Speech level in rooms with very low and very high reverberation

Arianna Astolfi[1], Giuseppina Emma Puglisi[1], Umberto Fugiglando[1],

Alessio Carullo[2], Antonella Castellana[2]

[1]Politecnico di Torino, Department of Energy
(arianna.astolfi@polito.it, giuseppina.puglisi@polito.it)
[2]Politecnico di Torino, Department of Electronics and Telecommunications
(alessio.carullo@polito.it, antonella.castellana@polito.it)

Abstract

Numerous studies have dealt so far with changing in speech production for talkers due to different acoustic environments, but mostly have been focused on the effect of noise or distance from the listeners and not many on the effect of reverberation. Reverberation has an influence on the voice production, supporting talkers as well as increasing speech level towards an audience. In spite of its positive effects, excessive reverberation influences talkers making them use erroneous vocal behaviors, which may be cause of discomfort and a risk for vocal health, especially in the case of prolonged speech, as for teachers in classrooms.

The present work explores the differences in sound pressure level and sound power level of running speech from several speakers, in semi-anechoic and reverberant rooms. Two types of spontaneous speech, a free monologue and the description of a map, have been addressed with a communicative intent to a listener seated at a fixed distance of 6 m. Measurements have been carried out with a headworn microphone and with Voice Care, a voice dosimeter based on a contact microphone. Subjective impressions were caught after subjects have spoken in the two rooms.

Keywords: speech level, reverberation, semi-anechoic room, reverberant room, voice dosimeter.

PACS no. 43.55.Hy, 43.70.Jt, 43.72.Ar.

1 Introduction

When background noise is present in a room, a global increase of speech intensity occurs, leading to Lombard reflex [1]. Lombard reflex is highly variable from speaker to speaker, leading to a significant inter-speaker variability [2]. Even in the absence of masking noise, level increases are observed in speech produced at distance, perhaps as a form of compensation for perceived listener difficulties [3-4].

The increase in vocal intensity is accompanied by a raised fundamental frequency and a flatter spectral tilt, resulting in enhanced speech energy in medium and high frequencies [5-6]. These modifications can simply be considered as direct consequences of an increase in vocal effort [7], since it is well known that intensity and frequency are not independently controlled in the human vocal system and speakers tend to raise their voice in pitch when they raise their voice in loudness [8].

Despite the extensive literature on the effects of noise on speech, few, if any, data have been published reporting details of the acoustic changes at global level that take place when a speaker modifies his vocal output while speaking in the presence of reverberation.
Black [9] investigated the effect of size, shape and reverberation time of different rooms on vocal intensity and found that vocal intensity was higher in small rooms than in large ones and among the large rooms the intensity was higher in live rooms (difference of 3.11 dB on average, significant at a confidence level of 99%).
Pelegrín-García *et al*. [4] by means of a headworn microphone analyzed the effect of acoustical environment on the natural speech produced by 13 male talkers, evoked by means of a map task in the absence of background noise. At 6 m from the speaker's mouth the sound power level increased by 2.4 dB in anechoic chamber compared to reverberation chamber. Speakers said that both the anechoic chamber and reverberation rooms were unnatural and uncomfortable places to speak in.
Recently developed devices estimate the vocal sound pressure level from the skin vibration at the position of the vocal folds, through the use of contact microphones glued to the speaker's neck, at the jugular notch. These devices have been produced with the intent of long-term voice monitoring. Few information on the uncertainty related to voice monitoring with these devices have been published. Ŝvec *et al.* [10] found that the SPLs at 30 cm from the speaker's mouth could be estimated with a 95% confidence interval of ± 5 dB in female and ± 6 in males. Hillman *et al.* [11] compared sound pressure levels extracted from a microphone placed 15 cm from the mouth and from a small accelerometer. The average errors in the estimation of the sound pressure level from the acceleration signal were 3.2 ± 6.2 dB. Carullo *et al.* [12] found that the standard uncertainty on the sound pressure level estimated at 16 cm from the speaker's mouth with an electret condenser microphone (ECM) is not greater than 4.2 dB.
The present work investigates differences in sound pressure level and sound power level of running speech in semi-anechoic and reverberant rooms, with a speaking style similar to the one used in everyday life. Two types of spontaneous speech, a free monologue and the description of a map, have been addressed with a communicative intent from some speakers to a listener seated at a fixed distance of 6 m. Measurements have been carried out with a headworn microphone and with Voice Care [13], a voice dosimeter based on a ECM. Subjective impressions were also acquired after the subjects had spoken in the two rooms by means of a short close-answers questionnaire.

2 Experimental set-up

Experiments were carried out in the semi-anechoic and reverberant rooms of the National Institute of Metrological Research (INRiM), in Turin (Italy). Table 1 shows the volume, the mid-frequency average reverberation time and the background noise levels measured in the unoccupied rooms. The task of each speaker was to tell a short continuous free speech based on a topic they knew well to a listener sat in front of him/her, firstly in the semi-anechoic room and then in the reverberant room. Describing a map was another speaking task accomplished by a subgroup of subjects. Changings in the vocal parameters were detected through the use of the omnidirectional MU-55HN headworn microphone by Mipro (Chiayi, Taiwan) and a vocal analyzer, the Voice Care, developed at the Politecnico di Torino (Italy).

Table 1 - Volume, average reverberation time at the octave band center frequencies 0.5 kHz, 1 kHz and 2 kHz, and A-weighted equivalent background noise level measured in the semi-anechoic and reverberant rooms of INRiM.

	Volume (m^3)	$T_{0.5\text{-}2\ \text{kHz}}$ (s)	$L_{\text{Aeq,bn}}$ (dB)
Semi-anechoic room	384	0.11 (s.d. 0.01)	24.5
Reverberant room	294	7.38 (s.d. 1.61)	30.3

2.1 The subjects

The sample of speakers was composed of 44 subjects, 21 males and 23 females, with the majority of them being MSc or PhD students of the Politecnico di Torino. None of the participants had any severe visual impairment or any vocal or hearing disorders. Table 2 shows the number of subjects who undertook the various experiments in both the semi-anechoic and reverberant rooms. Typically, each speaker ran the speech task wearing the two devices contemporarily.

2.2 Instructions

The speakers were asked to make a continuous 5 minute-long free speech, with the aim of transmitting information on something they knew well (e.g. the research topic they dealt with, a recipe, the rules of a game, the path from their house to the workplace, etc.), while standing 6 m away from a young listener, sat on-axis in front of them such as to enable eye-contact.
In order to evoke another form of natural speech in a very specific mode of communication, part of the subjects were also asked to describe a map. The map contained twelve landmarks (e.g., "school bus," "shop," and "yacht club"), starting and ending point marks, and a path connecting these two points. The speakers were instructed to describe the route from the starting to the ending points, indicating the landmarks along the path, while trying to enable visual-contact with the listener, who were tracking the route on a mute map.
Before they started speaking, each subject was instructed to repeat the vowel /a/ and to hit the ECM of the Voice Care device at the same time, so that the signal acquired by the contact microphone and the signal acquired by the Mipro MU-55HN headworn microphone were synchronized.

Table 2 - Number of subjects who undertook the experiments with the Mipro MU-55HN headworn microphone and the Voice Care device, for the speech tasks of free speech and describing a map. Distinction between female (F) and male (M) is also reported.

	Voice Care			Mipro MU-55HN		
	F	M	Overall	F	M	Overall
Free speech	8	15	23	16	13	29
Describing a map	5	10	15	14	14	28
Overall	13	25	**38**	30	27	**57**

2.3 Questionnaires

A questionnaire was administered to the talkers at the end of each talk in order to find out the preferred acoustic settings between the rooms. It held 7 statements, partly taken from a previous study by Pelegrín-García and Brunskog [14]. It is based on a 5-point discrete scale in which each step is labelled from 1 to 5 and the first and last has opposite semantic descriptors.
The rated questions or statements were the following: degree of perceived reverberance in the venue (from 'very low' to 'very high'); the venue is good to speak in; the venue enhances and supports my speech; I must raise my voice in order to be heard in the venue; the acoustics makes my voice sound unnatural; how you perceive your voice at the end of the experiment (from 'no voice problem' to 'extremely severe voice problems').

3 Measurement of the vocal parameters

The overall mean sound pressure levels of the voiced speech, L_m, the mode, L_{mode}, and the overall equivalent sound pressure levels, L_{eq}, were obtained for each talk. In the case of Voice Care, the equivalent vocal sound pressure level estimated from the skin vibration, according to Ŝvec *et al.* [10], is calculated as the average of the voiced energy over all the frames, including the unvoiced ones, whose energy was set to zero. In the case of the headworn microphone Mipro MU-55HN, the sound power level (L_W) of the speaker's voice was estimated according to a procedure described below.

3.1 Voice Care

Before starting the speech task, the speakers were supplied with Voice Care [13], a portable vocal analyzer that consists of an Electret Condenser Microphone (ECM AE38 [Alan Electronics GmbH (Dreieich, Germany)]), fixed at the jugular notch by means of a surgical band, connected to a data-logger. The acquired samples were subdivided into frames of 30 ms, which correspond to the inter-syllabic pauses.
The device provides an estimation of the voiced sound pressure levels (SPLs) at a 16 cm from the speaker's mouth after a calibration vs a reference microphone, which consists in repeating the vowel /a/ at increasing levels. Such procedure had been performed before starting the experiments in both the semi-anechoic and reverberant rooms. In order to determine the difference in the vocal output between the two rooms, the same calibration function of the semi-anechoic chamber was also applied for both the monitoring. The results are usually shown as histograms of occurrences related to estimated SPLs.

3.2 Headworn microphone Mipro MU-55HN

The acoustic speech signal was picked up with the omnidirectional headworn microphone Mipro MU-55HN placed at a distance of about 2.5 cm from the lips' edge of the talkers, on-axis. The microphone exhibits a flatness of ±3 dB in the range from 40 Hz to 20 kHz. It was connected to the bodypack transmitter ACT-30T that transmits to a wireless microphone system Mipro ACT 311. The wav signals were recorded with the handy recorder ZOOM H1 (Zoom Corp., Tokyo, Japan) in 16 bits/44.1 kHz format and later processed with MATLAB. The sound pressure levels were sampled with a logging interval of 1 s.
An arbitrary calibration of the headworn microphone was carried out against the reference sound level meter XL2, by NTi Audio. The characterization was performed in the anechoic chamber of Politecnico di Torino, where a background noise level of 26.2 dB was measured. Initially, the sound level meter was calibrated by coupling to the pressure calibrator B&K 4230, which provides a nominal pressure of 1 Pa @ 1 kHz. Then, both the sound level meter and the headworn microphone were placed at a distance of 2.5 cm, on-axis, from the B&K type 4128 Head and Torso Simulator (B&K, Nærum, Denmark) equipped with a loudspeaker in the mouth. The HATS was connected through the amplifier ALPINE MRP T222 (Alpine Electronics, Inc., Tokyo, Japan) and the audio device TASCAM US-144 (TEAC America, Inc., Montebello, CA) to a notebook PC. The software DIRAC 5 was run to generate different sound pressure levels of ICRA noise, i.e. an artificial noise signal with speech-like spectral properties, in the usual range observed in professional voice users (55-72 dB @ 1 m) [6]. The output signals from the headworn microphone and the reference microphone were simultaneously acquired and post-processed by means of MATLAB scripts. Equivalent, mean and mode sound pressure level were calculated using the calibration wav file as a reference. The difference between the equivalent sound pressure levels estimated by the two devices was added to the headworn microphone levels in order to obtain calibrated values.

A correction factor, G_{refl}, due to the increase of the overall equivalent sound pressure level at the headworn microphone in the reverberant room compared to the semi-anechoic room, was also obtained. For this measurements the HATS emitting ICRA noise, equipped with the headworn microphone, was used. The equivalent sound pressure level reading from the semi-anechoic room was subtracted to the reading in the reverberant chamber in order to obtain G_{refl}, which resulted to be 0.34 dB (st. dev. 0.05). It is useful to underline that being the microphone close enough to the source so that only the direct field is present, i.e. the signal to noise ratio is assumed to be so good, and according to Brunskog *et al.* [15] the G_{refl} can be still neglected.
The correction factor G_{refl} was then subtracted to each overall equivalent sound pressure level of the speakers measured in the reverberant room in order to get the corrected overall equivalent sound pressure level from reflections. It should be noted that being the dead room a semi-anechoic room, a reflection from the floor could have been occurred compared to a full anechoic room. In order to suppress this reflection thick sound absorbing panels were placed on the floor of the room.
The difference between the corrected overall equivalent sound pressure level and the overall power level, G_{dist}, was then determined by performing sound power level measurements in the reverberant room, in a similar way as described by Brunskog *et al.* [15]. In particular, the HATS was placed in the reverberant chamber equipped with the headworn microphone; an ICRA noise signal was fed to the loudspeaker and measured simultaneously by the headworn microphone and by calibrated 1/2" microphone, B&K type 4943, located in the reverberant field of the room, according to the sound power level standard measurements ISO 3743-2 [16].
It was then assumed that all the speakers had the same directivity, equal to that of the HATS. It was thus considered that G_{dist}, which resulted equal to 23.3 dB (st. dev. 0.05), was constant for each speaker. The overall sound power level of each subject in the reverberant room was finally estimated as the difference between the corrected overall equivalent speech sound pressure level in the reverberant chamber and G_{dist}; while the overall speech sound power level of each subject in the semi-anechoic room was estimated as the difference between the overall equivalent speech sound pressure level in the semi-anechoic room and G_{dist}.

4 Results

The results concern the comparison of the speech sound pressure levels between the different device in the two rooms, the different rooms with the same device and the different tasks. Very important: the sound power level difference between the two rooms and the sound pressure level difference when the calibration function of the semi-anechoic chamber is applied to both the monitorings, are measures of increased vocal effort due to the environment.
Due to the different sample frequency [10] and distance from the mouth [17], the comparison between the devices can be affected by other causes of uncertainty and has been considered as a preliminary confirmation of the correctness of the adopted procedure.

4.1 Statistical analysis

The statistical analysis was carried out with a MATLAB script and the results compared with IBM SPSS statistics package (version 21.0, Armonk, NY). The outcomes of two conditions were initially compared using the one-tailed Wilcoxon signed-rank test [18]. The monitorings in the two rooms of the same subject were considered dependent and a test based on paired samples was applied. The test requires two related samples or repeated measurements on a single sample, taken in pairs, without any specific assumptions on the distributions, and tests whether the median of the difference between the pairs is greater than zero. In order to apply the test, the equivalent, mean and mode of the SPL

distributions were calculated in the two different rooms for each subject involved in the study, and a single pair was thus obtained for each subject.
A different approach was then applied considering that the SPL distributions of the monitorings for every single subject over the two rooms can be considered independent, as long as the speech made by the subject was different in the two rooms. In this case, the Mann-Whitney U test was adopted in its unilateral version. It assesses, for the Voice Care data and individually for each subject, whether the two different sets of samples in the two rooms come from populations that are one stochastically larger than the other, without making any assumptions on the type of the distributions.

4.2 Results from Voice Care monitoring

Table 3 shows the average values of the equivalent, mean and mode sound pressure levels estimated with Voice Care at 16 cm from the speaker's mouth, in the semi-anechoic (sa) and reverberant (r) rooms of INRiM, and the level differences between the rooms. Sound pressure levels were obtained either applying the calibration function acquired in the room where the monitoring took place or applying the calibration function of the semi-anechoic room to both the monitoring in the two rooms. Lower p-values than a significance level of 0.05 of the one-tailed Wilcoxon signed ranks test are also shown in Table 3. A tendency of higher equivalent, mean and mode sound pressure level is shown in the reverberant room compared to the semi-anechoic room in the case of free speech, as expected when two different calibration functions for the two rooms were used. The same behaviour resulted not significant in the case of describing a map. Higher mean values of 2.5 dB, 3.5 dB and 2.7 dB were found in the reverberant room compared to the semi-anechoic room.
A tendency of higher equivalent, mean and mode sound pressure level is also shown in the semi-anechoic room compared to the reverberant room in the case of describing a map, when the same calibration function has been applied for the two rooms. The same behaviour resulted not significant in the case of free speech. Higher mean values of 2.1 dB, 1.8 dB and 2.4 dB were found in the semi-anechoic room compared to the reverberant room. This finding supports an increase in the vocal intensity in a dead room compared to a reverberant room of about 2 dB, as found in literature [4], but only in the case of the specific task of describing a map.
The same result has been supported by the Mann-Whitney U test for the same speaker in the two rooms, when the same calibration function has been applied. Table 4 shows that in the case of describing a map, the alternative hypothesis H_1 is accepted in 13 cases out of 15, i.e. the distributions in the semi-anechoic chamber are significantly larger than in the reverberant room. In the case of free speech there is not a tendency, as only 10 out of 23 subjects support the previous finding.
The average value of the equivalent sound pressure level over the overall sample was 79.8 dB (st. dev. 1.5) and 82.3 dB (st. dev. 1.5) in the semi-anechoic and the reverberant room respectively, in the case of free speech; a speech equivalent sound pressure level of 79.0 dB (st. dev. 2.1) and 80.3 dB (st. dev. 1.9) has been found in the case of describing a map, respectively. Higher values in reverberant room are expected, but results are comparable among them, and correspond to a vocal effort, in terms of equivalent sound pressure level @ 1 m from the speaker's mouth (-15.9 dB), between "normal" and "raised" according to the ANSI S3.5 [19].
In the case of free speech, for both the rooms, the equivalent sound pressure level is greater than the mode, which is greater than the mean sound pressure level. In the case of describing a map the mode sound pressure level is greater than the equivalent sound pressure level, which is greater than the mean. These outcomes corroborate the results by Švec et al. [10], who confirmed that the equivalent sound pressure level is higher than the mean, in the case of reading with loud voice.
In all the cases the mode is greater than the mean, remarking therefore a non-normal distribution of the sound pressure level occurrences. The differences between mode and mean are greater in the case of describing a map than in the case of free speech and in the semi-anechoic room than in the reverberant room, hence supporting increased fatigue in the case of describing a map and speaking in a dead room.

Table 3 - Average value (upper cells) and standard deviation of the average (lower cells) of equivalent, mean and mode sound pressure level (dB) estimated with Voice Care at 16 cm from the speaker's mouth, in the semi-anechoic (sa) and reverberant (r) rooms of INRiM, and level differences between the rooms (ΔL). Sound pressure levels were obtained either applying the calibration function acquired in the room where the monitoring took place, L_{sa} and L_r, or applying the calibration function of the semi-anechoic room to both the monitoring in the two rooms, L_{sa} and $L_{r(sa)}$. Results are shown in the case of free speech and for the speech task of describing a map for a number of female (F) and male (M) subjects and overall for each speech task. The *p*-values of the one-tailed Wilcoxon signed ranks test of the paired lists of the parameters related to the overall sample in two rooms are at the bottom. Values lower than a significance level of 0.05, reported in bold and italic style, indicate the acceptance of the alternative hypothesis H_1 which states both that $H_{1,left}$:M_D<0 and $H_{1,right}$:M_D>0, where M_D is the median of the population of the differences between the paired sample data in the two rooms. In the case of reverberant room the two different calibration functions were considered.

Type of speech	Sample	Subj.	L_{eq}			L_m			L_{mode}			ΔL_{eq}		ΔL_m		ΔL_{mode}	
			sa	r	r(sa)	sa	r	r(sa)	sa	r	r(sa)	sa-r	sa-r(sa)	sa-r	sa-r(sa)	sa-r	sa-r(sa)
Free speech	F	8	79.6	80.6	79.5	76.3	79.1	76.6	78.0	80.5	78.5	-1.0	0.1	-2.9	-0.3	-2.5	-0.5
			3.5	3.1	4.0	3.2	2.8	3.3	3.3	3.0	3.4	2.8	2.0	2.6	1.8	2.4	1.8
	M	15	79.9	83.3	78.0	78.1	81.9	76.0	80	82.7	77.5	-3.4	1.9	-3.8	2.1	-2.7	2.5
			1.4	1.7	2.1	1.4	1.8	2.3	1.5	1.8	2.7	1.8	1.6	1.5	1.5	1.5	2.0
	Overall	23	**79.8**	**82.3**	**78.5**	**77.4**	**80.9**	**76.2**	**79.3**	**82.0**	**77.9**	**-2.5**	**1.3**	**-3.5**	**1.3**	**-2.7**	**1.4**
			1.5	1.5	1.9	1.4	1.5	1.9	1.5	1.6	2.1	1.5	1.2	1.3	1.1	1.3	1.5
	p-value			***0.049***			***0.011***			***0.025***							
				$H_{1,left}$			$H_{1,left}$			$H_{1,left}$							
Map	F	5	75.5	77.8	73.8	74	77.6	73.6	78	80	76.0	-2.3	1.7	-3.6	0.4	-2	2.0
			5.2	4.9	4.8	4.5	4.4	4.2	5.7	4.8	4.9	2.0	1.1	1.4	0.6	1.9	1.1
	M	10	80.8	81.5	78.6	81	81.7	78.5	84	83.1	81.4	-0.7	2.2	-0.7	2.6	0.9	2.6
			1.9	1.6	1.2	1.9	1.4	1.0	2.4	1.3	1.3	2.3	0.9	2.0	1.1	2.4	1.2
	Overall	15	**79.0**	**80.3**	**77.0**	**78.7**	**80.3**	**76.8**	**82.0**	**82.1**	**79.6**	**-1.2**	**2.1**	**-1.7**	**1.8**	**-0.1**	**2.4**
			2.1	1.9	1.8	2	1.7	1.6	2.5	1.7	1.8	0.7	0.7	1.4	0.8	1.7	0.9
	p-value				***0.004***			***0.007***			***0.004***						
					$H_{1,righ}$			$H_{1,righ}$			$H_{1,righ}$						

4.3 Results from Mipro headworn microphone

Table 5 shows the average values of equivalent, mean and mode sound pressure level and sound power level, estimated with the headworn microphone Mipro MU-55HN at about 2.5 cm from the speaker's mouth, in the semi-anechoic and the reverberant rooms of INRiM, and the level differences between the rooms. An increase in the mean sound pressure level for the case of describing a map in the reverberant room compared to the semi-anechoic room has been observed as well as in the sound power level in the semi-anechoic room compared to the reverberant room, for the case of free speech.
The average value of the equivalent sound pressure levels over the overall sample was 94.3 dB (st. dev. 0.8) and 93.5 dB (st. dev. 0.9) in semi-anechoic and reverberant room respectively, in the case of free speech; a speech equivalent sound pressure level of 94.7 dB (st. dev. 0.9) and 94.3 dB (st. dev. 0.9) has been found in the case of describing a map, respectively. Results are comparable among them, and correspond to a "normal" vocal effort, in term of equivalent sound pressure level @ 1 m from the speaker's mouth (-32 dB), according to the ANSI S3.5 [19].
For both the speech tasks and the rooms, the equivalent sound pressure level is very similar to the mode sound pressure level, and both are higher that the mean sound pressure level. In all the cases the

mode is greater than the mean, remarking therefore a non-normal distribution of the sound pressure level occurrences. The differences between mode and mean are greater for the case of describing a map than for the case of free speech, hence supporting increased fatigue for the case of describing a map.

Table 4 - One-tailed Mann-Whitney U-test p-values of the significance of the difference between the two medians (M) of the distributions of sound pressure level (SPL) occurrences estimated with Voice Care at 16 cm from the speaker's mouth, for several female (F) and male (M) speakers, in the semi-anechoic (sa) and reverberant (r) rooms of INRiM, in the cases of free speech and when a map was described. SPLs were obtained applying the calibration function of the semi-anechoic room to both the monitorings in the two rooms, SPL_{sa} and $SPL_{r(sa)}$.Values lower than 0.05 are reported in bold and italic style and indicate the acceptance of the alternative hypothesis H_1:M_{sa}>$M_{r(sa)}$.

Subj.	F01	F02	F03	F04	F05	F06	F07	F08	M01	M02	M03	M04
Free speech	*1.000*	***0.000***	*1.000*	*1.000*	***0.000***	*1.000*	*1.000*	***0.000***	***0.000***	*1.000*	***0.000***	*1.000*
Map	-	-	***0.000***	***0.000***	***0.000***	-	*1.000*	***0.000***	-	-	-	-
Subj.	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	
Free speech	***0.000***	***0.000***	***0.006***	*1.000*	*1.000*	***0.000***	*1.000*	***0.000***	*1.000*	*1.000*	*1.000*	
Map	***0.000***	-	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	***0.000***	*1.000*	

Table 5 - Average value (upper cells) and standard deviation of the average (lower cells) of equivalent, mean and mode sound pressure level and sound power level (dB), estimated with the headworn microphone Mipro MU-55HN at a distance of about 2.5 cm from the speaker's mouth. Data are related to the task of free speech and of describing a map, for a number of female (F) and male (M) subjects and overall for each speech task, in the semi-anechoic (sa) and reverberant (r) rooms of INRiM. Sound pressure level and sound power level differences between the rooms (ΔL) are also reported. The p-values of the one-tailed Wilcoxon signed ranks test of the paired lists of the parameters related to the overall sample in the two rooms are at the bottom. Values lower than a significance level of 0.05, reported in bold and italic style, indicate the acceptance of the alternative hypotheses H_1, which are $H_{1,left}$:M_D<0 and $H_{1,right}$:M_D>0. M_D is the median of the population of the differences between the paired sample data in the semi-anechoic and reverberant rooms.

Type of speech	Sample	Subj.	L_{eq}		L_m		L_{mode}		$L_{w,mipro}$		$\Delta L_{eq,sa-r}$	$\Delta L_{m,sa-r}$	$\Delta L_{mode,sa-r}$	$\Delta L_{w,sa-r}$
			sa	r	sa	r	sa	r	sa	r				
Free speech	F	23	92.5	91.0	90.4	89.4	92.3	91.4	69.1	67.4	1.4	1.0	0.9	1.8
			0.8	0.8	0.8	0.7	0.9	0.8	0.8	0.8	0.6	0.6	0.6	0.6
	M	19	96.6	96.5	94.0	94.6	96.3	96.9	73.3	72.8	0.1	-0.6	-0.6	0.5
			1.4	1.2	1.4	1.3	1.5	1.0	1.4	1.2	0.9	1.1	0.9	0.9
	Overall	42	**94.3**	**93.5**	**92.0**	**91.7**	**94.1**	**93.9**	**71.0**	**69.8**	**0.8**	**0.3**	**0.2**	**1.2**
			0.8	0.9	0.8	0.9	0.9	0.8	0.8	0.9	0.5	0.6	0.5	0.5
	p-value $H_{1,right}$								***0.009***					
Map	F	13	92.5	92.2	87.9	88.9	93.0	92.7	69.2	68.6	0.3	-1.0	0.3	0.6
			1.0	1.2	1.2	1.0	1.2	1.2	1.0	1.2	0.6	0.7	0.8	0.6
	M	12	96.8	96.4	91.6	92.8	96.6	97.7	73.5	72.8	0.4	-1.2	-1.1	0.7
			1.4	1.0	1.3	1.1	1.4	1.0	1.4	1.0	1.1	1.2	1.1	1.1
	Overall	25	**94.7**	**94.3**	**89.8**	**90.9**	**94.8**	**95.2**	**71.3**	**70.7**	**0.3**	**-1.1**	**-0.4**	**0.7**
			0.9	0.9	0.9	0.8	1.0	0.9	0.9	0.9	0.6	0.7	0.7	0.6
	p-value $H_{1,left}$				***0.033***									

4.4 Results from the questionnaire

Table 6 shows the results from the questionnaire, where it can be shown that neither the semi-anechoic nor the reverberant rooms are venues good to speak in. People reported they had been raising their voice when speaking in the semi-anechoic room more than in the reverberant room. None of the rooms enhance and support the speech and in the reverberant room the voice sounds unnatural.

Table 6 - Percentage of the answers associated to the points 1+2, 3 and 4+5 on the 5-point discrete scales related to the questions submitted after the speech tasks in the semi-anechoic (sa) and reverberant (r) rooms of INRiM. The highest percentages for each question are written in bold.

Quest. N.	Room	Answers (%) 1+2	3	4+5	Question issue
1	sa	**100**	0	0	Degree of reverberance in the venue (1 = very low; 5 = very high)
	r	0	0	**100**	
2	sa	23.1	23.1	**53.8**	The venue is good to speak in (1 = totally disagree; 5 = totally agree)
	r	**94.2**	5.8	0.0	
3	sa	**40.4**	28.8	30.8	The venue enhances and supports speech (1 = totally disagree; 5 = totally agree)
	r	**73.1**	15.4	11.5	
4	sa	30.8	21.2	**48.1**	I must raise my voice in order to be heard in the venue (1 = totally disagree; 5 = totally agree)
	r	**61.5**	9.6	28.8	
5	sa	**63.5**	19.2	17.3	The acoustics makes my voice unnatural (1 = totally disagree; 5 = totally agree)
	r	23.1	17.3	**59.6**	
6	sa	**86.5**	13.5	0.0	Perception of voice problems at the end of my speech in the venue (1 = no voice problem; 5 = extremely severe voice problems)
	r	**71.2**	17.3	11.5	

5 Conclusions

The main conclusions of this study can be summarized as follow:

- In the case of describing a map, a tendency of increase the speech intensity in the semi-anechoic room compared to the reverberant room is shown;
- The mode is much greater than the mean of the SPL distributions in the case of describing a map compared to the free speech, supporting increased fatigue in the case of describing a map;

Uncertainty must be evaluated for these measures in order meaningful compare data.

Acknowledgements

The Italian National Institute for Occupational Safety has funded this work. The kind cooperation of students from the Politecnico di Torino made this work possible. Last, particular thanks are extended to the INRiM researchers who made their laboratories available for the experiments.

References

[1] Lombard, E. Le signe d'élévation de la voix, *Annales des maladies de l'oreille et du larynx* Vol. 37, 1911, pp 101-109.

[2] Junqua, J. C. The Lombard reflex and its role on human listener and automatic speech recognizers, *J. Acoust. Soc. Am*. Vol. 93, 1993, pp 510-524.

[3] Cheyne, H.; Kalgaonkar, K.; Clements, M.; Zurek, P. Talker-to-listener distance effects on speech production and perception, *J. Acoust. Soc. Am.*, Vol. 126 (4), 2009, pp 2052-2060.

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

[5] Summers, W. V.; Pisoni, D. P.; Bernacki, R.H.; Pedlow, R. I.; Stokes, M. A. Effects of noise on speech production: acoustic and perceptual analyses, *J. Acoust. Soc. Am.*, Vol. 84 (3), 1988, pp 917-928.

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

[7] Traunmuller, H.; Eriksson, A. Acoustic effects of variation in vocal effort by men, women and children, *J. Acoust. Soc. Am*., Vol. 107 (6), 2000, pp 3438-3451.

[8] Titze, I. R. On the relation between subglottal pressure and fundamental frequency in phonation, *J. Acoust. Soc. Am.*, Vol. 85, 1989, pp 901-906.

[9] Black, J. The effect of room characteristics upon vocal intensity and rate, *J. Acoust. Soc. Am*., Vol. 22, 1950, pp 174-176.

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

[11] Hillman, R.E.; Heaton, J.T.; Masaki, A., Zeitels; S.M.; Cheyne, H.A. Ambulatory monitoring of disordered voices, *Ann. Otol. Rhinol. Laryngol.*, Vol. 115(11), 2006, pp 795-801.

[12] Carullo, A.; Vallan, A.; Astolfi, A.; Pavese, L.; Puglisi, G.E. Validation of calibration procedures and uncertainty estimation of contact-microphone based vocal analyzers, *Measurement* Vol. 74, 2015, pp 130-142.

[13] Carullo, A.; Vallan, A.; Astolfi, A. Design Issues for a Portable Vocal Analyzer, *IEEE T. Instrum. Meas.*, Vol 62 (5), 2013, pp 1084-1093.

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

[15] Brunskog, J.; Gade, G.; Payà-Ballester, G.; Reig-Calbo, L. Increase in voice level and speaker comfort in lecture rooms, *J. Acoust. Soc. Am.*, Vol. 125, 2009, pp 2072–2082.

[16] ISO, International Standard 3743-2: *Acoustics - Determination of sound power levels of noise sources using sound pressure - Engineering methods for small, movable sources in reverberant fields - Part 2: Methods for special reverberation test rooms*, Genève, 1994.

[17] Šrámková , H; Granqvist, S; C. T. Herbst, C; Švec, J. G. The softest sound levels of the human voice in normal subjects, voice level and speaker comfort in lecture rooms, *J. Acoust. Soc. Am.*, Vol. 137(1), 2015, pp 407–418.

[18] Siegel, S.; Castellan Jr., N. J. *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill, New York (USA), 1988, pp 1–399.

[19] ANSI, American National Standard Institute Standard S3.5: *Methods for Calculation of the Speech Intelligibility Index*, New York, 2002.