtual environments. *J. Acoust. Soc. Am.* **2022**, *152*, 3384–3395. [CrossRef] [PubMed]

https://doi.org/10.3390/buildings16193788