# Intelligibility Prediction in Scholar Classrooms

**Samantha Di Loreto – Università Politecnica delle Marche, Italy – s.diloreto@pm.univpm.it**
**Fabio Serpilli – Università Politecnica delle Marche, Italy – f.serpilli@univpm.it**
**Valter Lori – Università Politecnica delle Marche, Italy – v.lori@univpm.it**
**Costanzo Di Perna – Università Politecnica delle Marche, Italy – c.diperna@univpm.it**

## Abstract

In recent Italian Law, the DM 11/01/2017 about Environmental criteria, reference values for the acoustic indoor quality descriptors of public buildings are imposed. These refence values are in compliance with the national standards UNI 11532-1 and UNI 11532-2. Part two of the series standard, in particular, describes the procedures and gives limit values for the acoustic comfort descriptors for schools. Regarding schools, adequate acoustic comfort targets are required in terms of indoor noise level and acoustic quality. Indoor acoustic quality targets refer to reverberation time (RT), Clarity (C50) and/or speech intelligibility (STI). The limit values for these indoor acoustic quality parameters, established by the national standards, are related to the measurement methods results; however, it is necessary to use prediction methods to estimate these parameters during the design phase. The aim of this study is to verify the prediction method accuracy used to determine intelligibility score. The study was developed to model the existing calculation method of speech transmission index (STI) in Matlab software to determine the acoustic speech intelligibility in school classrooms. A school building located in central Italy, in the Marche Region, was taken as a case study. This research aims to determine a correlation factor between the results of predictions and measurement speech intelligibility methods.

## 1. Introduction

The theme of the acoustic comfort (ambient noise, sound insulation, reverberation time, speech intelligibility) in primary school classrooms, in secondary school classrooms, as well as in university classrooms, has been the focus of several studies all around the world (Sala & Viljanen, 1995; Zannin et al., 2009). High noise levels in classrooms cause students to tire early, their cognitive abilities to decline, and they do not understand the content of the lessons. Excessive noise, too high reverberation, or the combined presence of both these effects in a classroom could reduce speech intelligibility, which is defined as the percentage of a message understood correctly.

The standard UNI EN ISO 9921 (UNI, 2004) specifies the requirements for the performance of speech communication and recommends the level of speech communication quality required for conveying comprehensive messages in several case studies. In (Pickett, 2005) many measurements of the intelligibility of speech were made to calculate the disturbance produced by different amounts of vocal force. The results of this case study show less than 5 % deterioration in intelligibility over the range, from a moderately low voice to a very loud voice (55 to 78 dB in a free field at one m from the lips). Other studies (Bradley et al., 1999; Yang & Bradley, 2009; Yang & Mak, 2018) have shown that speech intelligibility is influenced by reverberation time (RT), as well as by signal-to-noise ratios (SNR).

In (Choi, 2020), speech intelligibility tests were carried out in 12 university classrooms in Korea; the test results indicate that young adult listeners at university have a mean score of 95 % correct at a signal-to-noise ratio (SNR) value of +3 dB(A), which is a considerably lower SNR value than for the younger students in elementary schools. As a result, much attention to the development of effective objective indicators of quality and/or intelligibility are of particular interest, the measured parameters include reverberation time, early decay times, energy ratios, and STI values. The STI is a physical metric related to the intelligibility of speech degraded by additive noise and reverberation (Goldsworthy et al., 2004). Scientists nowadays consider the STI to be the parameter that best reflects the

Part of
Pernigotto, G., Patuzzi, F., Prada, A., Corrado, V., & Gasparella, A. (Eds.). 2023. Building simulation applications BSA 2022. bu,press.
https://doi.org/10.13124/9788860461919

intelligibility of speech (in a sound transmission system) (Steeneken & Houtgast, 1980). Consequently, the STI measure correlates well with subjective intelligibility scores for stimuli distorted by linear filtering, reverberation, and additive noise. Experiments in literature evaluate the effectiveness of the prevision method at predicting speech intelligibility.

In (Peters, 2020) the potential binaural effect of reducing reflection and reverberation was studied. These conditions create a reduction in intelligibility because echoes and strong discrete reflections, arriving late, lead directly to a wrong assessment when using the STI. Similarly, in (Schwerin & Paliwal, 2014) the STI approach was revisited and a variation was proposed which processes the modulation envelope in short-time segments, requiring only an assumption of quasi-stationarity (rather than the stationarity assumption of STI) of the modulation signal. Based on the tests in (Hongshan et al., 2020), the corresponding relation between STI and speech intelligibility in large spaces was modified, and a new rating threshold of STI was also proposed.

This paper aims to determine a correlation factor between the results of prediction and measurement speech intelligibility methods. The study was developed to model the existing calculation method of speech transmission index (STI) to determine the acoustic speech intelligibility in some classrooms at the Faculty of Engineering of the Università Politecnica delle Marche, Italy.

In this work, two sections are included. In the first, STI values are evaluated and calculated with the calculation method described in the annex L of BS EN 60268–16 (BSI, 2020). In the second, the result of the simulations is compared to the objective intelligibility measures in the same classes.

## 2. Material And Methods

### 2.1 Reference Values For Speech Transmission Index (STI)

The STI aims to objectively quantify speech intelligibility at a specific location in one environment when speech is produced through a normalized signal at another specific location in the same environment.

The STI index is based on the measurement of the Modulation Transfer Function (MTF). MTF quantifies the reduction in the modulation index of a test signal, depending on the modulation frequency. For each modulation frequency, the MTF is determined by the ratio between the modulation index of the signal at the listener, $m_0$, and the modulation index of the test signal, $m_i$. A family of MTF curves is determined, in which each curve is relative to each octave band of speech emission and is defined by the values that the modulation index reduction factor m assumes for each modulation frequency present in the envelope of a natural speech signal. For the STI index measurement, 7 octave bands, from 125 Hz to 8 kHz, and 14 modulation frequencies, between 0.63 Hz and 12.5 Hz at one-third octave intervals, are considered. The 98 (7x14) *m-values* are finally summarized in a single index, the STI, varying between 0 and 1, which represents the effect of the transmission system on intelligibility.

The STI quantifies the combined effect of background noise interference and reverberation on speech intelligibility reduction, with or without sound amplification systems.

The UNI EN ISO 9921 standard (UNI, 2004) establishes a relationship between STI value and their subjective assessment in terms of intelligibility for a normally hearing user. The values are shown in Table 1:

Table 1 – Relation between STI and Speech Intelligibility according to UNI EN ISO 9921:2004, Table F.1

| **Intelligibility rating** | **Sentence score %** | **STI** |
|---|---|---|
| Excellent | 100 | > 0.75 |
| Good | 100 | 0.60 to 0.75 |
| Fair | 100 | 0.45 to 0.60 |
| Poor | 70 to 100 | 0.30 to 0.45 |
| Bad | < 70 | < 0.30 |

Another classification of speech intelligibility is provided in of BS EN 60268–16 (BSI, 2020); the standard defines qualification intervals for the levels of STI obtained, as shown in the following Fig. 1. The typical STI requirements for dedicated applications are also provided in Fig. 2.

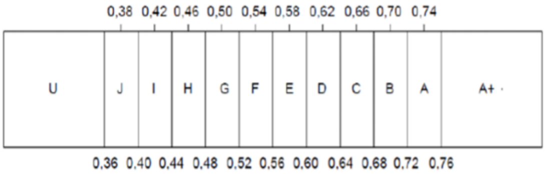


Fig. 1 – Qualification intervals for STI levels

| Category | Nominal STI value | Type of message information | Examples of typical uses (for natural or reproduced voice) | Comment |
|---|---|---|---|---|
| A+ | >0,76 | | Recording studios | Excellent intelligibility but rarely achievable in most environments |
| A | 0,74 | Complex messages, unfamiliar words | Theatres, speech auditoria, parliaments, courts, Assistive Hearing Systems (AHS) | High speech intelligibility |
| B | 0,7 | Complex messages, unfamiliar words | | |
| C | 0,66 | Complex messages, unfamiliar words | Theatres, speech auditoria, teleconferencing, parliaments, courts | High speech intelligibility |
| D | 0,62 | Complex messages, familiar words | Lecture theatres, classrooms, concert halls | Good speech intelligibility |
| E | 0,58 | Complex messages, familiar context | Concert halls, modern churches | High quality PA systems |
| F | 0,54 | Complex messages, familiar context | PA systems in shopping malls, public buildings' offices, VA systems, cathedrals | Good quality PA systems |
| G | 0,5 | Complex messages, familiar context | Shopping malls, public buildings' offices, VA systems | Target value for VA systems |
| H | 0,46 | Simple messages, familiar words | VA and PA systems in difficult acoustic environments | Normal lower limit for VA systems |
| I | 0,42 | Simple messages, familiar context | VA and PA systems in very difficult spaces | |
| J | 0,38 | | Not suitable for PA systems | |
| U | <0,36 | | Not suitable for PA systems | |

Fig. 2 – Value for STI qualification bands and typical applications

There are two measurement methods for STI: the direct and indirect method. The direct method uses modulated (speech-like) test signals to directly measure the modulation transfer function. Typically modified Pink Noise with modulation frequencies was used. In this case, the measurement signal is either applied as an electric input to the system or through a "human speaker" loudspeaker to a microphone. The indirect method uses impulse response and forward energy integration (Schroeder integral) to derive the modulation transfer function. STI can be measured at the same time as other room acoustic parameters. This means that speech intelligibility will normally be measured using an omnidirectional speaker.

## 2.2 Room Descriptions and Measurements

The university building is in a suburban area of Ancona city, away from road traffic and other environmental noise sources. In addition, the classrooms are located at the rear of the building in relation to the access road. The external SPL during the daytime period is between 45 and 55 dB(A).

For the assessment of speech intelligibility, the AT2 classroom, belonging to the Engineering Faculty of the Marche Polytechnic University, was chosen as a case study. Classroom AT2 has a volume of 378 m$^3$, an average height of 3 m and a base area of 126 m$^2$.

The classroom has a sound-absorbing acoustic ceiling, wooden chairs, and tables. The windowed surface occupies 1/3 of the total surface of the concrete perimeter walls. Fig. 3 shows AT2 classroom, and the measurement positions, as required by UNI 11532-2 (UNI, 2020).

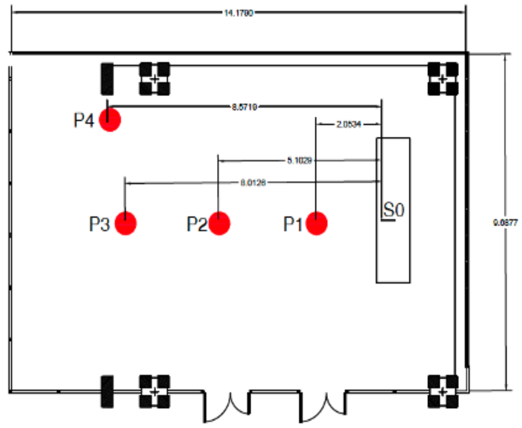


Fig. 3 – Plan of classroom AT2

The measurements in the classroom were done at four measurement points, chosen in compliance with UNI 11532 standard. Three positions were selected along the imaginary line traced on the longitudinal axis of the classroom, between the sound source and the back of the classroom, and a position was selected as representative of the most unfavorable listening condition (due to background noise, distance from the speaker, etc.). The STI measurements were derived from the impulse response measures and background noise measures with the indirect methodology proposed by BS EN 60268–16 (BSI, 2020).

Table 3 shows the results of STI for each measurement point and the STI mean value, without and with measurement uncertainty.

Table 3 – Value of STI for single point of measure, STI mean and STI mean with measurement uncertainty

| STI (P1) | STI (P2) | STI (P3) | STI (P4) | STI mean |
|---|---|---|---|---|
| 0.61 | 0.60 | 0.58 | 0.56 | 0.56 |
| **STI mean with measurement uncertainty** | | **Speech quality in accordance with CEI EN 60268-16** | | |
| 0.53 | | FAIR | | |

## 3. STI Prediction Using Indirect Method

Prediction of the STI of a sound system may be based on the MTF matrix that is calculated from the predicted room acoustic and electro-acoustic parameters and from the measured or estimated background noise levels, for each octave band contributing to the STI version chosen. The STI measure uses artificial signals (e.g., sinewave-modulated signals) as probe signals to assess the reduction in signal modulation in several frequency bands and for a range of modulation frequencies (0.6–12.5 Hz).

As requested in the reference standard, the speech spectrum at 1 meter in front of the mouth of a male speaker with the ambient noise spectrum reported in the Table H.1 of UNI EN ISO 9921:2004, see Table 4 and Table 5 was concatenated.

Table 4 – Speech spectrum at 1m in front of the mouth of a male speaker to UNI EN ISO 9921:2004, Table H.2

| **Octave band (Hz)** | **125** | **250** | **500** | **1000** | **2000** | **4000** | **8000** |
|---|---|---|---|---|---|---|---|
| **SPL@1m (dB)** | 62.9 | 62.9 | 59.2 | 53.2 | 47.2 | 41.2 | 35.2 |

Table 5 – Ambient noise spectrum according to UNI EN ISO 9921:2004, Table H.1

| **Octave band (Hz)** | **125** | **250** | **500** | **1000** | **2000** | **4000** | **8000** |
|---|---|---|---|---|---|---|---|
| **SPL@1m (dB)** | 41 | 43 | 50 | 53,2 | 47 | 42 | 39 |

The STI was calculated based on modulation transfer function (MTF) and the calculations used the method of Houtgast and Steeneken (1973).

In (UNI, 2020) for the calculation of the STI in classrooms without amplification system and with volumes > 250 m$^3$, an emission signal at 1m in axis to the source equal to 70 dB is required. So, for the calculation of the predictive STI, the reference signal of the speech was increased by 10 dB.

The modulation transfer function of the transmission path may be quantified by comparing the ratio of the modulation depth at the output and input of the test signal, and it was be written as Eq. (1):

$$m(fm) = \frac{|\int_0^{\infty} h(t)^2 e^{-j2\pi fm^t} dt|}{\int_0^{\infty} h(t)^2 dt} \cdot [1 + 10^{-\frac{SNR}{10}}]^{-1} \qquad (1)$$

where:

- $m(fm)$ is the modulation transfer function of the transmission channel

- $h(t)$ is the impulse response of the transmission channel

-$SNR$ is the signal-to-noise ratio in dB

Considering a diffuse reverberant field, the impulse response was written as Eq. (2):

$$h(t) = \frac{Q}{r^2} \cdot \delta(t) + \frac{13{,}8\ Q}{r_c^2 T} e^{\frac{-13{,}8\ t}{T}} \qquad (2)$$

where:

- $Q$ is the directivity factor for the sound source (talker)

-$r$ is the talker-to-listener distance

-$T$ is the reverberation time of the room space

The reverberation time was calculated with the method described in UNI EN 12354-6 (UNI, 2006), starting from the acoustic absorption of the room. The impulse response of the classroom was calculated in the four different positions of the room.

The standard UNI 11532-2:2020 in Paragraph 4.5 defines an optimal reverberation time, $T_{ott}$ , corresponding with a conventional occupation of the environment equal to 80 % for categories A1, A2, A3, A4. The categories of the environment, in relation to the destined use, are reported in Table 6.

Table 6 – Categories of the environment in relation to the destined use according to UNI 11532-2:2020

| **CATEGORY** | **Activities in the environment** | **Methods of intervention** |
|---|---|---|
| A1 | Music | Objective achieved with integrated design of geometry, furniture, residual noise control |
| A2 | Spoken / conference | |
| A3 | Lesson / communication as speech and lecture | |
| A4 | Special classroom lecture / communication | |
| A5 | Sport | Objective achieved with sound absorption and residual noise control |
| A6 | Areas and spaces not intended for learning and libraries | |

The reference values for optimal reverberation time for A1-A4 categories are reported in Table 7.

Table 7 – Categories of the occupied environment in relation to the destined use according to UNI 11532-2:2020

| **CATEGORY** | **Occupied environment 80 %** |
|---|---|
| A1 | $T_{ott}$ = (0.45Log(V) + 0.07) <br> (30 m$^3$ < V < 1000 m$^3$) |
| A2 | $T_{ott}$ = (0.37Log(V) - 0.14) <br> (50 m$^3$ < V < 5000 m$^3$) |
| A3 | $T_{ott}$ = (0.32Log(V) – 0.17) <br> (30 m$^3$ < V < 5000 m$^3$) |
| A4 | $T_{ott}$ = (0.26Log(V) - 0.14) <br> (30 m$^3$ < V < 500 m$^3$) |

In Fig. 4, the graph of the simulated reverberation time vs measured reverberation time, for a conventional occupation of the environment equal to 80 %, is reported.

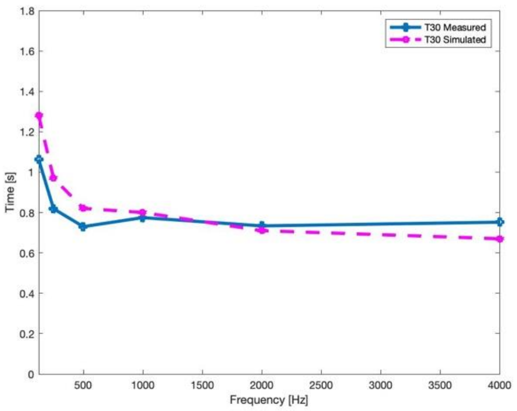


Fig. 4 – Reverberation time value in the octave bands between 125 Hz and 4000 Hz simulated (empty room) and measured

A constant MTF over the modulation frequencies indicates that speech intelligibility is mainly determined by background noise. A continuously decreasing MTF indicates an important influence of the reverberation and an MTF that decreases first and then increases again indicates the presence of an echo. Fig. 5 shows the result of the simulation of the modulation transfer function in the 7 octave bands calculated for P1.

The STI index can be finally obtained by using the weighted average method for the modulation transmission index on the considered octave bands Eq. (3):

$$STI = \sum_{k=1}^{7}(a_k \, x \, MTI_k) - \sum_{k=1}^{6}\beta_k \, x \, (MTI_k \, x \, MTI_{k+1})^{1/2} \qquad (3)$$

Where:

- $\alpha_k$ is the weight coefficient of octave band $f_m$
- $\beta_k$ is the redundancy factor between octave band $k$ and octave band $k + 1$.

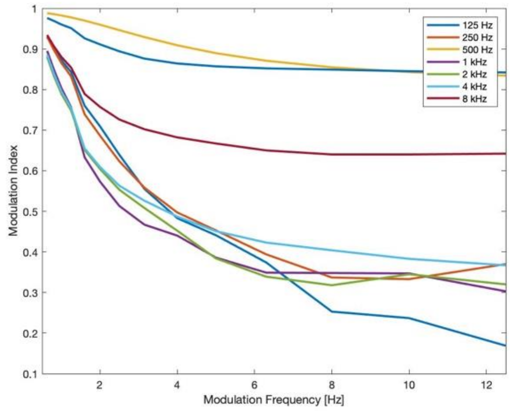


Fig. 5 – Modulation transmission ratio in the 7 octave bands

Table 8 shows the relationship between $\alpha_k$, $\beta_k$ and $MTI_k$ to determine the STI for P1.

Table 8 – Result of the calculation for the P1

| **Frequency [Hz]** | **125** | **250** | **500** | **1000** | **2000** | **4000** | **8000** |
|---|---|---|---|---|---|---|---|
| **$a_k$ male** | 0.085 | 0.127 | 0.230 | 0.233 | 0.309 | 0.224 | 0.173 |
| **Combined $MTI_k$ x $a_k$ weighting** | 0.054 | 0.090 | 0.163 | 0.157 | 0.207 | 0.118 | 0.071 |
| **$\beta_k$ male** | 0.085 | 0.078 | 0.065 | 0.011 | 0.047 | 0.095 | 0.000 |
| **Combined $MTI_k$ x $\beta_k$ weighting** | 0.054 | 0.055 | 0.046 | 0.007 | 0.032 | 0.005 | 0.000 |
| **sum $a_k$ * $MTI_k$** | 0.860 | | | | | | |
| **sum $\beta_k$ * $MTI_k$** | 0.244 | | | | | | |
| **STI (P1)** | 0.62 | | | | | | |

The same calculation was carried out for all the positions and the STI simulation results are shown in Table 9.

Table 9 – Results of the calculation of STI for P1, P2, P3 P4 and STI mean

| STI (P1) | STI (P2) | STI (P3) | STI (P4) | STI mean |
|---|---|---|---|---|
| 0.62 | 0.55 | 0.53 | 0.53 | 0.55 |
| **STI mean with measurement uncertainty** | | **Speech quality in accordance with CEI EN 60268-16** | | |
| 0.53 | | FAIR | | |

## 4. Results

From the comparison between the results of STI obtained between measured and simulated values, it can be seen that the difference is very low. This attests that the predictive model turns out to be very effective to ensure a good internal quality of the classrooms during the design phase.

In particular, the STI mean, simulated and measured, is equal and, in both cases, speech intelligibility is FAIR in accordance with the reference standard.

Considering the results of simulations and according to the background literature, a statistical analysis for the case study was carried out.

The proposed correlation model between the measurements of STI versus the simulations of STI is based on a polynomial function, according to the following Eq. 4.

$$y=ax^3+bx^2+cx+d \quad (4)$$

where $y$ is the response variable and $a$, $b$, $c$, $d$ represents partial correlation coefficients (coefficients with 95 % confidence bound).

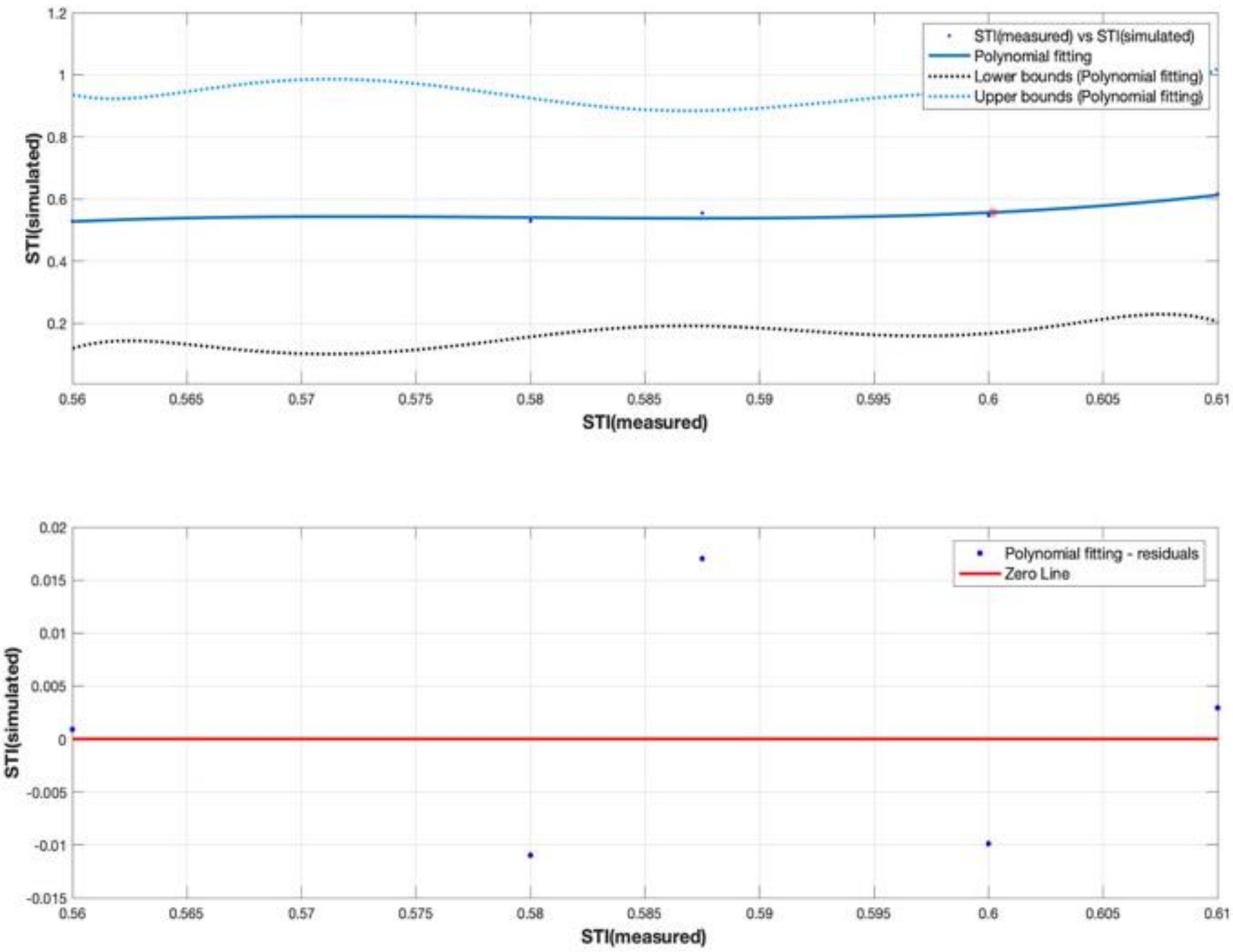


Fig. 6 – Best fit polynomial curve and residuals of the STI_m vs STI_p considered for each point of measure

Table 10 – Results of the polynomial regression

| DFE | SSE | $R^2$ | R | RMSE |
|---|---|---|---|---|
| 1 | 5.18 | 0.89 | 0.60 | 0.02 |

The result of the correlation shows the statistical significance is indicated by the $R^2 = 0.89$ and this represents a good correlation between the variables (Fig. 6; Table 10).

## 5. Conclusions

This paper systematically provides the flow of STI indirect test method specified in BS EN 60268-16 and introduces in detail the calculation formula involved in the indirect method, with reference to Schroeder's Frequency analysis and therefore to the limits of validity of the sound equations of classical theory, associated with the simulation of the room.
The study highlighted that the one of the major problems when developing this type of prediction is represented by the error generated by a low signal-to-noise ratio. Therefore, the choice of the speech spectrum, as well as the residual noise setting, represents an important choice in order for overestimation errors of the STI not to be incurred.
Although the standard is clear in recommending standard spectra, a possible solution could be to simulate the environment impulse response using a commercial room acoustic software and enter, in the input phase, an environmental noise that could be representative of the acoustic scene of the room.

## Nomenclature

### Symbols

| | |
|---|---|
| STI | Speech transmission index |
| IS | Intelligibility score |
| MTF | Modulation transfer function |
| SPL | Sound pressure level |
| STI_m | Speech transmission index (measured) |
| STI_p | Speech transmission index (predicted) |

## References

Bradley, J. S., R. Reich, and S. G. Norcross. 1999. "On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility." *The Journal of the Acoustical Society of America* 106: 1820.

BSI. 2020. BS EN 60268–16:2020. Sound system equipment - Objective rating of speech intelligibility by speech transmission index.

Choi, Y.-J. 2020. "The intelligibility of speech in university classrooms during lectures." *Applied Acoustics* 162: 107211. doi: https://doi.org/10.1016/j.apacoust.2020.107211

Goldsworthy, R. L., and J. E. Greenberg. 2004. "Analysis of speech-based speech transmission index methods with implications for nonlinear operations." *The Journal of the Acoustical Society of America* 116: 3679.

Hongshan, L., H. Ma, J. Kang, C. Wanga. 2020. "The speech intelligibility and applicability of the speech transmission index in large spaces." *Applied Acoustics* 167: 107400. doi: https://doi.org/10.1016/j.apacoust.2020.107400

Houtgast, T., and H. J. M. Steeneken. 1973. "The Modula-tion Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility." *Acta Acustica* united with *Acustica*: 66-73.

Peters, R. 2020. Uncertainty in Acoustics. Boca Raton: CRC Press.

Pickett, J. M. 2005. "Effects of Vocal Force on the Intelligibility of Speech Sounds." *The Journal of the Acoustical Society of America* 28(902): 1956.

Sala, E., and V. Viljanen. 1995. "Improvement of acoustic conditions for speech communication in classrooms." *Applied Acoustic* 45: 81-91. doi: https://doi.org/10.1016/0003-682X(94)00035-T

Schwerin, B., and K. Paliwal. 2014. "An improved speech transmission index for intelligibility prediction." *Speech Communication* 65: 9-19. doi: https://doi.org/10.1016/j.specom.2014.05.003

Steeneken, H., and T. Houtgast. 1980. "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria." *The Journal of the Acoustical Society of America* 318-326.

UNI. 2004. EN ISO 9921:2004. Assessments of speech communication.

UNI. 2006. EN ISO 12354-6:2006. Building Acoustics - Estimation of acoustic performance of buildings from the performance of elements - Part 6: Sound absorption in enclosed spaces.

UNI. 2020. UNI 11532-2:2020. Internal acoustical characteristics of confined spaces - Design methods and evaluation techniques - Part 2: Educational sector.

Yang, D., and C. M. Mak. 2018. "An investigation of speech intelligibility for second language students in classrooms." *Applied Acoustic* 134: 54-149. doi: https://doi.org/10.1016/j.apacoust.2018.01.003

Yang, W. Y., and J. S. Bradley. 2009. "Effects of room acous-tics on the intelligibility of speech in class-rooms for young children." *The Journal of the Acoustical Society of America* 125: 922-933.

Zannin, P. H. T., D. Petri, and Z. Zwirtes. 2009. "Evaluation of the acoustic performance of classrooms in public schools." *Applied Acoustics* 70: 626-635. doi: https://doi.org/10.1016/j.apacoust.2008.06.007