Comparison Between Simulated and In-Situ Measured Speech Intelligibility in the Multilingual Context of the Free University of Bozen-Bolzano

Simone Torresin – Free University of Bozen-Bolzano – simone.torresin@natec.unibz.it
Chiara Visentin – University of Ferrara - vsnchr@unife.it
Nicola Prodi – University of Ferrara– prdncl@unife.it
Francesca Cappelletti – University IUAV of Venezia – francesca.cappelletti@iuav.it
Andrea Gasparella – Free University of Bozen-Bolzano – andrea.gasparella@unibz.it

Abstract

Classrooms acoustics can affect students´ speech intelligibility and learning performance depending on its background noise level and/or reverberation. Speech intelligibility is usually assessed in real classrooms through a subjective approach, by performing speech intelligibility tests, or through an objective approach, by evaluating speech transmission index (STI) from impulse response, speech and noise level measurements. An acoustic simulation technique makes it possible to assess acoustical conditions for speech reception in virtual environments, thus allowing for predicting intelligibility before a classroom is built or renovated. However, in order to obtain reliable results, the simulation model needs to be calibrated and validated with in-situ measurements.
The aim of this work is to compare tests performed in-situ on a group of people, with tests performed on the same people by reproducing the auralized test signal through headphones, in terms of intelligibility scores (IS), response times (RT), listening efficiency values (DE) and related STI values. Simulations have been carried out using the room acoustic software Odeon version 14.01. The investigation focused on a university classroom, which is part of the Classroom Spaces Living Lab of the Free University of Bozen-Bolzano, currently equipped with devices for monitoring energy and indoor comfort conditions, as well as detailed external weather conditions. By exploiting the bilingual context in South Tyrol, Diagnostic Rhyme Tests (DRT) in the Italian language were administered to both Italian and German native speaker students, the latter with an Italian level at least equal to B2, according to the common European framework of reference for languages. In this way, speech reception performance of the two groups has been investigated and compared.

1. Introduction

The effect of adverse acoustical conditions in learning environments has been widely investigated in literature and turned out to be detrimental for students´ performance (Shield and Dockrell, 2003; Klatte et al., 2010).
Excessive reverberation and/or background noise lead to poor speech intelligibility, and affect negatively students' listening, learning and behaviour. Acoustical comfort in classrooms can be achieved when both teacher´s speech results highly intelligible and students' effort is minimized.
Speech intelligibility can be estimated in real classrooms through an objective approach, by evaluating the speech transmission index (STI) from measurements of impulse response, and speech and noise level (IEC, 2011), or through a subjective approach, by performing on people speech intelligibility tests based on sentences, isolated words or non-sense items.
In this latter case, the speech recognition accuracy is measured by the percentage of correctly understood items, defined as intelligibility score (IS). Beside IS, the effort paid by the listener in the recognition of the speech material needs to be considered, which can be tracked by the time required to the participant to give a response, defined as response time (RT). Intelligibility scores and response time have thus been combined into a joint metric called listening efficiency (DE), which describes both accuracy and effort put in the speech recognition process (Prodi et al., 2010).

Part of
Pernigotto, G., Patuzzi, F., Prada, A., Corrado, V., & Gasparella, A. (Eds.). 2018. Building simulation applications BSA 2017. bu,press.
https://doi.org/10.13124/9788860461360

Besides objective or subjective in-situ evaluations, the acoustic simulation technique makes it possible to assess acoustical conditions for speech reception in virtual environments, even before a classroom is built or renovated. Through computer simulation, subjective speech reception performance can be assessed by reproducing the auralized listening conditions through loudspeakers in an anechoic chamber or via headphones, after convolving the anechoic speech material with simulated binaural room impulse responses (BRIRs). Similarly, the objective evaluation metric STI can be derived from simulated impulse responses, noise and signals.

Acoustic models are usually calibrated based on a comparison between measured and simulated room-acoustical parameters. However, there are many perceptual features of a sound field which may not be completely described by standard room acoustical parameters (Postma and Katz, 2016). This may result in significant differences in speech reception between real and auralized conditions.

Several studies were intended to validate the assessment of speech intelligibility based on auralized signals (Yang and Hodgson, 2007; Hodgson et al., 2008; Zhu et al., 2015). Objective metrics and IS obtained from the simulated sound field were compared with the ones obtained from in-situ measurements and direct listening. However, to the authors´ knowledge, no study has so far integrated RT and DE metrics in the validation procedure.

The main objective of the present study is thus to investigate how accurately STI values can be predicted by acoustic simulation and how auralized listening conditions can lead to a reliable assessment of speech recognition performance, in terms of IS, RT and, as a consequence, DE values.

By exploiting the multilingual context of South Tyrol, a region in northern Italy, listening tests in the Italian language have been proposed both to Italian and German native speaker students and academic staff. A comparison of effort and speech recognition performance between the two groups is thus provided.

2. Methods

One existing university classroom, part of the Classroom Spaces Living Lab of the Free University of Bolzano-Bozen, has been selected for in-situ and virtual speech intelligibility tests. The room is currently equipped with devices for monitoring energy and comfort indoor conditions, as well as detailed external weather conditions. The classroom has a parallelepiped shape with dimensions of 7.29 m (width), 7.62 m (length) and 3.55 m (height), resulting in a volume of 197 m^3.

The façade of the room has a concrete painted wall with two large windows of approximately 6 m^2 each. Partitions with adjacent classrooms are made of a double layer of painted plasterboard on each side, with insulation in the cavity, whereas the partition with the corridor is acoustically treated with a Topakustik® 6/2 type finishing. The ceiling is made of unpainted concrete and the floor is raised, with a linoleum finishing. Twenty-four not-upholstered wooden chairs are distributed in the seating area.

Reverberation times measured in the fully occupied condition (25 persons) are shown in Fig. 1.

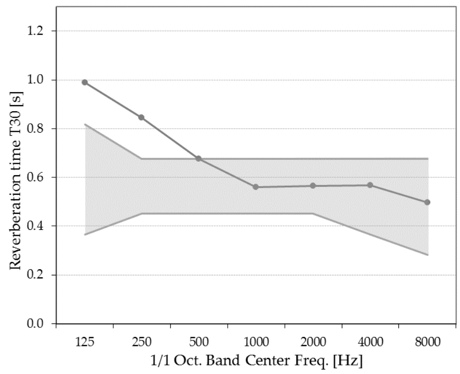

Fig. 1 - Reverberation times T_{30} [s] measured in the fully occupied classroom compared with target values suggested by the DIN 18041 standard (DIN, 2016)

Values were obtained from impulse response measurements and are the spatial average of 6 microphone positions evenly distributed over the seating area.

The mid-frequency reverberation time is equal to 0.62 s, close to the target value suggested by the DIN 18041 standard, T_{soll}=0.56 s (DIN, 2016).

In Fig. 1 measured reverberation times are compared with the tolerance range derived, according to the standard, from the target value T_{soll}. Measured octave-band values above 500 Hz comply with the standard requirements.

2.1 In-Situ Intelligibility Tests and Measurements

This section describes speech intelligibility tests performed in the university classroom and the objective characterization of listening conditions. This allowed the collection of reference values for the calibration of the corresponding virtual model and the validation of the auralized listening conditions.

2.1.1 Participants, test material and procedures

A Diagnostic Rhyme Test (DRT) in the Italian language was used for this study (Bonaventura et al., 1986). Test words were meaningful disyllabic words, with a phoneme distribution representative of the Italian language. The speech material, consisting in the target words embedded in a carrier phrase, was read by an Italian native female speaker and recorded in a silent room. The material was split into 6 lists of 18 words and played back inside the classroom by a B&K type 4720 artificial mouth with a level of 63 dBA measured at 1 m in front of the loudspeaker. The signal source was located at 1.5 m above the floor at the teacher´s position and oriented towards the seating area. The test setup is shown in Fig. 2.

Listening tests were performed under three acoustic conditions, called, respectively, “actual ambient noise” (A), “stationary noise” (S), and “fluctuating noise” (F).

In the first condition, no disturbing noise was added. The background noise consisted in the emission from the classroom projector and from the ventilation system. The two masking signals were a speech-shaped stationary noise (S) and a fluctuating noise (F), created by processing a signal according to the ICRA instructions (Dreschler et al., 2001). During the test session, S and F were played back with a B&K type 4292-L omnidirectional source located on floor, exactly below the speech source, as to screen the direct noise path towards the listeners and thus guaranteeing a diffuse noise condition. Noise levels were varied as to obtain, at 1 m in front of the signal source, a null signal to noise ratio (SNR) for the stationary noise and a SNR equal to 1 dBA for the fluctuating noise.

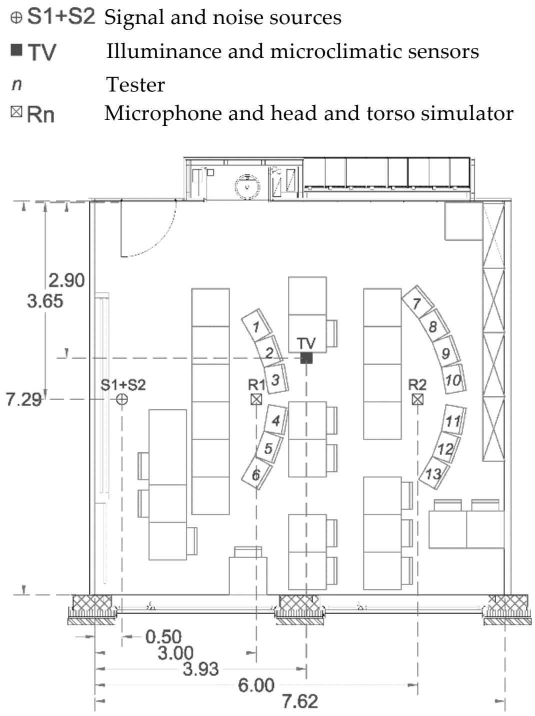

Fig. 2 – Test and measurement setup during in-situ speech intelligibility tests

The tests were administered to 26 normal-hearing young adults, half of them Italian native speakers, and, the other half, German native speakers with an Italian level at least equal to B2, according to the Common European Framework of Reference for Languages. Participants were selected from students and academic staff on the basis of their self-declared language skills.

The tests were administered to the two sample groups in two subsequent one-hour sessions. In each session, the subjects sat around two receiver positions, R1 and R2, as illustrated in Fig. 2, located, respectively, at 3.9 m and 7.6 m from the classroom front-end wall. After a training session, the participants completed three tests, each under a different noise (A, S, F). The subjects were then asked to listen again to the three listening conditions and to rate on three Visual Analogue Scales the effort paid in speech comprehension. The participants were then invited to change their sitting positions: the testers sitting on the back of the classroom were asked to move frontward, and vice

versa. With this new arrangement, the three listening tests and related effort evaluations were repeated, under the same noise conditions but with a different exposure order (S, F, A) and with different word lists. In this way, each participant experienced each of the three noise conditions at each receiver position. The two-phase experiment was then repeated for the second sample group.
The tests were administered through *Intelligo*, a system that manages the synchronous audio playback and response collection, also enabling automatic recording of response times (Prodi et al., 2012). The participants, after hearing a carrier phrase followed by a target word, had to select the word they had heard on a mobile phone touchscreen by choosing between three options, two-word alternatives or the "none of the above" option, as shown in Fig. 3.
Following the tests, German native speakers were asked to fill out a brief questionnaire aimed at a deeper understanding of their language background. The answers regarding their age of acquisition of the German and Italian languages, their language use at home, with friends and at university, their country of origin, their language proficiency certifications and their parents´ mother tongues were collected. The questionnaire confirmed that all German native speakers had been exposed to the German language from birth and started the study of the Italian language at primary school (around 6 years of age).

2.1.2 Objective measurements

At the end of each one-hour experiment, the objective description of the test acoustic conditions was performed. With participants sitting quietly, monaural and binaural impulse responses were measured, and signal and noise levels collected at the two receiver positions by means of two B&K type 4128-c head and torso simulators and two omnidirectional microphones located on the top of each manikin. Speech level measurement was carried out recording a continuous speech sample.
Indoor microclimatic parameters – i.e. air temperature, relative humidity, air velocity – were monitored during the test sessions and used as input data for the subsequent acoustic model. A DeltaOHM Thermal Microclimatic logger type HD 32.1 was positioned as shown in Fig. 2, as to collect a mean value inside the space area. Even if a denser measurement grid would be required, a rough estimation of thermal comfort conditions can be performed by calculating Fanger's comfort indices PMV (Predicted Mean Vote) and PPD (Percentage of People Dissatisfied). The average PMV value calculated for the test period is equal to 0.0 (PPD = 5 %), considering people in sedentary activity (1.2 met) while wearing 0.8 clo. This confirms that testers were in thermo-neutral conditions and that during the tests the classroom could be included in the I Category of comfort according to UNI EN 15251 prescriptions (CEN, 2008).
In addition, illuminance level measurements were conducted at the same monitoring position by the use of a LI-210 Photometric Sensor (LI-COR), in order to obtain a rough estimation of the horizontal illuminance level on the task area (i.e. the surface at the desks´ height). An average illuminance level of 522 ± 60 lx was measured.

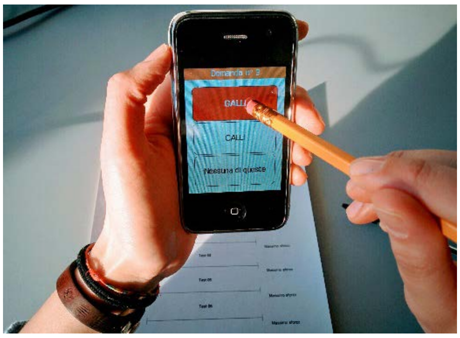

Fig 3 – Example of Intelligo test screen

2.2 Acoustic Simulation and Intelligibility Tests

A classroom model was created and calibrated based on measured room-acoustical parameters. Auralized speech intelligibility tests were then created and reproduced via headphones, in such a way to virtually replicate the listening conditions experienced in the real room as accurately as possible.

2.2.1 Room modelling and model calibration

The space was modelled using the room acoustics software Odeon version 14.01 (Odeon A/S, 2016). The software employs a hybrid approach that combines, below a selected reflection transition order, a mixture of the image source method and ray-tracing and, above this transition order, a special ray tracing process that generates secondary sources radiating energy locally from the surfaces (ray-radiosity).

A geometric model, made of 261 surfaces, was created in SketchUp and imported in Odeon. Desks and unoccupied chairs were simulated as suspended planes. Occupied chairs were modelled as parallelepipeds with sides of 0.6 m (depth), 0.5 m (width) and 0.4 m (height), corresponding to the seat and parallelepipeds with sides of 0.2 m (depth), 0.5 m (width) and 0.4 m (height) as the seatback.

For room-acoustical parameter predictions, the virtual source was defined as to represent the directivity pattern of the speech source used in real classroom measurements. The *Tlknorm* source available in Odeon was used, having the directivity of a human talker, but it was shaped with a typical female spectrum according to IEC 60268-16 (IEC, 2011).

Concerning the noise source, it was modelled as an omnidirectional source (*Omni*), as to replicate the dodecahedral source employed in real classroom measurements.

Two receiver points were set corresponding to R1 and R2 receiver positions. Simulations were performed with a transition order of 2, 2000 early rays and 16000 late rays. Air temperature and relative humidity were set according to average values measured during the tests (T=23 °C, RH=23 %).

Mid-frequency scattering coefficients were assigned to surfaces, taking into account only scattering due to surface roughness, being diffraction phenomena handled directly by Odeon. A scattering coefficient of 0.50 was applied for occupied chairs and of 0.70 for radiators. A coefficient of 0.05 was assigned to all other surfaces.

The initial absorption coefficients were assigned to surfaces based on Odeon´s material library and on data available from literature. These values were then adjusted to calibrate T30 values. The calibration procedure was performed with slight adjustments in order to keep physically realistic values for material properties. The calibrated octave-bands absorption coefficients are shown in Table 1.

2.2.2 Participants, test material and procedures

Auralized listening conditions were created by convolving anechoic signals with the binaural impulse responses (BRIRs) calculated in Odeon. The BRIRs were calculated by introducing in Odeon the signal and noise sources with a white spectrum, in order not to include twice the overall frequency response after convolution with anechoic signals.

The virtual listener was defined by the head-related transfer function (HRTF), previously measured for the B&K type 4128-c manikin. The generic headphone equalisation filter *Subject_021Res10deg_diffuse.wav* was employed as to compensate for a non-linear frequency response of the headphones.

Table 1 – Octave-band absorption coefficients of the main linings in the classroom

Material	125	250	500	1000	2000	4000
Painted concrete wall	0.02	0.02	0.03	0.04	0.05	0.05
2 · 13 mm plasterboard on steel frame, mineral wool in cavity, surface painted	0.12	0.08	0.06	0.04	0.04	0.05
Rough concrete ceiling	0.02	0.03	0.03	0.03	0.04	0.07
Desks and furniture	0.04	0.05	0.05	0.05	0.03	0.01
Perforated wood panel, Topakustik 6/2	0.30	0.32	0.42	0.66	0.70	0.45
Audience on wooden chairs	0.22	0.25	0.56	0.69	0.81	0.78

The auralization procedure involved creating separate BRIRs at each receiver position (R1 and R2), for both speech and noise sources. The BRIRs were then convolved through AudioMulch software with the anechoic speech and noise signals to recreate a subset of the listening conditions proposed during the in-situ experiment. Specifically, the two conditions (position R1 and R2) with the stationary masker (S) were selected.

Listening tests with auralized material were proposed in a quiet environment, by reproducing signals over Audio Technica type ATH-m50x headphones.

In order to calibrate the reproduction level of the signals and measure the STI of the virtual conditions, auralized speech and noise signals were measured with headphones placed on the head and torso simulator.

The test in virtual conditions was proposed to 21 listeners who took part in the in-situ experiment. Details are given in Table 2.

Following a training session, participants completed two tests respectively at the R1 and R2 virtual receiver positions, under stationary noise (S). Tests were administered again through the Intelligo system and self-evaluation of effort was performed through visual analogue scales after each test completion.

Table 2 – Participants to the auralized speech intelligibility test

Italian native speakers	Age
Female (n=5)	23.4 (± 0.8)
Male (n=5)	25.4 (± 1.6)
All (n=10)	24.4 (± 1.6)
German native speakers	Age
Female (n=6)	23.0 (± 2.5)
Male (n=5)	29.4 (± 9.7)
All (n=11)	25.9 (± 7.5)

3. Results

3.1 Room Acoustical Parameters

Measured and simulated T30 and C50 values are provided in Fig. 4 and 5, as mean values over the two receiver positions. Differences are related to the just noticeable differences (JND) indicated in EN ISO 3381-1 standard, assuming for C50 the same threshold as for C80 (CEN, 2009).

As a consequence of the model tuning, good agreement is found between simulated and in-situ measured T30 values. T30 differences are within the one-JND threshold, except for 8000 Hz where the difference reaches 2 JNDs (10 %).

Simulated C50 values are within the one-JND threshold from measured ones, except for 1k, 2k and 8k octave-bands where C50 differences are, respectively, equal to 1.28 dB, 1.45 dB and -2.55 dB.

3.2 Speech and Noise Levels

Fig. 6 presents a comparison between simulated and measured speech and noise levels, as mean values between the two R1 and R2 receiver positions. Simulated values were derived from recordings of speech and noise signals reproduced over headphones on a head and torso simulator.

Differences are generally above the JND threshold. Signal prediction is up to 5 dB above measurement at 8k, while noise prediction is up to 5 dB above measurements at 125 Hz.

3.3 STI

STI values were derived from in-situ measured impulse responses and signal and noise levels, according to IEC 60268-16 standard (IEC, 2011). Similarly, STI values related to auralized sound fields were calculated from simulated BRIRs and from recordings of speech and noise signals reproduced over headphones on a head and torso simulator. Final STIs were obtained as average values between the two ears. Female STI values are considered. As illustrated in Table 3, good agreement is found between simulated and measured STI values.

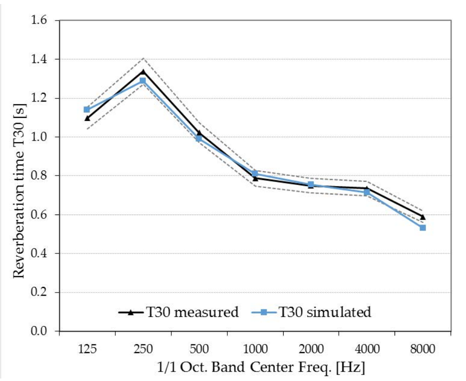

Fig. 4 – Comparison between mean measured (± 1 JND) and simulated T30

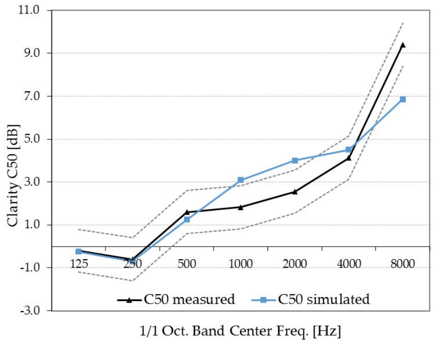

Fig. 5 – Comparison between mean measured (± 1 JND) and simulated C50

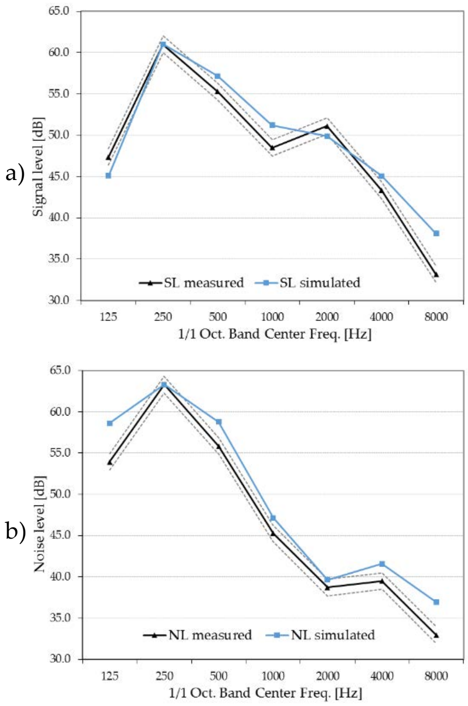

Fig. 6 - Comparison between mean measured (± 1 JND) and simulated signal (a) and noise (b) levels

Table 3 – Comparison between STI values obtained from in-situ measurements and from auralization

STI female	In-situ	Auralized
R1	0.55	0.54
R2	0.46	0.46

3.4 Intelligibility scores

Single IS values were calculated as follows:

$$IS = (\,raw\ score + 1) \cdot 0.5 \qquad (1)$$

where the raw score is equal to "+1" for correct answers, "-1" for incorrect answers and "-0.5" for the "none of the above" option (Prodi et al., 2010).

Pooled IS values were averaged, for each acoustical condition, separately for the Italian (n=10) and German (n=11) native speakers and for in-situ and auralized tests. Results are analyzed by comparing IS, along with their standard deviations, with STI values, as shown in Fig. 7.

Even though a statistical analysis of data has not been yet performed, some tendencies can be outlined and preliminarily discussed, relying on mean values and standard deviations.

Differences between measured and simulated listening conditions are always within the standard deviations, suggesting that the real and auralized conditions cannot be discriminated by the IS metric. As expected, IS at position R1, closer to the speech source, are higher than IS at the rearward position R2. In addition, data show a tendency towards higher IS for the Italian group compared to the German one, at both listening positions.

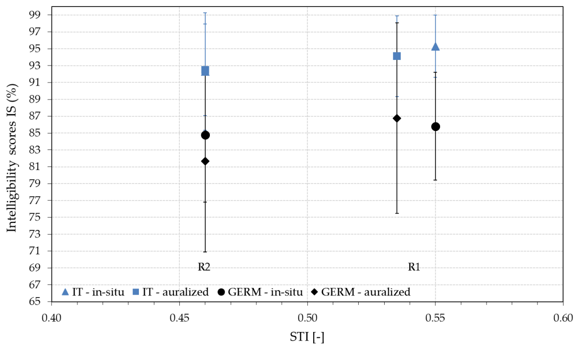

Fig. 7 – Comparison of intelligibility scores (IS) and speech transmission index (STI) from in-situ and auralized tests, for Italian (IT) and German (GERM) native speakers

3.5 Response Time

RT values from in-situ and auralized tests are presented in Fig. 8, by distinguishing the two sample groups (IT and GERM) and the two listening conditions (position R1 and R2).

Differences between in-situ and auralized tests are lower than individual standard deviations. A tendency can be observed towards shorter RTs for the Italian group compared to the German one and for position R1, closer to the sound source, compared to position R2.

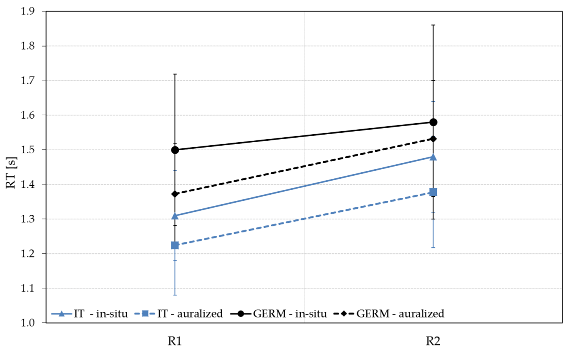

Fig. 8 - Comparison of response times (RT) from in-situ and auralized tests, for Italian (IT) and German (GERM) native speakers

3.6 Listening Efficiency

DE values were calculated as the ratio of IS and RT values. Values from in-situ and auralized tests for the two sample groups are provided in Fig. 9.

By comparing differences between real and simulated conditions with their standard deviations, it could be concluded that no differences are found in the DE metric. Differences between Italian and German native speakers are, in this case, larger than standard deviations, suggesting a difference between the two groups, with a higher efficiency for Italians.

In accordance with IS and RT results, a higher listening efficiency is reached at position R1.

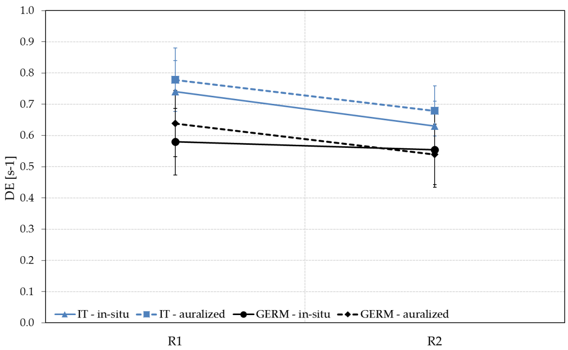

Fig. 9 – Comparison of listening efficiency values (DE) from real and auralized tests, for Italian (IT) and German (GERM) native speakers

3.7 Self-Evaluated Effort

Data collected from subjective effort assessments were normalized on a 0-10 scale and averaged for the two sample groups and for in-situ and auralized tests, as shown in Fig. 10.

Differences between real and simulated conditions are lower than standard deviations for individual results, suggesting a lack of differences. However, it should be noticed that the more scattered nature of data makes it difficult to recognize differences clearly also between IT and GERM groups. The only clear tendency is towards higher effort paid by listeners at the furthest position.

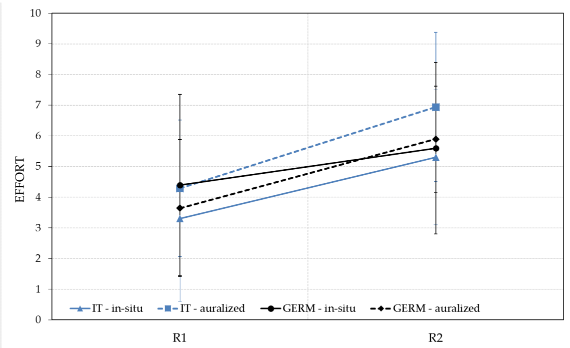

Fig. 10 – Comparison of subjective effort assessments from in-situ and auralized tests, for Italian (IT) and German (GERM) native speakers

4. Conclusion

The present study investigates the effectiveness of a calibrated acoustical model in providing valid auralized sound fields for speech-intelligibility evaluation. Listening tests at two positions inside a real and an auralized university classroom were proposed both to Italian and German native speakers, in the presence of a masking stationary noise.

Descriptive statistics results suggest good agreement between STI, IS, RT, and DE values and thus an overall consistency of the two procedures.

Moreover, the GERM group showed a tendency to achieve lower IS, longer RT, and lower DE values compared to their Italian peers. To assess the significance of the trends outlined by data a detailed statistical analysis is under course for both IS, RT, and DE and the perceived subjective effort.

Acknowledgements

This study has been funded by the project "Klimahouse and Energy Production" in the framework of the programmatic "financial agreement with the Autonomous Province of Bozen-Bolzano of Research Capacity Building and by the internal project "Human Centred Design of the Built Environment: definition of a methodology for the experimental assessment of the overall Indoor Environmental Quality" of the Free University of Bozen-Bolzano.

The authors gratefully acknowledge students and academic staff for their kind collaboration in the listening tests. Part of the instrumentation used for this research has been provided by FISTEC laboratory of University IUAV of Venezia. Additional thanks go to Lorenza Pistore for her kind support during tests and microclimatic measurements.

References

Bonaventura, P., A. Paoloni, F. Canavesio, P. Usai. 1986. *Realizzazione Di Un Test Di Intelligibilità per La Lingua Italiana. Rapporto Tecnico 3C1286*. Rome, Italy: Fondazione Ugo Bordoni.

CEN - European Committee for Standardization. 2008. *EN 15251 - Indoor Environmental Parameters for Assessment of Energy Performance of Buildings, Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics*. Brussels, Belgium: CEN.

CEN - European Committee for Standardization, . 2009. *EN ISO 3382: Acoustics - Measurement of Room Acoustic Parameters - Part 1: Performance Spaces*. Brussels, Belgium: CEN.

DIN - German Insitut for Standardization. 2016. *DIN 18041: Acoustic Quality in Rooms - Specifications and Instructions for the Room Acoustic Design*.

Dreschler, W.A., H. Verschuure, C. Ludvigsen, S. Westermann. 2001. "ICRA Noises: Artificial Noise Signals with Speech-like Spectral and Temporal Properties for Hearing Instrument Assessment". *Audiology* 40(3): 148–57. doi: 10.3109/00206090109073110.

Hodgson, M., N. York, W Yang, M. Bliss. 2008. "Comparison of Predicted, Measured and Auralized Sound Fields with Respect to Speech Intelligibility in Classrooms Using CATT-Acoustic and ODEON". *Acta Acustica United with Acustica* 94(6): 883–90. doi:10.3813/AAA.918106.

IEC - International Electrotechnical Commission. 2011. *IEC 60268 Sound System Equipment - Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index*.

Klatte, M., T. Lachmann, M. Meis. 2010. "Effects of Noise and Reverberation on Speech Perception and Listening Comprehension of Children and Adults in a Classroom-like Setting". *Noise and Health* 12(49): 270–82. doi:10.4103/1463-1741.70506.

Odeon A/S. 2016. Odeon - User´s Manual.

Postma, B.N.J., B.F.G. Katz. 2016. "Perceptive and Objective Evaluation of Calibrated Room Acoustic Simulation Auralizations". *The Journal of the Acoustical Society of America* 140:4326. doi:10.1121/1.4971422.

Prodi, N., C. Visentin C. Bellettini. 2012. "Listening Efficiency Testing". In: *Proceedings of the AES International Conference*.

Prodi, N., C. Visentin A. Farnetani. 2010. "Intelligibility, Listening Difficulty and Listening Efficiency in Auralized Classrooms". *The Journal of the Acoustical Society of America* 128(1): 172–81. doi:10.1121/1.3436563.

Shield, B.M., J.E. Dockrell. 2003. "The Effects of Noise on Children at School: A Review". *Building Acoustics* 10(2): 97–116. doi:10.1260/135101003768965960.

Yang, W., M. Hodgson. 2007. "Validation of the Auralization Technique: Comparative Speech- Intelligibility Tests in Real and Vir¬tual Classrooms". *Acta Acustica United with Acustica* 93(6): 991–99.

Zhu, P., F. Mo, J. Kang, G. Zhu. 2015. "Compar-isons between Simulated and in-Situ Meas-ured Speech Intelligibility Based on (Binaural) Room Impulse Responses". *Applied Acoustics* 97: 65–77. doi:10.1016/j.apacoust.2015.04.005.