Euronoise 2018
Crete

Speech reception in kindergarten and primary school classrooms: effects of age and mother tongue

Chiara Visentin, Nicola Prodi
Dipartimento di Ingegneria, Università di Ferrara, Italia.

Alessandro Peretti
Scuola di Specializzazione in Medicina del Lavoro, Università di Padova, Italia.

Alberta Pierobon
Dipartimento di Scienze Cardiologiche, Toraciche e Vascolari, Università di Padova, Padova, Italia.

Jacopo Griguolo
Tecnico Competente in Acustica Ambientale, Rovigo, Italia.

Giorgio Marcuzzo
UOC Medicina Preventiva e Valutazione del Rischio, Azienda Ospedaliera di Padova, Italia.

Giovanni Battista Bartolucci
Dipartimento di Scienze Cardiologiche, Toraciche e Vascolari, Università di Padova, Padova, Italia

Summary

A correct hearing and understanding of the verbal message is an essential requirement in the learning process. When listening in presence of background noise or too long reverberation, a complex interplay of perceptual and cognitive processes has to be deployed: the youngest pupils, due to their still immature auditory processing, are especially vulnerable. The aim of the present study is to examine how adverse acoustic conditions modulate differences in speech reception performance during the lesson period, for children aged 5 to 7 having different mother tongue.

Word Identification by Picture Identification (WIPI) tests in the Italian language were proposed to 154 children. A subset of 33 sequential bilinguals children was identified, who started the acquisition of the Italian language after 3 years of age. The tests were conducted in four real classrooms, with reverberation time in occupied conditions varying between 0.43 and 1.31 s. In each room, the tests were presented at two listening conditions: ambient noise (SNR=15 dB) and speech-shaped stationary noise (SNR=0 dB). During the experiment data on the number of words correctly identified (speech intelligibility, IS) and on the response time (RT) were acquired for each participant. It was assumed that an increase in RT results would reflect greater cognitive processing. The results show that for 5 to 7 years old pupils the presence of noise always causes lower IS and longer RT with respect to a baseline condition. Under similar listening conditions a disadvantage was found for the youngest pupils (5 years) with respect to the oldest (6, 7 years), showing up as both an IS decrease and a RT increase. For the proposed task, the effect of the mother language was apparent for the youngest children alone, and was limited to the accuracy results.

PACS no. 43.55.Hy, 43.71.Gv

1. Introduction

Noise at school can be harmful to speech perception and listening comprehension of children [1], so much so that the sound environment in classrooms can be inappropriate for learning [2, 3]. Over the years, it has been shown that children's performance in diverse tasks is affected by noise and that, as a consequence, their academic attainments can be mined [4, 5]. Different trends according to age can be outlined and it was found that, in the same listening condition, younger pupils achieve a lower speech intelligibility (number of words correctly understood, IS) due to a less mature auditory system and cognitive development [6, 7].
In any case, when dealing with the problematic experience of listening in the classrooms, besides IS also the relevant dimension of listening-related effort should be considered. Listening-related effort was recently defined as "the level of processing resource allocation to overcome obstacles in goal pursuit when performing a listening task" [8]. Several factors may require that an increased amount of cognitive resources is called for during speech reception, either related to the listener (e.g. hearing impairment, language proficiency, etc.) or to the transmission chain (e.g. suboptimal listening conditions). As a result, an outcome of effort is often elicited [9, 10], especially so in presence of sustained listening demands, as during lessons.
Several methods have been proposed over the years to qualify the listening effort construct (see [8] for a review). Similarly to previous works investigating effort in children [7, 11, 12], in this study the behavioral measure of response time (RT) implemented with a single task paradigm was used. The underlying assumption is that single-task response time is effective in tracing working memory operations involved with speech processing in children [13]. Then, RT can be assumed as a behavioral marker of effort: a longer response time would indicate an increase of cognitive load and hence of processing effort.
The single task approach presents advantages with respect to the dual task paradigm, where a primary task is paired with a secondary task, whose performance is monitored and used to describe the effort [14]. First, concerns on the reliability of dual-task paradigms for children have been raised [15] due to the difficulty of pupils to control the allocation of resources according to instructions. Then, the single-task paradigm allows the implementation of the approach in the framework of field experiments inside real classrooms, allowing for an ecological evaluation. It has to be noticed that the main shortcoming of the single task approach is that the attention of participants might depend on the task difficulty itself, so that the relationship between mental effort and response time could change depending on the experimental conditions [9, 10].
In the present study, listening effort resulting from an extended behavioral monitoring has been investigated under ecological classroom conditions with a panel of 5 to 7 years old normal hearing pupils. In this sense, the experiment extends the analysis outlined in [7] to a group of younger children, in presence of different background noises.

2. Materials and methods

2.1. Classrooms description and measurement set up

The study took place in Padova (Italy) in the second half of the school year, and involved two kindergartens (K1, K2) and two primary schools (P1, P2). The schools were located either in the city center (K2 and P1) or in a residential area in the outskirts of the city (K1 and P2). One classroom was selected in each school, to be used as a laboratory for the speech-in-noise tests and the acoustical measurements. All classrooms were box-shaped and characterized by flat surfaces. Here it is important to notice that whereas in schools K1, P1 and P2 the ceiling was acoustically treated, no treatment was present in school K2.
During the experiment, the classrooms were set up as for regular lessons with wooden desks and chairs, closets and shelves. A Grass 44AB mouth simulator was placed at the teacher's desk, 1.5 m high from the ground and used to deliver the speech signal. An additional loudspeaker (Quested® F11) was used to playback interfering noise during the experiment. It was placed on the floor, on the side of the mouth simulator, facing in the opposite direction of the audience to minimize the impact of the direct sound and obtain a background noise with no recognizable directional characteristics. Two measurement positions were defined in the audience area, at the center of two of the four ideal sectors in which the area was divided. Two

omnidirectional B&K type 4189 1/2 inch microphones were set at a height of 1.1 m and used for the objective description of the acoustical conditions during the experiment.

2.2. Participants

The experiment was presented to kindergarten and primary school children, normal hearing and aged 5 to 7 years old. The group of 5 years old children (5Y in the following) was composed by pupils attending their last year of kindergarten whereas the group of older children (6Y and 7Y) was composed by pupils attending their first or second grade at primary school. Written consent forms granting permission for children to participate in the experiment were obtained from the parents.

The children's linguistic background was investigated using a questionnaire, compiled by either teachers or parents. The answers pinpointed that schools K2 and P1 were mainly attended by pupils of Italian mother tongue (L1), which represented the 88.5% of the students. For these schools only L1 students were then considered in the data analysis. On the other hand, in schools K1 and P2 a large group of pupils (61.1%) spoke Italian as a second language (L2). In these cases, both linguistic groups were considered in the analysis. The characteristics of the children participating in the experiment are detailed in Table I.

Table I. Children participating in the experiment dived according to school, age and mother tongue (L1: native, L2: non-native).

School	Age	# of participants	# L1	# L2
K1	5Y	14	6	8
K2	5Y	18	18	-
P1	6Y	43	43	-
	7Y	39	39	-
P2	6Y	21	9	12
	7Y	19	6	13
All		154	121	33

2.3. Speech material

Speech reception was assessed using the Word Intelligibility by Picture Identification (WIPI) test, in its implementation in the Italian language named TIPI [16]. It is a closed-set, picture-pointing test, based on 96 disyllabic word belonging to the everyday vocabulary of a 4 years old child and that are graphically represented with hand-drawn pictures. The words are organized in 50 groups of six-items.

For the experiment, sentences composed by a carrier phrase and a target word were created (e.g. "Ora diremo la parola *barba*" which is Italian for "Now we will say the word *beard*"). The sentences were recorded by an adult, native Italian, female speaker in a silent room. The recordings were then edited as to match the long-term spectrum of a female speaker [17] and set to the same root mean square value. The recordings were then organized into five lists of 10 words each.

2.4. Listening conditions

For the experiment, two listening conditions were set, named "baseline" (B) and "working classroom" (W). In both conditions the speech signal was calibrated to a level of 63 dB(A), measured at 1 m in front of the speech source. The level corresponds to a speaker talking with a vocal effort intermediate between "normal" and "raised" [18]. In condition B only the actual ambient noise of the classrooms, primarily made up by noises coming from the adjacent classrooms, where pupils were engaged in quiet activities (teaching time or individual work) was present. On the contrary, in condition W, a stationary speech-shaped noise was played back by the additional loudspeaker, producing an energetic masking of the signal. The noise level was calibrated similarly to the speech signal to obtain a signal-to-noise ratio (SNR) of +3 dB at 1 m in front of the mouth simulator.

The objective description of the listening conditions during the experiment was achieved by measuring the reverberation time and the long-term levels of the ambient noise, the speech signal, and the stationary noise (reproduced as proposed during the tests) at the end of the experiment, within the occupied classroom. The listening conditions are summarized in Table II.

Table II. Listening conditions in the four schools (K1, K2, P1, P2) in the two listening conditions (B: baseline, W: working classroom): reverberation times ($T_{20,mid}$ averaged across 500–2000 Hz octave bands), signal-to-noise ratios (SNR), speech transmission index (STI).

School		K1	K2	P1	P2
$T_{20,mid}$ [s]		0.43	1.31	0.54	0.52
SNR (dB)	B	+17.6	+16	+15.7	+15.8
	W	-0.1	+0.8	+0.3	-1.8
STI	B	0.73	0.56	0.73	0.71
	W	0.47	0.42	0.53	0.47

It has to be noted that, even though in condition B the measured SNRs were always greater than the

limit of +15 dB defined in [19], owing to the presence of acoustical treatment, the resulting sound environments differed between the classrooms. Then, the measured STI values corresponded to an intelligibility rated as “Good” (K1, P1 and P2) or only “Fair” (K2) [18].
Condition B is the most comfortable that pupils could experience within the classroom during lessons, and hence it was assumed as the baseline, against which the condition with stationary noise added was compared. This latter condition was set as to reproduce a working classroom, where the pupils carry out quiet educational activities with their teachers [20].

2.5. Procedures

The experiment took place in whole-class groups. Upon entering the laboratory classroom, each child was given a touchscreen handset, to be used for response selection, and randomly assigned a seating position. Then, the experiment was presented to the class, displaying the speech material and getting practice with the response system. A wireless test bench was used to manage the experiment [7]. The children listened to a target word embedded in the carrier phrase; at the offset of the audio playback, six pictures were displayed on the touchscreen and the children were instructed to select the picture corresponding to the word they heard. A maximum of 20 s was allowed to give the response and only once all participants responded (or reached the timeout) the next target word was automatically presented.
A training session was firstly proposed; afterwards, they completed four listening tests (2 test lists x 2 listening condition). The listening conditions were alternated during the presentation, blocking the order within the repetitions. Presentation order and test lists were counterbalanced across the classes. The entire experimental session, from the training phase to the final listening test, lasted approximately 30 minutes. The children were instructed to pay attention to the task, and asked to respond as accurately as possible. They were not informed that RT data were also acquired during the experiment nor they were given any recommendation to respond as quickly possible.
The data collected during the experiment were the picture scores and the response times, defined as the time elapsed between the audio offset and the response selection on the touchscreen.

2.6. Statistical analysis

Data analysis was performed using generalized mixed-effects models (GLMM), chosen on the account of the repeated measures design of the experiment and the not-normal distribution of the response variables.
The software R was used for the analysis (packages *lme4*, *lsmeans*); a significance level α=0.05 was always set. In particular, a GLMM with a binomial distribution was used to analyze IS data, whereas RT results were analyzed using a Gamma distribution with a log-link function. Model selection was based on a forward procedure using the likelihood ratio test, and the statistical assumptions of the final models were verified by checking the normality of the residuals. When appropriate, planned pairwise comparisons were performed, correcting for the test multiplicity using a Benjamini-Hochberg procedure.
Prior to data analyses, RT data greater than 10 s were removed from the dataset (1.2% of the data) as possibly due to participants’ inattention.

3. Results

Due to the differences in the room acoustics conditions realized within the classrooms, for the analysis of the results three separate models were set up; in all of them child was considered a random factor.
The first model comprised data of school K2 and investigated the effect of listening conditions on 5Y, L1 pupils. The second model, including data from school P1, investigated the effects of age and listening condition (and their interaction) for 6Y and 7Y, L1 pupils. Finally, the third model grouped the data from schools K1 and P2 and included age, mother tongue, listening conditions and their interactions as fixed factors.
The results of the statistical analyses are summarized in Table III.

4. Discussion

The aim of this study was related to the presence or not of markers of listening effort in 5 to 7 years old pupils. In particular, the behavioral measure of RT was compared from B to W conditions, and an increase in RT results was taken as an indicator of a heavier processing load.

Table III. Results of the statistical analyses for the three models including: (a) 5Y, L1, (b) 6Y and 7Y, L1, and (c) 5Y and 6Y-7Y, both L1 and L2. The dashes indicate that the corresponding effect is not significant; only significant interactions are reported. *p<0.05, **p<0.01, ***p<0.001.

Model	Effect	IS	RT
5Y, L1	Listening condition	$IS_B > IS_W$ ***	$RT_B < RT_W$ **
6Y, 7Y L1	Listening condition	$IS_B > IS_W$ ***	$RT_B < RT_W$ ***
	Age	-	-
5Y, (6Y,7Y) L1 and L2	Listening cond. x Age	$IS_B > IS_W$ ***	5Y: - 6Y, 7Y: $RT_B < RT_W$ ** A: $RT_{5Y} > RT_{6Y,7Y}$ ** SSN: -
	Mother lang. x Age	5Y: $IS_{L1} > IS_{L2}$ * 6Y, 7Y: - L1: - L2: $IS_{5Y} < IS_{6Y,7Y}$ *	-

4.1. Effects of conditions: 5 years old pupils

As regards the group of the L1, 5Y pupils (school K2) the descriptive statistics of the collected data are showed in Figure 1. The statistical analysis returned a significant main effect of listening condition for IS indicating, as expected, a large decrease in task accuracy when noise was played back. It is worth noticing that, owing to the absence of acoustical treatment, the 100% accuracy in the speech reception task was never reached by the 5Y children, even in the most favorable baseline condition. The IS change was paired by a significant increase in RT results: performing the speech reception task in presence of stationary noise required to the 5Y pupils an additional processing time of 301 ms, interpreted as an outcome of listening effort.

4.2. Effects of conditions: 6, 7 years old pupils

The descriptive statistics of the results of 6 and 7 years pupils of Italian mother tongue are reported in Figure 2.

First, no age effect was outlined in the statistical model, neither for IS nor RT. The result differs from previous findings: in [7], where a panel of 6-7 years old pupils was tested, older pupils were found to display higher IS results and faster RTs than 6Y, even in the quiet condition. In that study the same equipment, procedure and speech material

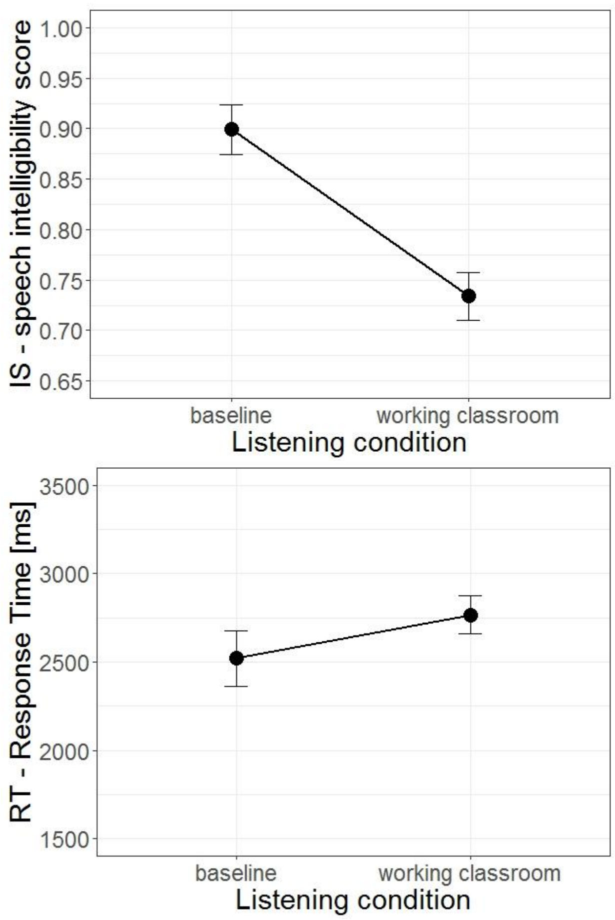

Figure 1. Mean results (IS and RT) of the speech-in-noise tests in school K2, averaged across the 5 years old pupils, for the two listening conditions. The error bars represent the 95% confidence intervals between participants.

as in the present study was used but the background noises (traffic, activity, tapping) had different characteristics. In particular, the noises had both a changing state and a degree of salience [1] that helped in stretching the performance gap due to age. On the other hand, the present stationary noise had at the least such characteristics, and for this reason, no age effects showed up between 6Y and 7Y.

For both age groups the statistical analysis indicated a significant effect of the listening condition, showing that, disregarding the lesser disrupting potential of the stationary noise, its presence causes lower IS and longer RT with respect to the more favorable listening condition. In particular, a RT increase of 90 ms was found due to the presence of noise; the finding witnesses an increased involvement of cognitive resources even with a background noise having basic spectro–temporal features.

4.3. Effects of listening condition, age and mother tongue

The effects of age (5Y, 6Y and 7Y), mother tongue (L1, L2) and listening condition on IS and RT were assessed based on the data collected in schools K1 and P2, where the same listening conditions were realized. For the analysis, owing to absence of differences between 6 and 7 years old pupils (see Section 4.2), 6Y and 7Y were aggregated within a single group. The descriptive statistics are represented in Figure 3.

First, as regards the effect of listening condition, it was found that all children (independently on the age) achieved greater IS results in condition B than in condition W. When data were averaged across mother tongue, the IS gap was equal to 8.8% for 5Y and to 13.5% for the oldest pupils. The RT metric returned a more complex patter than IS, showing a significant interaction between listening conditions and age of participants. In fact, no difference between the listening conditions was found for the youngest pupils, whereas the primary school children showed a significant RT increase of 435 ms when noise degraded the speech signal. Furthermore, 6Y-7Y pupils responded significantly faster than 5Y in the baseline condition; the finding indicated higher speed of information processing for the oldest pupils, stemming for the auditory and cognitive development. The difference was not carried out in condition W, where similar RT results were found for the two groups of children. The finding could be interpreted as a better ability of 5Y

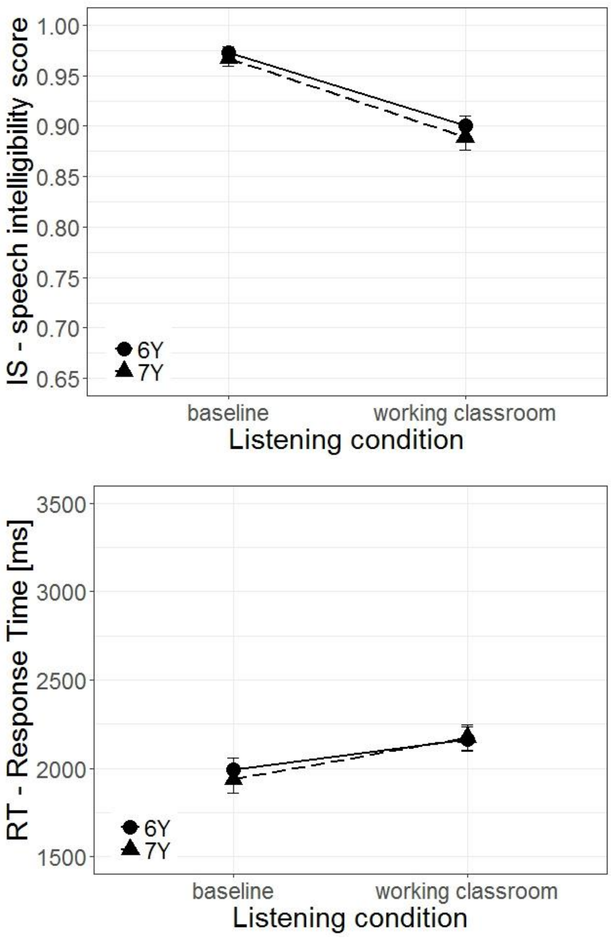

Figure 2. Mean results (IS and RT) of the speech-in-noise tests in school P1, averaged across pupils, for the two listening conditions; the results are divided according to age (6 vs. 7 years). The error bars represent the 95% confidence intervals between participants

to cope with more challenging listening conditions. The youngest children due to a daily exposure to noisier environments [21] would indeed develop strategies for tuning out the disturbing noise, thus achieving the same RT results as in baseline conditions. Alternatively, it could be hypothesized that in the more challenging condition W, 5Y pupils put higher attention on the speech reception task; the increased arousal would then alter the relationship between mental effort and response time, masking the effect of the listening conditions. However, it has to be noticed that, due to the small number of children participating in the experiment in schools K1 and P2, the statistical analysis may have a limited power in detecting the effects of the factors here considered.

Finally, the effect of mother tongue was considered. Interestingly, the effect was disclosed by IS alone, indicating a significant interaction between age and mother tongue. In particular, the

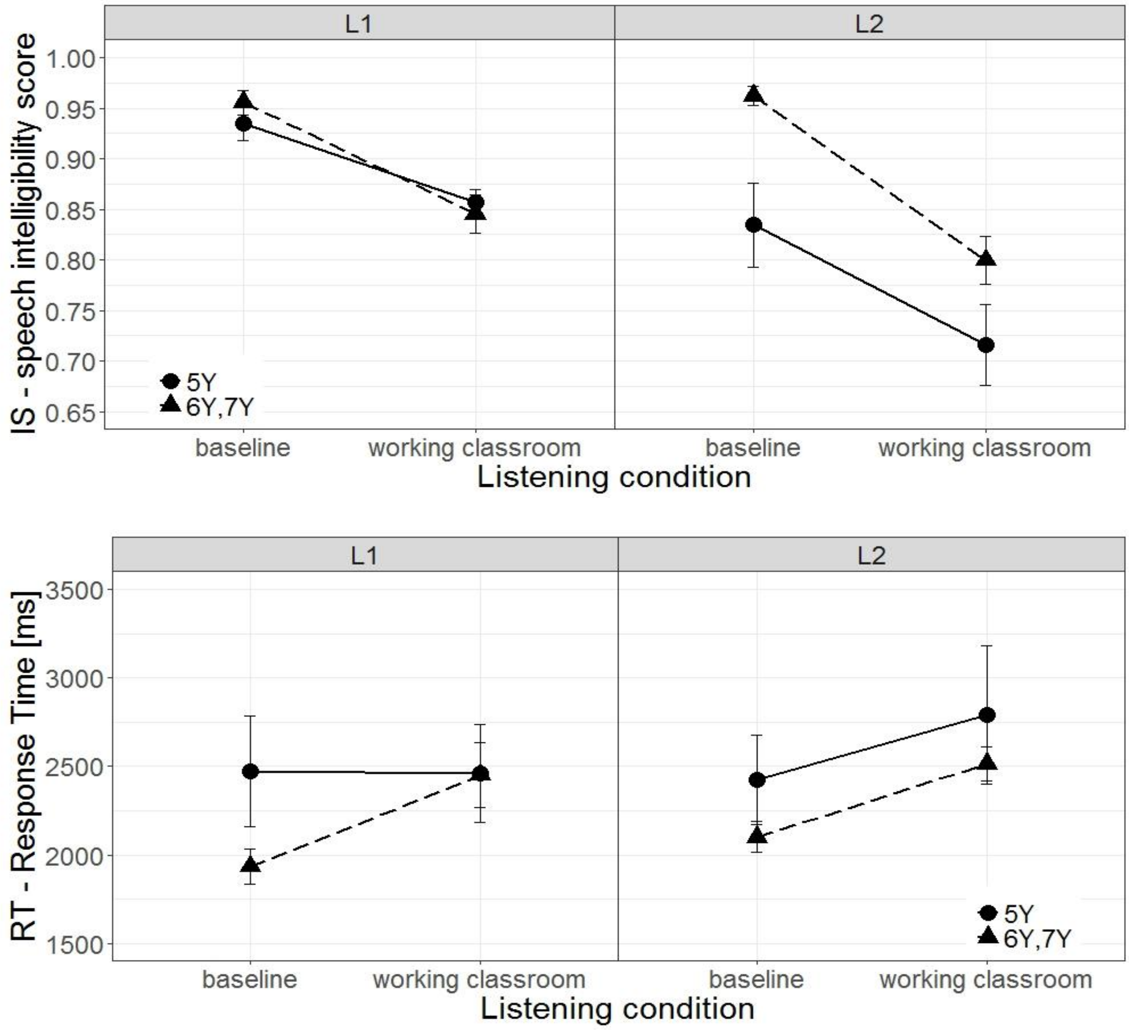

Figure 3. Mean results (IS and RT) of the speech-in-noise tests in schools K1 and P2, averaged across pupils, for the two listening conditions; the results are divided according to age (5 *vs.* 6, 7 years) and mother tongue (L1: native, L2: non-native listeners). The error bars represent the 95% confidence intervals between participants.

results indicate a disadvantage of L2 listeners on the accuracy results only for the youngest pupils (ΔIS=-11.2%), due to an inaccurate perceptual processing of the non-native words. The absence of differences between L1 and L2 listeners for the group of older pupils might suggest that, for the specific task here selected (based on simple words belonging to the everyday vocabulary of a 4 years old child) the accuracy gap is already recovered at the end on the first grade of primary school.

On the contrary, the RT metric did not show any significant effect of the pupils' mother tongue, which was instead expected on the basis of the interference from the native language of L2 on the lexical or phonetic processing of the words. The L2 listeners participating in the experiment can be generically defined as "early sequential bilingual", that is children who learn a language from birth, and then begin to learn a second language during early childhood. However, differences exist within this heterogeneous group, related to the individual proficiency of the L2 pupils in the Italian language. This aspect was not assessed in the present study (e.g. using vocabulary testing) but potentially it could be relevant for explaining part of the variability observed in the RT results of L2 participants.

5. Concluding remarks

Based on the results of the study, three main observations can be made:

(1) The single-task response time paradigm is effective in tracing changes in the cognitive load of the participants, and it can thus be considered informative of the listening effort. The metric can be successfully employed even with the youngest pupils to investigate the effects of the listening

conditions within the classrooms. In particular, it was found that the RTs significantly slowed down when a stationary noise was added, witnessing an increased involvement of cognitive resources compared to a baseline condition. The finding is relevant in the context of noisy classrooms, as the need to exert greater processing resources may exhaust children's energy and, besides short-term effort, eventually lead to mental fatigue, decreased attention and poorer academic performance.

(2) With reference to the specific background noise, no differences were found in the speech reception performance of 6 and 7 years old pupils. On the contrary, an increase of RTs was found for the youngest children in the baseline condition.

(3) As regards the effects of mother tongue, it was observed that the youngest L2 listeners had a disadvantage on the perceptual level, which was not observed for the older pupils. The effect was not paired by changes in the RT results, probably due to individual differences in the non-native language proficiency.

Acknowledgement

The Authors would thank the teacher and the pupils of the schools participating in the study.

References

[1] M. Klatte, K. Bergström, T. Lachmann: Does noise affect learning? A short review on noise effects on cognitive performance in children. Front. Psychol. 4 (2013).

[2] K. Anderson: The problem of classroom acoustics: the typical classroom soundscape is a barrier to learning. Semin. Hear. 25 (2004) 117–129

[3] D. L. Valente, H. M. Plevinsky, J. M. Franco, E. C. Heinrichs-Graham, D. E. Lewis: Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children. J. Acoust. Soc. Am. 131 (2102) 232–246.

[4] B. M. Shield, J. E. Dockrell: The effects of noise on children at school: a review. Building Acoustics 10 (2003) 97–116.

[5] J. E. Dockrell, B. M. Shield: Acoustical barriers in classrooms: The impact of noise on performance in the classroom. Br. Educ. Res. J. 32 (2006) 509–525.

[6] L. Leibold, A. Y. Bonino, E. Buss, Masked speech perception in infants, children and adults. Ear Hear 37 (2016) 345–353.

[7] N. Prodi, C. Visentin, A. Feletti: On the perception of speech in primary school classrooms: Ranking of noise interference and of age influence. J. Acoust. Soc. Am. 133 (2013) 255–268.

[8] M. K. Pichora-Fuller, S. E. Kramer, M. A. Eckert, B. Edwards, B. W. Hornsby, L. E. Humes, et al.: Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). Ear Hear. 37 (2016) 5S-27S.

[9] F. H. Bess, B. W. Hornsby: Commentary: Listening can be exhausting—Fatigue in children and adults with hearing loss. Ear Hear. 35 (2014) 592–599.

[10] R. McGarrigle, K. J. Munro, P. Dawes, A. J. Stewart, D. R. Moore, J. G. Barry, S. Amitay: Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'white paper'. Int. J. Audiol. (2014).

[11] S. Gustafson, R. McCreery, B. Hoover, J. G. Kopun, P. Stelmachowicz: Listening effort and perceived clarity for normal hearing children with the use of digital noise reduction. Ear Hear 35 (2014) 183–194.

[12] D. Lewis, K. Schmid, S. O'Leary, J. Spalding, E. Heinrichs-Graham, R. High: Effects of noise on speech recognition and listening effort in children with normal hearing and children with mild bilateral or unilateral hearing loss. J. Speech Lang. Hear. Res. 59 (2016) 1218–1232.

[13] N. Cowan, J. N. Towse, Z. Hamilton, J. S. Saults, E. M. Elliott, J. F. Lacey, ... G. J. Hitch: Children's working-memory processes: a response-timing analysis. J. Exp. Psychol. Gen. 132 (2003) 113–132.

[14] J. P. Gagné, J. Besser, U. Lemke: Behavioral assessment of listening effort using a dual-task paradigm: A review. Trends Hear. 21 (2017) 1–25.

[15] S. Choi, A. Lotto, D. Lewis, B. Hoover, P. Stelmachowicz: Attentional modulation of word recognition by children in a dual-task paradigm. J. Speech Lang. Hear. Res. 51 (2008) 1042–1054.

[16] E. Arslan, E. Genovese, E. Orzan, M. Turrini: Valutazione della percezione verbale nel bambino ipoacusico: Test di Identificazione di Parole Infantili 1-2 (Evaluation of the verbal perception in hearing impaired children: Infant Words Identification Test 1-2). Ecumenica, Bari (Italy), 1997.

[17] IEC 60268-16:2011. Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index.

[18] ISO 9921:2003, Ergonomics – Assessment of speech communication

[19] H. Sato, J. S. Bradley: Evaluation of acoustical conditions for speech communication in working elementary school classrooms. J. Acoust. Soc. Am. 123 (2008) 2064–2077.

[20] B. Shield, J. E. Dockrell: External and internal noise surveys of London primary schools. J. Acoust. Soc. Am. 115 (2004) 730–738.

[21] G. W. Evans, S. J. Lepore: Nonauditory effects of noise on children: A critical review. Children's environments 10 (1993) 31–51.