



MADRID
inter.noise 2019
June 16 - 19

NOISE CONTROL FOR A BETTER ENVIRONMENT

Spatial benefits on speech intelligibility in real classroom acoustics under energetic and informational masking noise

Puglisi Giuseppina Emma¹, Minelli Greta², Astolfi Arianna³
Politecnico di Torino, Department of Energy
Corso Duca degli Abruzzi, 24, 10129, Torino, Italy

Warzybok Anna⁴, Kollmeier Birger⁵
Cluster of Excellence Hearing4All and Department of Medical Physics and Acoustics, University Oldenburg
D-26111 Oldenburg, Germany

ABSTRACT

Challenging acoustic conditions, i.e. high noise and long reverberation time, negatively affect speech intelligibility. This is particularly true for school environments where learning is delivered, and students of every age develop their cognitive abilities. Research has primarily focused on the effect of reverberation and noise on speech intelligibility and on the spatial release from masking under laboratory conditions, whereas few studies considered these aspects ecologically. Also, the effect of noise on speech intelligibility was widely investigated considering its energetic rather than its informative content.

This work deepens the extent to which the spatial release from masking is affected by reverberation and noise under real classroom acoustics, in order to help the design of learning environments to enhance speech intelligibility. Binaural room impulse responses were acquired at increasing speaker-to-listener distances, with noise sources at 0°, 120° and 180° from the listener's head, in classrooms with reverberation times ranging from 0.4 s to 3.5 s, as to represent the typical conditions of Italian schools. Then, listening tests were performed: the impulse responses were convolved with speech and noise anechoic stimuli, and presented via headphone to a selected panel of normal hearing adults.

Further analyses are now in progress; preliminary results reveal that speech intelligibility is worse under higher reverberation times and, averagely, under informational masking noise, as expected. As far as the spatial release from masking is concerned, when longer reverberation times are present in the room there is a tendency to have greater benefits under informational noise.

Keywords: Noise in Classrooms, Classroom Acoustics, Speech Intelligibility
I-INCE Classification of Subject Number: 61

¹ giuseppina.puglisi@polito.it

² greta.minelli@polito.it

³ arianna.astolfi@polito.it

⁴ a.warzybok@uni-oldenburg.de

⁵ birger.kollmeier@uni-oldenburg.de

1. INTRODUCTION

Excessive noise and long reverberation times can degrade the ability of a listener to discriminate useful sounds (e.g. target voices) in everyday life environments. This is particularly true for school settings where learning is delivered, and students of every age and background develop their cognitive abilities. Detrimental effects are especially strong in those environments as they affect both the speaking and the listening tasks [1-6]. Most classrooms in Italy do not comply with national or international standards, as they are settled in historical buildings with big volumes and vaulted ceilings that generate unfavourable environments. Furthermore, the multicultural and inclusive settings (i.e. classrooms that include students with different mother tongues and backgrounds, and with cognitive deficit too) in which listeners are commonly immersed nowadays, make it necessary to develop strategies for the enhancement of speech intelligibility that account for different premises at the same time.

To answer to the need of having comfortable teaching and learning environments, research has focused on the assessment of speech intelligibility using accurate and multilanguage tests [7], and on the investigation on the benefits that the architectural design and spatial settings give to listening [8-9]. So far, research has primarily focused on the study of the effect of reverberation on speech intelligibility [10] and on the spatial benefits, i.e. spatial release from masking (SRM) and the so-called “cocktail party phenomenon” Cherry and revised by Bronkhorst [11-13], under laboratory conditions, whereas few studies considered these aspects ecologically. In particular, Dirks and Wilson [14] showed that to understand the 50% of a speech a release from masking of up to 10 dB occurs when the target and noise sources are spatially separated. On the opposite, when the sources are colocated, that is, when they are oriented in the same direction, the binaural hearing cues give a small gain (~1 dB, which is comparable to the sound pressure level just noticeable difference), which can therefore be considered as negligible. Platte and Vom Hövel (1980), Plomp and Mimpen (1981), Duquesnoy (1983), Bronkhorst and Plomp (1988), Peissing and Kollmeier (1997) [15-19], proved in several studies that the largest release from masking (~12 dB) occurs when the masking noise is located behind the interaural axis, that is, at about 120° from the listener’s head orientation.

Furthermore, the effect of noise on speech intelligibility was widely investigated considering its energetic rather than its informative content. From a cognitive point of view, energetic masking noise (EM) and informative masking noise (IM) correspond to peripheral and central sites of origin within the auditory pathway, respectively, as defined by Pollack (1975) [20]. As far as the energetic content of both them is concerned, the spectral distribution may differ significantly since EM is generally similar to white noise or to a speech-modulated noise, whereas IM has a spectrum that changes based on the characteristics of the talker(s) it refers to. Also, IM has an informational content that EM does not have, therefore it requires a listener to segregate the target speech stream to be understood from other speech streams and to direct attention to it, which may be challenging since the target and the masker have very similar characteristics [21].

This work investigates the influence of reverberation and noise in real classrooms on speech intelligibility, which was measured adaptively converging to signal-to-noise ratio yielding 80% correct recognition scores (i.e. Speech Reception Threshold, SRT80). To make the study accurate, listening tests were built on a well-established methodology, that is, the “Matrix Sentence Test”. The Matrix Sentence Test for the Italian language (ITAMatrix) was optimized and validated to be used at research and diagnostic purposes with adults [22], and hereby adapted for the investigation under real acoustics.

2. METHODOLOGY

Based on the available literature that focuses on competitive listening under realistic acoustic conditions (either simulated or measured), five experiments were designed to study the effects of acoustics on speech intelligibility in two representative Italian classrooms, one with acoustical treatment and one without, where binaural room impulse responses (BRIRs) were measured at a head and torso simulator ears.

The description of the case studies, of the measurement procedure in-field and of the listening tests design and administration is given in the following paragraphs.

2.1 Case studies and classroom acoustic parameters

A traditional Italian primary school was involved in the study. It is located in a residential area of the city of Torino (northern Italy), in a building that dates back to the end of the XIX century. Classrooms generally present big volumes, i.e. more than 200 m³ on average, wide windows and no acoustic treatments either on the ceilings or on the walls. They can face local streets with low traffic noise or the inner courtyard, therefore noise from the outside should be properly controlled in order to avoid annoyance and disturbance to the teaching activities. In 2010, one classroom of the school was acoustically treated within a cooperation project with an industry, which has sponsored such intervention. The new classroom has then been adopted by the school as “reading room”, making it available to all the students. The acoustic treatment included a false-ceiling in rock-wool panels, a reflective panel placed above the teacher’s desk in order to re-direct the useful reflections to the rear side of the room too, and a mix of absorbent and vibrating panels on the lateral walls. The detailed description of the intervention is available on a manual published by Rockwool®, curated by Astolfi and Giovannini [23]. So, to this work’s aim, two classrooms with different acoustics were considered: classroom A, with the acoustic treatment introduced above, and classroom B, without any acoustical treatment. Table 1 shows the main geometrical characteristics and the features of the two classrooms.

Table 1. Geometrical characteristics and features of the classrooms

	Classroom A	Classroom B
Volume [m³]	171	282
School floor	First	Ground
Space faced by the room	Low-trafficked street	Low-trafficked street
Acoustic treatment	Yes	No

In both classrooms, acoustic parameters were measured in unoccupied conditions, in compliance with UNI EN ISO 3382-2 standard [24] applying the integrated impulse response method. A sweep signal that was emitted in the room by a Head and Torso Simulator (HaTS, model 4128 by Brüel&Kjær) that was used as signal generator, whereas an omnidirectional calibrated microphone (model XL2 by NTi Audio) was used as a receiver. The main parameters that were extracted from the impulse responses were the reverberation time (T30, s), the clarity (C50, dB) and the early decay time (EDT, s). For T30 and EDT measurements were performed with two sources and six microphone positions, and the results were averaged in order to obtain a mean spatial value; C50 was measured in the central position of the room. Frequency averaging in the range of 0.250÷2 kHz for T30 and in the range of 0.5÷1 kHz for EDT and C50 were calculated based on the compliancy with the German standard DIN 18041 [25] and the UNI EN ISO 3382-2 standard [24], respectively. The noisiness of the classrooms was also measured, and was

evaluated as the A-weighted equivalent background noise level ($L_{A,eq}$, dBA) was monitored during the measurement session. Table 2 reports the results from the classroom acoustic measurements that were considered to characterise the rooms.

Table 2. Measured classroom acoustic parameters (values in italic represent those that agree with the optimal values reported in the reference standards)

	Classroom A	Classroom B
Reverberation time $T30_{0.25-2kHz}$ (s) Optimal value by reference [25]: volume dependent	<i>0.41</i>	3.47
Clarity in the central position $C50_{0.5-1kHz}$ (dB) Optimal value by reference [24]: > 0 dB	9.52	<i>-4.87</i>
Early decay time $EDT_{0.5-1kHz}$ (s) Optimal value by reference [24]: 0.30 ÷ 0.70 s	<i>0.30</i>	2.95
Noise level $L_{A,eq}$ (dBA) Optimal value by reference [26]: ≤ 35 dBA	29.0	48.9

2.2 Measurement setup in-field for the listening tests preparation

Figures 1 and 2 show the pictures of the selected classrooms A and B, respectively, and the spatial positioning of receiver, noise-source and speech-source. The receiver consisted in the same Head and Torso Simulator (HaTS, model 4128 by Brüel&Kjær) introduced in the previous paragraph, the speech-source consisted in a TalkBox (by NTi Audio) that has the same polar directivity diagram of the human voice, and the noise-source consisted in an omnidirectional dodecahedron (by Brüel&Kjær). Measurements were performed at the end of the school year when desks and chairs were moved for a general cleaning of the rooms, therefore only shelves on the walls were present at that time and it should be considered that all the following results are related to fully unoccupied classroom condition.

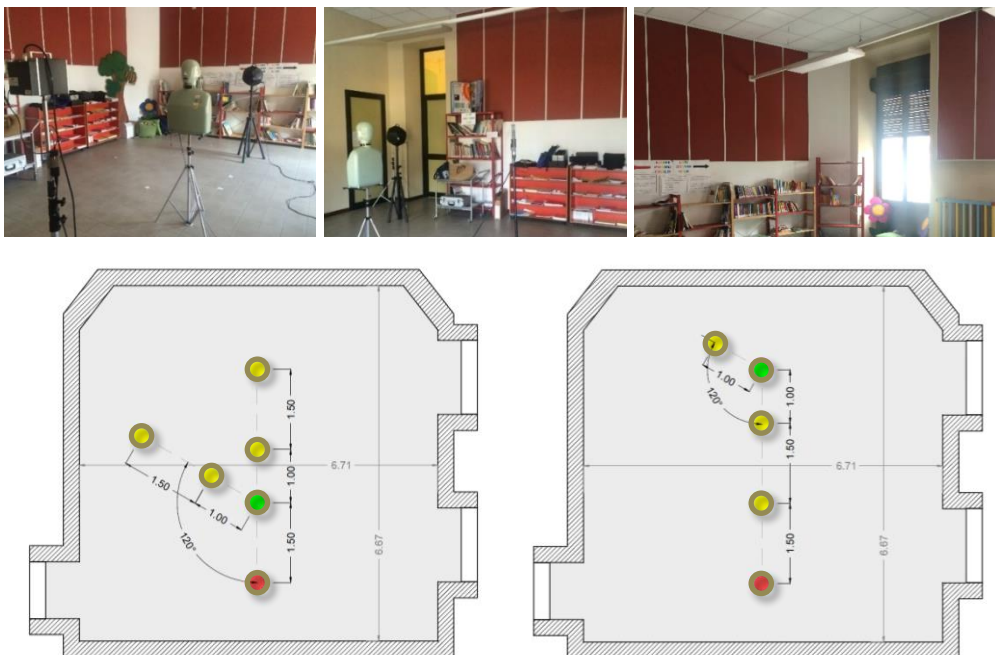


Figure 1. Measurement set-up in room A (good acoustics). The dots indicate the position of the speech-source (red), of the receiver (green) and of the separate noise-sources (yellow)

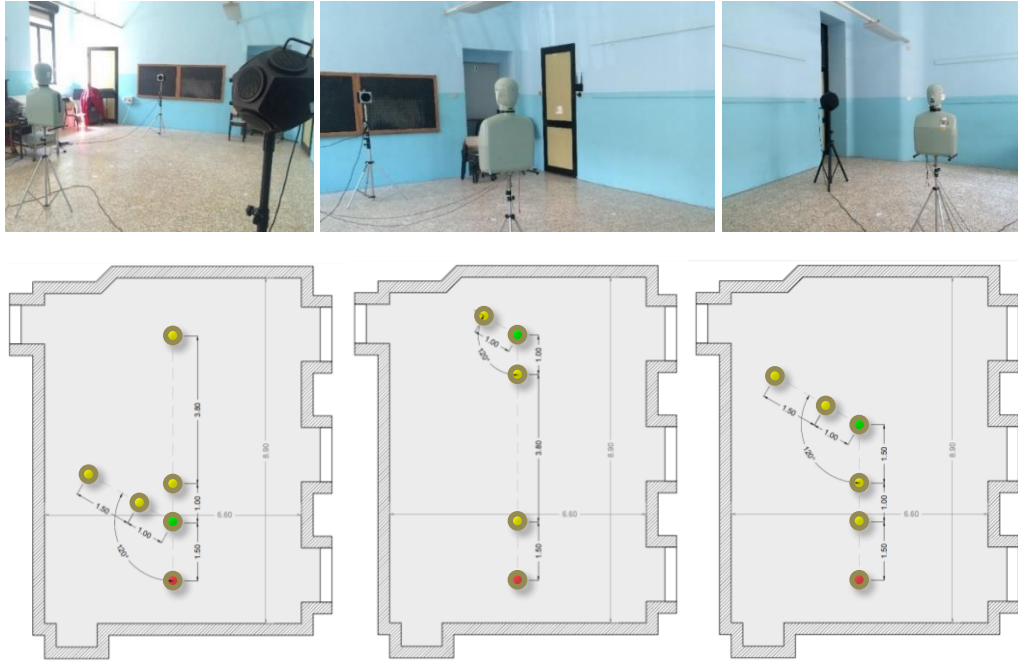


Figure 2. Measurement set-up in room B (bad acoustics). The dots indicate the position of the speech-source (red), of the receiver (green) and of the separate noise-sources (yellow)

In all the experiments, the receiver position was fixed in axis with the speech-source, as to represent the situation of students listening to the teacher in axis and from the central part of the room. In particular, the receiver was moved at two distances from the speech-source in room A (good acoustics), i.e. at 1.5 m and at 4 m, and at three distances room B (poor acoustics), i.e. at 1.5 m, 4 m and 6.3 m. The noise-source was moved around the receiver's head at several angles and distances, in order to investigate on the spatial release from masking due to the binaural cues in the cocktail party phenomenon, namely at 0° , 120° and 180° .

2.3 Listening tests design and administration

Listening tests were based on the procedure and material described in Puglisi et al. [22], particularly on the open-set format. To ease the duration of the tests and to avoid deconcentration bias, a short version of the test was used that was previously evaluated and optimized to make results as accurate as in the extended version [27]. In summary, the test consists in the administration of syntactically correct but semantically unpredictable 3-words sentences randomly built from a 7×3 matrix of words (verbs, numerals, names). Such sentences, which were uttered in anechoic conditions, were then convolved with the room impulse responses acquired in the different positions and classrooms in order to obtain ecological sentences. The processed sentences could then be delivered to listeners who had to repeat aloud the words s/he understood, and an experimenter could check on a laptop the correctly heard words of the sentence.

Based on such a methodology, speech intelligibility was evaluated in terms of Speech Reception Thresholds (SRTs) with an adaptive procedure converging to signal-to-noise ratio to yield 80% correct recognition scores (SRT80 expressed in dB SNR). To clarify, SRTs can be considered as the signal to noise ratio at which a sentence is to be said in order to make a listener understand a fixed percentage of the entire speech (e.g. in the case of SRT80, the search is for the 80% of understanding). In summary, to this work's aim the measurement conditions can be listed as follows:

- In room A (good acoustics)
 - SRT to yield 80% of speech intelligibility
 - Fixed noise level (60 dB)
 - 2 receiver positions on axis with the speech source
 - 7 noise-source positions based on the most and less advantageous release from masking conditions
- In room B (poor acoustics)
 - SRT to yield 80% of speech intelligibility
 - Fixed noise level (60 dB)
 - 3 receiver positions on axis with the speech source
 - 11 noise-source positions based on the most and less advantageous release from masking conditions

The listening tests were performed in the anechoic room of Politecnico di Torino, in the Department of Energy, where listeners were provided with headphones (model HDA200 by Sennheiser). Forty-three volunteer listeners (mean age equal to 28 ± 6 years) were involved in total. On average, the tests lasted about 35 minutes with a short break in-between. All the subjects were self-reported normal hearing; however, a short test for the assessment of the hearing threshold was performed before each listening test using the smartphone app uHear (version 2.0.2). The administration of the listening tests was organized in two main steps: the first part was oriented to familiarise with the test procedure and was self-conducted by the listener; the second part was experimenter-conducted (i.e. an experimenter was in the anechoic room together with the listener to acquire the answers) and was oriented to the collection of the results as the ITAMatrix test was presented in the open-set format.

As introduced above, the main outcome of the listening tests consisted in the acquisition of the SRTs for each noise/reverberation condition (i.e. separately in the two rooms, considering both IM and EM) and for each spatial configuration. Then, based on the acquired absolute values of SRTs, the Speech Release from Masking (SRM) could be calculated to evaluate the spatial benefits of speech and noise separation in angle and in distance under the different acoustic conditions. In such a way, variations in speech intelligibility under different acoustics and noise-maskers could be assessed.

3. RESULTS AND CONCLUSIONS

Starting from the acquired SRTs, the actual release from masking could be evaluated under different masking noises and acoustics in the spatial conditions that are represented in Figure 3. SRM was calculated as the difference between an SRT value measured with noise in the colocated position and the SRT value measured with noise in the separated position, as suggested in Westermann and Buchholz [28]. With such a calculation, higher values of SRM indicate a better condition, thus a major benefit in the spatial separation of noise.

Although the normalized error calculation [29] reveals a non-significant difference between SRTs measured in the colocated vs separated condition, SRM can be considered as tendencies that demonstrate strong detrimental effects of noise on speech intelligibility when its path is oriented in the same direction of the target to be understood. Then, in highly reverberant conditions (i.e. room B) there is a tendency to have an effect of spatial separation of speech and noise only when an informational masker is present. Interestingly, moving the noise source in an advantaged position (120°) enhances speech intelligibility for informational masker, particularly in the case of high reverberation and

close noise-listener distance. In good classroom acoustics (i.e. room A) there is a tendency to have spatial release from masking (≈ 3 dB) under the presence of an energetic masker.

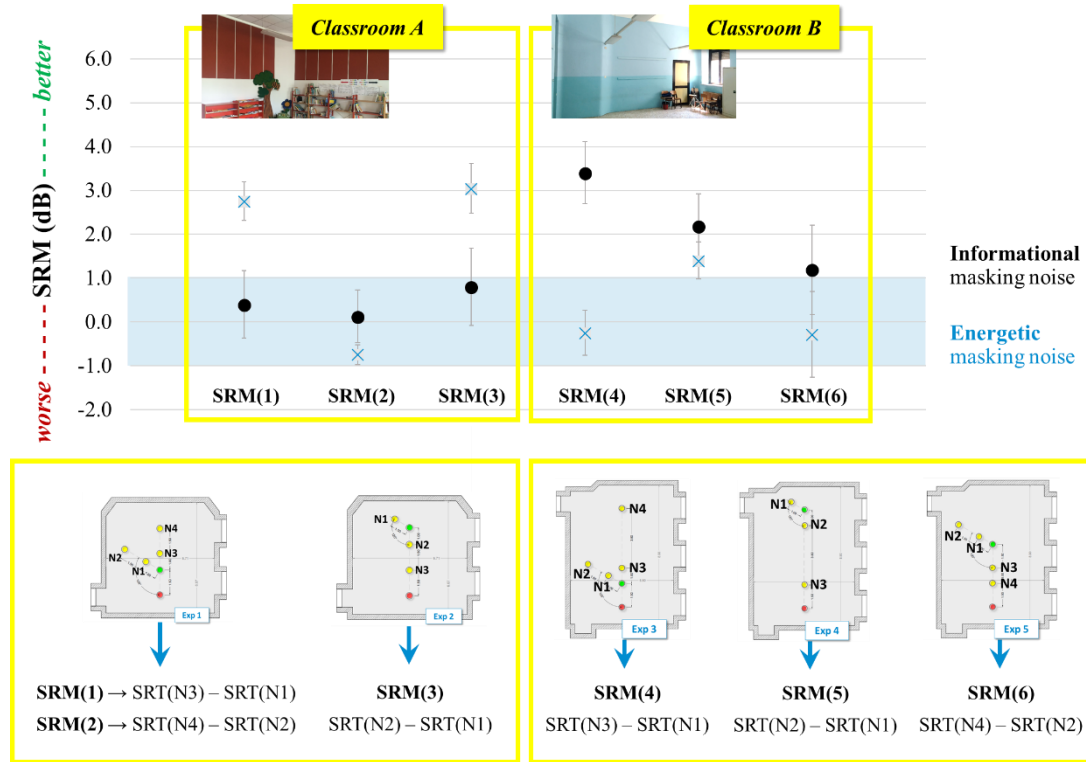


Figure 3. Results of Speech Release from Masking (SRM) under different spatial conditions and acoustics, for both types of speech-maskers

The presented preliminary results show that lower (better) SRT80s were measured under shorter reverberation time conditions, indicating the detrimental effect of reverberation on speech intelligibility, as expected. A major effect of the reflections was also proved in the increase of SRT80s when the speaker-to-listener distance increased. These results may help in the acoustic design of classrooms as they are an evidence to support the need of reducing reverberation in learning environments as it degrades speech intelligibility. Furthermore, geometrical and architectural features in the design of classrooms need to be accounted for in the designing process as they can negatively enhance detrimental effects (e.g. too big classrooms imply greater speaker-to-listener distances, which were seen as worst conditions).

With respect to the effect of masker type on intelligibility, so far unforeseen SRM was found under very low reverberation with energetic masking, and under very high reverberation with informational masking. Also, a large variability across listeners was observed when considering the difference between colocated and separated noise, especially under informational noise. Such aspects still need to be deepened to understand the mechanisms underlying speech perception in real complex auditory scenes.

5. ACKNOWLEDGEMENTS

This work is part of a cooperation between the Applied Acoustic group of Politecnico di Torino and the Cluster of Excellence “Hearing4All” of the University of Oldenburg. The authors would like to thank “Scuola primaria Fontana” in Torino (Italy)

for making the classrooms available for the measurements, and all the listeners that have taken part to the tests with passion and enthusiasm. Special thanks are extended to Filippo Bolognesi for his participation in the listening tests administration.

6. REFERENCES

- [1] A. Astolfi, P. Bottalico and G. Barbato, “*Subjective and objective speech intelligibility investigations in primary school classrooms,*” J. Acoust. Soc. Am. 131(1), 247-257 (2012)
- [2] N. Prodi, C. Visentin and A. Feletti, “*On the perception of speech in primary school classrooms: Ranking of noise interference and of age influence,*” J. Ac. Soc. Am. 133(1), 255-268 (2013)
- [3] L. C. Cantor Cutiva, G. E. Puglisi, A. Astolfi and A. Carullo, “*Four-day-follow-up study on the self-reported voice condition and noise condition of teachers: Relationship between vocal parameters and classroom acoustics,*” Journal of Voice 131(1), 120.e1-120.e8 (2017)
- [4] G. E. Puglisi, L. C. Cantor Cutiva, A. Astolfi and A. Carullo, “*Four-day-follow-up study on the voice monitoring of primary school teachers: Relationships with conversational task and classroom acoustics,*” J. Ac. Soc. Am. 141(1), 441-452 (2017)
- [5] C. Visentin, N. Prodi, and F. Cappelletti, S. Torresin and A. Gasparella “*Using listening effort assessment in the acoustical design of rooms for speech,*” Building and Environment 136, 38-53 (2018)
- [6] G. E. Puglisi, A. Prato, T. Sacco and A. Astolfi, “*Influence of classroom acoustics on the reading speed: A case study on Italian second-graders,*” J. Ac. Soc. Am. EL 144(2), EL144-EL149 (2018)
- [7] B. Kollmeier, A. Warzybok, S. Hochmuth, M.A. Zokoll, V. Uslar, T. Brand and K.C. Wagner, “*The multilingual matrix test: Principles, applications, and comparison across languages: A review,*” Int. J. Audiol. 54(2), 3-16 (2015)
- [8] Y.-J. Choi, “*Effects of periodic type diffusers on classroom acoustics,*” Applied Acoustics 74, 694-707 (2013)
- [9] G. E. Puglisi, F. Bolognesi, L. Shtrepi, A. Warzybok, B. Kollmeier and A. Astolfi, “*Optimal classroom acoustic design with sound absorption and diffusion for the enhancement of speech intelligibility,*” proceedings of Acoustics '17 – Boston (2017)
- [10] J. RENNIES, T. Brand and B. Kollmeier, “*Prediction of the influence of reverberation on binaural speech intelligibility in noise and in quiet,*” J. Ac. Soc. Am. 130(5), 2999-3012 (2011)
- [11] E. C. Cherry, “*Some experiments on the recognition of speech, with one and with two ears,*” J. Acoust. Soc. Am. 25(5), 975-979 (1953)
- [12] A. W. Bronkhorst, “*The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions,*” Acta Acustica United with Acustica 86, 117-128 (2000)
- [13] A. W. Bronkhorst, “*The cocktail-party problem revisited: Early processing and selection of multi-talker speech,*” Atten. Percept. Psychophys. 77, 1465-1487 (2015)
- [14] D. D. Dirks R. H. and Wilson, “*The effect of spatially separated sound sources on speech intelligibility,*” J. Speech Hear. Res. 12, 5-38 (1969)

- [15] H. J. Platte and H. Vom Hövel, “Zur Deutung der Ergebnisse von Sprachverständlichkeitsmessungen mit Störschall im Freifeld,” *Acustica* 45, 139-150 (1980)
- [16] R. Plomp and A. M. Mimpen, “Effect of the orientation of the speaker’s head and the azimuth of a noise source on the speech reception threshold for sentences,” *Acustica* 48, 325-328 (1981)
- [17] A. J. Duquesnoy, “Effect of a single interfering noise or speech source on speech intelligibility,” *J. Acoust. Soc. Am.* 74, 739-743 (1983)
- [18] A. W. Bronkhorst and R. Plomp, “The effect of head-induced interaural time and level differences on speech intelligibility in noise,” *J. Acoust. Soc. Am.* 83, 1508-1516 (1988)
- [19] J. Peissing and B. Kollmeier, “Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners,” *J. Acoust. Soc. Am.* 101, 1660-1670 (1997)
- [20] I. Pollack, “Auditory informational masking,” *J. Acoust. Soc. Am.* 57(1), S5 (1975)
- [21] G. Kidd, C. R. Mason, A. Brughera and W. M. Hartmann, “The role of reverberation in release from masking due to spatial separation of sources for speech identification,” *Acust. Acta Acust.* 91, 526-536 (2005)
- [22] G.E. Puglisi, A. Warzybok, S. Hochmuth, C. Visentin, A. Astolfi, N. Prodi and B. Kollmeier, “An Italian matrix sentence test for the evaluation of speech intelligibility in noise,” *Int. J. Audiol.* 54(2), 44-50 (2015)
- [23] A. Astolfi and M. Giovannini, “*Acustica delle aule scolastiche – Requisiti prestazionali, soluzioni di progetto, verifiche a calcolo e in opera,*” edited by Rockwool® (2010)
- [24] UNI EN ISO 3382-2, “*Acoustics - Measurement of room acoustic parameters – Part 2: Reverberation time in ordinary rooms,*” International Organization for Standardization, Genève (2008)
- [25] DIN 18041, “*Hörsamkeit in Räumen - Anforderungen, Empfehlungen und Hinweise für die Planung*”, Deutsche Institut für Normung, Berlin (2016)
- [26] Building Bulletin 93, “*Acoustic design of schools: Performance standards,*” Department for Education and Education Funding Agency (2015)
- [27] A. Warzybok, E. S. Garbaruk, M. Goykhburg, Z. Merza, G. E. Puglisi, C. Montuschi, F. di Bernardino, D. Zanetti, F. Sellami, K. C. Wagener, I. Holube, A. Astolfi, R. Albera, P. A. Pavlov, G. Tavartkiladz and B. Kollmeier, “Evaluation of internationally compatible speech test in noise for the pediatric population,” proceedings of 13th Congress of the European Federation of Audiology Societies, *Journal of Hearing Science* 7(2), 46 (2017)
- [28] A. Westermann and J.M. Buchholz, “The effect of spatial separation in distance on the intelligibility of speech in rooms”, *J. Ac. Soc. Am.* 137(2), 757-767 (2015)
- [29] G. E. Puglisi, A. Astolfi, L. C. Cantor Cutiva and A. Carullo, “Assessment of indoor ambient noise level in school classrooms,” *Proc. Euronoise*, 715-720 (2015)