Effectiveness of acoustic treatments and PA redesign by means of student activity and speech levels



*Published Version:*
De Salvio, D., D'Orazio, D. (2022). Effectiveness of acoustic treatments and PA redesign by means of student activity and speech levels. APPLIED ACOUSTICS, 194, 1-12 [10.1016/j.apacoust.2022.108783].

*Availability:*
This version is available at: https://hdl.handle.net/11585/899842 since: 2024-09-12

*Published:*
DOI: http://doi.org/10.1016/j.apacoust.2022.108783

# Effectiveness of acoustic treatments and PA redesign by means of student activity and speech levels


Domenico De Salvio, Dario D'Orazio*

*Department of Industrial Engineering, University of Bologna*

*Viale del Risogimento 2, Bologna, 40136, Italy*



## Abstract

The assessment of the acoustic quality of learning spaces traditionally concerns objective parameters. The most recent standards consider the occupancy of students in the design process. However, the dynamical behaviour of students during lectures is not considered by these two approaches. The student activity (SA) - i.e. the noise generated by students during lessons - was proposed as a metric to assess this dynamical context. Moreover, teachers' speech level (SL) may depend on the acoustic properties of the room and speech reinforcement system (Public Address PA). In the present work, SA and SL measured in two historical university lecture halls are used to assess the quality of the acoustic environment before and after renovation works. Restoration includes both acoustic treatments and PA redesign. Both measurements were carried out in a pre-COVID19 scenario. Clustering techniques, Gaussian Mixture Model and K-means, were used to measure the student activity and the speech levels after long-term monitoring of active classrooms. Outcomes show lower mean levels of both SA and SL and lower signal-to-noise ratios, suggesting the achieving of quieter environments. Differences of behaviour by both students and teachers have been detected through Lombard slopes and correlation coefficients. These preliminary analyses suggest that clustering techniques seem to be a valid tool to assess acoustic quality in dynamic contexts.



*Keywords:* Classroom acoustics, Machine learning, Student activity, PA system, Gaussian Mixture Model, K-means clustering


## 1. Introduction

Well-being in learning spaces is a basic condition to provide an effective and successful learning process. Concerning this, acoustic comfort has a strong impact on the attention span of the students as well as on the vocal effort of


*Corresponding author

*Email address:* dario.dorazio@unibo.it (Dario D'Orazio)

the teachers [1, 2]. The most important aspect in educational spaces regards communication quality, thus speech intelligibility. The outline of this concern is assumed made up of two factors: the reverberation and the signal-to-noise ratio [3]. Intelligibility parameters can be related to the sound energy in the room, e.g. the early-to-late index $C_{50}$, or the modulation transfer function, e.g. the Speech Transmission Index STI [4]. Both kinds of metrics can consider the background noise to link the two factors of speech intelligibility. In this case, the early-to-late index $C_{50}$ is defined as $U_{50}$ [5].

However, the background noise is not only due to mechanical systems or traffic but there is a dynamic component as well, represented by the human noise. This includes both the activities carried out in neighbouring spaces and, mostly, the noise generated by the students attending lectures, called student activity [6]. Background noise leads the speakers to increase their voice because of the Lombard effect, as well as increases both energetic and informational masking [7]. The increasing of the vocal intensity of the teacher is affected by the room acoustics characteristics besides the background noise. In fact, the feedback given by the room amplification has an approximately linear relationship to the voice level. Hence the vocal support plays an essential role in the teachers' comfort [8]. Performance tasks of students are influenced too, having detrimental effects on reading comprehension and vocabulary-learning tasks [9].

In light of these considerations, occupancy plays a key role in the intelligibility issue. Student activity can be measured via clustering methods used in machine learning also, like Gaussian Mixture Model, besides the more typical equivalent and percentile levels [6, 10, 11, 12]. Further unsupervised clustering algorithm, specifically K-means clustering, was used to detect whether classrooms were occupied or empty [13]. Previous work conducted by the authors compared all these methods in three university lecture halls [14].

The enhancement of acoustic environments involves both objective and subjective aspects. Students prefer treated spaces, indeed [15]. Despite the subjective impression and the listening effort depend more on reverberation than intelligibility besides other factors like sentence complexity, age, and linguistic abilities, the location of treatments is fundamental to properly increase the acoustic quality of classrooms [16, 17, 18]. In fact, in spaces with high intelligibility scores, high comprehension by the students is achieved even for low signal-to-noise ratios [19]. Improvements are observed on the speech levels as well. Good acoustic conditions help the communication quality between teachers and students; however, a reinforcement system can improve the teachers' comfort, keeping down their voice levels whether or not they have a quiet voice [20]. A public address (PA) facilitates achieving high SNRs preventing the trigger of the Lombard effect [11]. This issue concerns most of all large lecture halls. Thus, a PA does not regard all seats in a large room but direct sound must be reached adequately up to the back [21, 22]. Nevertheless, intelligibility improvements are influenced by the adaptation of the listening effort and binaural processes, thus it is important to understand to what extent a PA helps, together with the acoustic characteristics of the room, the perceptual enhancement of the comfort of listeners [23, 24].

In this work, the student activity was measured before and after the acoustic restoration of two lecture halls. Both measurements were performed in a pre-COVID19 scenario. The outcomes are focused on detecting changes in students' and teachers' behaviour during lectures. Thus, the Lombard effect and correlation among occupancy and the levels of the identified sound sources have been analyzed. Further discussions have been carried out about the spectra of the sound sources. Finally, considerations regarding the variations of student activity and speech levels have been made through 15-minutes samples monitoring of each lesson.

## 2. Method

### *2.1. Lecture halls*

Two university lecture halls have been renovated to enhance the acoustical conditions. The renovation works have been concerning the acoustic treatments of the surfaces and the re-design of the PA system. Following, a brief description of the spaces under study:

- Hall I has a rectangular plan, wooden and terraced seats that return a typical amphitheatre shape of the space, reflective surfaces, and an articulated false ceiling. The volume is about 1000 $m^3$, and the maximum occupancy is about 250 students.

- Hall II has a rectangular plan, wooden and terraced seats, reflective surfaces, and a flat false ceiling. The volume is about 900 $m^3$, and the maximum occupancy is about 200 students.

The two halls are quite similar in geometry and material properties; and vary essentially by the shape of the false ceilings, besides an extra volume on the rear part of the room in Hall I.

### *2.2. Design of acoustic treatments*

The renovation works regarded the installation of acoustic treatments made by wooden absorber panels, placed to cover the rear wall and the overhanging beams of the halls. Moreover, the PA systems have been replaced with line arrays (L-Acoustic Syva) located behind the teachers' desk and supplementary loudspeakers (L-Acoustic 5XT) as fillers for the first and the last rows of the audience area. The high similarity of the halls allowed to design same treatments for both halls. Due to the different sensibility of the loudspeakers, the gain and the phases between line arrays and fillers were set to reach a homogeneous coverage of the direct sound in the seating area with the help of a numerical model (Odeon Room Acoustics v.15.0 and Soundvision v.3.5.1).

Impulse response measurements have been carried out before and after the restoration in both halls according to ISO 3382 in an unoccupied state. Monoaural receivers were used to acquire Exponential Sine Sweep – 512 K length, sampled at 48 kHz – signals sent from an omnidirectional source. The latter was a high SPL custom dodecahedron calibrated in a reverberation chamber according to ISO 3741. Two source positions, one on the axis in the middle of the room behind the desk, and the other asymmetrical near the desk were used. Receivers were located homogeneously in the seating area. The same source-receiver positions were used before and after the restoration. The reverberation time, early-to-late index, and the Speech Transmission Index before and after the treatments are shown in Table 1.

The acoustic characteristics of both halls are shown in Figure 1. Reverberation time $T_{30}$ and the sound strength $G$ before (black lines) and after (red lines) the treatments are plotted on top for Hall I and on the bottom for Hall II. The treatments had the best effectiveness on mid and high frequencies of reverberation time. The sound strength $G$ shows how the treatments affect the sound decay through space. PA system is a directional source; thus, the offset of the $G$ values is arbitrary and set to be comparable with the omnidirectional sources. Both halls show a faster decrease of SPLs after the renovation; this underlines the need for a PA system, with an optimized directivity, to spread the sound more homogeneous up to the bottom of the room. The reliability of the PA is shown with the red dashed lines.

### *2.3. Student activity and speech levels measurements*

The active lectures were measured through the student activity SA and the speech levels SL. Two sound level meters monitored 9 lessons before (lessons A-I) and 9 lessons after (lessons J-R) the renovation works. The method previously described in [14] was used (please, see previous work for further details).

Sound level meters were placed in the middle of the students' area. Thus, the source-receiver distance was about 10 m. The lessons were entirely recorded, then the pauses were cut in post-processing in order to analyze only the lecture time. Equivalent continuous sound levels were recorded every 0.1 s to ensure the recording of the pauses among syllables and words [25]. An operator attended lectures to check for possible unforeseen events and give broader context to the analysis. Indeed, it could be possible to have different kinds of lessons, depending on the involvement of the students in the discussion or on the use of multimedia files.

After the cut of the unwanted parts of the recordings in post-processing, the sound pressure levels SPLs of each sound level meter were processed via two unsupervised clustering techniques, the Gaussian Mixture Model and the K-means clustering. The first technique, the Gaussian Mixture Model (GMM), decomposes the probability distribution function of the recorded SPLs as a sum of Gaussian curves. In Fig. 2 an example of the recorded lecture is shown. The bars show the recorded occurrences and the solid curve of the obtained probability density function to process via clustering methods. Each SPL is assigned to each curve basing on the likelihood function. In this work, each sound source, the student activity, and the speech level are associated with each mean obtained by the algorithm. The second technique, the K-means clustering, assigns each recorded SPL basing on the minimization of the geometrical distance among them. Each centre of gravity of the obtained clusters is called "centroid". Similar to the means for GMM, each centroid is associated with each sound source.

In order to run these algorithms, the number of clusters to obtain must be set. Since only two sound sources are expected, a fixed number of clusters k = 2 were set for both algorithms. This is confirmed by checking on the detectable peaks of the occurrences curves [6]. The lower values, i.e. the lower mean of the Gaussian mixture or the lower centroid of the clusters, were attributed to the student activity, whereas the higher values were to the speech level of the teachers.

It is worth noting that all lessons analyzed in this work were carried out through the PA system.

## 3. Results

The 18 lessons were recorded and analyzed via Gaussian Mixture Model, K-means clustering, percentile, and equivalent levels. Table 2 shows the outcomes of the clustering processes averaged over the two sound level meters. The lectures from A to I in the upper part of the Table refer to previous outcomes obtained in [14] and were measured before the restoration, whilst the lectures from J to R refer to the measurements carried out after the renovation. The mean values are shown at the end of each series of lectures, ante and post. The standard deviations of the outcomes obtained by the two receivers are shown in brackets. Other information in this Table is provided by the hall, the occupancy with the respective percentage over the maximum capacity, and the equivalent absorption area $A_{occ}$ taking into account the contribution of the students for each lesson.

The measured A-weighted SA and SL values before the treatments lie respectively in the range of 47.5 − 61 dB and 63.3 − 75.5 dB for GMM, 48.8 − 56.5 dB and 64.2 − 76 dB for KM, 45.8 − 61.6 dB and 64.8 − 79.2 for percentile and equivalent levels. The measured SA and SL levels after the treatments lie respectively in the ranges 47.2 − 53.9 dB and 59 − 72.1 dB for GMM, 49.7 − 54.1 dB and 61.2 − 72.7 dB for KM, 45.9 − 53.3 dB and 61.1 − 74.4 dB for percentile and equivalent levels. Before the restoration work, the standard deviations between the two receivers, respectively for SA and SL, lie in the ranges 0 − 2 and 0.8 − 4.6 dB for GMM, 0 − 1.9 and 0.8 − 4.5 dB for KM, 0.5 − 9.5 and 3.6 − 7.1 dB for percentile and equivalent levels. Concerning the measured s.d. of SA and SL after the treatments, values lie respectively in the ranges 0.3 − 3.1 and 0.1 − 1.9 dB for GMM, 0.2 − 1.5 and 0.1 − 2.1 dB for KM, 0.4 − 1.1 and 0 − 3 dB for percentile and equivalent levels. The measured A-weighted mean values of SA and SL and their standard deviations in brackets before the treatments are respectively 52.1 (1.0) and 68.3 (2.6) dB for GMM, 53.3 (0.8) and 69.6 (2.5) dB for KM, 53.1 (4.4) and 72.2 (4.7) dB for percentile and equivalent levels. The same parameters measured after the treatments are respectively 50.8 (1.0) and 65.6 (0.8) dB for GMM, 51.7 (0.8) and 67.6 (1.1) dB for KM, 50 (0.7) and 67.8 (1.1) for percentile and equivalent levels. The measured values in the present work are both higher if compared to SA and SL values measured by Hodgson [6] in

American university classrooms and lecture halls were respectively in the range 30 – 50.2 dB and 43 – 59 with average values of 41.9 and 50.8 dB. Choi [12] measured university active classroom in Korea with average values of noise – SA was not detected – and SL respectively of 43.7 and 51.4. It is worth noting that, unlike the cited works and as stated above, all lectures were carried out with a PA system.

The occupancy is similar before and after the treatments, considering the weight of outliers. Before the treatments, lesson E and lesson I were attended by respectively 250 and 200 people, whilst lesson M and lesson R, measured after the works, were attended by 80 and 95 people.

## 4. Discussions

### *4.1. Signal-to-noise ratio and assessments about the treatments*

The intelligibility issue is outlined by two main factors, as stated above: the reverberation time and the signal-to-noise ratio. During lectures both of these factors change, the first according to the occupancy, the second as a function of SA and SL. The signal-to-noise ratio SNR can be defined as the difference between SA and SL, indeed. Whereas the occupancy can be determined by counting the number of students, the detection of SA and SL is more complicated. Table 2 shows the outcomes obtained via GMM, KM, percentile, and equivalent continuous levels.

Comparing the techniques, the mean values of SA and SL decreased after the treatments by all methods. It could means that a quieter environment has been achieved, particularly for teachers who show an average reduction of about 2.3 dB for the unsupervised methods, GMM and KM, and about 4.4 dB for the equivalent continuous levels and the 90th percentiles. The student activity has decreased as well, respectively an average of about 1.4 for the unsupervised methods and 3.1 dB for the 90th percentile level. Moving the analysis among methods after the restoration works (lessons from J to R), it is worth noting that the KM gives back higher values of SA with respect to the other in most lectures whilst the $L_{90}$ the lower values. Regarding the SL, the $L_{eq}$ returns the highest values in the greater part of lectures whereas the GMM is the lowest in all cases. Thus, the comparison among techniques seems to confirm the results obtained in previous work [14]. Percentile and equivalent continuous levels seem to overestimate SL whereas returns the lowest values for SA after the works. This could mean that the human chatter does not fit with the 90th percentile of exceeded time. Differences between GMM and KM are due to the heteroscedasticity of the measured values, i.e. the different variances among data return different clustering outcomes [26].

Further considerations are needed about the initial hypotheses of the two algorithms. As pointed out in [27], the mean difference between GMM and KM regards the cluster distribution of data. In GMM, a single data point can belong to more than one cluster with an assigned probability whereas in KM this is not possible. In this latter technique single data point can be assigned only to one cluster. The ability to assign one point to one or more clusters is the difference between hard and soft clustering [28].

To better explain to what extent the heteroscedasticity affects the results, it is useful to visualize two different cases. In Fig. 3a and 3c, the recorded probability density functions - solid line - and the two Gaussian components - dotted for SA and dashed for SL - are shown on the left for two lessons (G and D). The data points belonging to fuzzy borders, i.e. the data in common between the two clusters weighted by an assigned probability, are highlighted with the dashed area. The latter is a function of the number of clusters and the variances of the Gaussian curves, thus the larger the variance, the larger the overlap area. Looking at the shape of the SA curve, it can be seen how large its variance is unlike the SL curve, and thus the effect of heteroscedasticity on the mean values. On the right, in Fig. 3b and 3d, the same lessons are plotted post-processed via KM. The clusters associated with SA are in dark grey, whereas the ones associated with SL are in light grey. Here, the distinction of the clusters is sharp, each data point belongs to one and only one cluster. The consequences of the fuzzy borders and the heteroscedasticity of the clusters are particularly evident in lesson G, which returns a difference of SA of 5.3 dB between GMM and KM. Conversely, it is worth noting that the difference of SL of the same lesson for both GMM and KM is 0.6 dB. Concerning lesson D, the difference of SA calculated via GMM and KM is 3.1 dB, whereas is 1.8 dB regarding SL.

It is worth noting that the means of the s.d. are quite different, deeply for $L_{90}$ and $L_{eq}$ which show a reduction of 3.7 dB of student activity and 3.6 dB of speech level. The mean obtained by the unsupervised methods measured lower decreases but just for the speech level, respectively of 1.8 for GMM and 1.4 for KM. However, looking at the single lessons, it is evident how the s.d. of the student activity are lower after the treatments; thus the lack of difference of the mean values is due to the outliers. The decrease of s.d. may suggest that the sound field after the renovations is more diffuse and more homogeneous coverage is achieved with the redesign of PA.

Besides the values obtained for each lesson and the means, the correlations between student activity and speech levels were plotted to evaluate changes in the Lombard effect. The relationship between the student activity and the speech levels is shown in Figure 4. The outcomes obtained before the restoration are plotted in black, whereas the others are in red. Regression lines show the tendencies of the results. As stated above, both SA and SL have been decreased but not proportionally. As a consequence, in Fig. 4 the change of slope of the relationship between the sound sources can be noticed. This could mean that the vocal effort of teachers is less affected by the babble of the students and thus by the Lombard effect.

### *4.2. Effects of occupancy*

The occupancy plays a key role in the dynamical context of the acoustics of a lesson. While this parameter (occupancy) is fixed in each room in general school grades, it continuously changes during university lectures. Equivalent absorption area increases and decreases within university lecture halls. In fact, the absorption area of the occupied room is one of the main features affecting all the parameters proposed by Hodgson et al in their predictive model [6]. As highlighted by Choi, the effect of the students depends on the acoustical characteristics of the room and their distribution through space [21, 29].

In Figure 5, the relationship of the occupancy and signal-to-noise ratio, student activity, and speech levels was investigated. As stated above, the before and after states have similar occupancies despite the outliers. Black and red lines represent respectively the relationship before and after the restoration in all plots. In the left part, the relationship between occupancy and the signal-to-noise ratio is shown. The enhancement of the acoustic conditions of the halls seems to make the correlation more sensitive. SNR increases linearly with the occupancy, indeed. This could mean that the bigger the audience, the quieter the environment. Being the SNR the difference between SA and SL, following analyses concerning the relationship between SA, SL, and occupancy allow to deepen this outcome. It is well known that a SNR equal or greater to 15 dB does not affect the intelligibility scores and values around 20 dB are considered as “ideal” targets in classrooms [30, 4, 31]. In this work, the occupancy of about 120 students seems to be a threshold for the behaviour of listeners. In fact, crowded lectures seem to trigger a psychoacoustical effect which leads the students to achieve the best SNR without affecting the intelligibility. Lower occupancies keep the SNR lower than 15 dB whilst higher occupancies reach SNR values greater than 15 dB. In correspondence with the occupancy of about 120 people, the records are more variable, and the measured SNRs span from about 9 to 20 dB. None of the recorded lessons exceed the value of 20 dB, as expected for the reasons stated above. Results suggest that in large lecture halls, despite the optimal acoustic characteristics, a PA is necessary to achieve a SNR of +15 dB without affecting the vocal effort of the teachers.

In the middle of Fig. 5, the graph shows the correlation between occupancy and student activity. It is possible to notice how, after the acoustic treatments, the SA seems to keep a constant tendency regardless of the number of students attending the lectures. It could mean that the SA is independent of the occupancy. This is particularly true for GMM, less for KM which preserves a descending tendency as measured before the treatment, even though with different slopes. This result could mean that the treatment has been particularly effective on the listening effort, in fact before the works, crowded lectures measured lower values of SA.

The right part of Fig. 5 shows the relationship between occupancy and speech levels. In this correlation, the acoustical treatment seems to have the most important effect. Tendencies changed from descending to ascending, indeed. This means that teachers tend to increase their voice levels with rising occupancy but it could not be strictly related to the vocal effort since all lectures were carried out with a PA system. The reasons to explain this behaviour could be multiple. High occupancies before the restoration helped to reach lower reverberation, more similar to values obtained after the treatment. Furthermore, black SA tendencies decrease with respect to the occupancy thus, before works, combining reverberation and noise conditions, the more crowded the quieter the environment. Black and red lines cross each other in correspondence with an occupancy of about 150 students. The main differences between before and after states are noticeable below this value. Red tendencies change slopes mainly because lower SL values were measured with half-empty halls.

### *4.3. Spectral analysis*

Spectral analysis can prove whether the clustering works similarly to the expected speech signal recorded during lectures. In the present work, the aid of the PA to the teacher's voice turns out to be very useful to assess the spectral distribution of the clusters. Indeed, it is expected to have an SL spectrum very similar to the reference standards [32, 33, 34]. The clustering was broadened over the octave bands from 125 to 4000 Hz. In Figure 6 the average relative spectra of each lesson are shown. On the left, the plot presents the outcomes of SA (dashed line) and SL (solid line) obtained before the renovation, while on the right, there are the results obtained after the works. Relative spectra were evaluated setting the 1 kHz octave band equal to 0. Values are averaged over all measured lectures and shown both before and after the acoustic treatments of the halls. The shape are clearly attributable to speech sources. In fact, being produced by a single source, SL has a sharp tendency in agreement with [6], whereas SA is slightly different, as expected, since it is more diffuse and affected by the noise which modify the shape, especially in the low frequencies where the greatest uncertainties are [35, 36, 37].

Different outcomes concern the spectra obtained after the treatments, which result more flattened even for SL, as shown in Fig. 6b. It is worth recalling that treatments regarded the redesign of the PA besides the surfaces, thus SL seems to be deeply affected by the equalization of the new loudspeakers. But the most interesting result concerns the shape of SA from the middle to the high frequencies 1-4 kHz. In fact, despite the differences underlined above about the PA, the shapes of the spectra from 125 up to 500 Hz are quite similar before and after the treatments. From 1 kHz up, the behaviour of SA is not as expected since it cannot be influenced by the PA like SL. In this frequencies, the clustering seems to be less reliable. The reasons why this happened could be hypothesized on multiple levels. On the side of the algorithms, if the peaks of the occurrences curves are not so clear, then it may be difficult to characterize the difference between the two sources. On the side of the renovation works, the treatments regarded mainly the absorption of mid and high frequencies, thus il could be possible that the formants of the speech are more affected losing more energy by the treatments than the fundamentals. On the side of the behaviour of the students, it has been noticed in previous sections how quieter environments have been achieved, thus it could mean that the detection of SA is more difficult.

### *4.4. Variations of SA and SL during lectures*

In order to increase the statistical significance of the study, a power analysis was conducted. First, each lesson has been analyzed by slots of 15 minutes increasing the sample size up to 45. Then, the huge effect size of about 3.8

calculated for the populations of student activity and speech levels allows the analysis to reach a significance level of $p < 0.001$ and a statistical power of 100%.

Besides the statistical significance, the data sample augmentation leads to analyze the temporal fluctuations of the SNR during lessons. Irrelevant speech noise, like student activity, can affect speech intelligibility with informational or energetic masking [38, 39]. Thus, it is important to consider to what extent the SNR varies during lessons [40].

The temporal tendencies of each lecture before and after the treatments for GMM and KM are shown in Fig. 7. SL are plotted with solid lines, whereas SA with dashed lines. The results between the methods seem to be quite coherent and with similar tendencies. Few exceptions are represented by lesson G before the treatments and lesson O after the restoration. Differences in this kind of analysis are strictly related to the shape of the occurrences curve and the consequent variance of data, as seen in the previous section concerning Fig. 3. In these two lessons, the occurrences curves are particularly skewed, thus the SA curve has a high variance and the heteroscedasticity is broadly pronounced. The small differences obtained for lesson O in Table 2 can be due to the average calculated on the whole recording and the smaller standard deviation measured between the two sound level meters after the acoustic treatments. Despite both techniques produce similar results, the KM seems to be less sensitive to oscillations. This stability may be due to the non-fuzzy borders of KM clusters.

Correlation coefficients before and after the renovation have been evaluated between student activity and speech levels for both unsupervised methods besides occupancy and equivalent absorption area in occupied state. In Table 3 the correlation matrix is shown. Before and after correlations are indicated respectively in regular and bold. The first interesting results concern the decrease of the correlation among SA and SL for both techniques GMM and KM. This is particularly evident for KM which lowers the correlation from 0.71 to 0.29 whereas for GMM from 0.85 to 0.61. The drop of the coefficient states that SA and SL keep on having a growing and related tendency but weaker, with scattered data. The matrix highlights as GMM and KM compute differently SA. The acoustic treatments affect deeply the anti-correlation between the equivalent absorption area and SA which is completely lost for GMM and strengthen for KM. Regarding SL, despite its weak correlation with the equivalent absorption area, it is worth noting how the regression slope changes for both GMM and KM as seen in Fig. 5. More in general, the equivalent absorption area loses its anti-correlation with almost all parameters except for SA calculated via KM.

## 5. Conclusions

The effectiveness of acoustic treatments and audio redesign has been assessed by the measurement of student activity and speech levels during lectures. Eighteen lessons were monitored through two sound level meters in two halls before and after renovation. Then, the student activity SA and the speech levels SL were analyzed through two clustering algorithms - Gaussian Mixture Model and K-means -, broadly used in machine learning to find patterns among data. Results show a decrease in the mean values of both SA and SL and the standard deviations between the two receivers. Thus, quieter environments and more diffuse sound fields have been achieved. Moreover, the relationships between the student activity and the speech levels highlight changes in the Lombard effect. A drop in the slopes of tendencies between the two parameters has been detected after the acoustic treatments. Further relationships have been explored among the occupancy, the student activity, the speech levels, and the signal-to-noise ratio to investigate the role of occupancy during lectures. Outcomes reveal different behaviours ors by both students and teachers during lessons after the restoration. The signal-to-noise ratio, i.e. the difference between SL and SA, seems to be more sensitive to the occupancy, despite the student activity seems to be quite constant aside from the number of students attending lectures. Thus, it has been shown how the increase of the signal-to-noise ratio depends on the increase of the speech levels with the occupancy. Before the treatments, both student activity and speech levels tended to decrease with the increase of occupancy. Increasing the statistical power of the discussions, the comparison between the two clustering techniques shows that the Gaussian Mixture Model is more sensitive to oscillation than K-means clustering. The correlation matrix among all the parameters taken into account in this work shows how the treatments affect the dependency among student activity, speech levels, and occupancy. Lastly, the average relative spectra evaluated before and after the treatments show the reliability of the clustering methods. Weak points highlight the need of digging into these techniques to broaden the available features and make the use of these algorithms more robust.

## References


[1] J. Brunskog, A. C. Gade, G. P. Bellester, L. R. Calbo, Increase in voice level and speaker comfort in lecture rooms, The Journal of the Acoustical Society of America 125 (4) (2009) 2072–2082.

[2] M. Cipriano, A. Astolfi, D. Pelegrín-García, Combined effect of noise and room acoustics on vocal effort in simulated classrooms, The Journal of the Acoustical Society of America 141 (1) (2017) EL51–EL56.

[3] M. Hodgson, Experimental investigation of the acoustical characteristics of university classrooms, The Journal of the Acoustical Society of America 106 (4) (1999) 1810–1819.

[4] S. R. Bistafa, J. S. Bradley, Reverberation time and maximum background-noise level for classrooms from a comparative study of speech intelligibility metrics, The Journal of the Acoustical Society of America 107 (2) (2000) 861–875.

[5] L. Nijs, M. Rychtáriková, Calculating the optimum reverberation time and absorption coefficient for good speech intelligibility in classroom design using u50, Acta Acustica united with Acustica 97 (1) (2011) 93–102.

[6] M. Hodgson, R. Rempel, S. Kennedy, Measurement and prediction of typical speech and background noise levels in univiersity classrooms during lectures, The Journal of the Acoustical Society of America 105 (1) (1999) 226–233.

[7] V. Best, C. Conroy, G. Kidd Jr, Can background noise increase the informational masking in a speech mixture?, The Journal of the Acoustical Society of America 147 (2) (2020) EL144–EL150.

[8] D. Pelegrín-García, B. Smits, J. Brunskog, C.-H. Jeong, Vocal effort with changing talker-to-listener distance in different acoustic environments, The Journal of the Acoustical Society of America 129 (4) (2011) 1981–1990.

[9] D. Connolly, J. Dockrell, B. Shield, R. Conetta, C. Mydlarz, T. Cox, The effects of classroom noise on the reading comprehension of adolescents, The journal of the Acoustical Society of America 145 (1) (2019) 372–381.

[10] H. Sato, J. S. Bradley, Evaluation of acoustical conditions for speech communication in working elementary school classrooms, The Journal of the Acoustical Society of America 123 (4) (2008) 2064–2077.

[11] J. Peng, H. Zhang, D. Wang, Measurement and analysis of teaching and background noise level in classrooms of chinese elementary schools, Applied Acoustics 131 (2018) 1–4.

[12] Y.-J. Choi, Evaluation of acoustical conditions for speech communication in active university classrooms, Applied Acoustics 159 (2020) 107089.

[13] L. C. Brill, L. M. Wang, Comparison of occupied and unoccupied noise levels in k-12 classrooms, The Journal of the Acoustical Society of America 139 (4) (2016) 1979–1979.

[14] D. D'Orazio, D. De Salvio, L. Anderlucci, M. Garai, Measuring the speech level and the student activity in lecture halls: Visual-vs blind-segmentation methods, Applied Acoustics 169 (2020) 107448.

[15] J. Peng, S.-K. Lau, Y. Zhao, Comparative study of acoustical indices and speech perception of students in two primary school classrooms with an acoustical treatment, Applied Acoustics 164 (2020) 107297.

[16] S. R. Bistafa, J. S. Bradley, Predicting speech metrics in a simulated classroom with varied sound absorption, The Journal of the Acoustical Society of America 109 (4) (2001) 1474–1482.

[17] C. Visentin, N. Prodi, F. Cappelletti, S. Torresin, A. Gasparella, Using listening effort assessment in the acoustical design of rooms for speech, Building and Environment 136 (2018) 38–53.

[18] N. Prodi, C. Visentin, E. Borella, I. Mammarella, A. Di Domenico, Using speech comprehension to qualify communication in classrooms: Influence of listening condition, task complexity and students' age and linguistic abilities, Applied Acoustics 182 (2021) 108239.

[19] Y.-J. Choi, The intelligibility of speech in university classrooms during lectures, Applied Acoustics 162 (2020) 107211.

[20] S. M. Kennedy, M. Hodgson, L. D. Edgett, N. Lamb, R. Rempel, Subjective assessment of listening environments in university classrooms: Perceptions of students, The Journal of the Acoustical Society of America 119 (1) (2006) 299–309.

[21] Y.-J. Choi, Effect of occupancy on acoustical conditions in university classrooms, Applied Acoustics 114 (2016) 36–43.

[22] M. Hodgson, Case-study evaluations of the acoustical designs of renovated university classrooms, Applied Acoustics 65 (1) (2004) 69–89.

[23] G. M. Ellis, P. Zahorik, A dissociation between speech understanding and perceived reverberation, Hearing research 379 (2019) 52–58.

[24] P. Zahorik, Adaptation to Room Acoustics and Its Effect on Speech Understanding, Universitätsbibliothek der RWTH Aachen, 2019.

[25] D. Wang, S. S. Narayanan, Robust speech rate estimation for spontaneous speech, IEEE Transactions on Audio, Speech, and Language Processing 15 (8) (2007) 2190–2201.

[26] D. Barber, Bayesian reasoning and machine learning, Cambridge University Press, 2012.

[27] D. De Salvio, D. D'Orazio, M. Garai, Unsupervised analysis of background noise sources in active offices, The Journal of the Acoustical Society of America 149 (6) (2021) 4049–4060.

[28] M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot, C.-A. Deledalle, Machine learning in acoustics: Theory and applications, The Journal of the Acoustical Society of America 146 (5) (2019) 3590–3628.

[29] Y.-J. Choi, Effects of the distribution of occupants in partially occupied classrooms, Applied Acoustics 140 (2018) 1–12.

[30] J. S. Bradley, Predictors of speech intelligibility in rooms, The Journal of the Acoustical Society of America 80 (3) (1986) 837–845.

[31] M. Hodgson, E.-M. Nosal, Effect of noise and occupancy on optimal reverberation times for speech intelligibility in classrooms, The Journal of the Acoustical Society of America 111 (2) (2002) 931–939.

[32] IEC, 60268-16: Sound system equipment-part 16: Objective rating of speech intelligibility by speech transmission index, International Electrotechnical Commission, Switzerland.

[33] ANSI, 3.5,(1997), Methods for the Calculation of the Speech Intelligibility index (SII).

[34] ISO, 3382-3: Acoustics measurement of room acoustic parameters. part 3: Open plan offices, Geneva, Switzerland.

[35] G. Leembruggen, J. Verhave, S. Feistel, L. Holtzem, P. Mapp, H. Sato, T. Steinbrecher, S. Van Wijngaarden, The effect on sti results of changes to the male test-signal spectrum, Proc. IOA 38 (2016) 78–87.

[36] J. H. Rindel, C. L. Christensen, A. C. Gade, Dynamic sound source for simulating the lombard effect in room acoustic modeling software, in: INTER-NOISE and NOISE-CON Congress and Conference Proceedings, Vol. 2012, Institute of Noise Control Engineering, 2012, pp. 954–966.

[37] BB93, Guidance on computer prediction models to calculate the speech transmission index for bb93. version 1.0, Building Bulletin 93.

[38] A. W. Bronkhorst, The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions, Acta Acustica united with Acustica 86 (1) (2000) 117–128.

[39] G. E. Puglisi, A. Warzybok, A. Astolfi, B. Kollmeier, Effect of reverberation and noise type on speech intelligibility in real complex acoustic scenarios, Building and Environment (2021) 108137.

[40] B. Shield, J. E. Dockrell, External and internal noise surveys of london primary schools, The Journal of the Acoustical Society of America 115 (2) (2004) 730–738.

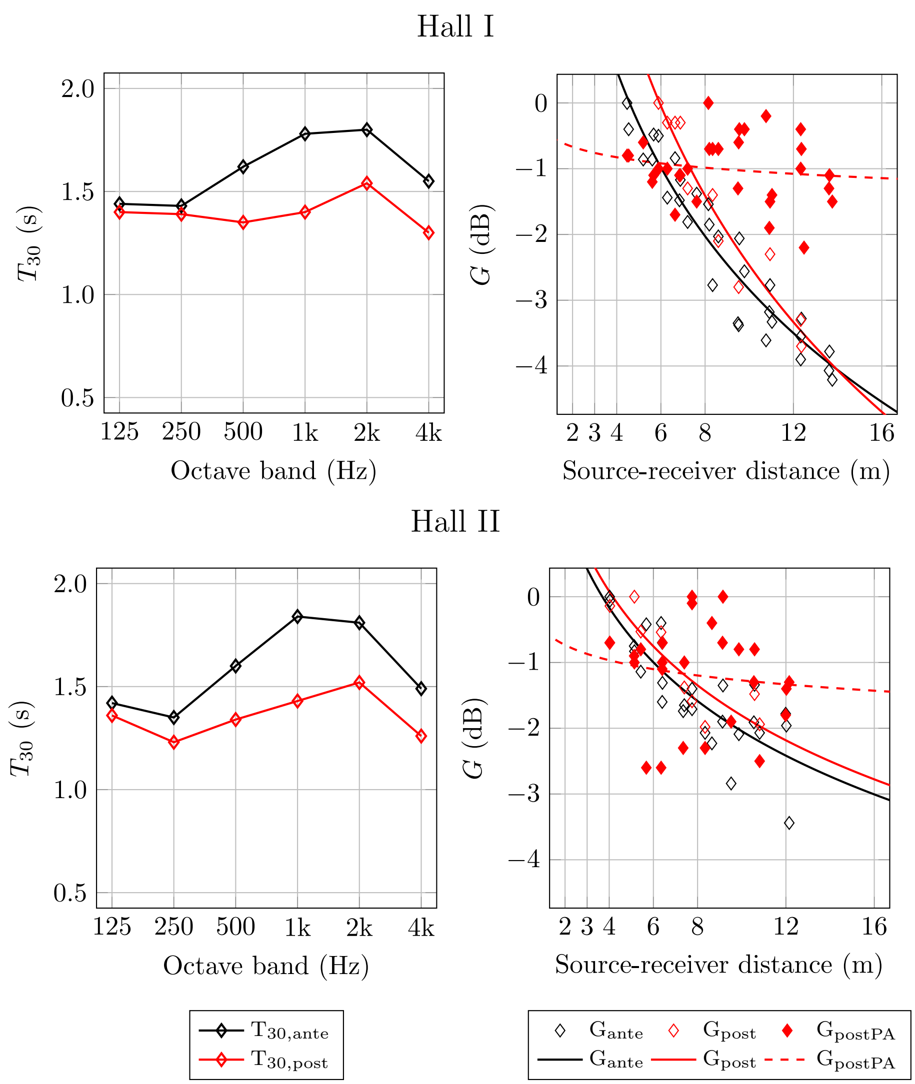



*Figure 1: Room acoustic properties of the two halls before and after the renovation works, respectively the Hall I is plotted on top whilst the Hall II on bottom. On the left the reverberation time as function of frequency octave band. On the right the sound strength G as function of the source-receiver distance. Values of* $G_{PA}$ *have to be taken only qualitatively because they do not refer to an omnidirectional source.*

*Table 1: General and acoustic data of the halls under study before and after the restoration, respectively indicated as "ante" and "post". Besides the shape of the inner space, it is shown the volume "V", the maximum occupancy "N", the reverberation time in unoccupied state "T", the early-to-late index "*$C_{50}$*", the Speech Transmission Index STI and the equivalent absorption area* $A_0$ *of the lecture halls in unoccupied state. The subscript "M" states a value averaged over all the receivers in the octave bands of* $500-1000$ *Hz, whereas "3" over the octave bands of* $500-2000$ *Hz.*

| Hall | Shape | Volume ($m^3$) | Occupancy | $T_M$ (s) | | $C_{50,3}$ (dB) | | STI | | $A_0$ ($m^2$) | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | V | N | Ante | Post | Ante | Post | Ante | Post | Ante | Post |
| I | Amphitheater | 1000 | 250 | 1.70 | 1.37 | -2.8 | -1.4 | 0.49 | 0.52 | 94 | 117 |
| II | Amphitheater | 900 | 200 | 1.72 | 1.38 | -2.4 | -1.0 | 0.47 | 0.54 | 84 | 105 |

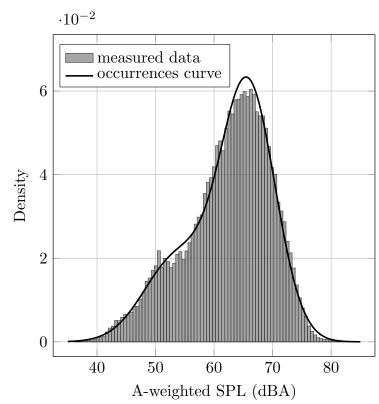
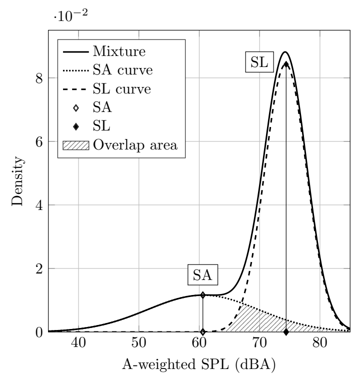


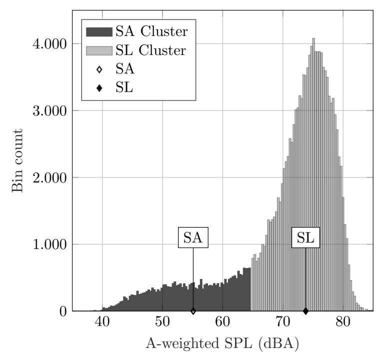

*Figure 2: Example of the occurrences curve obtained by a recorded lesson. The histogram shows the normalized bin count of the A-weighted recorded SPLs, whereas the curve indicates the probability distribution function to process via clustering techniques.*

*Table 2: Overview of the recorded lessons. For each lesson, the corresponding room, the number of people, the percentage of occupancy, and the equivalent absorption area taking into account the contribution of the people are shown. Measured A-weighted values of student activity SA, received speech level SL extracted through, Gaussian Mixture Model GMM, K-means clustering KM, equivalent continuous* $L_{eq}$ *and percentile levels* $L_{90}$ *are reported. Values are averaged over the two receiver positions. All values of SA and SL are in dBA.*

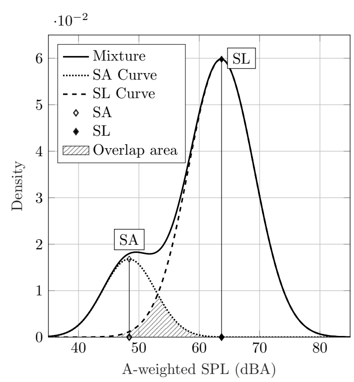
| Lesson | Hall | Occupancy (%) | $A_{occ}$ ($m^2$) | GMM | | KM | | $L_{90}$ | $L_{eq}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | | | SA (s.d.) | SL (s.d.) | SA (s.d.) | SL (s.d.) | (s.d.) | (s.d.) |
| A | I (Ante) | 145 (60%) | 156 | 48.2 (1.2) | 65.0 (4.0) | 52.2 (1.2) | 68.3 (4.0) | 48.0 (0.5) | 69.9 (4.8) |
| B | I (Ante) | 200 (80%) | 179 | 47.5 (1.5) | 63.3 (4.6) | 48.8 (1.3) | 64.2 (4.5) | 45.8 (0.6) | 64.8 (4.9) |
| C | I (Ante) | 100 (50%) | 137 | 53.3 (1.8) | 66.3 (4.1) | 55.8 (1.9) | 68.4 (3.9) | 53.0 (1.7) | 68.9 (4.5) |
| D | I (Ante) | 150 (60%) | 158 | 51.2 (2.0) | 67.2 (4.5) | 52.7 (1.9) | 68.4 (4.4) | 47.6 (1.3) | 69.7 (4.6) |
| E | II (Ante) | 250 (125%) | 190 | 48.4 (0.3) | 67.5 (1.5) | 49.1 (0.1) | 68.0 (1.6) | 52.1 (9.5) | 72.3 (7.1) |
| F | II (Ante) | 160 (80%) | 152 | 50.3 (1.5) | 66.5 (1.5) | 53.1 (0.2) | 68.5 (1.0) | 55.0 (8.1) | 71.9 (4.0) |
| G | II (Ante) | 120 (60%) | 135 | 61.0 (0.6) | 75.5 (1.4) | 55.7 (0.7) | 74.9 (1.4) | 61.6 (5.4) | 78.6 (5.3) |
| H | II (Ante) | 150 (75%) | 164 | 55.3 (0.1) | 75.3 (0.8) | 55.8 (0.0) | 76.0 (0.8) | 56.4 (7.0) | 79.2 (3.6) |
| I | II (Ante) | 200 (100%) | 188 | 53.4 (0.0) | 68.0 (1.0) | 56.5 (0.3) | 69.7 (0.8) | 58.8 (5.9) | 74.6 (3.7) |
| Mean | Ante | 164 | 162 | 52.1 (1.0) | 68.3 (2.6) | 53.3 (0.8) | 69.6 (2.5) | 53.1 (4.4) | 72.2 (4.7) |
| J | I (Post) | 130 (50%) | 172 | 51.3 (1.1) | 64.6 (0.7) | 52.7 (0.7) | 66.1 (0.5) | 49.7 (0.8) | 66.5 (0.4) |
| K | I (Post) | 185 (75%) | 195 | 49.9 (0.4) | 69.6 (0.4) | 50.9 (0.4) | 70.2 (0.3) | 48.1 (0.5) | 71.5 (0.1) |
| L | I (Post) | 130 (50%) | 172 | 47.2 (0.5) | 64.6 (0.5) | 50.9 (0.6) | 70.2 (0.5) | 45.9 (0.4) | 65.8 (0.4) |
| M | I (Post) | 80 (30%) | 151 | 49.9 (0.4) | 59.0 (0.3) | 51.0 (0.4) | 61.2 (0.2) | 48.3 (0.5) | 61.1 (0.0) |
| N | I (Post) | 190 (75%) | 197 | 52.7 (0.4) | 68.9 (0.1) | 51.0 (0.2) | 69.0 (0.1) | 48.8 (0.5) | 70.3 (0.1) |
| O | II (Post) | 110 (55%) | 168 | 53.9 (3.1) | 72.1 (1.1) | 54.1 (1.3) | 72.7 (2.0) | 53.3 (1.0) | 74.4 (2.4) |
| P | II (Post) | 125 (65%) | 175 | 51.4 (0.8) | 60.8 (1.3) | 50.4 (0.9) | 62.0 (2.1) | 46.7 (0.7) | 62.0 (1.3) |
| Q | II (Post) | 120 (60%) | 173 | 48.8 (0.3) | 66.4 (1.1) | 49.7 (0.9) | 69.8 (2.1) | 58.4 (0.5) | 71.6 (3.0) |
| R | II (Post) | 95 (50%) | 161 | 52.1 (1.8) | 64.9 (1.9) | 54.4 (1.5) | 67.0 (2.1) | 51.4 (1.1) | 67.5 (2.3) |
| Mean | Post | 129 | 174 | 50.8 (1.0) | 65.6 (0.8) | 51.7 (0.8) | 67.6 (1.1) | 50.0 (0.7) | 67.8 (1.1) |



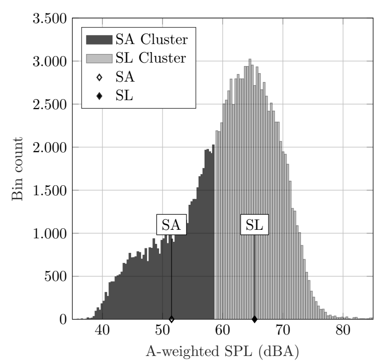

(a) Lesson G - GMM




(b) Lesson G - KM




(c) Lesson D - GMM




(d) Lesson D - KM

Figure 3: *Lessons G (on top) and lesson D (on bottom). On the left, the Gaussian mixtures and the respective components are shown. The solid lines indicate the probability density function recorded during the lecture. The dotted and the dashed lines show respectively the Gaussian curves associated to the SA and SL. On the right, the recorded SPLs are shown as function of their occurrences distribution. The SA and SL clusters are shown respectively in dark and light grey. Markers indicate the mean values (for GMM) and the centroids (for KM) which identify the SPL of each sound source.*

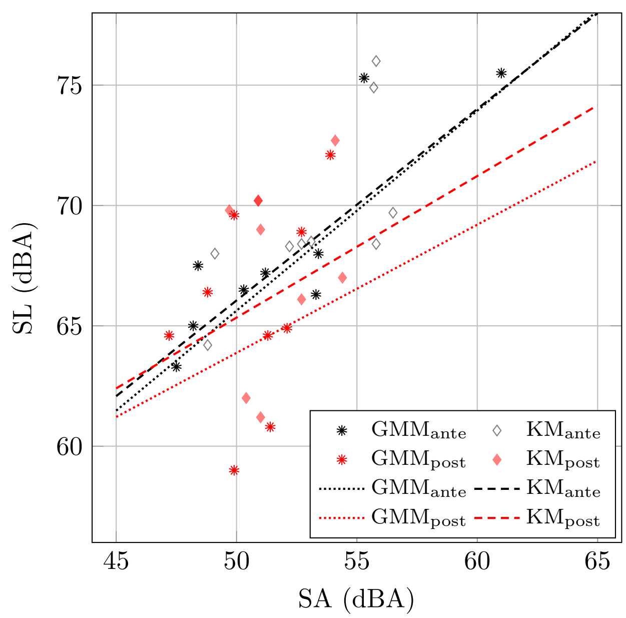



*Figure 4: Relationship between SA and SL measured values via GMM and KM. Black and red markers indicate respectively before and after acoustic treatments. Each marker refers to a whole lesson.*

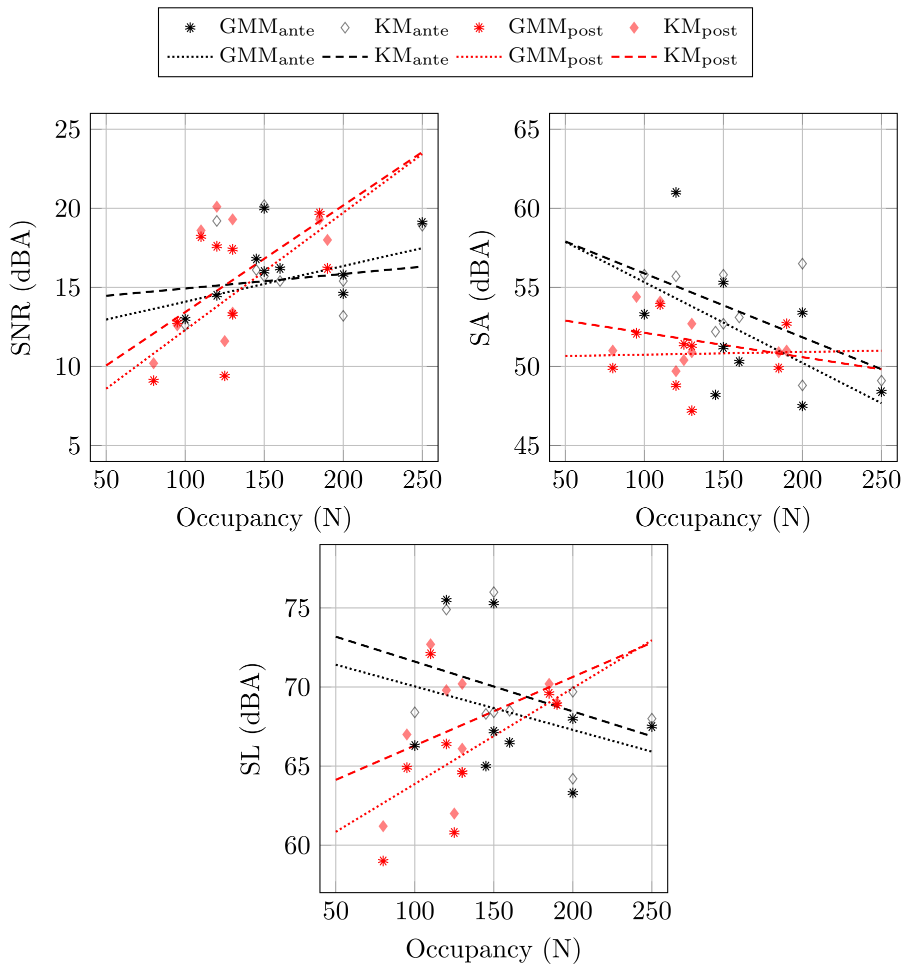
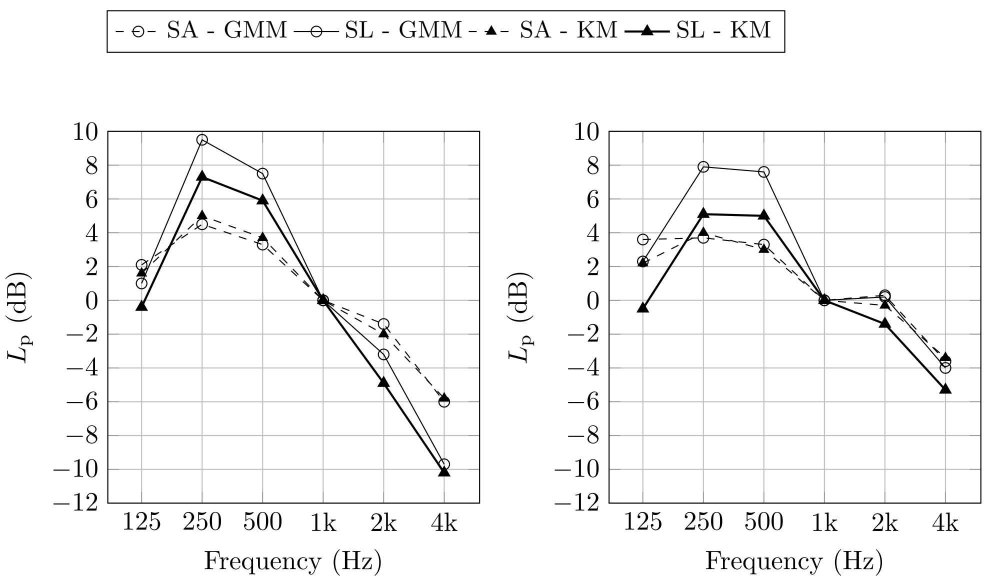



*Figure 5: Relationship between Occupancy and SNR, SA and SL measured values via GMM and KM. Black and red markers indicate respectively before and after acoustic treatments. Each marker refers to a whole lesson.*




(a) Ante - relative spectra of SA and SL

(b) Post - relative spectra of SA and SL

*Figure 6: Average relative spectra of student activity SA and speech levels SL obtained via GMM and KM. On the left SA and SL obtained before the acoustic treatments of the halls are shown, on the right the after state. Values are averaged over all measured lectures.*

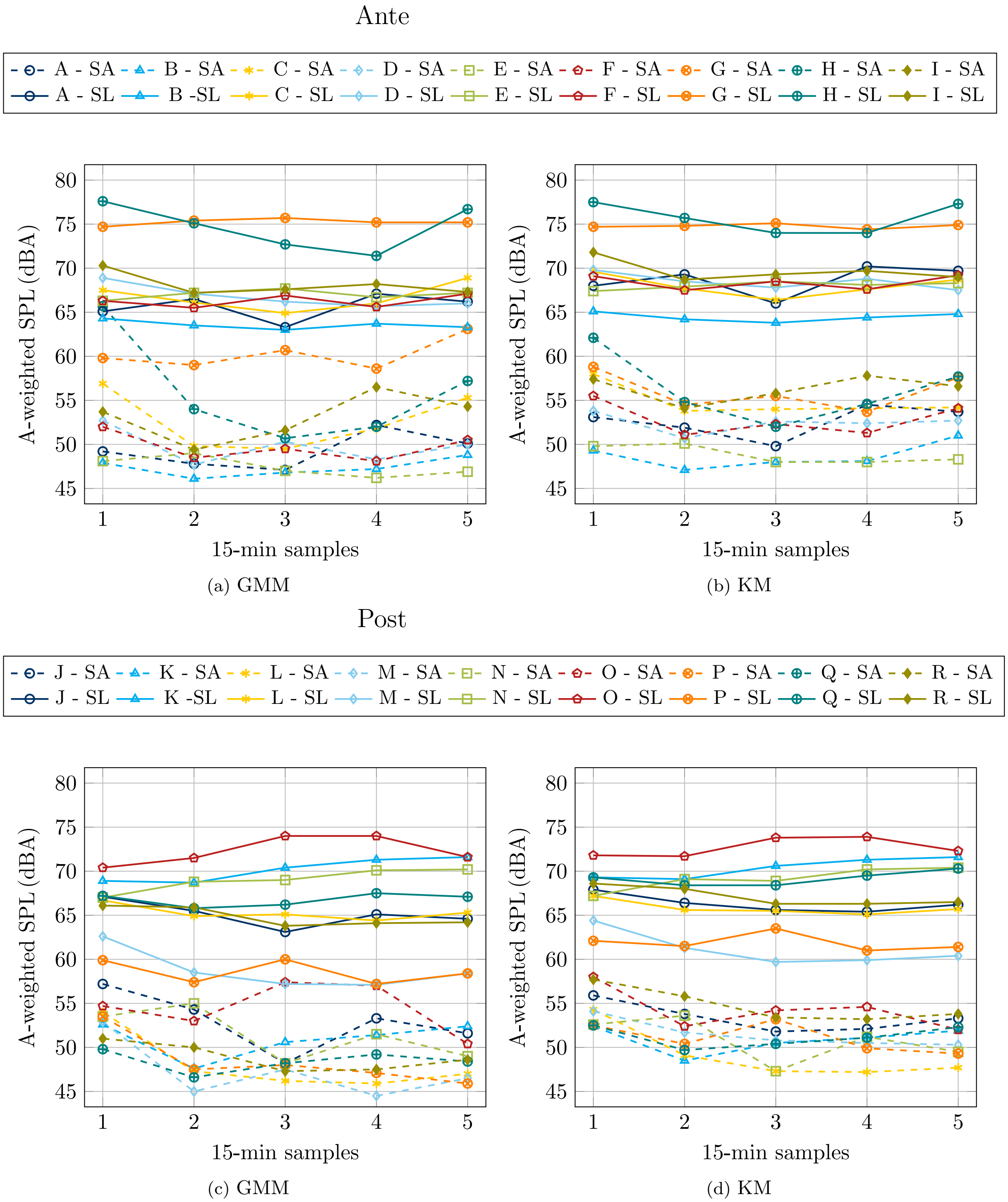



*Figure 7: 15-minutes samples of SA and SL for each lecture before and after the acoustic treatments of the halls. SA and SL are indicated respectively with dashed and solid lines.*

*Table 3: Correlation matrix among main parameters of this study before and after the acoustic treatments of the halls. The main parameters are: the student activity SA and speech levels SL for both methods Gaussian Mixture Model GMM and K-means clustering KM, occupancy N, and equivalent absorption area in occupied state* $A_{occ}$. *Before and after correlations are indicated respectively in regular and bold.*

| Correlation coefficients - Ante/**Post** | | | | | | |
|---|---|---|---|---|---|---|
| | SA - GMM | SA - KM | SL - GMM | SL - KM | N | $A_{occ}$ |
| SA - GMM | 1 | - | - | - | - | - |
| SA - KM | 0.85/**0.66** | 1 | - | - | - | - |
| SL - GMM | 0.85/**0.61** | 0.65/**0.21** | 1 | - | - | - |
| SL - KM | 0.81/**0.57** | 0.71/**0.29** | 0.97/**0.98** | 1 | - | - |
| N | -0.48/**0.13** | -0.52/**-0.54** | -0.32/**0.34** | -0.36/**0.14** | 1 | - |
| $A_{occ}$ | -0.44/**0.07** | -0.38/**-0.54** | -0.31/**0.33** | -0.32/**0.16** | 0.90/**0.99** | 1 |