

## Article

# An Acoustic Way to Support Japanese Children's Effective English Learning in School Classrooms

Naoko Evans <sup>1,2</sup>, Miki Kaneko <sup>1</sup>, Ivan Seleznov <sup>1</sup>, Taiki Shigematsu <sup>1</sup> and Ken Kiyono <sup>1,\*</sup>

<sup>1</sup> School of Engineering Science, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka 560-8531, Japan; evans@bpe.es.osaka-u.ac.jp (N.E.); kaneko@bpe.es.osaka-u.ac.jp (M.K.); ivan.seleznov@bpe.es.osaka-u.ac.jp (I.S.); shigematsu@bpe.es.osaka-u.ac.jp (T.S.)

<sup>2</sup> TOA Corporation, 7-2-1 Minatojima-Nakamachi, Chuo-ku, Kobe 650-0046, Japan

\* Correspondence: kiyono@bpe.es.osaka-u.ac.jp

**Abstract:** In this paper, the importance of implementing good acoustic conditions in classrooms using sound amplification systems is investigated to support more effective English education for elementary school children. To date, the failure of educating English as a second language at Japanese schools has been demonstrated by poor English conversation ability of those who completed a compulsory six-year English language course at Japanese junior-high and high schools (age 12–18). To amend the situation, teaching English became compulsory at grade three (age 8–9) and above at most Japanese elementary schools in the 2020 academic year. We conducted acoustic measurements of two types of sound amplification systems, a pair of PC loudspeakers and another with a loudspeaker array, in a typical classroom at an elementary school in Japan. We also analysed English listening test results of 216 Japanese native children (age 11–12) who were learning English in their usual classes in Japan, to compare the effects of those two systems. Results of logistic regression analysis adjusted by the discrimination difficulty of word pairs demonstrated the statistically significant association between correct answer rate of the English tests and classroom acoustic factors. Although, on average, upgrading the sound amplification system had positive effects on the correct answer rate, it also had a negative impact when the word pairs had English phoneme contrasts that do not appear in Japanese phoneme structure. Combined with the acoustic measurements' results, it was also revealed that heterogeneous sound fields that depend on seat positions could be compensated using sound amplification systems with loudspeaker arrays. Our findings suggest that improvement of both acoustic quality and teaching methods is required for children to acquire English communication skills effectively in their classroom.

**Keywords:** classroom acoustics; second and foreign language education; hearing in children; impulse response



**Citation:** Evans, N.; Kaneko, M.; Seleznov, I.; Shigematsu, T.; Kiyono, K. An Acoustic Way to Support Japanese Children's Effective English Learning in School Classrooms. *Appl. Sci.* **2021**, *11*, 6062. <https://doi.org/10.3390/app11136062>

Academic Editor: Francesc Aliás

Received: 31 March 2021

Accepted: 5 May 2021

Published: 29 June 2021

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



**Copyright:** © 2021 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

English language education at Japanese public schools has been criticized. The average score of Test of English as a Foreign Language in Japan has been around the lowest among Organisation for Economic Co-operation and Development countries [1]. Many Japanese people cannot make a simple daily conversation in English after completing a compulsory six-year English language course in Japan [2]. In April 2020, the age to start teaching English at most Japanese elementary schools was officially lowered to resolve the matter. Grade three and four (age 8–10) have some English lessons to get familiar with the language, whereas grade five and six (age 10–12) learn English as a subject, thus their skills get evaluated through tests [3].

Previous research has shown evidence of poor acoustics found in school classrooms and their negative influence on children in various aspects [4–10]. Children are often too exhausted from trying to comprehend speech contents; subsequently they have little

energy left to perform any cognitive tasks using the obtained information [11,12]. Not only language related skills, such as reading, writing, speaking and listening, but also others, including numeracy, working memory and motivation of children were deteriorated as a consequence [13–21].

The further the distance between a speaker and a listener, the less speech is heard [22]. In a closed space like a classroom, sound reflection and absorption by the ceiling, floor, walls, furniture as well as people present, all integrate together with the sound directly emitted by a speaker to shape diverse characteristics of the sound field [23,24]. Consequently, the intelligibility of a teacher's speech signals fluctuates depending on where children are, even in the same classroom [25]. However, such effects on human perception in real rooms of sizes often seen in classrooms are not well understood [26], compared to those in large indoor spaces for performance (e.g., concert halls and auditoria).

The length of time in which children are required to listen to their teachers, peers, or audio materials accounts for the majority of learning time children spend in their classrooms [9], taking up to 45–75% in total [7,27]. With the teaching style shift towards what is sometimes called "active learning" [28], children are demanded to listen more and more in classrooms that are noisier than ever. The introduction of more group work also leads to children's learning abilities being more linked with their hearing skills than ever [29].

It is also important to note that hearing in children is not as sophisticated as that in adults, especially in noisy environments [14,30]. Thus, adults are often unable to perceive the hearing difficulties experienced by young children even in the same classroom [7,31]. The way one speaks usually varies under noisy conditions. This, known as the "Lombard effect", is recognised to reduce intelligibility but also cause strains on speakers. Such deteriorated speech signals are inadequate for children to learn anything, especially languages [32,33]. Moreover, for those aged between 5 and 15 years in particular, it is common to contract middle ear infectious disease Otitis Media, which causes temporary hearing loss of up to 40 dB hearing level and often lasts more than one months [8].

Many teachers have also expressed their discontent with poor acoustics at schools [6,34,35]. Teachers are more likely to report voice disorders than those in any other professions [34]. Absent teachers due to voice related health problems impose non-negligible costs. Since much information is delivered in speech form in classrooms, good acoustics there would benefit a much larger population than just children with already-diagnosed hearing impairments.

Although the above mentioned problems have been identified, there is no clear guidelines regarding the quality of acoustic signals to be presented to children, to protect and support both children and teachers in Japanese school classrooms, other than the school building design guidelines [36] written from an architectural perspective. Northern and Downs stated [31] that solving such problems with acoustics in classrooms is not too hard. Rather than the absence of funding nor knowledge of its solution, missing awareness of both the problem and the potentially achievable answers is the main issue. In addition, the importance of evidence-based education research has been promoted in recent years [37,38]. Education practices should be based on the best available scientific evidence, rather than tradition, personal judgement, or other influences. However, evidence in the real-life situations collected at real schools are still not enough.

The aim of this study is to demonstrate the importance of good acoustic conditions in the second and foreign language education. We present our findings using data collected in a real-life environment at a typical elementary school in Japan, with help from Japanese native children who were learning English as their second language. Two kinds of sound amplification systems are utilised; one is what had been used at the school and another is what we prepared as an example alternative for a comparative study. By statistically analysing English listening test results, combined with some acoustic measurements, effect of more adequately treated classroom acoustics on children's learning is illustrated.

## 2. Methods

### 2.1. Participants

In total, 216 children from six classrooms at a typical public mixed-gender elementary school in Japan, aged between 11 and 12 years (grade six), participated over two consecutive weeks (212 in the week one, and 201 in the week two) in December 2019. The majority of Japanese elementary schools are public schools. The school was selected after consulting the local education authorities in a city while searching for a school that can cooperate in our study. The number of children who attended each lesson is presented in Table 1.

**Table 1.** Number of children attended each lesson.

	Week One		Week Two	
	Day One (Monday)	Day Two (Tuesday)	Day One (Monday)	Day Two (Tuesday)
Lesson 1	N/A	36	N/A	34
Lesson 2	35	36	34	35
Lesson 3	36	33	35	27
Lesson 4	N/A	36	N/A	36

(N/A: not applicable).

No prior information was available regarding children's English language skills, English word familiarity, nor hearing level, apart from the following information:

- At the school, the grade 6 children had been taking three English lessons every fortnight, taught mainly by a Japanese native teacher in the English learning room, dedicated for English lessons of the grade five and six, for roughly 17 months before the current study began.
- They also had an English native teacher assisting in some lessons, typically once a week.
- Each lesson was 45-min-long.

We also had no knowledge regarding their intellectual abilities nor their socio-economic characteristics.

### 2.2. English Listening Test

English listening tests using recorded speech audio signals emitted by two sorts of sound amplification systems (Figure 1) were given to the participants. Participants were asked to indicate whether they thought the speech sounds of each word pair were the same or not by circling their choice between "the same" and "different" (both written in Japanese) for each question on provided sheets. The task consists of 42 English word pairs, such as "walk-walk" and "ten-pen". The questions comprise clear speech recordings of either a minimal pair or the same word spoken twice with at least one second (s) interval, voiced by several English native speakers in a randomised order. The speech audio materials were prepared by editing the sound source provided in the accompanying CDs of the two books [39,40] with permission of the licensor through PLSclear (Publishers' Licensing Services, London, UK), using WaveLab Pro 9.5 (Steinberg, Hamburg, Germany) [41]. Restrictions on letter and word selections were determined according to the availability of the sound source. The maximum amplitude for each word was adjusted to be the same. For the speech audio materials, the balance to encompass varieties in terms of vowels vs. consonants, minimal pair vs. the same word, and male voice vs. female voices was all decided manually during the preparation period in an attempt to make it fair wherever possible within all restrictions applied.

Each participant was given four tests in total as illustrated in Figure 2. Each test had 21 questions with approximately four-min-long audio materials. Two tests were integrated into each lesson, one at the beginning and another at the end of each lesson. Two lessons

were scheduled over two consecutive weeks, with one week interval in between. The first test was given usually after an introduction song and a small talk session. After another small talk, often followed by some group work, the second test was completed with some closing remarks added by their Japanese native teacher. All lessons were held during their regular teaching hours in their English room at the school. During the two week period, lessons were scheduled according to a weekly timetable set by the school. Thus, a class that had an English lesson on Monday starting at 9:30 in the first week, had another one on Monday at 9:30 in the following week, for example. Tests were designed so that the same word appeared twice, once a week, and it was played by both sound amplification systems one after another over the two weeks. The orders of questions were randomised using a pseudo random number generator written in Python [42]. The sound amplification system was alternated after every four questions.

To avoid causing too much stress to the children, it was determined in advance that roughly 10 min is the maximum length of time allowed for the test procedures in each lesson in our study. To protect their privacy, we asked each participant to place a sticker on each answer sheet before starting a test. These stickers came in five colours, signifying their seat groups (Figure 3, left), determined in advance to be utilised in our analyses.



Figure 1. Two sound amplification systems: System A (left), and System B (right).

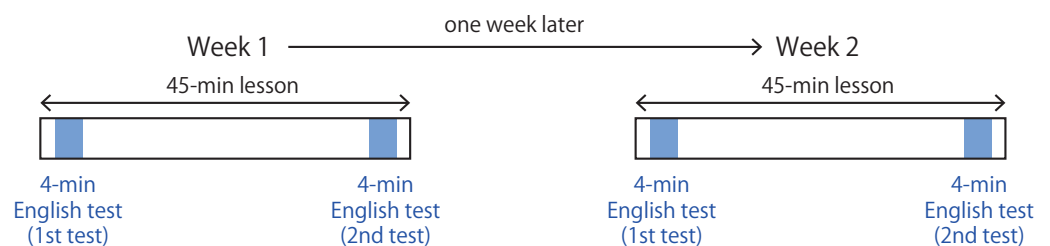
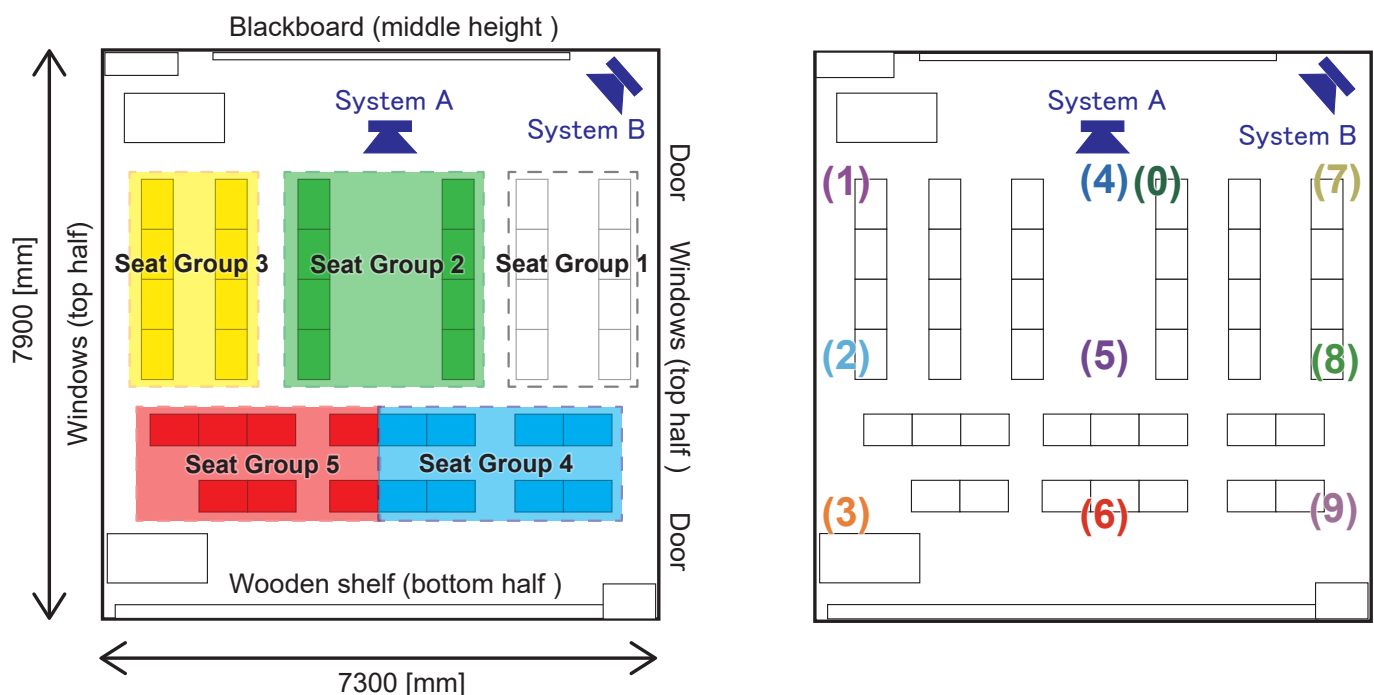


Figure 2. Test Protocol.



**Figure 3.** Layouts of the English room with positions of two sound amplification systems: (left) with dimensions and colour coded seat groups, and (right) with 10 measurement points, shown as numbers in brackets.

### 2.3. Sound Amplification Systems for a Classroom

Two different sound amplification systems, “System A” and “System B”, were used for a comparison. System A (Figure 1 left), was a pair of typical PC loudspeakers (S-0264C, Logitech, Lausanne, Switzerland) which had been used during regular English lessons at the school for a long time before the current study began. We kept the System A where it had been placed for their previous English lessons, on the bottom shelf of a metal rack with a projector on the top (Figure 1 left), placed in the centre front of the room (shown as “System A” in Figure 3 right). It enabled us to compare between “a regular system” found at a school and another brought in with a little more care. The System A generates point-source-like signals, thus direct signals would spread and decay by 6 dB per doubling the distance in a free field. Sound emitted by such simple sound amplification systems are usually deteriorated by their representational features.

System B (Figure 1 right) consisted of a signal processor, an amplifier and a loudspeaker array (DP-SP3, DA-250F, and SR-H3L, respectively; TOA, Hyogo, Japan). The System B was placed at the front corner of the room (shown as “System B” in Figure 3 right) so that the emitted signals could cover the whole room with little interference caused by sound reflected by the walls whilst maintaining easier installation, smaller additional costs, and less interruptions to learning activities, compared with many other alternatives. The System B is designed to realise sound fields that are more evenly spread in a room, and has a relatively flat response. The advantage of loudspeaker arrays is that emitted acoustic signals become like a line source as the distance to the listening point widens for a while; therefore, they would only decay 3 dB per doubling the distance in a free field.

Levels of both systems were adjusted so that the A-weighted equivalent continuous sound pressure level over 10 s, denoted by  $L_{Aeq,10s}$  [43], was 75 dBA at position (0) (Figure 3 right), using pink noise. This level was determined based on the exposure limit for young children, suggested by the World Health Organization, as 75 dBA 8-h  $L_{EX}$  in order to protect children from risks of noise-induced hearing loss [44].

#### 2.4. Acoustic Measurements

Multiple impulse response measurements in the unoccupied furnished English room were carried out on an additional day, to gain a better understanding of the difference in the acoustic characteristics realised by the two sound amplification systems. Both systems were placed in the same ways as where they were during English tests (Figure 3, right). The measurements were carried out while no lessons in and around the school building were taking place.

The Exponential Sine Sweeps (ESS) method [45] was used. Signals were recorded by using a microphone of a calibrated class two sound level meter [46] (NL-42, RION, Tokyo, Japan) set at a height of 1 m with a portable recorder (PCM-D100, SONY, Tokyo, Japan), and sampled at 44.1 kHz, 16-bit. ESS signals (30 Hz to 22 kHz, also sampled at 44.1 kHz) were played and recorded several times at each of 10 positions ((0) to (9) in Figure 3, right). Albeit imperfect, the method with averaging and deconvolution in the frequency domain was adopted to obtain each impulse response, calculated with Python, so that non-linear and time-variant behaviours of both the room acoustics and the systems were compensated to a certain extent [47]. The transfer function as well as Reverberation Time (RT) for each octave band and Speech Transmission Index (STI) were estimated by Systune pro (AFMG, Berlin, Germany) [48] in accordance with ISO 3382-2 [49] and IEC 60268-16 [50], respectively. STI is often included in guidelines or standards for classroom acoustics [27], and it was developed as an objective method to predict speech intelligibility due to transmission channels [51]. T60 is estimated from the decay time found between  $-5$  dB and  $-25$  dB (T20), whereas STI is calculated as a weighted average of seven coefficients corresponding to seven octave bands, with centre frequencies from 125 Hz to 8 kHz [52]. Note that these measurement results including STI have known limitations and they only help prediction of true speech intelligibility, which should only be quantified via listening tests employing human perception [50].

The unoccupied noise level was measured in the furnished room, and calculated using the same sound level meter as above.

#### 2.5. Statistical Analyses

The English test results were analysed after the following pre-processing steps. Firstly, data of children who were absent in any lessons were removed. Secondly, weekly numbers for the data, obtained in week two of the study from children who were absent in the week one, were replaced by those from week one in order to separately examine the impact of children getting used to the type of tests after week one.

To evaluate the discrimination difficulty of the word pairs for Japanese native children, the correct answer rate  $p_c$  of each word pair was estimated as the percentage of the correct answers. The question difficulty levels were defined as shown in Table 2 and were employed for adjustment in the following analyses.

Both univariate and multivariate logistic regression analyses [53] were used to evaluate the associations between the correct answer rate  $p_c$  and independent factors, such as question difficulty level, habituation effects (repetition of tasks), type of the sound amplification system and seat group (Table 3). In the multivariate logistic regression analysis, the Akaike Information Criteria (AIC) was used for variable selection. To estimate the bias due to overfitting (called “optimism”), we also carried out bootstrap validation of the multivariate models using 200 repetitions of bootstrapping from the original data [54]. As a performance metric, we used C-statistic, defined as the area under the receiver operating characteristic curve for the prediction of the correct answer of a task. The optimism for the C-statistic was estimated as the difference between the mean C-statistic of the bootstrap sample based models and that of the original model. The optimism-corrected C-statistic was given by subtracting the optimism from the original C-statistic.

The relationship between seat positions and acoustic characteristics realised by the two sound amplification systems (System A and System B) was explored using yet another sub data set using difficulty level 1 to 3 words. In this analysis, English test score (expressed

in percentages) of each participant who attended both lessons were used. Seat groups were categorised into two groups for simplicity: Front (Seat Group 1–3) and Back (similarly, 4–5). Normality of the English test scores in each sub set was assessed using the Shapiro–Wilk test. Because the test score distributions showed non-normality, the Wilcoxon signed-rank test was applied for paired comparison of medians, and the Mann–Whitney U test for unpaired comparison. As a multiple comparison correction method, Bonferroni-corrected  $p$ -values were employed.

$p < 0.05$  was considered to be statistically significant. All statistical analyses were performed by using R 3.6.0 [55].

**Table 2.** Question Difficulty.

Question Difficulty Level	Correct Answer Rate $p_c$
Level 4	$p_c \leq 40\%$
Level 3	$40\% < p_c \leq 60\%$
Level 2	$60\% < p_c \leq 80\%$
Level 1	$80\% < p_c$

**Table 3.** Explanatory variables assumed in the study.

	Explanation of the Explanatory Variable
Difficulty Level (1, 2, 3, 4)	difficulty of word pair in question
Test Identifier (1st, 2nd)	habituation effects within each lesson
Week (1, 2)	habituation effects over two weeks
Sound Amplification System (A, B)	acoustic characteristics
Seat Group (1, 2, 3, 4, 5)	Seat position defined in Figure 3

## 2.6. Ethical Notes

The English listening test study at the school was conducted with the approval of the Human Research Ethics Committee of the Graduate School of Engineering Science, Osaka University (protocol code R1-16, 3 December 2019). The participation of the study was agreed by the school’s head teacher and the guardians of children were informed about the study in advance.

## 3. Results and Discussions

### 3.1. Acoustic Measurements

In the classroom, the measured unoccupied noise level,  $L_{Aeq,10s}$ , was 31.3 dBA with the air conditioners turned off. Occupied noise level varied around 40 dBA to 50 dBA during English tests with no one speaking inside the room although lessons in other rooms and playground were carried on. The American Speech–Language–Hearing Association recommends to ensure unoccupied classroom noise level  $\leq 35$  dBA [56]. Correspondingly, the Japanese school building guidelines [36] recommend  $L_{Aeq,T} \leq 40$  dBA, where T indicates the measurement period that can be determined depending on the situation.

Table 4 shows the estimated RT and STI, calculated by the methods explained in Section 2.4. Likewise, the recommended RT for unoccupied furnished classrooms is 0.6 s at maximum, as average values over octave bands with centre frequencies of 500 Hz, 1 kHz, and 2 kHz (according to the American standard [56,57], for classrooms  $\leq 283$  m<sup>3</sup>) or similarly for octave bands with centre frequencies of 500 Hz and 1 kHz (according to the Japanese guidelines [36], for classrooms  $\leq 200$  m<sup>3</sup> approximately). Unlike the common practice of simply measuring the room reverberation and STI, often seen in architectural acoustics, the results we obtained informs us the combined characteristics of the unoccupied furnished room and either of the utilised sound amplification systems. As shown in Table 4, the RT values of both systems are almost within the recommended range with little variation between the two sets although System B is slightly better than System

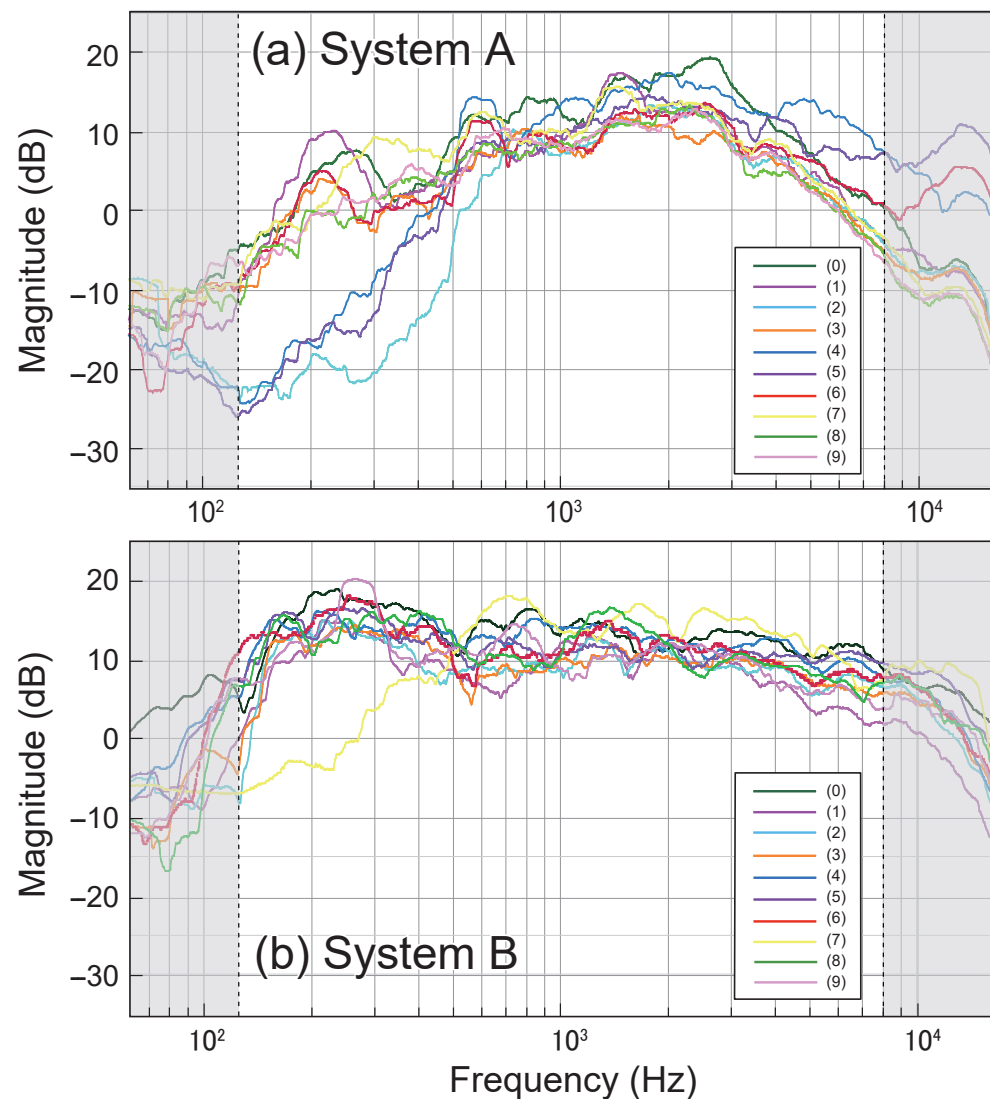
A. The STI, on the other hand, are better with the System B at all measurement positions, compared to that with System A. According to an indication provided in the standard [50],  $STI > 0.76$  observed at all positions with System B imply “Excellent intelligibility”, but  $STI > 0.66$  observed at all positions except at position (2) with System A imply “High speech intelligibility”, and  $STI > 0.62$  at position (2) with System A still implies “Good speech intelligibility”.

**Table 4.** RT (s) for octave bands with centre frequencies 500 Hz, 1 kHz, and 2 kHz (in accordance with the American standard [57]), and STI for the two systems. The measurement positions are indicated in Figure 3 (right).

System A				
Position	500 Hz	1 kHz	2 kHz	STI
(0)	0.66	0.48	0.41	0.794
(1)	0.57	0.44	0.39	0.762
(2)	0.6	0.48	0.41	0.622
(3)	0.6	0.52	0.41	0.717
(4)	0.53	0.46	0.42	0.736
(5)	0.64	0.47	0.41	0.702
(6)	0.6	0.47	0.42	0.745
(7)	0.6	0.48	0.42	0.775
(8)	0.61	0.51	0.42	0.715
(9)	0.67	0.46	0.41	0.739
System B				
Position	500 Hz	1 kHz	2 kHz	STI
(0)	0.57	0.45	0.41	0.848
(1)	0.58	0.48	0.42	0.775
(2)	0.56	0.48	0.43	0.781
(3)	0.53	0.49	0.44	0.801
(4)	0.57	0.5	0.43	0.814
(5)	0.55	0.51	0.44	0.784
(6)	0.67	0.55	0.63	0.792
(7)	0.57	0.52	0.43	0.805
(8)	0.59	0.53	0.43	0.796
(9)	0.57	0.5	0.45	0.797

In comparison with that in RT values, the difference in transfer functions between the two systems was more clearly observed as depicted in Figure 4. The frequency bands between 125 Hz and 8 kHz are understood to contain the most elements of English speech sound [31]. Therefore, signals over this range should be clearly audible (e.g.,  $SNR \geq +15$  dB is recommended in [56]) when English speech stimuli are presented to children in their classrooms. The frequency response of System A reveals that the signals arriving at each listening point diverge more than that of System B. It indicates the difficulty of using System A to ensure that everyone can hear speech sound very well at wherever they are in the same room without harming hearing of no one, including those close to the loudspeakers. System B exhibits its relatively flat response at least over the frequency bands between 200 Hz and 10 kHz at many measurement positions. The shortcomings observed at some positions below 500 Hz for both systems (e.g., System B at Position (7)) are possibly due to the room modes although the cause is not fully identifiable with the collected information only.





**Figure 4.** Transfer functions realised by two systems ((a) System A, (b) System B), both measured by using ESS method (Section 2.4) in the unoccupied furnished English room on a day in December 2019. x-axis: frequency (62.5 Hz to 16 kHz, but unshaded for 125 Hz to 8 kHz), y-axis: magnitude (−35 dB to 25 dB), depicted with 1/3 octave windowing function in Systune pro (AFMG, Berlin, Germany).

### 3.2. English Test

The correct answer rate  $p_c$  is shown in Table 5, of which the list is sorted according to the difficulty levels in ascending order. The mean correct answer rates were 94.2% in Level 1, 73.7% in Level 2, 50.6% in Level 3 and 22.3% in Level 4. The word pairs categorized into the Level 4 indicated extremely low correct answer rates. Note that chance level in a two-choice forced task, like our English test, is 50%. Nevertheless, the Japanese native children yielded much lower,  $p_c < 50\%$ , with some word pairs. This finding would imply the existence of a bias effect caused by the difference between Japanese and English phonemes and/or lack of English knowledge.

All explanatory variables in Table 3 were selected for the multivariate logistic regression model after applying variable selection based on the AIC minimisation. The Odds Ratio (OR) and its 95% Confidence Interval (CI) of each explanatory variable in univariate and multivariate logistic regression analyses are shown in Table 6. An OR greater than 1 implies a positive association with an increased correct answer rate, whereas an OR less than 1 implies a negative association. The bootstrap validation result is shown in Table 7.

**Table 5.** Correct answer rate  $p_c$  of each word pair.

Word 1	Word 2	Correct Answer Rate $p_c$ (%)	Number of Samples	Difficulty Level
light	right	17.2	412	4
long	wrong	20.6	413	4
tin	tin	21.8	412	4
pull	pool	33.4	413	4
tennis	dennis	41.9	413	3
track	truck	48.9	413	3
top	top	52.4	412	3
coat	goat	59.1	411	3
cap	cab	63.8	412	2
class	glass	69.6	411	2
men	men	71.2	413	2
sheep	cheap	73.1	412	2
fool	fool	74.3	412	2
lock	look	74.8	412	2
cot	pot	75.3	413	2
bin	bin	77.7	412	2
car	cow	78.2	413	2
cart	card	78.9	413	2
bag	bug	83.1	413	1
snore	snow	83.5	412	1
vine	wine	84.3	413	1
pot	pot	87.1	412	1
bell	bell	88.6	413	1
could	could	88.9	413	1
pear	bear	90.8	413	1
Bobby	Bobby	93.0	413	1
cab	cab	94.2	413	1
pig	peg	94.7	413	1
wood	good	96.1	413	1
ten	pen	96.6	412	1
pat	pat	96.6	413	1
cup	cup	96.6	413	1
fan	fan	96.6	413	1
jeer	jeer	97.8	412	1
hill	hill	98.1	413	1
sand	sand	98.8	413	1
walk	walk	99.0	413	1
wig	fig	99.3	412	1
cut	cut	99.3	413	1
toy	toy	99.3	413	1
jam	yam	99.3	413	1
rack	rock	99.8	413	1

Obviously, the most powerful predicting factor for the correct answer rates  $p_c$  was the difficulty level of word pairs. In the multivariate analysis, correct answer rates  $p_c$  of the second week were better than those of the first week, indicating that many children did indeed get used to the task over those two weeks. Additionally, sound amplification system type also improved the correct answer rates  $p_c$  to a certain extent. In regards to seat positions, those who sat on the seats located at the back of the classroom (Seat Group 4–5) performed worse than those whose seats were located in Seat Group 1 area. The habituation effects during single lesson is uncertain in the results at this point.

**Table 6.** Logistic regression for all data.

	Univariate		Multivariate	
	OR	95% CI	OR	95% CI
Difficulty Level				
Level 2 vs. Level 1	0.17 ***	(0.15–0.19)	0.17 ***	(0.15–0.19)
Level 3 vs. Level 1	0.06 ***	(0.06–0.07)	0.06 ***	(0.06–0.07)
Level 4 vs. Level 1	0.02 ***	(0.01–0.02)	0.02 ***	(0.02–0.02)
Test Identifier of a lesson				
2nd Test vs. 1st Test	1.08 *	(1.00–1.16)	0.88 **	(0.80–0.96)
Week				
Week 2 vs. Week 1	1.15 ***	(1.07–1.24)	1.23 ***	(1.13–1.34)
Sound Amplification System				
System B vs. System A	1.06	(0.99–1.15)	1.11 *	(1.01–1.21)
Seat Group				
Group 2 vs. Group 1	0.96	(0.86–1.08)	0.94	(0.82–1.08)
Group 3 vs. Group 1	0.98	(0.87–1.10)	0.97	(0.85–1.12)
Group 4 vs. Group 1	0.91	(0.81–1.02)	0.87 *	(0.76–1.00)
Group 5 vs. Group 1	0.83 **	(0.73–0.93)	0.76 ***	(0.66–0.88)

(OR = Odds Ratio. Significance codes: \*\*\*  $\leq 0.001$ , \*\*  $\leq 0.01$ , \*  $\leq 0.05$ ).

**Table 7.** Bootstrap-validated C-statistic for the multivariate logistic regression model shown in Table 6.

Original C-Statistic	Optimism	Corrected C-Statistic
0.8314	0.0002	0.8312

Multivariate analysis excluding the effects of the difficulty level can be beneficial to evaluate the effects of other factors. Then, we divided the data into two subsets, without and with Level 4 words, and carried out logistic regression analyses for each subset. Results are shown in Tables 8 and 9, where only variables selected by AIC minimisation for the multivariate logistic regression models are included. Bootstrap validation results are shown in Table 10.

The sub-analysis using the subset containing only difficulty levels 1–3 words (Table 8) showed similar tendencies with our first analysis results (Table 6). On the contrary, the sub-analysis results for the Level 4 words (Table 9) appeared to be different, especially regarding the sound amplification system and the seat groups.

Speech recognition is known to be governed by acoustic signals as well as the language experience of listeners [31]. Generally speaking, for those learning second languages (L2), differentiating phonemes that do not exist in their first language (L1) is more challenging [58]. The English word pairs categorised into the difficulty level 4 have no difference in pronunciation when they are transcribed in Japanese characters like Katakana. Therefore, one of the causes for the extremely low correct answer rate would be due to lack of exposure to such phoneme contrasts. In addition, elementary school children may not have knowledge of English words in the test. Therefore, it is extremely difficult for such children to suppose that words in those pairs may be different.

Early studies suggested that such non-native phonemes may be incorporated into listeners' L1 categories, which typically had been developed before their first birthdays [59,60]. For example, the lowest question difficulty of 17.2% was for the word pair “light–right” (Table 5). Japanese people are particularly known to have tendency of getting confused between /l/ and /r/ sound [61,62]. The third formant ( $F_3$ ) [63] is reported to be the critical cue that English native listeners use when differentiating between /l/ and /r/, but many Japanese native listeners frequently fail to employ it [64]. Although speech perception is not just about simply detecting the formant, being able to hear such useful acoustic speech cues is helpful for listeners to infer phonemes contained in speech signals [63].

Special care or perhaps separate training is required when teaching L2 phonemes that do not appear in their L1 phoneme structure. Research showed that mothers often communicate with their infants with a special form of speech, known as “parentese”, with higher pitch, overemphasised intonation contour, spoken slowly [65]. Additionally, albeit incomplete, many have attempted to develop methods to train Japanese natives (including adults) to learn abilities to differentiate /l/ and /r/, as in [62], for example. Above results suggest sound amplification systems equipped with functions allowing flexible control over realised frequency response could be valuable in second and foreign language education.

**Table 8.** Logistic Regression: sub analysis 1 (without Level 4 data).

	Univariate		Multivariate	
	OR	95% CI	OR	95% CI
Difficulty Level				
Level 2 vs. Level 1	0.17 ***	(0.15–0.19)	0.17 ***	(0.15–0.19)
Level 3 vs. Level 1	0.06 ***	(0.06–0.07)	0.06 ***	(0.06–0.07)
Test Identifier of a lesson				
2nd Test vs. 1st Test	0.72 ***	(0.66–0.78)	0.92	(0.83–1.01)
Week				
Week 2 vs. Week 1	1.20 ***	(1.10–1.31)	1.26 ***	(1.14–1.38)
Sound Amplification System				
System B vs. System A	1.13 **	(1.03–1.23)	1.17 **	(1.06–1.28)
Seat Group				
Group 2 vs. Group 1	0.90	(0.78–1.03)	0.88	(0.75–1.02)
Group 3 vs. Group 1	0.93	(0.80–1.07)	0.91	(0.78–1.07)
Group 4 vs. Group 1	0.81 **	(0.71–0.92)	0.77 ***	(0.67–0.90)
Group 5 vs. Group 1	0.72 ***	(0.63–0.83)	0.67 ***	(0.58–0.78)

(OR = Odds Ratio. Significance codes: \*\*\* ≤ 0.001, \*\* ≤ 0.01).

**Table 9.** Logistic Regression: sub analysis 2 (Level 4 (the hardest) data only).

	Univariate		Multivariate	
	OR	95% CI	OR	95% CI
Test Identifier of a lesson				
2nd Test vs. 1st Test	0.62 ***	(0.46–0.81)	0.61 ***	(0.46–0.81)
Sound Amplification System				
System B vs. System A	0.79 *	(0.63–1.00)	0.78 *	(0.62–0.99)
Seat Group				
Group 2 vs. Group 1	1.40	(0.96–2.05)	1.41	(0.97–2.06)
Group 3 vs. Group 1	1.41	(0.96–2.07)	1.42	(0.97–2.08)
Group 4 vs. Group 1	1.76 **	(1.23–2.53)	1.77 **	(1.23–2.55)
Group 5 vs. Group 1	1.63 *	(1.12–2.40)	1.64 *	(1.12–2.41)

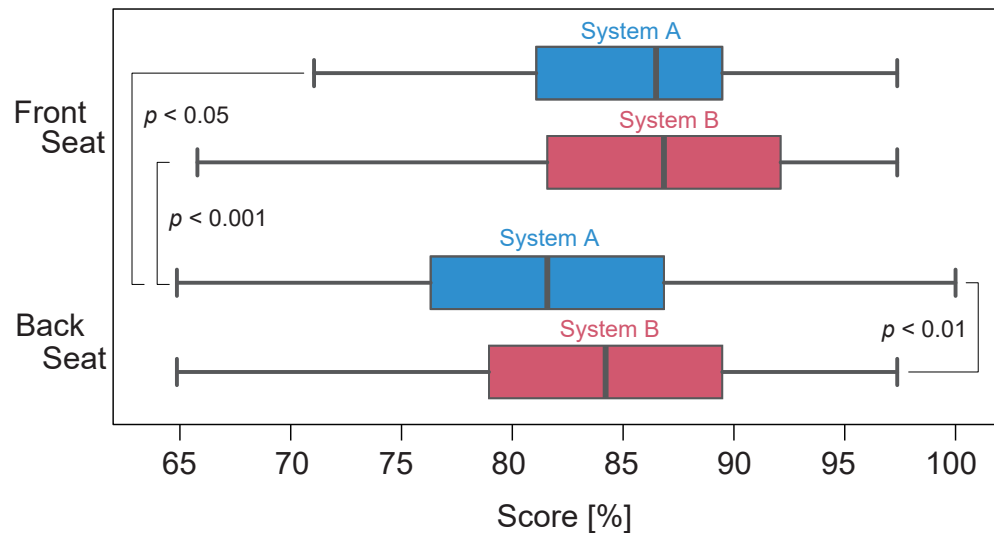
(OR = Odds Ratio. Significance codes: \*\*\* ≤ 0.001, \*\* ≤ 0.01, \* ≤ 0.05).

**Table 10.** Bootstrap-validated C-statistic for the multivariate logistic regression models shown in Tables 8 and 9.

	Original C-Statistic	Optimism	Corrected C-Statistic
Difficulty Level 1–3	0.7757	0.0006	0.7750
Difficulty Level 4	0.5931	0.0077	0.5854

Finally, to assess the association between the test score and the heterogeneous sound fields in the classroom, we divided the subset with difficulty levels 1–3 words further into four groups according to sound amplification systems (A or B) and seat locations (Front or

Back), then compared the medians of test scores (38 questions per system). As shown in Figure 5, the lowest median of the test score is observed in the sub group “System A at the back seats”, and it is significantly lower than that of “System B at the back seats”. With System A, the difference between front and back seats is noticed, indicating heterogeneous sound fields in the classroom realised by using such a system.



**Figure 5.** Box-and-whisker plots of sub groups divided according to sound amplification systems and seat locations. The centre line of the box is the median, the ends of the box are the quartiles, and the ends of the whiskers extend to 1.5 times the inter quartile range, in Seat vs. System analyses.

Collective findings from both the acoustic measurements (Section 3.1) and our statistical results exposed that the difference in the acoustic quality realised by the two sound amplification systems had an impact on, at least, Japanese children’s English learning, and the inequality in their classroom.

#### 4. Conclusions

The current study has demonstrated the importance of delivering good acoustic conditions in classrooms to help children learn English more effectively through a scientific approach, using data collected in a real-life school classroom in Japan.

The statistical results of English listening tests taken by 216 Japanese children (aged between 11 and 12 years) using two sound amplification systems revealed the following three points. (1) There is association between the acoustic conditions in classrooms and the correct answer rate of the English tests taken by children in the classroom. (2) Using sound amplification systems with loudspeaker arrays can reduce the existing inequality in a classroom due to the heterogeneous sound fields that depend on the seat positions. (3) English word combinations including phonemes that do not exist in L1 are particularly hard for Japanese children to discriminate, and they negatively affect the correct answer rate of the English tests. Hence, teaching those words particularly requires extra care. Furthermore, the joint results with the multiple acoustic measurements show that even when the difference in acoustic characteristics realised by the two systems is at the magnitude of what has been shown, its impact on children’s learning is still noteworthy. The observed differences in our acoustic measurements were not so noticeable in the RT values but they were evident in the transfer functions.

Although improvement can be made to the English listening test materials for any future studies of similar kinds, as a consequence, we illustrated that more adequately adjusted acoustics could play a significant role to contrive children’s learning at their schools. More research should be encouraged, for example, to set guidelines regarding acoustic quality in classrooms at schools in any countries without them.

**Author Contributions:** Conceptualization, N.E.; methodology, N.E. and K.K.; measurement, N.E., M.K., I.S. and K.K.; formal analysis, N.E. and K.K.; writing—original draft preparation, N.E.; writing—review and editing, N.E., M.K., I.S., T.S. and K.K.; visualization, N.E. and K.K.; project administration, N.E. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by JSPS KAKENHI Grants No. 19H04282.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Human Research Ethics Committee of the Graduate School of Engineering Science, Osaka University, Japan (protocol code R1-16, 3 December 2019).

**Informed Consent Statement:** Written informed consent was obtained from the principal of the participating school. In addition, an opportunity to reject participation or use of the children’s data was provided to the children’s legal guardians using a document specifically prepared for this purpose as an opt-out method.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical issues.

**Acknowledgments:** N.E. gratefully acknowledges guidance on designing the English tests provided by associate Takayuki Nakamori at Kyoto University (Japan).

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

## References

1. EF Education First. *EF English Proficiency Index: A Ranking of 100 Countries and Regions by English Skills*; Technical Report; EF Education First Ltd.: Lucerne, Switzerland, 2020. Available online: <https://www.ef.com/epi/> (accessed on 28 June 2021).
2. Nuttall, L. Comparative Education: Why Does Japan Continue to Struggle with English? *Read. Matrix Int. Online J.* **2019**, *19*, 74–92.
3. Okumura, S. Homeroom Teachers or Specialist Teachers? Considerations for the Workforce for Teaching English as a Subject at Elementary Schools in Japan. *Asian J. Educ. Train.* **2017**, *3*, 1–5. [[CrossRef](#)]
4. Houtgast, T. The effect of ambient noise on speech intelligibility in classrooms. *Appl. Acoust.* **1981**, *14*, 15–25. [[CrossRef](#)]
5. Bradley, J.S. Speech intelligibility studies in classrooms. *J. Acoust. Soc. Am.* **1986**, *80*, 846–854. [[CrossRef](#)] [[PubMed](#)]
6. Hétu, R.; Truchon-Gagnon, C.; Bilodeau, S.A. Problems of noise in school settings: A review of literature and the results of an exploratory study. *J. Speech-Lang. Pathol. Audiol.* **1990**, *14*, 31–39.
7. Rosenberg, G.G.; Blake-Rahter, P.; Heavner, J.; Allen, L.; Redmond, B.M.; Phillips, J.; Stigers, K. Improving Classroom Acoustics (ICA): A three-year FM sound field classroom amplification study. *J. Educ. Audiol.* **1999**, *7*, 8–28.
8. Nelson, P.B.; Soli, S. Acoustical Barriers to Learning: Children at Risk in Every Classroom. *Lang. Speech Hear. Serv. Sch.* **2000**, *31*, 356–361. [[CrossRef](#)] [[PubMed](#)]
9. ASHA Working Group on Classroom Acoustics. *Acoustics in Educational Settings: Technical Report*; American Speech-Language-Hearing Association: Manchester, MD, USA, 2005.
10. Sato, H.; Bradley, J.S. Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *J. Acoust. Soc. Am.* **2008**, *123*, 2064–2077. [[CrossRef](#)]
11. Rabbitt, P. Recognition: Memory for words correctly heard in noise. *Psychon. Sci.* **1966**, *6*, 383–384. [[CrossRef](#)]
12. Key, A.P.; Gustafson, S.J.; Rentmeester, L.; Hornsby, B.W.Y.; Bess, F.H. Speech-Processing Fatigue in Children: Auditory Event-Related Potential and Behavioral Measures. *J. Speech Lang. Hear. Res. JSLHR* **2017**, *60*, 2090–2104. [[CrossRef](#)]
13. Evans, G.W.; Lepore, S.J. Nonauditory Effects of Noise on Children: A Critical Review. *Child. Environ.* **1993**, *10*, 31–51.
14. Anderson, K.L. Kids in noisy classrooms: What does the research really say. *J. Educ. Audiol.* **2001**, *9*, 21–33.
15. Shield, B.M.; Dockrell, J.E. The Effects of Noise on Children at School: A Review. *Build. Acoust.* **2003**, *10*, 97–116. [[CrossRef](#)]
16. Anderson, K. The Problem of Classroom Acoustics: The Typical Classroom Soundscape Is a Barrier to Learning. *Semin. Hear.* **2004**, *25*, 117–129. [[CrossRef](#)]
17. Shield, B.M.; Dockrell, J.E. The effects of environmental and classroom noise on the academic attainments of primary school children. *J. Acoust. Soc. Am.* **2008**, *123*, 133–144. [[CrossRef](#)] [[PubMed](#)]
18. Tiesler, G.; Machner, R.; Brokmann, H. Classroom Acoustics and Impact on Health and Social Behaviour. *Energy Procedia* **2015**, *78*, 3108–3113. [[CrossRef](#)]
19. Magimairaj, B.; Nagaraj, N.; Benafield, N. Children’s Speech Perception in Noise: Evidence for Dissociation From Language and Working Memory. *J. Speech Lang. Hear. Res.* **2018**, *61*, 1294–1305. [[CrossRef](#)]
20. Peng, J.X.; Jiang, P. The Effects of the Noise and Reverberation on the Working Memory Span of Children. *Arch. Acoust.* **2018**, *43*, 123–128.
21. Cucharero, J.; Hänninen, T.; Lokki, T. Influence of Sound-Absorbing Material Placement on Room Acoustical Parameters. *Acoustics* **2019**, *1*, 644–660. [[CrossRef](#)]

22. Leavitt, R.; Flexer, C. Speech degradation as measured by the Rapid Speech Transmission Index (RASTI). *Ear Hear.* **1991**, *12*, 115–118. [CrossRef]
23. Iannace, G.; Trematerra, A.; Trematerra, P. Acoustic correction using green material in classrooms located in historical buildings. *Acoust. Aust.* **2013**, *41*, 213–218.
24. Berardi, U.; Iannace, G.; Trematerra, A. Acoustic Treatments Aiming to Achieve the Italian Minimum Environmental Criteria (CAM) Standards in Large Reverberant Classrooms. *Can. Acoust. Can.* **2019**, *47*, 73–80.
25. Kuusinen, A.; Saariniemi, E.; Sivonen, V.; Dietz, A.; Aarnisalo, A.A.; Lokki, T. An exploratory investigation of speech recognition thresholds in noise with auralisations of two reverberant rooms. *Int. J. Audiol.* **2021**, *60*, 210–219. [CrossRef]
26. Kaplanis, N.; Bech, S.; Lokki, T.; van Waterschoot, T.; Holdt Jensen, S. Perception and preference of reverberation in small listening rooms for multi-loudspeaker reproduction. *J. Acoust. Soc. Am.* **2019**, *146*, 3562–3576. [CrossRef] [PubMed]
27. Mealings, K. Classroom acoustic conditions: Understanding what is suitable through a review of national and international standards, recommendations, and live classroom measurements. In Proceedings of the Acoustics 2016: The Second Australasian Acoustical Societies' Conference, Brisbane, Australia, 9–11 November 2016.
28. Freeman, S.; Eddy, S.L.; McDonough, M.; Smith, M.K.; Okoroafor, N.; Jordt, H.; Wenderoth, M.P. Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8410–8415. [CrossRef] [PubMed]
29. Rowe, K. Effective teaching practices for students with and without learning difficulties: Constructivism as a legitimate theory of learning AND of teaching? *Stud. Learn. Process.* **2006**, *10*, 1–24.
30. Johnson, C.E. Children's phoneme identification in reverberation and noise. *J. Speech Lang. Hear. Res.* **2000**, *43*, 144–157. [CrossRef]
31. Northern, J.L.; Downs, M.P. *Hearing in Children*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2014.
32. Lyberg-Åhlander, V.; Haake, M.; Brännström, J.; Schötz, S.; Sahlén, B. Does the speaker's voice quality influence children's performance on a language comprehension test? *Int. J. Speech Lang. Pathol.* **2015**, *17*, 63–73. [CrossRef]
33. Rantala, L.; Sala, E. Associations between Classroom Conditions and Teacher's Voice Production. *Energy Procedia* **2015**, *78*, 3120–3125. [CrossRef]
34. Cantor Cutiva, L.C.; Vogel, I.; Burdorf, A. Voice disorders in teachers and their associations with work-related factors: A systematic review. *J. Commun. Disord.* **2013**, *46*, 143–155. [CrossRef]
35. Puglisi, G.; Cutiva, L.C.; Pavese, L.; Castellana, A.; Bona, M.; Fasolis, S.; Lorenzatti, V.; Carullo, A.; Burdorf, A.; Bronuzzi, F.; et al. Acoustic Comfort in High-school Classrooms for Students and Teachers. *Energy Procedia* **2015**, *78*, 3096–3101. [CrossRef]
36. Architectural Institute of Japan Environmental Standard. *Standard and Design Guidelines for School Environment in School Buildings*; Architectural Institute of Japan: Tokyo, Japan, 2020.
37. Slavin, R.E. Evidence-Based Education Policies: Transforming Educational Practice and Research. *Educ. Res.* **2002**, *31*, 15–21. [CrossRef]
38. Odom, S.L.; Brantlinger, E.; Gersten, R.; Horner, R.H.; Thompson, B.; Harris, K.R. Research in Special Education: Scientific Methods and Evidence-Based Practices. *Except. Child.* **2005**, *71*, 137–148. [CrossRef]
39. Baker, A. *Tree or Three? Student's Book and Audio CD: An Elementary Pronunciation Course*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2006.
40. Baker, A. *Ship or Sheep? Book and Audio CD Pack: An Intermediate Pronunciation Course*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2006.
41. Steinberg Media Technologies GmbH. The True Art of Mastering: WAVELAB. 2021. Available online: <https://new.steinberg.net/wavelab/> (accessed on 17 April 2021).
42. Python Software Foundation. Python 3.7.4. 2019. Available online: <https://docs.python.org/release/3.7.4/> (accessed on 17 April 2021).
43. International Organization for Standardization. *ISO 1996-1:2016 Acoustics—Description, Measurement and Assessment of Environmental Noise—Part 1: Basic Quantities and Assessment Procedures*; ISO: Geneva, Switzerland, 2016.
44. Neitzel, R. *Determination of Risk of Noise-Induced Hearing Loss Due to Recreational Sound: Review*; Technical Report; World Health Organization: Geneva, Switzerland, 2017.
45. Farina, A. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In Proceedings of the Audio Engineering Society Convention 108, Paris, France, 19–22 February 2000.
46. International Electrotechnical Commission. *IEC 61672-1:2013 Electroacoustics—Sound Level Meters—Part 1: Specifications*; IEC: Geneva, Switzerland, 2013.
47. Farina, A. Advancements in Impulse Response Measurements by Sine Sweeps. In Proceedings of the Audio Engineering Society Convention 122, Vienna, Austria, 5–8 May 2007.
48. AFMG Technologies GmbH. SysTune 1.2. 2020. Available online: <https://systune.afmg.eu/index.php/st-software-en.html> (accessed on 17 April 2021).
49. International Organization for Standardization. *ISO 3382-2:2008 Acoustics—Measurement of Room Acoustic Parameters —Part 2: Reverberation Time in Ordinary Rooms*; ISO: Geneva, Switzerland, 2008.
50. International Electrotechnical Commission. *IEC 60268 Sound System Equipment—Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index*; IEC: Geneva, Switzerland, 2020.
51. Houtgast, T.; Steeneken, H.J.M. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J. Acoust. Soc. Am.* **1985**, *77*, 1069–1077. [CrossRef]

52. AFMG Ahnert Feistel Media Group. *Systune 1.3 Manual*, 5th ed.; AFMG Technologies GmbH: Berlin, Germany, 2014.
53. Harrell, F.F.J. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, 2nd ed.; Springer International Publishing: Cham, Switzerland, 2015.
54. Steyerberg, E.W.; Harrell, F.E.J.; Borsboom, G.J.; Eijkemans, M.J.; Vergouwe, Y.; Habbema, J.D. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J. Clin. Epidemiol.* **2001**, *54*, 774–781. [[CrossRef](#)]
55. R Core Team. R: A Language and Environment for Statistical Computing. 2019. Available online: <https://www.r-project.org/> (accessed on 17 April 2021).
56. ASHA Working Group on Classroom Acoustics. *Acoustics in Educational Settings: Position Statement*; American Speech-Language-Hearing Association: Manchester, MD, USA, 2005.
57. Acoustical Society of America. *ANSI/ASA S12.60-2020/Part 1 American National Standard Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, Part 1: Permanent Schools*; Acoustical Society of America: New York, NY, USA, 2010.
58. Iverson, P.; Ekanayake, D.; Hamann, S.; Sennema, A.; Evans, B.G. Category and perceptual interference in second-language phoneme learning: An examination of English /w/-/v/ learning by Sinhala, German, and Dutch speakers. *J. Exp. Psychol. Hum. Percept. Perform.* **2008**, *34*, 1305–1316. [[CrossRef](#)] [[PubMed](#)]
59. Werker, J.F.; Tees, R.C. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* **1984**, *7*, 49–63. [[CrossRef](#)]
60. Best, C.C.; McRoberts, G.W. Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Lang. Speech* **2003**, *46*, 183–216. [[CrossRef](#)]
61. Strange, W.; Dittmann, S. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Percept. Psychophys.* **1984**, *36*, 131–145. [[CrossRef](#)]
62. Logan, J.S.; Lively, S.E.; Pisoni, D.B. Training Japanese listeners to identify English /r/ and /l/: A first report. *J. Acoust. Soc. Am.* **1991**, *89*, 874–886. [[CrossRef](#)]
63. Benesty, J.; Sondhi, M.M.; Huang, Y. (Eds.) *Springer Handbook of Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2008.
64. Gordon, P.C.; Keyes, L.; Yung, Y.F. Ability in perceiving nonnative contrasts: Performance on natural and synthetic speech stimuli. *Percept. Psychophys.* **2001**, *63*, 746–758. [[CrossRef](#)] [[PubMed](#)]
65. Kuhl, P.K.; Andruski, J.E.; Chistovich, I.A.; Chistovich, L.A.; Kozhevnikova, E.V.; Ryskina, V.L.; Stolyarova, E.I.; Sundberg, U.; Lacerda, F. Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science* **1997**, *277*, 684–686. [[CrossRef](#)] [[PubMed](#)]