Four-day-follow-up study on the voice monitoring of primary school teachers: Relationships with conversational task and classroom acoustics

Giuseppina Emma Puglisi, Arianna Astolfi, Lady Catherine Cantor Cutiva, and Alessio Carullo

Citation: J. Acoust. Soc. Am. **141**, 441 (2017); doi: 10.1121/1.4973805

View online: http://dx.doi.org/10.1121/1.4973805

View Table of Contents: http://asa.scitation.org/toc/jas/141/1

Published by the Acoustical Society of America

CrossMark

Four-day-follow-up study on the voice monitoring of primary school teachers: Relationships with conversational task and classroom acoustics

Giuseppina Emma Puglisi,[1,a)] Arianna Astolfi,[1] Lady Catherine Cantor Cutiva,[1] and Alessio Carullo[2]

[1]*Department of Energy, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy*

[2]*Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy*

(Received 17 June 2016; revised 9 November 2016; accepted 22 December 2016; published online 20 January 2017)

The present study has investigated the occupational voice use of 27 female primary school teachers over a four-day-follow-up. Sixty-one working-day voice samples were acquired with two contact sensor-based vocal analyzers in four schools with highly different classroom acoustics. The vocal parameters were compared with a conversational task that the teachers performed before each lesson and with the measured classroom acoustic parameters. The average equivalent sound pressure level at 1 m from the mouth, which refers to the teacher's vocal effort, and the voicing time percentage were 71.2 dB [standard error (SE) 1.0 dB] and 29%, respectively. The teachers' mean voice level and fundamental frequency were significantly higher in the occupational setting than in the conversational one, which is by 5.5 dB (SE 0.5 dB) and 50 Hz (SE 3 Hz), respectively. Higher voice levels were observed for higher background noise levels, at a rate of 0.53 dB/dB, and a tendency of the background noise to increase with increasing reverberation time was observed at a rate of 13 dB/s. An optimal reverberation time of 0.7 s was found to minimize the voice level, since teachers raised their voice at lower and higher reverberation times, the latter presumably due to higher background noise levels.

[http://dx.doi.org/10.1121/1.4973805]

[JFL] Pages: 441–452

I. INTRODUCTION

Voice is used by around one-third of the workers all over the world as a primary working tool.[1] The abuse of voice at work can be the cause of the onset of vocal pathologies at several levels, such as hoarseness, weak voice, sore throat, aphonia, nodules, and polyps.[2] The professional category of teachers, of any grade or level, has been reported to be one of the categories most frequently affected by voice disorders due to a sustained and continuous use of voice during their working activity,[3] thus practical actions of prevention should be introduced to preserve their vocal health. Since teachers cover a large percentage of the working population, being in the 2%–6% range of the labor force in industrialized countries,[4–6] their vocal behavior needs to be monitored repeatedly during the working hours and under realistic environmental conditions to investigate whether any significant changes occur that could make the vocal conditions worse over a follow-up period. Moreover, in order to understand whether teachers change the way they use their voice during the teaching hours, the monitoring of conversational voice samples in non-working periods should also be performed. Therefore, to meet the need of objectively assessing the teachers' voice use, research has focused on validating voice monitoring procedures by means of analyzers that are able to detect the vocal fold activity unobtrusively and that can be worn for long-terms.[7–9]

Gaskill *et al.*[10] monitored the voice use of two primary school teachers over two five-day workweeks using an Ambulatory Phonation Monitor (APM by KayPentax). They found an effectiveness in using the vocal dosimetry to reduce the teachers' vocal load, however no statistically significant changes in the vocal behavior were obtained.

Hunter and Titze[11] monitored the vocal activity of 57 teachers continuously for two weeks in occupational and non-occupational settings. In the occupational setting, the average of the most occurring voice intensity level, i.e., the mode of the sound pressure level (SPL_{mode}), was found to be 62.5 dB, which was 2.5 dB louder than the non-occupational level. They also found that the occupational voice use corresponded to an average value of the mode of the fundamental frequency ($F_{0,mode}$) of 194 Hz, which is 10 Hz higher compared to the non-occupational setting.

Cantor Cutiva *et al.*[12] investigated the changes in self-reported voice and noise conditions in relation to measured voice and noise parameters on the same teachers involved in this study. No significant differences in the self-reported voice condition were found in the monitored days, but a significant difference in the self-reported noise condition was found from day 1 to day 3.

The studies reported so far mainly refer to variations in voice parameters in longitudinal observations or under different settings, such as occupational and non-occupational

[a)]Electronic mail: giuseppina.puglisi@polito.it

voice use. However, voice disorders and vocal load can increase due to recurrent situations, such as the acoustic characteristics of the environments in which the voice is used. For example, people tend to increase their voice level in noisy conditions, and this effect is well-known as the Lombard reflex.[13] In general, high noise levels and very long or very short reverberation times may bring to negative effects on the teachers' vocal load demand[14–16] and on the academic attainment of pupils.[17]

Bottalico and Astolfi[18] used an APM 3200 to collect voice monitorings of 40 teachers from two primary schools that had different acoustic conditions. They investigated the relationship between voice and classroom acoustic parameters, with mid-frequency reverberation times in the classrooms that ranged between 0.4 and 1.6 s. When considering the voice monitoring of a traditional lesson, they found an increase in voice level by 0.72 dB and by 1 Hz in the pitch per 1 dB of increase in the background noise level. They also found that a reverberation time range of 0.75–0.85 s could offer a good support to voice, as it minimized the teachers' voice level.

Sato and Bradley[19] found an increase in the teachers' voice level during active lessons at a rate of 0.72 dB per 1 dB of increase in the noise level, after having conducted 27 speech measurements at four microphone positions per classroom, each by means of sound level meters located at a height of 1.2 m. The classrooms had a mid-frequency reverberation time that ranged between 0.3 and 0.7 s in unoccupied condition. The tendency in the voice use was found to be equal for each school grade, although the grade 1 teachers were able to better control their vocal emissions, with respect to the relative background noise level, since the standard deviation (SD) of the mean speech-to-noise ratio (dBA) was lower than those of any of the other investigated grades.

Durup *et al.*[20] monitored the vocal effort of 20 teachers with an APM 3200, aiming to find a relationship between occupational voice use and classroom acoustic parameters. The classrooms where the voice monitorings were performed had a mid-frequency reverberation time ranging between 0.3 and 1.1 s in unoccupied condition. The main result was a significant positive correlation between voice level and unoccupied ambient noise level, with a Lombard reflex at a rate of 0.69 dB/dB. No correlation was found between the teachers' voice levels and the classroom reverberation times, which can be explained by the fact that the classrooms in the study complied with the current standards on classroom acoustic design in the majority of the cases, and the range of reverberation conditions was not as wide as the range of noise levels.

Pelegrín-García *et al.*[21] investigated the combined effect of talker-to-listener distance and reverberation on the voice use of 13 male subjects under laboratory conditions. The effect of the background noise level was not considered, since it was below 45 dBA in all of the measured conditions and therefore did not affect the voice power levels, according to Lazarus.[22] They considered four reverberation time and room volume conditions ranging between 0.04 and 5.38 s and 410 and 1174 m^3, respectively. A major result was related to the effect of the talker-to-listener distance, which led to an increase in voice power level between 1.3 and 2.2 dB for each doubling in distance in the case of a reverberation room and an anechoic room, with a reverberation time of 5.38 and 0.04 s, respectively. They also found a relationship between the voice power level and the Room Gain (G_{RG}), which is a measure of the gain produced at the speaker's ears by the reflections in the room.[23,24] They observed a variation of -3.6 dB in the voice power level when the Room Gain increased by 1 dB. Furthermore, in another work Pelegrín-García *et al.*[25] also investigated the adaptation of voice levels to keep the autophonic level constant under different G_{RG} conditions. The curves that were determined allowed to predict the voice level variations in different environments that are only due to the Lombard reflex or sidetone compensation. In particular, they found that variations of voice level to maintain the autophonic level constant are not higher than 2 dB when G_{RG} is lower than 1 dB.

Lyberg-Åhlander *et al.*[26] studied the changes in voice use with respect to the Voice Support given by the classroom where the monitoring took place on a sample of 14 teachers with voice problems and 14 teachers with healthy voices. Voice Support is a measure of the strength of the reflected sound relative to the direct sound from one's own voice, which is positively correlated to the Room Gain. They found a tendency of teachers with voice problems to be more aware of classroom acoustics, since they lower their median voice sound pressure level when the Voice Support in the room increased, while teachers with healthy voice showed the opposite trend.

Pelegrín-García *et al.*[27] found that the mid-frequency reverberation time should be 0.6 s in full occupancy condition in classrooms with a maximum number of 40 students for flexible teaching methods. Higher reverberation times could have several consequences on the students' listening engagement and on the teachers' vocal comfort. In fact, it was pointed out that these conditions could negatively affect speech intelligibility and increase the activity noise levels, with a direct effect on the increase in the teachers' voice level due to the Lombard reflex. On the other hand, lower reverberation times could be detrimental, because of vocal comfort reasons.

Most of the studies conducted so far till lack as far as two main aspects are concerned. First, only a small number of works refer to long-term monitorings of speech for voice professionals, and to comparisons with a conversational voice use. Second, the measurement of voice parameters under realistic communication situations and their relationships with the acoustics of the rooms where the speech is measured still needs to be dealt with in depth since no clearly defined conclusions have yet been made. Therefore, this study focuses on three main aspects to explore: First, whether the teachers' voice varies significantly in a one-week equivalent follow-up, so that effective long-term monitorings can be planned to ensure occupational safety and health; second, whether teachers modify their voice production during an occupational vs conversational use; third, whether the measured voice parameters depend on the classroom acoustic characteristics, so that optimal conditions can be drawn for classroom acoustic design.

II. METHODOLOGY

The teachers' voice samples were acquired based on two voicing tasks. A natural continuous speech, here referred to as "conversational," was acquired before each entire monitoring (EM) of the teaching activity was started, so that a comfortable level, here defined as pre-monitoring (PM), could be obtained and compared with that of the EM. Also, the relationships between the measured classroom acoustic parameters and voice use in the EM were investigated. During all the voice EMs a number of pupils that varied between 17 and 23 was present in each classroom.

A. Participants

1. Subjects

Twenty-seven female teachers from four primary schools (grade 1 to 5, i.e., children aged 6–10 years) located in the provinces of Torino and Bolzano (Italy) were involved in this study. Teachers voluntarily participated after that the overall aims of the research activity were presented in a meeting, in which they were given information related to the scientific evidence of the prevalence of voice disorders on their professional category, of the planned monitoring methodology and duration. Their ages ranged from 31 to 60, with a mean age of 48.0 (SD 4.5 years) and the native language was Italian for 25 teachers, and German for two teachers. None of the teachers who participated in the monitoring campaign had reported having severe voice or hearing problems, although some of them had undergone speech therapies in the past to recover from unwanted voice disorders or to learn techniques to help them use better their voices. The therapy sessions were not supposed to influence the results of the work, since teachers themselves pointed out that they did not always make an aware use of voice due to the techniques they have learnt years before. Also, none of the teachers experienced professional singing or acting, thus they did not have specific knowledge to affect the way they spoke.

The years of experience of each monitored teacher was different and was classified in ranges as reported in Bottalico and Astolfi[18] (class 1 if $\leq$6 years, class 2 if 7 to 12 years, class 3 if 13 to 18 years, class 4 if 19 to 21 years, class 5 if $\geq$21 years). The working activity performed by the teachers was related to humanistic (H), scientific (S), or mixed (M) subjects. Gymnastic or handcraft teachers were not involved since they teach in rooms with architectural features, i.e., volumes and furniture, that make room acoustics not comparable to that of typical classrooms.

The vocal activity of teachers was monitored for one to four working days, which is equivalent to one working week, depending on their time-table and availability. The complete voice samples, which were acquired continuously over the teaching period for about 4 h by means of two portable vocal analyzers (the Voice Care device and the Ambulatory Phonation Monitor, model 3200), which are described in Sec. II C 1, were analyzed. Table I reports the own and working information, and the number of performed voice monitorings of each teacher.

2. Schools

The four primary schools where the voice monitorings took place differ in age of construction, location and architectural features, thus classroom acoustics is different from school to school. A common architectural aspect of the classrooms is that they were all plastered and that the floors were covered with ceramics tiles; bookshelves were usually present along the side walls.

School A is located in a residential area adjacent to Torino's city center, where vehicular traffic is not heavy, and it dates back to the late XIX Century. The classrooms did not present any acoustical treatment, had vaulted ceilings and an average height and volume of 4.9 m and 244 m^3, respectively. Schools B and C are located in the province of Torino, in a quiet area where several other school buildings are located, far from busy streets, and were both built in the second half of the XX Century. There were no acoustic treatments on either the ceilings or on the classroom walls. The height and volume of the classrooms were 3.5 m and 160 m^3 (SD 18 m^3), 3.5 m and 142 m^3 (SD 8 m^3) in schools B and C, respectively. School D is located in Bolzano in a mixed residential and commercial area facing onto a street, and it was constructed in the second half of the XX Century. The average height and volume of the classrooms were 3.5 m and the 144 m^3 (SD 4 m^3), respectively, and two types of acoustic treatment were present in some classrooms. Most of the classrooms had absorbent ceilings made of commercial tiles, while the use of expanded polyester tiles on the ceilings, which could be easily removed for cleaning and maintenance, was tested in one classroom. Table I shows the volume (V) of the classrooms of each school.

B. Acoustic parameters of the classrooms

The acoustic characteristics of each classroom were measured before starting the voice monitoring campaign, in the absence of children and teachers. The occupancy of the rooms was simulated by means of absorptive panels made of polyester fiber that were dimensioned in order to have the same absorptive properties as seated children, which has been set at about 0.35 m^2 at 1 kHz, according to Astolfi *et al.*[28] The methodology adopted to obtain the acoustic response of each classroom was the same in each school.

Differences exist between the listener-oriented and the speaker-oriented acoustic parameters of the rooms. The reverberation time ($T30_{0.250\text{–}2\,\mathrm{kHz}}$, s) can be considered a listener-oriented parameter that needs to be checked in order to guarantee a proper listening environment. It was measured in compliance with the UNI EN ISO 3382-2:2008 (Ref 29) standard, applying the backward integrated impulse response method. Two source types were used in the schools as impulse generators, namely, a "clapper-board," that is, a pair of wooden boards hinged together and clapped to generate impulsive signals, and a sweep signal that was emitted by a Bruel&Kjaer type 4128 Head and Torso Simulator (HaTS). Measurements were performed and the results were obtained from two sources and at three microphone positions each, therefore considering six source-receiver pairs in total. The

TABLE I. Description of each monitored teacher: Age, years of teaching category (1 refers to ≤ 6 years, 2 refers to 7 to 12 years, 3 refers to 13 to 18, 4 refers to 19 to 21 and 5 refers to ≥ 21 years), subject taught (humanities H, scientific S, mixed M), children's school grade (1 to 5), V of the individual teaching classroom and number of voice monitorings related to each acquisition device. The teachers are identified by an alphanumeric code, where the letter refers to the school (A to D) and the number corresponds to each specific teacher in the school. The reverberation time ($T30_{0.250\text{-}2kHz,occ}$), background noise level (L_{A90}), Room Gain ($G_{RG,0.5\text{-}2kHz,occ}$) and Decay Time at the ears ($DT_{40ME,0.5\text{-}2kHz,occ}$) that correspond to each voice monitoring are also reported. The values highlighted in bold are the cases of compliancy with the reference standards.

Teacher ID	Age (years)	Years of teaching category	Subject taught	School grade	V (m^3)	Entire monitorings (n) VC	Entire monitorings (n) APM	$T30_{0.250\text{-}2kHz,occ}$ (s)	L_{A90} (dB)	$G_{RG,0.5\text{-}2kHz,occ}$ (dB)	$DT_{40ME,0.5\text{-}2kHz,occ}$ (s)
A1	57	5	H	4	244	2	1	1.2/1.2/ 1.2	65.6/60.3/ 73.3	**0.4/0.4/ 0.4**	**0.9/0.9/ 0.9**
A2	44	5	H	3	244	3	1	1.4/1.4/ 1.4/1.4	68.2/NA/ 57.9/NA	**0.5/0.5/ 0.5/0.5**	**0.9/0.9/ 0.9/0.9**
A4	55	4	S	4	244	2	—	1.2/1.2	60.7/71.4	**0.4/0.4**	**0.9/0.9**
B1	49	5	H	4	180	3	—	0.8/0.8/ 0.8	47.3/NA/ NA	**0.5/0.5/ 0.5**	**0.6/0.6/ 0.6**
B2	50	5	S	2	160	1	1	0.7/0.7	55.6/NA	**0.5/0.5**	**0.5/0.5**
B3	52	5	H	3	122	—	1	**0.5**	NA	**0.4**	NA
B4	60	5	S	3	133	2	—	**0.5/0.5**	65.4/42.7	**0.4/0.4**	**0.4/0.4**
B6	49	3	H	2	160	3	1	0.7/0.7/ 0.7/0.7	55.6/NA/ 52.8/NA	**0.5/0.5/ 0.5/0.5**	**0.5/0.5/ 0.5/0.5**
B8	47	2	H	4	176	—	2	0.8/0.8	NA/NA	**0.4/0.4**	NA/NA
B9	40	1	S	1	160	1	—	0.7	NA	**0.4**	NA
C1	57	5	S	3	150	2	—	0.8/0.8	NA/NA	0.7/0.7	**0.7/0.7**
C2	59	5	H	2	150	3	—	1.0/1.0/ 1.0	NA/55.5/ 53.3	0.6/0.6/ 0.6	**0.8/0.8/ 0.8**
C3	57	5	H	5	135	2	—	0.7/0.7	NA/69.9	**0.5/0.5**	NA/NA
C4	53	5	H	4	135	3	—	0.9/0.9/ 0.9	NA/63.3/ 48.5	0.6/0.6/ 0.6	**0.6/0.6/ 0.6**
C6	32	2	S	2	150	1	—	1.0	NA	0.6	**0.8**
C7	34	2	S	4	135	2	—	0.9/0.9	NA/65.6	0.6/0.6	**0.6/0.6**
D1	31	2	H	3	149/140	2	—	**0.5/0.4**	55.0/47.7	**0.3/0.2**	NA/NA
D2	47	4	H	3	140	2	—	0.6/0.4	52.4/NA	**0.4/0.2**	NA/NA
D3	37	3	M	2	149	2	—	0.6/0.6	46.6/NA	**0.4/0.4**	NA/NA
D4	36	2	M	3	149	3	—	0.6/0.6/ 0.6	76.0/NA/ 57.7	**0.4/0.4/ 0.4**	NA/NA/ NA
D5	44	2	S	5	140	2	—	0.4/0.4	51.0/50.6	**0.2/0.2**	NA/NA
D6	46	5	H	3	140	3	—	0.4/0.4/ 0.4	52.7/38.4/ NA	**0.2/0.2/ 0.2**	NA/NA/ NA
D7	51	5	M	4	140	2	—	0.4/0.4	52.0/52.4	**0.2/0.2**	NA/NA
D8	43	4	S	1	149	2	—	0.6/0.6	51.3/55.8	**0.4/0.4**	NA/NA
D9	44	4	H	1	144	3	—	0.6/**0.5**/ **0.5**	48.2/47.0/ 49.7	**0.4/0.3/ 0.3**	**0.4**/NA/ NA
D10	43	5	S	3	140	2	—	0.4/0.4	61.3/53.5	**0.2/0.2**	NA/NA
D11	36	3	S	4	140	1	—	0.4	46.1	**0.2**	NA

results were averaged in order to obtain a mean spatial value, which means that the spatial average was acquired by considering the mean of the individual reverberation times for all of the independent source and microphone positions. The source was positioned in those places that were representative of those used by a given teacher in the classroom, at a height of 1.5 m from the floor. The microphones were evenly distributed over all the pupils' seating areas at ear height, 1.1 m above the floor, at a distance of 2 m from each other, not too close to the source and at least 1 m from any surface. Frequency averaging in the 0.250–2 kHz range and standard compliancy were done according to the German DIN 18041:2004.[30] Measurements were performed in unoccupied (with furniture but without persons) and occupied (simulated with polyester fiber panels) classrooms, although only occupied condition measurements were considered in the statistical analysis.

The speaker-oriented acoustic parameters of the rooms were: Voice Support, Room Gain, and Decay Time at the ears ($ST_{V,0.5\text{–}2kHz}$ and $G_{RG,0.5\text{–}2\,kHz}$, dB, and $DT_{40ME,0.5\text{–}2\,kHz}$, s, respectively). They were defined to account for the perceived room acoustics at the speaker's ears, according to the procedures outlined in Pelegrín-García *et al.*[23,24,27,31,32] Voice Support is a measure of the extent to which sound reflections at room boundaries amplify the voice of a speaker at his/her own ears. Room Gain is defined as the gain applied by the room to the voice of a speaker at his/her own ears. Decay Time at one's own ears is defined as the time it would take for the backward integrated energy curve of an oral-binaural room impulse response to decay 60 dB after the arrival of the direct sound, calculated from the initial decay of 40 dB and assuming a linear decay. The oral-binaural room impulse response (OBRIR) from the mouth to the ears of the Bruel&Kjaer 4128 Head and Torso Simulator (HaTS) was measured in two source positions inside each classroom to measure these parameters, with the HaTS being placed at a height of 1.5 m and at least 1 m from any surface. As suggested in the referenced studies, a speech-weighting was done in frequency and then an averaging was applied to the results in the 0.5–2 kHz range for all the parameters.

The background noise level was evaluated in terms of the A-weighted statistical level that was surpassed for 90% of the measuring time (L_{A90}, dB). It was measured using a class-1 sound level meter (either XL2 by NTi Audio or type 2222 by Brüel & Kjær), which was placed at 1.2 m from the ground, close to the teacher's desk, at a minimum distance of 1 m from any surface. Measurements were performed for the entire duration of a lesson, and the L_{A90} from the acquired wave signals were then calculated using *ad hoc* created MATLAB scripts.

C. Voice monitorings

1. Vocal activity monitoring of the teachers

The teachers' vocal activity was monitored during the working hours in the four-days of observation by means of the Voice Care (VC) and the Ambulatory Phonation Monitor (APM 3200, model 3200). The former is a low-cost vocal analyzer recently developed at the Politecnico di Torino by Carullo *et al*.,[33,34] while the latter is a commercial device made by KayPentax®. They both consist of a contact sensor, which is connected to a data logger. The sensor is placed at the jugular notch, and it detects the skin vibrations during phonation. An Electret Condenser Microphone (ECM AE38 [Alan Electronics GmbH (Dreieich, Germany)]) and a BU7135 accelerometer by Knowles Corp. (Itasca, IL), were used as contact sensors for VC and APM 3200, respectively. The acquired voice samples were grouped into 30 ms frames, which corresponded to the inter-syllabic pauses for VC, and into 50 ms frames for APM 3200. Since both devices were calibrated in laboratory, they provide measurements that are traceable to the same standards, thus allowing all the voice parameters to be collected in a single database.[34,35]

The devices provide an estimation of the voice sound pressure levels at a fixed distance from the speaker's mouth, after a calibration to a reference microphone (SPL in dB), the fundamental frequency (F_0 in Hz) and the voicing time percentage ($D_{t\%}$ in %), which is defined as the percentage of time spent phonating for the total monitoring period.[36] The calibration procedure was needed to accurately estimate the sound pressure levels from the voltage signals detected at the base of the neck. It is similar for the two devices and, to have an effective and accurate evaluation of voice parameters, it has to be performed in a quiet environment with the same room acoustics as the ambient where the subsequent voice monitoring will take place. In particular, in this work the reverberation characteristics of the rooms where the calibration procedure was performed were similar to those of the teaching classrooms, respectively, for each teacher, and noise, which was calculated as A-weighted equivalent sound pressure level (L_{Aeq}), was always lower than 40 dBA to guarantee a high signal-to-noise ratio. The calibration procedure consists of the vocalization of the vowel /a/ at increasing levels, wearing the contact sensor at the jugular notch and having a calibrated air microphone positioned 16 and 15 cm from the mouth, for the VC and APM 3200, respectively. Since the distance of the air microphone was different for the devices, all the SPL values obtained from the APM 3200 at 15 cm from the speaker's mouth were estimated at 16 cm according to the free-field sound propagation theory, in order to have comparable results with VC.

The results related to F_0 and SPL are usually shown as occurrence histograms, from which the following parameters can be obtained: Mean, mode, and SD of the SPL at 16 cm from the mouth ($SPL_{mean,16\,cm}$, $SPL_{mode,16\,cm}$, and $SPL_{SD,16\,cm}$), mean, mode and SD of the F_0 ($F_{0,mean}$, $F_{0,mode}$, and $F_{0,SD}$). The equivalent SPL ($SPL_{eq,16\,cm}$) can also be obtained from the acquired voice sample. It was defined as the total sound energy produced by the vibration of the vocal folds, estimated from the skin acceleration level, which is calculated as the average of the voiced energy over all the frames, including the unvoiced ones, whose energy was set to zero, as suggested by Švec *et al*.[8] The $SPL_{eq,16\,cm}$ values were estimated 1 m from the speaker's mouth ($SPL_{eq,1m}$) according to the free-field sound propagation theory, in order to obtain the teachers' vocal effort to comply with ANSI S3.5–1997.[37] The mean sound pressure level at 1 m from the speaker's mouth ($SPL_{mean,1m}$) was estimated based on the same procedure.

2. Conversational vs occupational tasks

Each entire voice monitoring (EM) was preceded by a pre-monitoring (PM), i.e., an interview carried out to obtain a voice sample of each teacher, which was taken as the daily conversational speech that had to be compared with the one measured over the working period. The PM sample consisted of a 5-min-long speech that each teacher had to perform in front of a listener seated at a distance of 1 m. This conversational speech was performed in a room with similar room acoustics of the subsequent EM, but with noise level lower than 40 dBA, that is the same room where the calibration procedure of the voice monitoring devices was taken. Teachers were asked to speak about a topic they knew well (e.g., a receipt, the road from home to school, the teaching content of the day, etc.), in order to produce a continuous speech pronounced at a comfortable and conversational pitch, with natural loudness and not using a singing voice. Although in some cases the voice monitoring started in the early morning, teachers' voice was warmed up since they had an informal conversation with the experimenter that lasted about 15 min as they arrived at school. Before performing the PM teachers had to perform the calibration of the monitoring device too, therefore they had to vocalize at increasing voice levels. After the PM, the EM, in which the working activity during entire morning or afternoon teaching hours was monitored (about 4 h), was started.

D. Statistical methods

The acquired data were statistically analyzed using the SPSS software (version 22; SPSS Inc., New York, NY). In order to understand whether the parameters related to the classroom acoustics and to the teachers' voices were normally or non-normally distributed, the Shapiro-Wilk test was applied to the complete dataset. All the calculations were performed assuming a 95% confidence interval. Note that, due to the in-field design of experiment, the individual variation of the results of the daily voice monitorings was not compensated for with respect to the changes in voice behavior related to noise and room acoustics.[21] Anyway, the variations of SPL_{mean} among teachers of the same school were previously reported in Cantor Cutiva *et al*.[12] and ranged

TABLE II. Average values of the measured acoustic parameters in the classrooms for each school and the optimal range. The number of classrooms in which the measurements were performed is reported for each school, as well as the number of total measurements that were carried out for each parameter. The mean values of each measured parameter are reported and its SD is indicated in brackets. The values highlighted in bold represent the cases of compliancy with the reference standards.

School (number of classrooms)	$T30_{0.250\text{-}2kHz,occ}$ (s)	$ST_{V,0.5\text{-}2kHz,occ}$ (dB)	$G_{RG,0.5\text{-}2kHz,occ}$ (dB)	$DT_{40ME,0.5\text{-}2kHz,occ}$ (s)	L_{A90} (dB)
A (2)	1.3 (0.02)	−**9.8** (0.2)	**0.45** (0.03)	**0.90** (0.03)	65.4 (2.3)
B (6)	0.7 (0.03)	−**9.8** (0.2)	**0.44** (0.02)	**0.52** (0.06)	53.2 (3.2)
C (4)	0.9 (0.02)	−8.4 (0.2)	0.60 (0.03)	**0.69** (0.03)	59.4 (3.3)
D (5)	**0.5** (0.03)	−**11.5** (0.6)	**0.31** (0.04)	**0.43** (NA)	52.3 (1.7)
Optimal range	0.5–0.6	−14 to −9	0.2–0.5	0.4–1.2	≤40
Reference	DIN 18041[30]	Pelegrín-García *et al.*[32]	Pelegrín-García *et al.*[32]	Pelegrín-García *et al.*[27]	BB93[39]
Number of total measurements	17	17	17	9	39

between 1.3 and 1.9 dB and 1.0 and 3.0 dB in terms of standard error for the EM and PM, respectively.

The correlation analysis was carried out to understand the mutual relationship between the classroom acoustic parameters and the teachers' voice parameters in EM and PM separately.

The regression analysis was used to further understand the dependency of the measured voice on the classroom parameters and between classroom acoustic parameters. In order to run the regression analysis, the data were grouped together on the basis of reverberation classes. The classrooms where the acoustic measurements took place were clustered into groups, according to the measured $T30_{0.250\text{–}2kHz,occ}$ and to its just noticeable difference (JND 5%, based on BS EN ISO 3382-2:2008; Ref. 29). This procedure allowed well defined and robust groups of data to be obtained.

The one-way analysis of variance (ANOVA) was applied to the average values to investigate the effect that fixed factors, such as the day of monitoring, the school, the school grade, the years of teaching and the subject taught, could have on the teachers' voice production. As a subsequent completion of the ANOVA, the Scheffé *post hoc* test[38] was applied to try to understand exactly what factors had an influence on voice changes (e.g., if the ANOVA suggested a significant effect of the "day of monitoring" on the variation of voice intensity, the Scheffé *post hoc* test revealed the specific day in which the intensity variation was observed).

III. RESULTS

A. Classroom acoustic parameters

Table I shows the $T30_{0.250\text{–}2kHz,occ}$, L_{A90}, $G_{RG,0.5\text{–}2kHz,occ}$, and $DT_{40ME,0.5\text{–}2kHz,occ}$ that correspond to each voice monitoring, while Table II shows the values averaged for each school.

The $T30_{0.250\text{–}2kHz,occ}$ is the classroom acoustic parameter that has to be checked from the listener's perspective to guarantee a proper listening environment. The German DIN 18041:2004 (Ref. 30) specifies the optimal values of $T30$ as a function of the room volume and the frequency in octave band for the teaching activity. As shown in Tables I and II, only five classrooms out of 18, and only one school out of four, complied with the reference standard. In particular, $T30_{0.250\text{–}2kHz,occ}$ was adequate in the classrooms where acoustical treatments had been carried out. With respect to the speaker-oriented parameters, the obtained results show compliance with the ranges proposed by Pelegrín-García *et al.*[27,32] in most of the classrooms, as it can be seen in Tables I and II.

In total, 39 noise monitorings were performed throughout entire lessons in the schools. The measured L_{A90} values were averaged for all the classrooms and grades, and consisted in 55.8 dBA (SD 8.7 dBA). Since the most commonly used standards such as the BB93 (Ref. 39) provide reference noise levels in the case of unoccupied classroom conditions, the measured L_{A90} values were first compared with the results obtained by Shield and Dockrell[40] who characterized primary school classrooms in London. In particular, they obtained an average L_{A90} value of 54.1 dBA, in occupied condition, considering various school locations and typologies, which is comparable to the average L_{A90} value obtained in this study. Also, Shield and Dockrell[40] found that L_{Aeq} in unoccupied condition was 7.1 dBA lower than the measured occupied L_{A90} value. Therefore, the L_{Aeq} in unoccupied condition measured by the authors can be supposed to be of about 48.8 dBA by applying the aforementioned difference; however, this predicted value exceeds the BB93 (Ref. 39) threshold for existing school buildings, that is, 40 dBA.

A positive association between L_{A90} on $T30_{0.250\text{–}2kHz,occ}$ was found and is reported in Fig. 1, where an increase of 13 dB/s is shown after a regression analysis (p-value = 0.005)

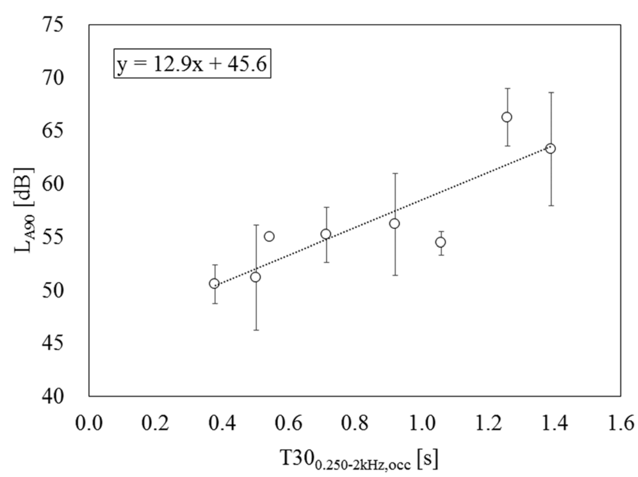

FIG. 1. Best-fit linear regression ($R^2 = 0.76$) between the background noise levels (L_{A90}) monitored during the working time and the reverberation time in the classrooms in occupied conditions ($T30_{0.250\text{-}2kHz,occ}$). Each experimental data in the graph represents the mean value of an average of six pairs. The error bars refer to the SD of the mean (standard error, SE).

has been obtained. This explains why school A was found to have on average a significantly higher L_{A90} than schools B and D, with $T30_{0.250–2kHz,occ}$ being the highest among all the schools (Table II).

The correlation matrix of the acoustic parameters of the classrooms and the volumes of the investigated classrooms is shown in Table III. The speaker-oriented parameters show a significant and highly positive correlation with $T30_{0.250–2kHz,occ}$, as well as with the volume (V), and thus corroborates the results of other studies.[27,32] A significant and positive correlation has also been found between L_{A90} and V, $T30_{0.250–2kHz,occ}$, and $DT_{40ME,0.5–2\,kHz}$.

B. Voice parameters of the teachers

The voice monitoring of 27 female teachers has been included in this study. These teachers were monitored for 1 to 4 working days, which is equivalent to one working week, but not all the performed monitorings were considered since some of them were based on erroneous calibration sessions. In particular, monitorings were discarded if the SPL range of the calibration did not correspond to the SPL range of the EM, as it is detailed in Carullo *et al.*[34] In such a way, by not considering those monitorings in which the SPL outcomes failed, a total number of 61 voice monitorings were considered, which corresponded to 2.3 days on average per each teacher.

1. Voice use in conversational vs occupational tasks

The voice parameters measured and averaged separately for each of the four schools was reported in Cantor Cutiva *et al.*[12] In particular, a significant difference between EM and PM was found for the $SPL_{mean,16\,cm}$ measured in schools A and B, after applying the Scheffé *post hoc* test[38] to the data reported in the aforementioned table. A significant decrease in $D_{t\%}$ was also found for the comparison between schools B and C during the EM, whereas $SPL_{mode,16\,cm}$ was found to be significantly higher in school B than in school C in the PM.

Table IV shows the correlation analysis between the teachers' voice parameters in the EM and PM conditions. As expected, strong positive correlations (p value < 0.01) between $SPL_{mean,16\,cm}$ and $SPL_{mode,16\,cm}$, and between $F_{0,mean}$ and $F_{0,mode}$ can be observed for both EM and PM. A strong positive correlation was found between $SPL_{mode,16\,cm}$ and $F_{0,mode}$, $F_{0,mode}$ and $F_{0,SD}$,and $F_{0,mean}$ and $F_{0,SD}$, but only in the case of EM. A strong negative correlation was observed in the EM between $SPL_{mean,16\,cm}$ and $D_{t\%}$, and a strong positive correlation was found between $SPL_{SD,16\,cm}$ and $F_{0,SD}$, but only in the case of PM.

An ANOVA was carried out to understand whether any of the fixed factors had a statistically significant effect on the teachers' voice parameters. No significant influence was found for the school grade, while a significant effect of the subject taught was found on $SPL_{mode,16\,cm}$ (p value = 0.037) and $F_{0,mean}$ (p value = 0.049) in the case of EM and PM, respectively, which was higher in the case of humanistic topics than in the case of scientific and mixed ones. As far as the years of teaching are concerned, in the case of EM, the $D_{t\%}$ was found to be significantly higher (p value = 0.049) for teachers with more than 21 years of teaching experience.

2. Day-by-day analysis

Table V shows the analysis of the collected data that were compared on the basis of the day in which the voice monitoring took place. An ANOVA was performed to understand whether there was a statistically significant difference in the values of the classroom acoustics and teachers' voice parameters averaged across the days (from day 1 to day 4). No significant difference was found in the day-by-day analysis between the classroom acoustic parameters, so this result supports the aim of investigating if significant differences in the voice parameters can be found that are only due to the prolonged and sustained voice use, and not due to the effect of acoustics.

In the case of EM, a general tendency to decrease intensity and the pitch parameters (SPL and F_0) from day one to day four can be observed, but this tendency was not statistically significant. In the case of PM, the ANOVA revealed a significant difference for the teachers' voice parameter, that is, of $SPL_{mode,16\,cm}$. In order to understand which days this

TABLE III. Correlation matrix of the classroom acoustic parameters and the volume of the investigated classrooms. Only correlation coefficients with a p value < 0.05 are reported, and the correlation coefficients with significance <0.01 are marked in bold. The number of pairs considered on the basis of the available measured data is also reported for each correlation.

	Volume (m^3)	log(V)	$T30_{0.250\text{-}2kHz,occ}$ (s)	$ST_{V,0.5\text{-}2kHz,occ}$ (dB)	$G_{RG,0.5\text{-}2kHz,occ}$ (dB)	$DT_{40ME,0.5\text{-}2kHz,occ}$ (s)	L_{A90} (dB)
Volume (m^3)	1						
N	61						
log (V)	**0.349**	1					
N	61	61					
$T30_{0.250\text{-}2kHz,occ}$ (s)	**0.818**	**0.335**	1				
N	61	61	61				
$ST_{V,0.5\text{-}2kHz,occ}$ (dB)			**0.670**	1			
N			61	61			
$G_{RG,0.5\text{-}2kHz,occ}$ (dB)			**0.651**	**0.981**	1		
N			61	61	61		
$DT_{40ME,0.5\text{-}2kHz,occ}$ (s)	**0.808**	**0.794**	**0.990**			1	
N	32	32	32			32	
L_{A90} (dB)	**0.484**		**0.561**	0.384	0.358	**0.580**	1
N	39		39	39	39	19	39

TABLE IV. Correlation matrix based on the 61 pairs of teachers' voice parameters measured during the entire monitoring (EM), based on 49 pairs for the pre-monitoring interview data (PM). Only correlation coefficients with a p value < 0.05 are reported, and the correlation coefficients with significance <0.01 are marked in bold.

	$SPL_{eq,16\,cm}$ (dB)	$SPL_{mean,16\,cm}$ (dB)	$SPL_{mode,16\,cm}$ (dB)	$SPL_{SD,16\,cm}$ (dB)	$F_{0,mean}$ (Hz)	$F_{0,mode}$ (Hz)	$F_{0,SD}$ (Hz)	$D_{t,\%}$ (%)
Teachers' voice parameters (EM)								
$SPL_{eq,16\,cm}$ (dB)	1							
$SPL_{mean,16\,cm}$ (dB)	**0.615**	1						
$SPL_{mode,16\,cm}$ (dB)	**0.597**	**0.866**	1					
$SPL_{SD,16\,cm}$ (dB)	**0.747**			1				
$F_{0,mean}$ (Hz)			0.254		1			
$F_{0,mode}$ (Hz)			**0.356**		**0.850**	1		
$F_{0,SD}$ (Hz)					**0.643**	**0.645**	1	
$D_{t,\%}$ (%)		**−0.360**			−0.280			1
Teachers' voice parameters (PM)								
$SPL_{eq,16\,cm}$ (dB)	1							
$SPL_{mean,16\,cm}$ (dB)	**0.649**	1						
$SPL_{mode,16\,cm}$ (dB)	**0.533**	**0.849**	1					
$SPL_{SD,16\,cm}$ (dB)	**0.394**			1				
$F_{0,mean}$ (Hz)					1			
$F_{0,mode}$ (Hz)					**0.892**	1		
$F_{0,SD}$ (Hz)	0.314			**0.418**	0.351		1	
$D_{t,\%}$ (%)								1

parameter changed significantly, the Scheffé *post hoc* test[38] was applied, which indicated a meaningful difference of $SPL_{mode,16\,cm}$ of 6.5 dB between day 4 and day 1, with mean values of 82.0 dB (SD of the mean, SE, equal to 1.9 dB) and 75.5 dB (SE equal to 1.5 dB), respectively. When considering the difference between EM and PM, a significant variation in the voice parameters in the weekdays was only found in the case of $SPL_{mean,16\,cm}$. However, the Scheffé *post hoc* test did not report the specific days in which $SPL_{mean,16\,cm}$ changed significantly.

C. Influence of classroom acoustics on vocal parameters

The relationship between the teachers' voice parameters and the acoustic parameters of the classrooms has been investigated through a regression analysis, which was based on data clustering of the reverberation classes, as explained in Sec. II D. Between $SPL_{mean,1m}$ and $T30_{0.250\text{–}2kHz,occ}$, a quadratic dependency has been obtained when considering the data clusters. As can be seen in Fig. 2, a value of 0.7 s has been assumed as the optimal $T30_{0.250\text{–}2kHz,occ}$ to minimize the $SPL_{mean,1m}$. Considering the same clusters of data, a 0.53 dB increase in speech level per 1 dB increase in noise level has been found, as can be seen in Fig. 3, thus confirming the presence of a Lombard reflex.[13] The robustness of the regressions has been fair for $SPL_{mean,1m}$ and $T30_{0.250\text{–}2kHz,occ}$ (p value $= 0.053$) and very high for $SPL_{mean,1m}$ and L_{A90} (p value $= 0.001$).

IV. DISCUSSION

A. Optimal acoustic conditions for speaking and listening in a classroom

Five classrooms out of 18 showed an adequate $T30_{0.250\text{–}2kHz,occ}$ value (Table I). This outcome can be explained by the fact that Italian schools are generally housed in historical buildings that do not have specific acoustic treatments, where large volumes and reflective surfaces negatively affect the classroom acoustics. As a main consequence of the long sound tail that exists in the majority of the considered classrooms, a significantly positive correlation was found between $T30_{0.250\text{–}2kHz,occ}$ and L_{A90} (Table III), as well as a strong linear relationship (Fig. 1).

Although the reverberation time does not meet the optimal values, the speaker-oriented parameters generally do respect the optimal ranges that Pelegrín-García *et al.*[27,32] found in their works, which can be considered consistent for several reasons. First, higher reverberation conditions are needed to enrich and support the voice of a talker. Second, a close relationship between $ST_{V,0.5\text{–}2\,kHz}$, $T30_{0.250\text{–}2kHz,occ}$, and room volume is defined in Pelegrin-Garcia *et al.*,[27] which allows medium-size classrooms (100 $m^3 < V < 250$ m^3) with different reverberation times to have similar speaker-oriented parameter values. As an example, a room of about 250 m^3 and 1.3 s of $T30_{0.250\text{–}2kHz,occ}$ would exhibit the same $ST_{V,0.5\text{–}2kHz}$ as a room of about 160 m^3 and 0.8 s of $T30_{0.250\text{–}2kHz,occ}$, that is, about -10 dB. As two prediction models for the $ST_{V,0.5\text{–}2kHz}$ and for the $DT_{40ME,0.5\text{–}2\,kHz}$, respectively, were developed by Pelegrín-García *et al.*,[27,32] it was investigated how the presented measured data fitted the predictions based on the formalized models to assess a relationship that can corroborate the past results. The predicted data varied at a rate of 1.4 dB/dB ($R^2 = 0.85$) and 1.3 s/s ($R^2 = 0.97$) for the $ST_{V,0.5\text{–}2kHz}$ and $DT_{40ME,0.5\text{–}2\,kHz}$, respectively.

B. Voice monitoring

1. Four-day-follow-up monitoring of the teachers' voice

A follow-up study on voice monitoring can help to understand to what extent the prolonged use of voice under

TABLE V. Mean values of the acoustic parameters of the classrooms and the teachers' voice parameters in/for the different monitored working days, referring to the EM, the PM and the difference between EM and PM. The SD of the mean (SE) was calculated for each parameter. The ANOVA significance across the days (p value < 0.05) is indicated with an asterisk (*). The Sheffé *post hoc* test reported a significant difference, which is marked in bold, between day 1 and day 4.

	Day 1 (n = 18)		Day 2 (n = 11)		Day 3 (n = 13)		Day 4 (n = 19)	
	m	SE	m	SE	m	SE	m	SE
Classroom acoustics parameters								
$T30_{0.250\text{-}2kHz,occ}$ (s)	0.7	0.08	0.7	0.09	0.8	0.09	0.8	0.06
$ST_{V,0.5\text{-}2kHz,occ}$ (dB)	−10.4	0.49	−10.8	0.52	−10.3	0.40	−10.0	0.41
$G_{RG,0.5\text{-}2kHz,occ}$ (dB)	0.42	0.04	0.37	0.04	0.41	0.03	0.45	0.03
$DT_{40ME,0.5\text{-}2kHz,occ}$ (s)	0.66	0.05	0.70	0.12	0.71	0.07	0.70	0.04
L_{A90} (dB)	56.2	2.1	52.8	6.3	55.5	2.4	57.4	2.9
Teachers' voice parameters (EM)								
$SPL_{eq,16\,cm}$ (dB)	86.6	1.8	85.6	2.9	86.9	2.2	89.0	1.8
$SPL_{mean,16\,cm}$ (dB)	85.8	1.4	83.8	1.4	83.0	1.3	82.8	1.3
$SPL_{mode,16\,cm}$ (dB)	87.6	1.6	86.2	2.0	86.2	1.4	83.8	1.6
$SPL_{SD,16\,cm}$ (dB)	7.8	0.7	8.8	1.4	10.1	1.1	10.5	0.7
$F_{0,mean}$ (Hz)	235	4.9	231	4.4	236	6.9	243	6.9
$F_{0,mode}$ (Hz)	223	5.5	218	6.1	233	11.5	242	12.9
$F_{0,SD}$ (Hz)	64	1.4	65	1.3	66	2.2	68	2.0
$D_{t,\%}$ (%)	30	1.9	27	1.4	31	1.6	29	1.5
Teachers' voice parameters (PM)								
$SPL_{eq,16\,cm}$ (dB)	80.4	1.5	78.2	1.2	78.8	1.0	79.3	2.0
$SPL_{mean,16\,cm}$ (dB)	80.3	1.8	79.6	1.5	77.7	1.0	76.0	1.1
$SPL_{mode,16\,cm}$ (dB)*	**82.0**	1.9	80.1	1.5	79.5	1.1	**75.5**	1.5
$SPL_{SD,16\,cm}$ (dB)	5.9	0.7	4.9	0.4	6.5	0.5	6.7	0.6
$F_{0,mean}$ (Hz)	188	6.4	176	5.1	185	4.5	193	6.6
$F_{0,mode}$ (Hz)	175	5.2	170	4.8	171	4.2	181	7.0
$F_{0,SD}$ (Hz)	46	2.4	42	1.4	48	2.5	44	2.6
$D_{t,\%}$ (%)	41	3.1	40	4.2	50	3.1	42	3.0
Teachers' voice parameters (EM-PM)								
$SPL_{eq,16\,cm}$ (dB)	5.9	1.4	3.9	1.2	6.3	1.4	9.2	1.7
$SPL_{mean,16\,cm}$ (dB)*	4.6	0.9	3.6	0.6	5.4	1.2	7.8	1.0
$SPL_{mode,16\,cm}$ (dB)	5.0	1.3	4.8	1.3	6.5	1.8	8.9	1.8
$SPL_{SD,16\,cm}$ (dB)	1.9	0.4	1.9	0.3	2.6	0.3	2.9	0.4
$F_{0,mean}$ (Hz)	41	5.9	55	3.4	53	5.6	53	5.7
$F_{0,mode}$ (Hz)	42	6.4	45	5.0	64	10.6	65	14.5
$F_{0,SD}$ (Hz)	19	2.9	24	1.2	17	3.1	24	3.7
$D_{t,\%}$ (%)	−9	3.8	−14	3.5	−20	2.6	−13	2.6

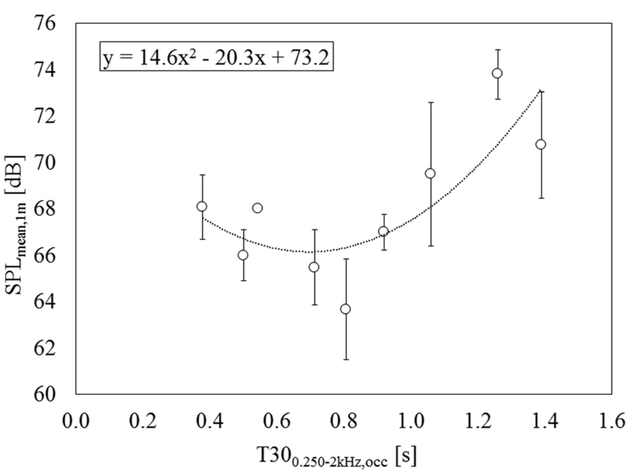

FIG. 2. Best-fit quadratic regression curve ($R^2 = 0.63$) between the voice levels of the teachers ($SPL_{mean,1m}$) and the reverberation times in the classroom in occupied conditions ($T30_{0.250\text{-}2kHz,occ}$). Each experimental data in the graph represents the mean value of an average of six pairs. The error bars refer to the SD of the mean (standard error, SE).

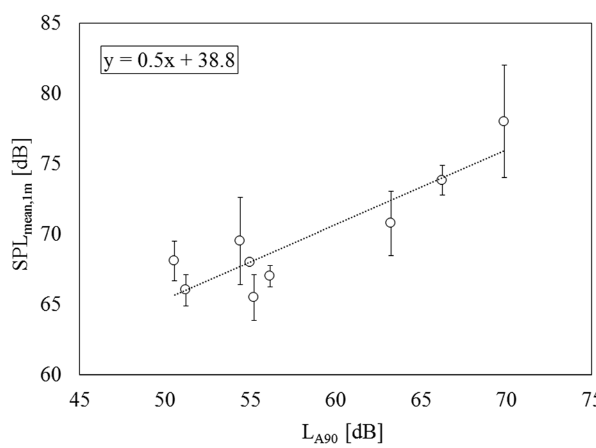

FIG. 3. Best-fit linear regression ($R^2 = 0.80$) between the vocal efforts of the teachers ($SPL_{mean,1m}$) and the background noise levels (L_{A90}) monitored during the working period. Each experimental data in the graph represents the mean value of an average of six pairs. The error bars refer to the SD of the mean (standard error, SE).

realistic communication situations is affected by the acoustic environment. Although a tendency to decrease the vocal sound pressure level and fundamental frequency was observed, the between day difference was not found to be statistically significant. This could be due, among several possible reasons, to the great variability in speech, which depends on the subject that is taught and also on the speech material.[41] The monitored teachers, in fact, were not asked to perform specific lessons (e.g., plenary or group lessons), but were asked to behave as they usually did during the teaching hours. Nevertheless, when comparing the mode of sound pressure level in the case of relaxed and comfortable speech (pre-monitoring, PM), a significant decrease can be observed from day one to day four, which can point out an increase in vocal fatigue due to voice overstraining, which leads the teachers to lower their normal comfortable speech in an attempt to keep an almost constant level when teaching. In fact, although comfortable speech is typical of close-distance conversations, teachers tend to talk during the working period so that they will be understood by each student in the classroom, even the furthest ones. Therefore, they will be likely to adjust their voice not only on the basis of the acoustics and on the perceived fatigue, but also in order to be intelligible by all of the pupils in the classroom. The dependency of the produced voice level on its intelligibility at several talker-to-listener distances was investigated in a laboratory study by Pelegrín-García *et al.*[21] They considered several acoustic conditions in terms of reverberation, but did not account for the presence of background noise, which was always absent, and found that the use of a certain voice level depends on the visually perceived distance from the listener, with a corresponding voice increase of 1.3 and 2.2 dB for double the distance in the case of rooms with reverberation times equal to 5.38 and 0.04 s, respectively.

As far as the monitoring of the repeated occupational use of the voice by teachers for subsequent days and under realistic acoustic conditions is concerned, the outcomes of this study can be compared with those of Gaskill *et al.*,[10] who monitored the vocal activity of two female teachers for 2 weeks by means of an APM, with the aim of comparing the weekly mean values of speech intensity (dB SPL); however, no relevant differences in the SPLs were found.

2. Vocal effort of the teachers

It has been possible to calculate the vocal effort, which was found on average to be 71.2 dB (SE 1.0 dB) $SPL_{eq,1m}$ for all the monitorings. According to ANSI S3.5–1997,[37] it is possible to classify that the teachers' vocal effort, on average, in the range between "raised" and "loud," that is 68–75 dB at 1 m from the mouth.

3. Conversational vs occupational voice use

The $SPL_{mean,16\,cm}$ and $SPL_{mode,16\,cm}$ of the teachers in the EM were found to be significantly higher than in the PM, that is, by 5.5 dB (SE 0.5) and 6.5 dB (SE 0.8 dB), respectively. The $F_{0,mean}$ and $F_{0,mode}$ of the teachers also showed an average increase in the occupational setting of 50 Hz (SE 3 Hz) and 55 Hz (SE 6 Hz), respectively. These findings are in agreement with various other studies that pointed out the normal behavior of teachers of talking louder during the working hours than in non-occupational settings. Hunter and Titze[11] reported an increase in the occupational SPL_{mode} of about 2.5 dB, as well as an increase in the occupational $F_{0,mode}$ of about 10 Hz. They also observed that the teachers had an average $D_{t\%}$ equal to 30% when teaching, which was double that of a non-occupational setting. In the present study, the voicing time percentage has been investigated as an indicator of the vocal load, and the obtained results corroborate the hypothesis of the aforementioned study in which the plenary teaching activity (EM) was characterized by an elevate voicing demand, which corresponded to an average $D_{t\%}$ of about 29% (SE 1%).

A significant positive correlation between $F_{0,SD}$ and both $SPL_{eq,16\,cm}$ and $SPL_{SD,16\,cm}$ was found in the PM. These relationships suggest that, in the case of conversational speech, teachers tend to have a better control of voice production due to a more controlled respiratory and phonatory coordination. Since respiratory and laryngeal components, such as the sub-glottal air pressure and the vocal fold adduction, affect the regulation of the voice level and intonation,[42] these elements may also affect the SDs of F_0 and SPL. Then, the sound pressure level was found to be significantly and positively related to the fundamental frequency in the EM, as observed by Hunter and Titze[11] and by Bottalico and Astolfi.[18] Moreover, the $D_{t\%}$ was found to be strongly and negatively related to the sound pressure level and the fundamental frequency. This suggests that the prolonged and sustained use of voice in the working hours can be a cause of overstrain, which makes speakers reduce phonation to ease the voice overload.

C. Association between classroom acoustic parameters and the voice use of teachers

An increase by 0.53 dB in the teachers' voice level has been found in the EM for every 1 dB of increase in classroom background noise. This result is in agreement with other studies on the Lombard reflex in primary school classrooms, such as the ones of Sato and Bradley,[19] Bottalico and Astolfi[18] and Durup *et al.*,[20] who found increases by 0.72, 0.72, and 0.69 dB, respectively, for every 1 dB increase in the noise level.

In a previous work by Pelegrín-García *et al.*,[21] the increase in the speaker-oriented parameters of Room Gain and Voice Support were found to correspond to a decrease in voice level; however, the presented data did not prove the same finding. It is reasonable to hypothesize a different trend, since the work of Pelegrín-García *et al.*[21] refers to laboratory conditions where noise was almost absent. The present study, instead, accounts for realistic acoustics, where noise was found to be significantly affected by reverberation (Fig. 1) and an evident Lombard reflex was present (Fig. 3). Therefore, the obtained results suggest that the teachers' voice level depended on both noise and reverberation, which allowed for side tone compensations as reported in Pelegrín-García *et al.*,[25] with the consequence of this level only being reduced under good classroom acoustics.

An optimal degree of reverberation was found, that is, the minimum value of the best fit quadratic regression curve between $T30_{0.250\text{–}2\text{kHz,occ}}$ and $\text{SPL}_{\text{mean,1m}}$, which corresponds to a reverberation of 0.7 s. This relation corroborates the results of the study by Bottalico and Astolfi,[18] in which the same quadratic curve was found for a monitored sample of 40 primary school teachers. They found that the minimum of the quadratic relation corresponded to 0.8 s of reverberation. Yang and Bradley[43] found that 0.3–0.9 s was an optimal reverberation range for good speech intelligibility in primary school classrooms, with an optimal peak value of 0.68 s. Recommended reverberation time values to preserve speech intelligibility and vocal comfort in fully occupied classrooms with volume below 210 m^3 and with less than 40 students, is in the range between 0.45 and 0.60 s, as reported by Pelegrín-García *et al.*[27] They also found a quadratic regression between the $\text{DT}_{40\text{ME},0.5\text{–}2\text{kHz}}$ and the perceived sensation of vocal comfort, which was defined as the average of the subjective impression related to several aspects of voice use in different acoustic environments. Particularly, they found the recommended $\text{DT}_{40\text{ME},0.5\text{–}2\text{kHz}}$ values in the range between 0.35 and 0.55 s to maximize the vocal comfort. The same regression model was applied to the measured data presented in this study, which was related to the quadratic regression between voice $\text{SPL}_{\text{mean,1m}}$ and $\text{DT}_{40\text{ME},0.5\text{–}2\,\text{kHz}}$, as shown in Fig. 4, so that a comparison between objectively measured voice sound pressure level and the perceived vocal comfort could be done, although the latter was not directly investigated in this work. Therefore, a recommended $\text{DT}_{40\text{ME},0.5\text{–}2\text{kHz}}$ value of 0.49 s and range between 0.29 and 0.53 s were found to minimize $\text{SPL}_{\text{mean,1m}}$ and maximize the vocal comfort, respectively, which is in good agreement with the referenced study. Given the relationship between $\text{DT}_{40\text{ME},0.5\text{–}2\text{kHz}}$ and $T30_{0.250\text{–}2\text{kHz,occ}}$ (Table III) and based on the prediction model in Pelegrín-García *et al.*,[27] a corresponding $T30_{0.250\text{–}2\text{kHz,occ}}$ range between 0.6 and 1.0 s can thus be found and corroborates all the findings.

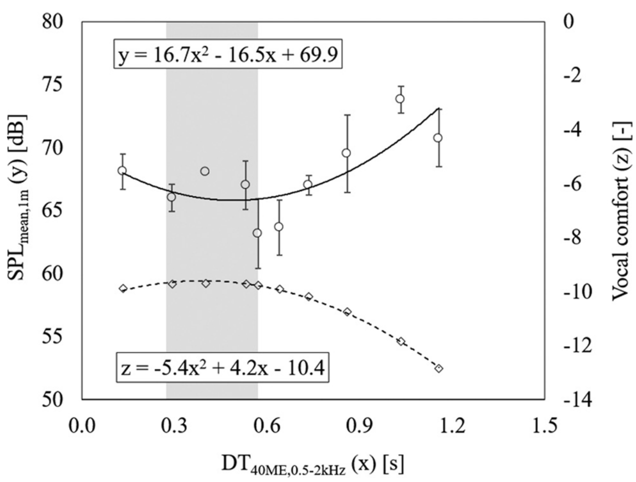

FIG. 4. Best-fit quadratic regression curve between the voice levels of the teachers ($\text{SPL}_{\text{mean,1m}}$, continuous black line with $R^2 = 0.60$), their predicted vocal comfort (dashed black line with $R^2 = 1.00$) and the Decay Time Measured at the speaker's Ears in occupied conditions ($DT_{40\text{ME},0.5\text{-}2\text{kHz,occ}}$). Each experimental data in the graph represents the mean value of an average of six pairs. The error bars refer to the SD of the mean (standard error, SE). The recommended range to minimize voice level and to maximize vocal comfort is highlighted in grey. Note that the curve referred to the predicted vocal comfort was shifted downward in the graph for representation needs of a constant quantity set at -10.

V. CONCLUSIONS

This study has investigated the vocal behavior of 27 female primary school teachers who were monitored in a four-day-follow-up, which is equivalent to one working week, in their classrooms, and the relationships between voice parameters and classroom acoustic parameters have been studied in detail. The teachers taught in classrooms that had greatly different reverberation times, when occupied. A total number of 61 voice monitorings were included in this work, with an average of 2.3 working days monitored per teacher. The teachers' voice parameters were analyzed in relationship to the measured classroom acoustic parameter reverberation time ($T30_{0.250\text{–}2\text{kHz,occ}}$), Voice Support ($ST_{\text{V},0.5\text{–}2\,\text{kHz}}$), Room Gain ($G_{\text{RG},0.5\text{–}2\,\text{kHz}}$), Decay Time at the ears ($DT_{40\text{ME},0.5\text{–}2\,\text{kHz}}$), and background noise level (L_{A90}).

First, the analysis of the teachers' occupational voice use outlined a slight tendency of SPL and F_0 to decrease from day one to day four, but it was not statistically significant. Nevertheless, it was found that the teachers tended to reduce significantly the mode of sound pressure level ($\text{SPL}_{\text{mode,16\,cm}}$) in the comfortable speech condition measured during the pre-monitoring interview (PM), which proves the need to rest and to reduce the load of their voice after teaching since they had overstrained it in the working hours. On the basis of these outcomes and being time-demanding, the occupational voice monitoring can be effectively planned to be longitudinal along at any weekday of an entire year of teaching to preserve the teachers' vocal safety, since the observation for continuous days in the same week does not give statistically different results. Second, the teachers' mean voice level and fundamental frequency were significantly higher in the occupational setting than in the conversational one (by 5.5 dB and 50 Hz, respectively), which represent the voice overstrain that teachers have to face every day. As far as the occupational voice use is concerned, the teachers' vocal effort ($\text{SPL}_{\text{eq,1m}}$) was found on average to be 71.2 dB, which is somewhere in the range between "raised" and "loud" according to ANSI S3.5–1997.[37] Last, it was investigated whether the teachers' mean SPL significantly changed due to changes in background noise and reverberation time, being strongly correlated themselves as acoustic parameters. A Lombard reflex was found at a rate of 0.53 dB/dB, and an optimal reverberation time of 0.7 s can be assumed to minimize the vocal effort. Also, a corresponding $\text{DT}_{40\text{ME},0.5\text{–}2\text{kHz}}$ in the range between 0.29 and 0.53 s was found to minimize the voice level and maximize the vocal comfort. These findings are in agreement with other studies that indicate similar values to guarantee good support to a talker and optimal speech intelligibility conditions, thus suggesting that an integrative approach in the acoustic design of classrooms should be used to meet the needs of both the talkers and the listeners.

Future research should explore other statistical analyses to investigate on the influence of intra-speaker variability in the relationships between classroom acoustics and voice parameters. The results obtained so far support the hypothesis that a strategic plan for the acoustical renovation of Italian primary school classrooms would be needed to meet the current optimal acoustic comfort requirements to preserve speech intelligibility and to ensure vocal comfort.

ACKNOWLEDGMENTS

This work has been funded by the Italian National Institute for Occupational Safety. The kind availability and cooperation of the administrators, children, and teachers of the involved primary schools have made this research possible. Finally, particular thanks are extended to Lorenzo Pavese, who helped in the first part of the monitoring campaign with the data collection and elaboration.

[1]E. Vilkman, "Voice problems at work: A challenge for occupational safety and health arrangement," Folia Phoniatr. Logop. **52**, 120–125 (2000).

[2]N. R. Williams, "Occupational groups at risk of voice disorders: A review of the literature," Occupational Med. **53**, 456–460 (2003).

[3]A. Astolfi, P. Bottalico, A. Accornero, M. Garzaro, J. Nadalin, and C. Giordano, "Relationship between vocal doses and voice disorders on primary school teachers," in *Proceedings of 9th Conference on Noise Control - Euronoise*, Prague, Czech Republic (June 10–13, 2012).

[4]B. Fritzell, "Voice disorders and occupations," Log. Phon. Vocol. **21**, 7–12 (1996).

[5]I. R. Titze, J. Lemke, and D. Montequin, "Populations in the U.S. workforce who rely on voice as a primary tool of trade: A preliminary report," J. Voice **11**(3), 254–259 (1996).

[6]National Center for Education Statistics, "Digest of Education Statistics: 2013," http://nces.ed.gov/programs/digest/d13/tables/dt13_208.20.asp (Last viewed February 15, 2016).

[7]H. A. Cheyne, H. M. Hanson, R. P. Genereux, K. N. Stevens, and R. E. Hillman, "Development and testing of a portable vocal accumulator," J. Speech Lang. Hear. Res. **46**(6), 1457–1467 (2003).

[8]J. G. Švec, I. R. Titze, and P. S. Popolo, "Estimation of sound pressure levels of voiced speech from skin vibration of the neck," J. Acoust. Soc. Am. **117**(3), 1386–1394 (2005).

[9]D. D. Mehta, M. Zañartu, S. W. Feng, H. A. Cheyne, and R. E. Hillman, "Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform," IEEE T. Biom. Eng. **59**(11), 3090–3096 (2012).

[10]C. S. Gaskill, S. G. O'Brien, and S. R. Tinter, "The effect of voice amplification on occupational vocal dose in elementary school teachers," J. Voice **26**(5), 19–27 (2012).

[11]E. J. Hunter and I. R. Titze, "Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings," J. Speech Lang. Hear. Res. **53**, 862–875 (2010).

[12]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," J. Voice **31**(1), 120.e1–120.e8 (2016).

[13]H. Lane and B. Tranel, "The Lombard sign and the role of hearing in speech," J. Speech Hear. Res. **14**(4), 677–709 (1971).

[14]A. Astolfi, A. Carullo, L. Pavese, and G. E. Puglisi, "Duration of voicing and silence periods of continuous speech in different acoustic environments," J. Acoust. Soc. Am. **137**(2), 565–579 (2015).

[15]P. Bottalico, A. Astolfi, S. Graetzer, and E. J. Hunter, "Silence and voicing accumulations in primary school teachers with and without voice disorders," J. Voice in press (2016).

[16]A. Astolfi, G. E. Puglisi, L. C. Cantor Cutiva, L. Pavese, A. Carullo, and A. Burdorf, "Associations between objectively-measured acoustic parameters and occupational voice use among primary school teachers," Energy Procedia **78**, 3422–3427 (2015).

[17]B. Shield and J. E. Dockrell, "The effects of environmental and classroom noise on the academic attainments of primary school children," J. Acoust. Soc. Am. **123**(1), 133–144 (2007).

[18]P. Bottalico and A. Astolfi, "Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms," J. Acoust. Soc. Am. **131**(4), 2817–2827 (2012).

[19]H. Sato and J. S. Bradley, "Evaluation of acoustical conditions for speech communication in working elementary school classrooms," J. Acoust. Soc. Am. **123**(4), 2064–2077 (2008).

[20]N. Durup, B. Shield, S. Dance, and R. Sullivan, "An investigation into relationships between classroom acoustic measurements and voice parameters of teachers," Build. Acoust. **22**(3+4), 225–242 (2015).

[21]D. Pelegrín-García, B. Smits, J. Brunskog, and C. Jeong, "Vocal effort with changing talker-to-listener distance in different acoustic environments," J. Acoust. Soc. Am. **129**(4), 1981–1990 (2011).

[22]H. Lazarus, "Prediction of verbal communication in noise—A review: Part 1," Appl. Acoust. **19**, 439–463 (1986).

[23]J. Brunskog, A. C. Gade, G. P. Bellester, and L. R. Calbo, "Increase in voice level and speaker comfort in lecture rooms," J. Acoust. Soc. Am. **125**(4), 2072–2082 (2009).

[24]D. Pelegrín-García, "Comment on 'Increase in voice level and speaker comfort in lecture rooms,'" J. Acoust. Soc. Am. **129**(3), 1161–1164 (2011).

[25]D. Pelegrín-García, O. Fuentes-Mendizàbal, J. Brunskog, and C. Jeong, "Equal autophonic level curves under different room acoustics conditions," J. Acoust. Soc. Am. **130**(1), 228–238 (2011).

[26]V. Lyberg-Åhlander, R. Rydell, A. Löfqvist, D. Pelegrín-García, and J. Brunskog, "Part Summary of the project 'Speakers' Comfort': Teachers' voice use in teaching environments," Build. Acoust. **22**(3+4), 209–224 (2015).

[27]D. Pelegrín-García, J. Brunskog, and B. Rasmussen, "Speaker-oriented classroom acoustics design guidelines in the context of current regulations in European countries," Acta Acust. Acust. **100**, 1073–1089 (2014).

[28]A. Astolfi, V. Corrado, and A. Griginis, "Comparison between measured and calculated parameters for the acoustical characterization of small classrooms," Appl. Acoust. **69**, 966–976 (2008).

[29]BS EN ISO 3382-2: *Acoustics - Measurement of Room Acoustic Parameters – Part 2: Reverberation Time in Ordinary Rooms* (International Organization for Standardization, Genève, Switzerland, 2008).

[30]DIN 18041: *Hörsamkeit in kleinen bis mittelgroßen Räumen* (*Acoustical Quality in Small to Medium-Sized Rooms*) (Deutsche Institut für Normung, Berlin, Germany, 2004).

[31]D. Pelegrín-García and J. Brunskog, "Speakers' comfort and voice level variation in classrooms: Laboratory research," J. Acoust. Soc. Am. **132**(1), 249–260 (2012).

[32]D. Pelegrín-García, J. Brunskog, V. Lyberg-Åhlander, and A. Lofqvist, "Measurement and prediction of voice support and room gain in school classrooms," J. Acoust. Soc. Am. **131**(1), 194–204 (2012).

[33]A. Carullo, A. Vallan, and A. Astolfi, "Design issues for a portable vocal analyzer," IEEE T. Instrum. Meas. **62**(5), 1084–1093 (2013).

[34]A. Carullo, A. Vallan, A. Astolfi, L. Pavese, and G. E. Puglisi, "Validation of calibration procedures and uncertainty estimation of contact-microphone based vocal analyzers," Measurement **74**, 130–142 (2015).

[35]A. Carullo, A. Penna, A. Vallan, A. Astolfi, L. Pavese, and G. E. Puglisi, "Traceability and uncertainty of vocal parameters estimated through a contact microphone," in *Proceedings of IEEE International Symposium on Medical Measurements and Applications*, Lisbon, Portugal (June 11–12, 2014).

[36]A. Nacci, B. Fattori, V. Mancini, E. Panicucci, F. Ursino, F. M. Cartaino, and S. Berrettini, "The use and role of the Ambulatory Phonation Monitor (APM) in voice assessment," Acta Otorhinolaryngol. Ital. **33**(1), 49–55 (2013).

[37]ANSI S3.5: *Methods for Calculation of the Speech Intelligibility Index* (Acoustical Society of America, New York, 1997).

[38]H. J. Seltman, "Experimental design and analysis – September 2015," http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf, Chap. 13 (Last viewed March 7, 2016).

[39]WSP, *Acoustic Design in Schools: Performance Standards* Building Bulletin, BB93 (Department for Education and the Educational Funding Agency, London, 2015).

[40]B. Shield and J. E. Dockrell, "External and internal noise surveys of London primary schools," J. Acoust. Soc. Am. **115**(2), 730–738 (2004).

[41]M. Cooke, S. King, M. Garnier, and V. Aubanel, "The listening talker: A review of human and algorithmic context-induced modifications of speech," Comput. Speech Lang. **28**(2), 543–571 (2014).

[42]K. K. Baker, L. O. Ramig, S. Sapir, E. S. Luschei, and M. E. Smith, "Control of vocal loudness in young and old adults," J. Speech Lang. Hear. Res. **44**(2), 297–305 (2001).

[43]W. Yang and J. S. Bradley, "Effects of room acoustics on the intelligibility of speech in classrooms for young children," J. Acoust. Soc. Am. **125**(2), 922–933 (2009).