

Evaluating Room Acoustics for Speech Intelligibility

Caitlin R. Kunchur

Dutch Fork High School, Irmo, SC, USA Email: crkunchur@gmail.com

How to cite this paper: Kunchur, C.R. (2019) Evaluating Room Acoustics for Speech Intelligibility. *Open Journal of Applied Sciences*, **9**, 601-612. https://doi.org/10.4236/ojapps.2019.97048

Received: June 25, 2019 **Accepted:** July 27, 2019 **Published:** July 30, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

Room acoustics play an important role in the intelligibility of speech. The main aspect of acoustics that is usually studied is the duration of the reverberation decay, since a long decay causes a blurring of phonemes. However, other parameters of the acoustics such as the strength of the reverberation can actually improve intelligibility. These factors do not receive the same attention. In many common practical situations such as classrooms and residential rooms, it would be of value to quantitatively study the acoustics to optimize the room's function, but this is not done routinely due to the expected expense or difficulty involved. This research explores inexpensive first-principle methods to quantitatively measure three key parameters of a room's acoustics: the reverberation decay time RT60, the reverberant intensity I_{R} , and the room's total absorption A. The required equipment includes two laptops installed with certain free softwares. Generation of the required noise signal and level detection are carried out using the REW software, and long-duration recordings are carried out using the Audacity software. The procedures are simple enough to be performed without specialized training and do not require specialized equipment, only commonly available household resources. This research also sheds light on the fact that not all reverberation is bad and that strong but short-duration reverberation can enhance communication. This information can be expected to benefit schools and other venues where speech intelligibility is vital.

Keywords

Acoustics, Reverberation, Speech, Music, Communication

1. Introduction and Background

Sound reaches a listener in two ways: through direct sound and reflected sound. Sound intensity *I* is defined as the power *P* per surface area *S*:

$$I = \frac{P}{S} \tag{1}$$

For direct sound that spreads out spherically, its intensity I_D decreases with distance *r* as expressed by the inverse square law:

$$I_D = \frac{P}{4\pi r^2} \propto \frac{1}{r^2} \tag{2}$$

If the sound is more concentrated in one direction, there will be an additional factor (>1) for that direction. This relationship applies to the direct sound and is altered by reflected sound.

Reflected sound falls into various categories of which three broad ones [1] [2] [3] [4] are: 1) early (<30 ms) indistinguishable single reflections, 2) delayed (>30 ms) single reflections, which are distinguishable as echoes (the minimum delay depends on the intensity), and 3) reverberation, which is the forest of slowly decaying multiple reflections within the room. The strength of reverberation is the reverberant intensity I_R and its duration is the decay time RT60. The reverberant intensity for a continuous sound source depends only on the source's power P and the absorption A of the room's surfaces:

$$I_R = \frac{P}{A} \tag{3}$$

where A refers to the fraction of sound energy absorbed by the room's total surface area per reflection. A is related to the absorption coefficient a, which is the fraction of sound energy absorbed per reflection per unit area by a given material, by the equation:

1

$$A = S\alpha_{\rm avg} \tag{4}$$

where *S* is the room's surface area and a_{avg} is the average absorption coefficient for the entire surface. Smooth and hard surfaces such as a mirror reflect almost all sound and thus have an absorption coefficient of almost zero. In contrast, a completely absorptive surface such as a mattress would have an absorption coefficient of almost one.

The total sound intensity in the room was modeled by adding the reverberant and direct intensities:

$$I = I_D + I_R = \frac{P}{4\pi r^2} + \frac{P}{A}$$
⁽⁵⁾

The reverberation decay time RT60 is defined as the number of seconds needed for the intensity to drop by a factor of 10⁶ (which corresponds to a drop in sound level of 60 dB). In the simplistic case of a symmetrical room with a low a_{avg} and uniformly distributed absorption, RT60 can be related to the volume of the room V and total absorption by the Sabine equation: RT60 = 0.161 V/A. This predicts that rooms with larger volumes will have longer reverberation times and that rooms with larger absorption values will have shorter reverberation times. For high values of a_{avg} , the Eyring-Norris equation: $RT60 = \frac{0.161V}{\left[-S\ln\left(1-\alpha_{avg}\right)\right]}$ works

slightly better. But, both equations are prone to error because of non-uniform distribution of absorption and anisotropic radiation pattern of the sound source [5]. Thus, in practice it is best to obtain RT60 empirically from the actual measured exponential decay of the reverberation after the sound has stopped, as this will apply most directly to a realistic situation. For example, in a classroom the sound source should be placed at the teacher's location in the front of the room and the microphone should be placed at the location of the audience.

A comparison between various professional methods for measuring RT60 can be found in [6] [7] [8]. These approaches use specialized equipment and involved analysis, and would therefore be less accessible to the average teacher or lay person. The first goal of this work was to explore simpler inexpensive first-principle methods to study room acoustics. The second goal was to measure not only RT60, as is usually the case, but also the other key parameters of a room's acoustics: A and I_R . The equipment consists of two laptops with a built-in microphone and speaker (or optionally an external microphone and speaker), and free software applications for measuring the sound level and generating signals.

2. Methods

2.1. Measurement of Reverberant Intensity and Absorption

This is carried out using the setup illustrated in **Figure 1** and applying Equation (5).

The first laptop plays white noise produced by the signal-generator application (app), within the REW software suite, through an external speaker. An external microphone was connected to the second laptop running the sound-level-meter (SLM) application, another part of the REW software suite. REW software is a free software [9]. Using white noise instead of single frequencies avoids standing waves and uneven distribution of sound level. For the detection, the SLM application gives a more accurate and stable reading of the sound level than using an oscilloscope software, which was attempted first. While the speaker played white noise continuously, the microphone was moved to different distances away from



Figure 1. Setup for measuring reverberant intensity I_R and absorption *A*. Laptop 1 plays white noise through a speaker that has been generated by the signal-generator application in the REW software. Laptop 2 receives the sound signal from a microphone and measures the sound level *L* using the sound-level-meter application in REW.

it and the sound level L (in dB) indicated by the SLM was recorded. The sound level is related to the intensity by:

$$L = 10 \log\left(\frac{I}{I_0}\right) \tag{6}$$

where $I_0 = 10^{-12} \text{ W/m}^2$ is the threshold of hearing.

Measurements were carried out in 3 rooms: a bathroom, a bedroom and a foyer. The bathroom has the smallest volume and least absorption compared to the other two rooms. The foyer has the largest volume, and the bedroom has the most absorption.

2.2. Measurement of the Reverberation Decay Time

There are two principal methods for measuring RT60: one uses impulsive sounds (pistol shot, balloon pop, electronically synthesized impulse, etc.) and the other uses continuously playing noise that is interrupted [6] [7] [8]. For home use, the ClapIR application for smart phones and the balloon-pop method (where the decaying sound of an actual balloon pop or its recording is captured on an oscilloscope application for a laptop) are examples of the impulse method. These popular home methods were found to be too unreliable, because such impulses are inconsistent and there are difficulties associated with the triggering of the oscilloscope. An impulse also has a very small amount of energy (unless very loud such as a pistol shot) leading to a poor signal-to-noise ratio.

After much experimentation, the setup shown in **Figure 2** was arrived at. The overall arrangement is similar to **Figure 1** except that the speaker faces away and is as far as possible from the microphone to minimize the effect of direct sound and emphasize the reverberant sound. Audacity (also a freeware) was chosen as the recording software instead of an oscilloscope software, because it has almost unlimited recording time making triggering unnecessary. (Audacity was set to record at the CD standard of 44.1 kHz sampling frequency and 16 bit vertical resolution). This setup proved to work well.



Figure 2. Setup for measuring reverberation decay time. Laptop 1 plays through a speaker, white noise generated by the signal-generator application in the REW software. The speaker faces away and is far from the microphone to fill the room with reverberation. Laptop 2 receives the sound signal from a microphone and makes an audio recording using Audacity software.

The Audacity recording was started and then the white noise signal was played for four seconds to allow the reverberation to fill the room and reach its maximum value. The exact level is unimportant since only logarithmic differences are needed in the analysis. The recording was continued for several seconds after the white noise was turned off, so that the entire decay was captured as seen in **Figure 2**.

The data downloaded from Audacity consists of the digitized sample values for the sound pressure P for each sampling time interval t. P is expressed as a fraction of Audacity's full scale value. During the seconds-long recording, there are hundreds of thousands of P values because the sampling frequency is 44.1 kHz. Squaring P gives a value that is proportional to the intensity. **Figure 3** shows an example of this intensity (normalized by its initial value) plotted versus time. There is a noticeable flat region while the signal is playing steadily and a noticeable decay when the sound is turned off. From the slope of the decay, the reverberation decay time was calculated for each room, the details of which are discussed in the "Data and Analysis" section.

2.3. Assessment of Speech Intelligibility

This last experiment conducted blind listening tests to correlate the previous physical measurements with how clearly speech could be heard in the various rooms with different values of A, I_R and RT60. A random sequence of two similar sounding phonemes "ch" and "j" was recorded. While the recording was played, each subject sat 1.5 m away and noted, in order, which phonemes they thought they heard. The percentage of phonemes judged correctly by the subjects provided an indication of how good speech intelligibility was in that room. This procedure was repeated for all three rooms. There were 261 blind tests of phonemes in total.

3. Data and Analysis

3.1. Reverberant Intensity and Absorption

Following the procedure outlined in Section 2.1, the sound level was recorded at different distances away from the speaker. This information, along with the calculated intensity (Equation (6)) is shown in **Tables 1-3** for the three rooms.

Figure 4 shows the plots of measured I vs $1/r^2$ data from the previous tables. At shorter distances, the inverse square law (straight line behavior) holds better because the direct sound is more dominant. As the distance from the speaker increases, the reverberation starts to have more influence and there is a bigger deviation from the inverse square law—as seen in the lower left portion of the graph, the intensity starts to saturate towards a constant value corresponding to I_R . The total intensity was modeled by adding the reverberant plus direct intensity as per Equation (5) (the functions with specific values for each room are shown within the figure). The plotted lines show the fitted functions. All parameters are known except for the total absorption values for each room, which

<i>r</i> (m)	<i>L</i> (dB)	$1/r^2$ (m ⁻²)	$I(W \cdot m^{-2})$
0.3	52.8	11.111	1.905E-7
0.325	52.1	9.467	1.622E-7
0.35	51.5	8.163	1.413E-7
0.375	50.7	7.111	1.175E-7
0.4	50.2	6.250	1.047E-7
0.425	49.9	5.536	9.772E-8
0.45	49.3	4.938	8.511E-7
0.475	48.6	4.432	7.244E-8
0.5	48.3	4.000	6.761-8
0.525	47.8	3.628	6.026E-8
0.55	47.5	3.306	5.623E-8
0.575	47.0	3.025	5.012E-8
0.6	46.6	2.778	4.571E-8
0.75	45.1	1.779	3.236E-8
0.9	44.4	1.235	2.754E-8
1.2	43.3	0.694	2.138E-8
1.8	41.1	0.309	1.288E-8
2.4	38.3	0.174	6.761E-9
3	36.4	0.111	4.365E-9

Table 1. Collected and calculated values for the foyer.

Table 2. Collected and calculated values for the bedroom.

<i>r</i> (m)	<i>L</i> (dB)	$1/r^2$ (m ⁻²)	$I(W \cdot m^{-2})$
0.3	52.3	11.111	1.698E-7
0.325	51.3	9.467	1.349E-7
0.35	51.0	8.163	1.259E-7
0.375	50.2	7.111	1.047E-7
0.4	49.6	6.250	9.120E-8
0.425	49.3	5.536	8.511E-8
0.45	48.7	4.938	7.413E-8
0.475	48.1	4.432	6.457E-8
0.5	47.4	4.000	5.495E-8
0.575	46.4	3.025	4.365E-8
0.75	44.5	1.778	2.818E-8
0.9	43.4	1.235	2.188E-8
1.075	42.0	0.865	1.585E-8

Table 3. Collected and calculated values for the bathroom.

<i>r</i> (m)	<i>L</i> (dB)	$1/r^2$ (m ⁻²)	$I(W \cdot m^{-2})$
0.3	53.2	11.111	2.089E-7
0.325	52.4	9.467	1.738E-7
0.35	51.7	8.163	1.479E-7
0.375	51.0	7.111	1.259E-7

Continued			
0.4	50.6	6.250	1.148E-7
0.425	50.2	5.536	1.047E-7
0.45	50.1	4.938	1.023E-7
0.475	49.9	4.432	9.772E-8
0.5	49.4	4.000	8.710-8
0.525	49.0	3.628	7.943E-8
0.55	48.7	3.306	7.413E-8
0.575	48.5	3.025	7.079E-8
0.6	48.3	2.778	6.761E-8
0.75	47.4	1.779	5.495E-8
0.9	46.2	1.235	4.169E-8
1.2	45.1	0.694	3.236E-8
1.5	44.9	0.444	3.090E-8
1.8	44.6	0.309	2.884E-8
2.1	44.4	0.227	2.754E-8



Figure 3. Example plot of normalized intensity vs time from data shown in Table 1.



Figure 4. Graph of intensity versus the inverse-square distance in three rooms with fitted functions for determining absorption values.

were adjusted until the function lines fit the data most closely. **Table 4** shows these absorption values determined from the fitted functions.

As seen in **Table 4**, the bathroom has the least absorption as expected because it has reflective surfaces like tiles and a mirror, and has the smallest surface area. The bedroom has the highest absorption because it has a thick carpet and a mattress. From these total absorption values, the average absorption coefficient was found by dividing it by the room's total surface area as per Equation (4). The foyer's volume and area were indeterminate due to the space being open and not having defined boundaries and therefore, the absorption coefficient could not be calculated.

3.2. Reverberation Decay Time

Following the procedure outlined in Section 2.2, the interrupted noise intensity was measured versus time using the Audacity recording software. These data are plotted for each of the three rooms in **Figures 5-7**.

As seen in the graphs, the intensity decreases exponentially (straight line portion on the log-linear plot) after the noise is stopped. Taking a pair of points (t_1 , I_1) and (t_2 , I_2) on the straight line portion of the decay, one gets the exponential decay constant τ (the time for *I* to drop by a factor of *e*) from:

$$\tau = \frac{t_1 - t_2}{\ln\left(I_2\right) - \ln\left(I_1\right)} \tag{7}$$

Table 4. Total room absorption values obtained from **Figure 4** and corresponding absorption coefficients a_{avg} . Volume and surface areas were calculated from the measured room dimensions.

	A (sabins)	$V(m^3)$	$S(m^2)$	$\alpha_{ m avg}$
Bathroom	8.33	15.2	37.0	0.23
Foyer	37.0			
Bedroom	55.6	53.5	103.4	0.54



Figure 5. Reverberation decay in bathroom.



Figure 6. Reverberation decay in foyer.



Figure 7. Reverberation decay in bedroom.

The RT60 (time for *I* to decay by a factor of 10⁶ or 60 dB) is obtained from this τ by multiplying it by the factor: $\frac{\log(10^6)}{\log(e)} = 13.8$. This procedure was applied for each room and the results are shown in Table 5. These measured RT60 values are in reasonable agreement with other measurements of residential rooms

allowing for differences that will depend on the furniture [10]. The bedroom, due to its high absorption (55.6 sabins), has the shortest reverberation time (0.35 s). The foyer has the largest volume and an intermediate ab-

beration time (0.35 s). The foyer has the largest volume and an intermediate absorption (37 sabins), and therefore has the longest reverberation time (0.62 s). The bathroom has the least absorption (8.33 sabins) but also the smallest volume; its RT60 (0.52 s) lies in between the values of the foyer and the bedroom.

3.3. Speech Intelligibility

Following the procedure outlined in Section 2.3, blind listening tests were conducted in each room on three subjects to assess speech intelligibility in each room. The results are shown in **Table 6** for a total of 261 blind tests of phonemes (87 blind tests per room). The statistics are sufficient to validate the conclusions since they overwhelmingly pass the chi-squared test, with a lowest value of χ^2 = 19.32 for any room; the probability (*i.e.*, *p*-value) of getting this or higher χ^2 value purely by chance is 1.10×10^{-5} . (In blind listening tests, the chi-squared value is given by $\chi^2 = (C - T/2)^2/(T/2) + (I - T/2)^2/(T/2)$ for a total number of trials *T*, number of correct judgements *C*, and number of incorrect judgements *I*. The commonly used critical value of χ^2 for one degree of freedom is 3.84 for a *p*-value of 0.05).

As seen in the results, the test conducted in the bathroom consistently produced the highest intelligibility rate for each subject. This is expected because the bathroom has a short RT60 because of its small volume, but equally important, it has a strong reverberant intensity (as seen in **Figure 4**) due to its low absorption. This confirms that simply having a low RT60 is not a sufficient criterion for best speech intelligibility, otherwise the bedroom would have had the highest speech-intelligibility score. Tests conducted in the foyer showed the worst intelligibility because of the foyer's long RT60 without the strong I_R of the bathroom. The speech intelligibility in the bedroom was in between that of the bathroom and the foyer because of its much shorter RT60 but similar I_R as the foyer. Most other studies of speech intelligibility have mainly focused on the effects of just RT60 and background noise [10] [11] [12]. In agreement with the present work, they too find that increasing just RT60 worsens speech intelligibility. However, a new aspect of the present work is to assess the influence of I_R in addition to RT60.

Room	RT60 (s)
Bathroom	0.52
Foyer	0.62
Bedroom	0.35

Table 5. Measured reverberation decay times.

Table 6. Percentage of phonemes judged correctly during blind listening tests.

Subject ——	Intelligibility Rate (%)			
	Bathroom	Bedroom	Foyer	-
1	86	83	69	
2	97	86	72	
3	97	79	79	
Average	93.3	82.7	73.3	

4. Summary and Conclusions

It is well known that room acoustics significantly influence speech intelligibility. In commercial auditoriums, the acoustics are studied by professionals using elaborate techniques. However, the expense and complexity of such methods preclude their routine use for spaces such as regular classrooms and residential rooms. This research explores simple but reliable methods to measure some principal acoustical parameters using equipment that is easily available (just a pair of laptops) and free software. The setup and procedures of the experiments are straightforward, as is their analysis and interpretation. Therefore, individuals without specialized training should be able to conduct such measurements.

Many acoustical studies only focus on the reverberation time RT60. However, as was shown here, the strength of the reverberation I_R can enhance speech intelligibility. Thus placing reflective surfaces close to the human speaker (*i.e.*, in the stage area) and more absorption elsewhere can help achieve the best speech intelligibility by allowing the sound to be louder at the listener location but not linger. The methods presented here can help to optimize this balance between I_R and RT60. The results also show why a smaller classroom tends to be more effective, aside from the more favorable teacher-to-student ratio. Using a microphone and electronic amplification is of course another way to boost sound intensity at the listener position without increasing the decay time.

Future extensions of this work can include comparative measurements of various classrooms and auditoria in schools to assess variability of speech intelligibility with size, shape, and materials of the room. Another extension is to use a larger variety of phonemes recorded in different voices to present an additional challenge for the subjects of the listening tests.

Acknowledgements

I would like to acknowledge D. Fogerty and A. Jurgens.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Everest, F.A. and Pohlmann, K.C. (2015) Master Handbook of Acoustics. McGraw-Hill, New York.
- [2] Olive, S.E. and Toole, F.E. (1989) The Detection of Reflections in Typical Rooms. *Journal of the Audio Engineering Society*, **37**, 539-553.
- [3] Lochner, J.P.A. and Burger, J.F. (1958) The Subjective Masking of Short Time Delayed Echoes by Their Primary Sounds and Their Contribution to the Intelligibility of Speech. Acustica, 8, 1-10
- [4] Meyer, E. and Schodder, G.R. (1952) On the Influence of Reflected Sound on Directional Localization and Loudness of Speech. *Nachrichten der Akademie der Wis*senschaften in Gottingen. II. Mathematisch-Physikalische Klasse, Illa, 6, 31-42.

- [5] Beranek, L.L. (2006) Analysis of Sabine and Eyring Equations and Their Applicationto Concert Hall Audience and Chair Absorption. *The Journal of the Acoustical Society of America*, **120**, 1399-1410. <u>https://doi.org/10.1121/1.2221392</u>
- [6] Volander, M. and Bietz, H. (1994) Comparison of Methods for Measuring Reverberation Time. *Acta Acustica United with Acustica*, **80**, 205-215.
- [7] Vigran, T.E. and Sorsdal, S. (1976) Comparison of Methods for Measurement of Reverberation Time. *Journal of Sound and Vibration*, 48, 1-13. https://doi.org/10.1016/0022-460X(76)90366-7
- [8] Schroeder, M.R. (1964) New Method of Measuring Reverberation Time. *The Journal of the Acoustical Society of America*, **37**, 409-412. https://doi.org/10.1121/1.1909343
- [9] Mulcahy J. (2003) REW: Room Eq Wizard. Version 5.18. <u>https://www.roomeqwizard.com/index.html</u>
- [10] Díaz, C. and Pedrero, A. (2005) The Reverberation Time of Furnished Rooms. Applied Acoustics, 66, 945-956. https://doi.org/10.1016/j.apacoust.2004.12.002
- [11] Gelfand, S.A. and Silman, S. (1979) Effects of Small Room Reverberation on the Recognition of Some Consonant Features. *The Journal of the Acoustical Society of America*, 66, 22-29. https://doi.org/10.1121/1.383075
- [12] Knecht, H.A., Nelson, P.B., Whitelaw, G.M. and Feth, L.L. (2002) Background Noise Levels and Reverberation Times in Unoccupied Classrooms: Predictions and Measurements. *American Journal of Audiology*, **11**, 65-71. https://doi.org/10.1044/1059-0889(2002/009)