DOI: 10.61782/fa.2025.0819

FORUM ACUSTICUM EURONOISE 2025

Speech Intelligibility Evaluation and Acoustic Performance of the Madrid Assembly Plenary Hall Combining Traditional and AI-based Methods

Esther García-Medina[1*] **María Centeno-Cerrato**[1] **José Trujillo**[2] **César Asensio**[3]

[1] Department of Mechanical Engineering, Instrumentation and Applied Acoustics Research Group, ETS Ingenieros Industriales, Universidad Politécnica de Madrid, Madrid, Spain
[2] LABENAC, Fundación para el Fomento de la Innovación Industrial, Madrid, Spain
[3] Grupo de Investigación en Instrumentación y Acústica Aplicada (I2A2), Departamento de Ingeniería Audiovisual y Comunicaciones, ETS Ingeniería y Sistemas de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain

ABSTRACT

From an architectural perspective, the Plenary Hall of the Assembly of Madrid is a remarkable venue featuring significant artistic value in its hanging structures and the mural that dominates the space. Due to the detection of certain issues with the room's acoustics, a descriptive analysis was conducted with the aim of identifying potential improvements in speech intelligibility for attendees of the plenary sessions. Reverberation time tests were conducted in accordance with ISO 3382 standards, revealing excessive reverberation compared to the requirements for this type of hall, which compromises the quality of speech intelligibility. The STIPA method (UNE-EN IEC 60268-16) was employed to evaluate this feature. However, the nature of interventions complicates the accurate capture of background noise during plenary sessions, hindering the adjustment of these estimates. To address this challenge, additional analyses have been conducted, integrating methods that leverage artificial intelligence. The results of this analysis highlight the challenges of making improvements once the hall is already in operation. While the sound reinforcement system has mitigated certain deficiencies, the scope for further action without substantial modifications to the architectural design is constrained.

*Corresponding author: esther.garcia.medina@upm.es.

Keywords: *room acoustics, intelligibility, reverberation time, STIPA.*

1. INTRODUCTION

The acoustics of buildings, such as conference rooms, theatres, auditoriums, and other similar spaces, is essential to ensure clear and effective communication. The architectural design of such venues focuses not only on aesthetics and functionality but also on the efficient transmission of sound, guaranteeing that speech remains clear and comprehensible for the audience. However, the integration of acoustics into architectural design can sometimes be compromised by factors including room geometry, materials, or furniture arrangement [1].

Difficulties in the Plenary Hall of the Madrid Assembly were experienced in understanding the interventions clearly during parliamentary debates. Therefore, a need to address these complaints regarding speech intelligibility appeared. Considering this situation, the Madrid Assembly decided to commission a detailed acoustic analysis of the space, aiming to identify factors affecting speech intelligibility and proposing technical solutions to mitigate the deficiencies.

In the context of parliamentary activity, speech intelligibility is crucial. The lack of correct communication not only hinders the fluency of dialogue between the deputies but can also lead to inaccuracies in the transmission of message, directly impacting the quality of the debate and the effectiveness of the legislative process [2].

However, the design criteria for a construction may sometimes prioritise other features, being acoustic factors relegated to a secondary role [1].
To assess the acoustic quality of such spaces, specific models and measurements methods have been developed. One of these is the Speech Transmission Index (STI), which rates the intelligibility in a room, defined as the proportion of verbal content that is comprehensible to the listener [3].
Speech can be considered as a modulated signal, because its level varies rapidly over time, leading to fluctuations in sound intensity. The preservation of the sound intensity envelope is essential for recognizing and understanding the speaker's vocalization [4].
The STI model is based on the principle that these fluctuations in the speech signal contain key information for intelligibility. Therefore, STI is dependent on the acoustic features of a room. Temporal domain distortions, such as reverberation, echoes, or background noise, can corrupt these fluctuations and deteriorate the intensity envelope of the speech signal, leading to the reduction of speech intelligibility [4].
Figure 1 illustrates an example of the impact of these phenomena (reverberation, echoes, background noise, etc.) on the signal envelope. The black line represents an ideal sinusoidal modulated signal, and the red line corresponds to a signal that has experienced all these previously mentioned time distortions. It can be seen that the envelope of the received signal is deteriorated.

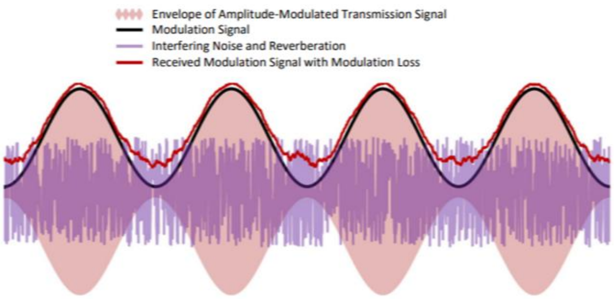

Figure 1. Effect of room acoustics on the degradation of the signal envelope [5].

STI values range on a scale from 0 to 1, where values closer to 1 indicate excellent intelligibility, while lower values suggest deficiencies in speech clarity.
Table 1 presents the recommended STI ranges according to the use of the space [3]. Spaces used for public speaking or academic purposes typically require higher STI values to guarantee clear communication. Parliaments are classified as environments that demand high speech intelligibility, preferably with STI values above 0.68.

Table 1. STI range classification based on the intended use of the space [3].

Category	STI Range	Comments	Exaples of typical uses
A+	> 0.76	Excellent intelligibility	Recording studios
A	0.72 - 0.76	High speech intelligibility	Theatres, speech auditoria, parliaments, courts
B	0.68 - 0.72	High speech intelligibility	Theatres, speech auditoria, parliaments, courts
C	0.64 - 0.68	High speech intelligibility	Teleconference, concert halls
D	0.60 - 0.64	Good speech intelligibility	Classrooms, concert halls
E	0.56 - 0.60	High quality PA systems	Concert halls, modern churches
F	0.52 - 0.56	Good quality PA systems	PA in shopping malls, pubilc offices, cathedrals
G	0.48 - 0.52	Target value for VA systems	PA in shopping malls, pubilc offices
H	0.44 - 0.48	Normal lower limit for VA systems	PA in difficult acoustic environments
I	0.40 - 0.44		PA in very difficult spaces
J	0.36 - 0.40		Not suitable for PA systems
U	< 0.36		Not suitable for PA systems

As previously mentioned, reverberation is a factor that influences the room intelligibility. An excessively high reverberation time (RT) can result in echo effects that hinder speech comprehension, while an RT that is too low can make sounds seem dry or unclear. Acoustic design must seek a balance that optimises intelligibility without compromising the overall auditory comfort [6].
The shape of the room, the seating arrangement and the choice of materials can directly influence the RT values. For instance, an excessive use of absorbent materials can reduce RT, negatively affecting speech clarity [7].
Consequently, recommendations are available that take into account the RT. These values are established as a function of the venue's volume. For instance, the DB-HR [8] — the regulatory document addressing acoustics in buildings and building elements in Spain — sets the requirements for RT values that new constructions, with volumes below 350 m^3, must legally meet under Spanish noise protection regulations.
In the case of conference rooms and seminar spaces it is recommended that the RT is kept within a range of 0.6 to 1.0 seconds to guarantee optimal intelligibility [9].
Additionally, in order to assess intelligibility from the STI, a proper characterisation of the background noise in the room is necessary. In this case, background noise characterisation in the Plenary Hall is challenging due to its functionality and intervention protocol, as sessions cannot be interrupted to conduct the measurements under the necessary conditions, that is with the Plenary Hall empty. This was reflected in the results obtained, as intelligibility values might be overestimated.

Hence, an alternative method based on the integration of artificial intelligence (AI) techniques is explored to complement the assessment of intelligibility.
With this context in consideration, the present study aims to address the attendees' complaints by identifying and quantifying the acoustic issues within the Plenary Hall of the Madrid Assembly through measurements of two acoustic parameters: RT and STI. Based on these analyses, technical and practical solutions are proposed to optimise speech intelligibility and ensure effective communication within the space during plenary sessions.

2. STUDY ROOM

The Plenary Hall of the Madrid Assembly serves as the central venue for parliamentary sessions and debates among the political representatives of the Community of Madrid. This glass-enclosed space houses a hemicycle designed to accommodate both Assembly members and the attending public. At the main platform, the Assembly Presidency and speakers are positioned, while the deputies' seats are arranged in a semi-circular layout, enhancing visual interaction among participants.
The hall features a total of 147 tiered seats, offering optimal acoustics and visibility towards the presidential table, which is slightly elevated at the front of the room. Additionally, an upper public gallery provides space for spectators, offering a clear view of the proceedings.
The architectural design of the hall plays a crucial role in both its functionality and aesthetic identity. For this reason, any potential acoustic modifications identified should be implemented with careful consideration not to compromise the architectural integrity of the space.
Figure 2 shows an overall view of the Plenary Hall.

Figure 2. Plenary Hall of the Madrid Assembly.

As it can be seen in Figure 2, wood is the predominant material within the hall, adorning the lateral and rear wall panelling, flooring, and the artistic mural that presides over the space. Beyond its aesthetic appeal, wood plays a crucial role in the hall's acoustics. In addition, there are perforated wooden panels, which enclose a rear absorbent material, designed to reduce reverberation and enhance speech intelligibility.
The hall is equipped with a conference system, enabling members of the Assembly to participate in discussions. This system equips each seat with an integrated loudspeaker directed towards each deputy, and an articulated microphone with a cardioid directivity pattern, designed to capture the participation of the deputy while minimising ambient noise and reverberation. Although these microphones offer height adjustability, they are often positioned too far from the speaker's mouth, affecting optimal sound capture.
Figure 2 also shows the perimeter electroacoustic sound reinforcement system, strategically distributed along the lateral walls and the back of the hall. This arrangement allows the wall behind the presidency to remain unobstructed, as it is designated for the display of an impressive artistic mural. The sound projection towards the presidency table is not optimal, leading to the installation of several audio monitors beneath the table with volume, bass and treble controls, complementing the general public address system.
To conclude this description, it is important to note that the subjective perception upon entering the room is one of excessive reverberation, particularly when entering from outside during plenary sessions. This is due to the hallways being equipped with a public address system that broadcasts the sessions, but with much less reverberation. Upon entering the Plenary Hall, there is a noticeable degradation in sound quality. The complaints from parliamentary members reveal the persistence of these issues regarding excessive reverberation.

3. METHODOLOGY

3.1 RT measurements

Acoustic measurements were conducted to characterise the behaviour of the Plenary Hall in terms of RT, following the procedures outlined in ISO 3382 [10]. This standard provides guidelines for assessing ordinary rooms through the integrated impulse response method.
During the measurement process, two positions for the impulsive sound source and four microphone positions were established, ensuring adequate spatial and statistical

coverage of the room (as specified in the engineering method). At each of these positions, the results from three measurements were averaged in accordance with the standard, to obtain more representative data.
The RT is characterised in the 1/3 octave frequency bands ranging from 250 Hz to 5000 Hz. This part of the spectrum is particularly relevant to conversational speech and to intelligibility [11], which is the focus of this study.
Measurements were performed with the room empty, except for four individual inside, so the results might be slightly underestimated. The presence of an audience could result in a slight improvement due to the additional sound absorption provided by the plenary attendees [12].

3.2 Intelligibility assessment

To complete the intelligibility analysis, measurements are conducted following two different procedures: a standardised method (STIPA) and an alternative approach based on the integration of AI techniques.

3.2.1 STIPA

STIPA is a simplified method for measuring STI based on emitting a modulated test signal, generated with a reduced number of modulation indices, and analysing how these modulations are deteriorated in the receiver by comparing a determined series of coefficients [3].
The test signal consists of a modulated sinusoidal pink noise within the octave band range of 125 Hz to 8 kHz, simulating the human speech spectrum. The signal modulation is carried out with a predefined set of two modulations per octave band generated simultaneously, which results in a total of 14 modulation indices.
To generate and emit this test signal a special sound source called talkbox is employed, which is a specialised loudspeaker designed to simulate a speaker's voice.
The talkbox is calibrated to emit 60 dBA of sound pressure level (SPL) at one metre along its propagation axis.
Each measurement requires between 15 and 25 seconds; however, to ensure the stability of the modulated signal, an approximate duration of 18 seconds is recommended [3].
The test signal is measured at the listener's position(s), and the STI is calculated based on the change in modulation depth. This refers to the extent of variation in the amplitude (though in other cases, it may involve frequency or phase), between the transmitted and received signals [13].
By analysing these variations, the STI can assess how well the speech signal has been preserved as it propagates through the room to the receiver point.
The "Full STI" method, which is the non-simplified version of the STIPA procedure, needs a total of 98 modulation indices and takes approximately 15 minutes per measurement [3]. This highlights the main advantage of STIPA, which is the significant time reduction.
Once the principles of STIPA have been explained, the next step is to discuss the placement of the speaker and the receiver points in the room. Figure 3 illustrates the positions selected for the study of the Plenary Hall, with the dashed lines indicating the various directions in which the speaker (sound source) was oriented. A total of one speaker position and four receiver positions are designated.

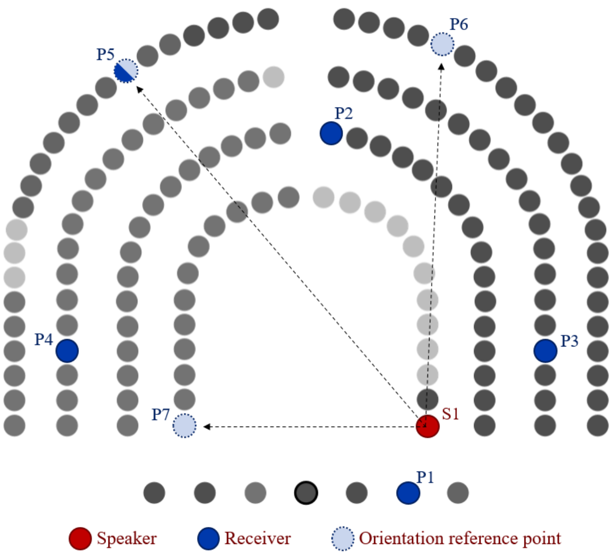

Figure 3. Positions of the speaker and receiver points selected in the Plenary Hall.

Only one speaker position is selected, with the talkbox collocated at a 1.65 metres height to accurately replicate the typical location and orientation of a talkers's mouth.
As noted in numerous session recordings, when government members are questioned and respond from their seats, they often turn towards their interlocutors, occasionally moving out of the microphone's capture range. Considering this, the simulated speaker source was placed at the most unfavourable point in the seating area due to its peripheral location within the Plenary Hall, where this issue is particularly critical. In addition, since this is the position of the President of the Community of Madrid, interventions from this speaker are more frequent and relevant.
Regarding the receiver microphones' positions, STI measurements were taken at four different seats, strategically selected based on the Assembly's request and to ensure a broader scope of analysis. At these points, the microphones were placed 1.20 metres above the floor to simulate the listening conditions of a seated person.

The measurement procedure consists of emitting a test signal using the talkbox, which is then captured by the microphone placed at the evaluated receiver position (P1 to P5), either directly or through the public address system.
The deterioration of the sound transmission chain is quantified by comparing the signal received by the measuring microphone with the original signal emitted by the loudspeaker. This degradation results from both the electroacoustic system and the room's acoustic characteristics, including the existing background noise.
Table 2 summarises the configurations of the STIPA measurements, specifying the name, speaker orientation, whether the sound reinforcement system (SRS) is switched on or off, and a comment explaining the intended evaluation scenario for each configuration.

Table 2. Summary of the measurement configurations for the STIPA method.

Name	Speaker orientation	SRS	Situation evaluated
A	P7	On	Most favourable case for speaker projection onto the microphone and assessment of effect at point P1
B	P5	Off	Room's intelligibility without sound reinforcement
C	P5	On	Improvement introduced by the complete reinforcement system under unfavourable microphone positioning conditions
D	P6	On	Speaker's orientation relative to the microphone
E	P6	Only desk monitors	How the auxiliary system complements the main reinforcement system (wall-mounted monitors)

It should be noted that in configuration A, the speaker projects their voice towards the microphone correctly, whereas in configurations B to E, the speaker is missaligned with the microphone. This results in the microphone's directivity not being able to capture the signal effectively.
For each speaker/receiver combination, three measurements were taken, with the result calculated as their average.
The tests were also conducted with the chamber empty. However, in a real situation, background noise could be higher, which could lead to an overestimation of the intelligibility characterised with this method.

3.2.2 Alternative procedure based on AI

In addition to the STIPA measurement, an alternative procedure for estimating intelligibility is explored, which could allow measurements to be conducted even in the presence of background noise during sessions.
The objective of introducing this alternative method was to explore whether an AI-based tool could enable the real-time processing of the room's sound during sessions, thereby facilitating both the collection and analysis of samples for intelligibility assessment under typical operating conditions.
The complementary procedure followed is detailed below.
Firstly, a binaural recording setup, consisting of a headphone/microphone unit and a data acquisition device, was used to emulate the experience of a plenary member.
One hour of a plenary session was recorded with the binaural setup placed in a corner of the room, as shown in Figure 4. This positioning ensured the most unfavourable intelligibility conditions in a reverberant field, avoiding direct sound from the monitors and early reflections.

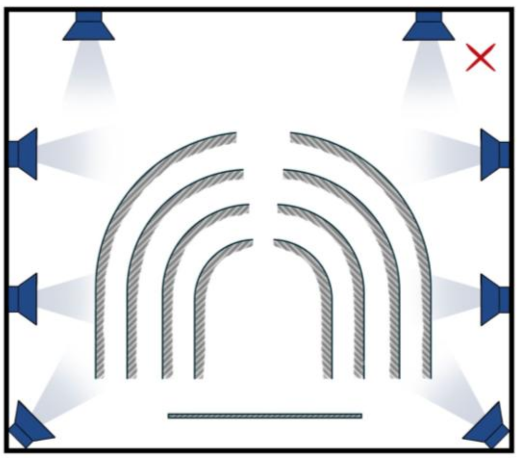

Figure 4. Position of the binaural recording setup in the Plenary Hall.

Additionally, the audio captured from the microphones installed in the seats of the plenary members is available

and serves as the reference for the direct sound that, ideally, each attendee should perceive during the sessions.
Once the recordings were obtained, two audio segments were extracted from the binaural recording, each lasting 20 seconds to ensure they contained a complete message.
In the selected audio segments, the speaker occupied the same seat as the talkbox in the STIPA measurements, therefore the sound source location is equivalent.
Then, the audio samples are processed using a text writing tool with speech recognition. This functionality, based on AI techniques, enables the identification of spoken words, converting them into a readable text format [14].
After the audio processing, a text was obtained for each one of the samples. Using this text, the intelligibility of the evaluated scenario could be determined by calculating the proportion of the dialogue understood by the AI.
To obtain an STI value, the number of words correctly recognised by the AI for the binaural recording sample is compared against the total number of words contained in the corresponding segment of the original speech.
Each text from the analysed samples was transcribed ten times, and an average was calculated to obtain more representative values of intelligibility.
To ensure the integrity of the audio signal during the transcription process, a virtual audio port is used, preventing potential degradation caused by additional cables or transmission systems.

4. RESULTS

4.1 RT

The results obtained for the RT demonstrate characteristic behaviour typical of spaces of this size, with a progressive reduction at higher frequencies attributed to air absorption. The analysis also reveals less common behaviour at lower frequencies (below 500 Hz), where an additional reduction in RT is noted. This could be explained by the presence of various types of wooden panels installed in the room, both perforated and non-perforated, acting as selective absorptive elements for this frequency range.
Despite this additional absorption at low frequencies, the RT across all analysed frequency bands remains significantly above the recommended values (1 second) for spaces where speech intelligibility is a critical factor.
Figure 5 presents the RT results in 1/3 octave bands, as specified by the ISO 3382 [10], in the Plenary Hall. It also includes the representation of typical values for venues with different uses, such as concert halls, liturgical music, and human speech, marked as dashed lines.

The RT in the Plenary Hall is considerably high, closer to spaces designed for musical appreciation rather than spoken communication. Such levels of reverberation would be suitable for venues like concert halls or auditoriums intended for musical performances, where prolonged reverberation and low sound absorption contribute to the richness of the sound. However, in spaces intended for verbal communication, these values of RT can be detrimental, as they cause overlapping speech and reduce message clarity. Therefore, ensuring a reverberation suited to the intended use of the space is crucial in guaranteeing that the room's acoustics promote speech intelligibility.

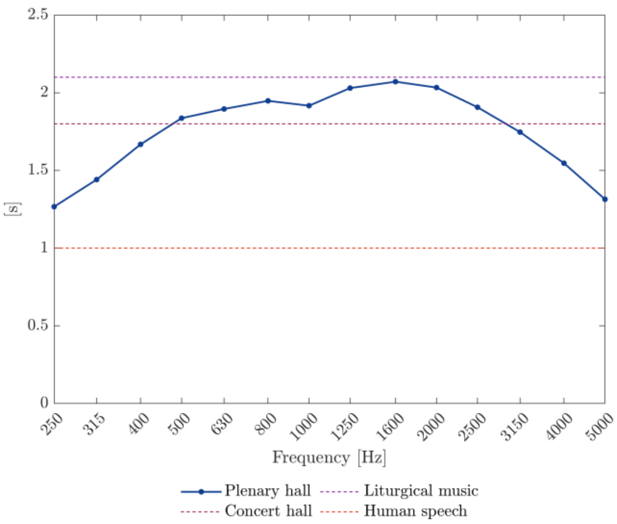

Figure 5. Average reverberation time results.

4.2 STIPA

Regarding STIPA measurements, Table 3 summarises the results obtained for the first configuration, where the talkbox is directed along the microphone axis and the electroacoustic system operates as it typically does during sessions. It was observed that, in the optimal position, the STI reached a maximum value of 0.67. Therefore, it can be concluded that when the speaker is perfectly aligned with the microphone, listening conditions remain satisfactory.

Table 3. STI results: STIPA configuration A.

Reciever position	Max	Min	Average	Deviation
P1	0.67	0.60	0.64	0.03

Table 4 presents the STI results calculated for the configurations B to E. The STI values obtained in these

configurations predominantly range between 0.45 and 0.60, which are below the recommended 0.68 for parliaments.
This highlights that there is a wide variety of listening conditions for each of the plenary members attending the sessions, which are generally less favourable than those evaluated in an ideal scenario, such as configuration A.

Table 4. STI results: STIPA configurations B to E.

Reciever Reposition	B	C	D	E
P1	0.45	0.58	0.50	0.46
P2	0.47	0.48	0.53	0.54
P3	0.60	0.46	0.65	0.64
P4	0.47	0.50	0.44	0.42
P5	0.45	0.49	0.45	0.46

Regarding configuration B, where the source is directed towards the upper-left part of the hemicycle (P5) and the electroacoustic system switched off, it was observed that the STI values dropped below 0.50 in most of the hall. However, an exception was identified in close proximity to the speaker, where STI reached 0.60.
These results indicate that, in the absence of sound reinforcement, intelligibility is compromised across most of the space, except in areas near the speaker.
When the public address system was activated, configuration C, a clear improvement in intelligibility was observed. However, STI values remained below from the optimal performance achievable when the speaker aligns correctly with the microphone installed in the seat.
As the speaker moves or turns to the upper-right part of the hemicycle (P6), as happens in configurations D and E, the directivity of the seat installed microphone becomes more critical for the intelligibility evaluation.
The STI calculated in P4, opposite direction of the speaker orientation in these two configurations, is close to 0.40. This is the minimum STI value registered in the study.
In addition, the results of configurations D and E are very similar even though in one of them the sound reinforcement system is completely operative while in the other case only has the desk monitors activated. Therefore, it can be concluded that improper microphone use is equivalent to deactivating the reinforcement system.

4.3 Alternative procedure based on AI

Moving on from the results obtained using the STIPA method, the next step is to review the intelligibility estimation obtained with the alternative procedure based on AI techniques. Table 5 presents the intelligibility results obtained through this alternative measurement procedure.

Table 5. Intelligibility results: alternative procedure based on AI.

Sample	Max	Min	Average	Deviation
1	0.58	0.30	0.42	0.09
2	0.62	0.09	0.35	0.13

The results obtained indicate poor intelligibility, anticipated due to the presence of background noise during the plenary session, which affects the accurate recognition of speech.
Additionally, significant variations between maximum and minimum values can be observed, highlighting the necessity of multiple transcriptions and averaging.

5. CONCLUSIONS

To address the requests made by the Madrid Assembly regarding difficulties in understanding the speakers during sessions, an acoustic study of the intelligibility conditions within the Plenary Hall was carried out.
Firstly, the RT was measured according to the ISO 3382 standard, confirming excessive reverberation in the venue. The RT values, ranging from 1.5 to 2 seconds, were significantly higher than the recommended for a room intended for oral communication. This prolonged RT is a determining factor in the conduct of plenary sessions.
Subsequently, the STIPA procedure was followed to analyse intelligibility in various configurations. The results show maximum STI values between 0.60 and 0.65, depending on the evaluation conditions. However, under less favourable conditions, STI values range between 0.40 and 0.60. These values are below the recommendations for parliaments, where the STI should be above 0.68.
The analysis of the different configurations indicates that the installed electroacoustic system contributes significantly to improving the overall acoustic conditions within the room, enhancing clarity in sound transmission. Nevertheless, to ensure its effectiveness, it is essential to maintain a low background noise level and proper microphone usage by parliamentary members. Failure to meet these factors reduces the system's effectiveness and compromises speech intelligibility in the Plenary Hall.
Regarding the results obtained with the alternative procedure based on AI techniques, the intelligibility values were also lower than those required for parliaments.
Several factors influence the AI's processing capabilities, including the prior training of the model in terms of vocabulary and linguistic structures, as well as the speaker's

specific pronunciation. Therefore, these results may not accurately reflect the real intelligibility conditions in the room due to the lack of control over the variables influencing AI performance, and it is recommended to consider the alternative procedure as a complementary method rather than a primary option for intelligibility assessment of a room.
To verify the proper functioning of the AI, two audio samples of the direct sound from the seat microphones were processed. These samples corresponded to the same fragments of the plenary session speech as those previously analysed. The transcription was 100% accurate, indicating that the binaural recordings serve as a representation of the effect of the reverberant field on intelligibility.
Despite the points mentioned above, as a potential line of future research, it is suggested to assess the transcription of isolated words in controlled scenarios and extend the sampling to identify possible limitations in AI recognition.
Once the acoustic analysis of the room is completed, some measures are proposed to contribute to the improvement of intelligibility and mitigate the deficiencies identified.
First of all, it is recommended to reduce the reverberation time as a primary measure to optimise acoustic quality. Furthermore, it is crucial to guarantee that seat microphones are used correctly, either through speaker training or by replacing existing microphones with hand-held or lapel microphones, independent of the speaker's movement.
In addition, increasing the direct sound received by the listener could be achieved by expanding or modifying the existing conference system or incorporating an auxiliary sound reinforcement. Ideally, in this case, each user should be able to adjust the sound level according to their needs, without applying this solution to the entire Plenary Hall.
In conclusion, the Madrid Assembly's request was addressed through an acoustic study of the intelligibility and performance of the hall, using traditional measurement procedures and exploring the feasibility of incorporating AI-based techniques as a complement.

6. ACKNOWLEDGMENTS

The authors acknowledge the Madrid Assembly for reaching out to the Fundación para el Fomento de la Innovación Industrial and the Research Group on Instrumentation and Applied Acoustics (I2A2) at the Universidad Politécnica de Madrid to carry out this study.

7. REFERENCES

[1] M. Long, *Architectural acoustics*. Elsevier, 2005.

[2] A. Foteinou, J. P. D. Cooper, and D. Murphy, "Acoustics, Audibility and Political Culture in the House of Commons, 1800–34," *Parliamentary History*, vol. 43, no. 3, pp. 316-337, 2024/10/01 2024, doi: https://doi.org/10.1111/1750-0206.12765.

[3] *IEC 60268-16:2020, Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index*, IEC, Geneva, Switzerland, 2020.

[4] *ISO 9921:2003, Ergonomics. Assessment of speech communication*, ISO, Geneva, Switzerland, 2003.

[5] J. Bedard. "STI, STIPA, and Smaart." https://support.rationalacoustics.com/support/solutions/articles/150000188813-sti-stipa-and-smaart (accessed 20th January 2025).

[6] G. E. Puglisi, A. Warzybok, A. Astolfi, and B. Kollmeier, "Effect of reverberation and noise type on speech intelligibility in real complex acoustic scenarios," *Building and Environment*, vol. 204, p. 108137, 2021/10/15/ 2021, doi: https://doi.org/10.1016/j.buildenv.2021.108137.

[7] M. Barron, *Auditorium acoustics and architectural design*. Spon Press, 2009.

[8] *Documento Básico HR: Protección frente al ruido*, M. De and M. Y. A. U. Transportes, 2019/12/19.

[9] S. Chiu, D. Noble, and E. Valmont, "9 - Acoustics in architectural fabric structures: The case of ETFE pillows," in *Fabric Structures in Architecture*, J. I. de Llorens Ed.: Woodhead Publishing, 2015, pp. 241-256.

[10] *ISO 3382-2:2008, Acoustics. Measurement of room acoustic parameters - Part 2: Reverberation time in ordinary rooms*, ISO, Geneva, Switzerland, 2008.

[11] R. Quam *et al.*, "Studying audition in fossil hominins: A new approach to the evolution of language?," 2012, pp. 1-37.

[12] C. Cha, H. Lee, and D. Jeong, "Measurements of sound absorption coefficients of raked audience seating in a rectangular scale model room," *Applied Acoustics*, vol. 217, p. 109872, 2024/02/15/ 2024, doi: https://doi.org/10.1016/j.apacoust.2024.109872.

[13] T. Houtgast, "Predicting speech intelligibility in rooms from the modulation transfer function, I. General room acoustics," *Acustica*, vol. 46, pp. 60-72, 1980.

[14] S. K. Gaikwad, B. W. Gawali, and P. Yannawar, "A review on speech recognition technique," *International Journal of Computer Applications*, vol. 10, no. 3, pp. 16-24, 2010.