Politecnico di Torino

Master of Science
Cinema and Media Engineering
A.A. 2020/2021
Graduation Session July 2021

Study and implementation of a methodology for measuring and assessing the Lombard effect in eating establishments during COVID-19 pandemic

The case study of the canteen employees of the Polytechnic of Turin

Supervisors:
Prof. Astolfi Arianna
Dott. Bottalico Pasquale
Dott. Puglisi Giuseppina Emma
Prof. Masoero Marco Carlo

Candidate:
Giorgio Carzedda

Abstract

Anthropic noise in highly attended environments has been observed and studied with increasing attention over the years. Very often the purpose of these studies has been finding several solutions to the acoustic discomfort linked to the noise generated by the presence of a large number of people, both indoors and outdoors. In particular, an attempt was made to analyze as thoroughly as possible the phenomenon whereby in very noisy environments a speaker unconsciously increases his vocal effort to always be understood by a listener: that is the Lombard effect. There are many researches in this regard, in which the relationship between noise and voice level has been examined in multiple conditions of both environments and social situations. Furthermore, starting from these observations there has been the attempt to formulate predictive models on noise produced in indoor environments by a large number of people speaking at the same time, in which the Lombard effect is one of the determining factors among others such as: room volume, reverberation time in unoccupied conditions, absorption coefficients, number of people present and ratio between the latter and the number of people who are actually speaking (group size).

Predictive algorithms identified could find a new interpretation following the COVID-19 pandemic that began in 2020 and is still ongoing globally. In fact, in order to combat the disease and its extremely high contagiousness, the main objective of all the measures adopted was to avoid the gathering of people and to promote the social distance of at least 1 meter between each of them. In this perspective, a research project was launched at the Politecnico di Torino in which, using a prototype device called Speech and Sound SEMaphore (SEM), developed by the Politecnico itself to report excessive noise due to the aggregation of people, it is possible to monitor and indicate the presence of gatherings in a closed environment starting from the anthropic noise measured. This could be achieved by reversing the paradigm of predictive models discussed earlier in which the noise produced becomes a measured factor while the number of people present becomes the object of prediction.

The subject of this thesis is to document the work I have done in this research project, in which my contribution has focused in particular on the development of a methodology useful for measurement, evaluation and parameterization of the Lombard effect. With this in mind, an experimental work was carried out in the field, during the course of 4 different days, carrying out noise measurements inside the staff canteen of the Politecnico di Torino. The measurements were made using both devices that recorded the individual voice levels of some participants in the research, and equipment that monitored the overall trend of the noise level within the environment. The data were subsequently processed through software by extrapolating the parameters useful for estimating the Lombard effect. Some considerations have been made on how to monitor and how to optimize the evaluation process, starting from a preliminary study on some cuts made on the recorded audio tracks and then automating the choice of these most representative segments. Finally, considerations were made on the results obtained and on the effectiveness of the methodology used, taking into account the exceptional circumstances in which it was applied (reduced canteen capacity, social distance).

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION: LOMBARD EFFECT IN EATING ESTABLISHMENTS

1.1 OVERVIEW

Noise conditions in eating establishments have been the subject of numerous studies over the years. This concern is due to the fact that these environments can be too noisy due to various factors related to both the structural nature of the premises (which influence the reverberation) and to the multiple elements that contribute to increasing the noise levels present. The latter can be substantially divided into two categories: noise derived from human activity and non-anthropic noise due to the presence of machinery (e.g. HVAC). The first category contains all the actions carried out by a human being that can cause some kind of sound (e.g. moving objects and cutlery on the table), of which the main factor is the speech: it is what most affects the level of anthropic noise in the environment. All these factors lead to the raising of noise levels and, consequently, to interfere on the conditions for having a verbal communication between the individuals present. In fact, in order to sustain a conversation and preserve the intelligibility of speech, they tend to unconsciously raise their level of voice with the aim of being understood by an interlocutor. This phenomenon is called the **Lombard effect**. In this introductory chapter it will be briefly described the nature of this phenomenon and the main studies that have deepened it. Particular attention will be paid to those where the Lombard effect has been studied in eating establishments and which have led to the formulation of noise predictive algorithms. Finally, reference will be made to the legislation governing measurements of the relationship between speech levels and noise levels.

1.2 BRIEF HISTORY OF THE LOMBARD EFFECT

In 1911 Etienne Lombard (1869-1920), a French otolaryngologist and surgeon, published a study entitled "Le Signe de l'élèvation de la voix" in which he reported some observations made on some patients of the Hôpital Lariboisière, where he worked [1]. He noted that by subjecting people with normal hearing to loud noise, they subconsciously increased the amplitude of their voice level, and then brought it back to normal levels when the noise was stopped. Despite his untimely death from a disease, his discovery strongly influenced scientific studies throughout the 20th century and early 21st century. For this reason, this effect was named Lombard in his honor. Today the Lombard effect is a valid tool in medicine, psychology and even ethology. Initially, the increase in speech amplitude due to noise was called the Lombard sign and, subsequently, the Lombard reflex being it an involuntary act. Among the first to use the term "effect" were Hanley and Harvey in 1965, terminology now universally accepted [2]. In 1971, sixty years after Lombard's discovery, Harland Lane and Bernard Tranel from University of Michigan reviewed all published papers on the phenomenon, most of which were related to hearing tests. They discussed the initial assumption made by Lombard that a speaker changes his voice's amplitude in a similar way according to the rising of the ambient noise level and the lowering of the level at which he hears his own voice (sidetone regulation) finding an effective relationship between the two functions [3]. Moreover, they found that the context in which the observations were made affected the behaviour of the subjects: the Lombard effect was less pronounced in those who performed a reading task rather than those who interacted with an examiner; human speakers need to be understood and are motivated by the search for intelligibility. The Lombard effect has also been studied in animals, especially mammals: they tend to have a vocal response to increased noise similar to human behavior. The magnitude of the Lombard effect varies within and among species and has been observed in various design modes of the experiment and by using different sources of noise [4]. Historically, human research has focused on medical and technological applications, while animal studies have focused on the evolution of acoustic communication and species conservation. Studies on the relationship between background noise level and human speech level have been numerous over the past century. In 1954, Korn indicated that noise levels below 45 dB do not affect speech power, but instead above 55 dB an increase of 0.38 dB can be observed in speech levels for each 1 dB increase in noise levels [5]. In 1958 Pickett studied the Lombard effect using an anechoic chamber and simulating free-field acoustical conditions, finding an increase of 8 dB in the average level of vocal effort with an increase in noise of 8 dB and, consequently, a slope of 1 dB/dB [6]. A particularly important study was carried out by Lazarus in 1986, who introduced the concept of predicting verbal communication in noise [7]. Through a review of a large number of investigations and measurements, he found that the Lombard effect began to manifest with noise levels around 45 dB and speech levels around 55 dB, resulting in an increase in voice levels of 0.3-0.6 dB for each increase in noise levels of 1 dB. Lazarus's studies were later taken up by the scientific community for the formulation of some predictive noise models that will be presented in more detail in the next paragraph. The Lombard effect has also been studied in children, especially in elementary schools by observing the interactions between students and teachers. The measured slope values are multiple: ranging from 0.22 dB/dB found by Dodd and Whitlock [8], to 0.82 dB/dB found by Sato and Bradley [9], to 1 dB/dB proposed by Sutherland [10], who analyzed previously published data on teacher's vocal effort. In addition, in recent years, experimental studies have been carried out to assess the magnitude of the Lombard effect at noise levels below 45 dB indicated by Lazarus. For example, in 2017, an article was published to investigate whether there was a precise starting point, in terms of noise levels, in which the Lombard effect began to manifest and what were the slopes at these low levels [11]. Also by the same author, Pasquale Bottalico, is a further 2018 publication in which the starting point for the Lombard effect was investigated in relation to restaurants and how it related to the perceived communication disturbance and the willingness to spend time and money for a meal [12]. Through the monitoring of 28 participants who had been instructed to read passages in a sound-attenuated booth in the presence of noise measured in restaurants from 35 to 85 dba, the slope derived was 0.54 dB/dB. This shows that the Lombard effect, although it was discovered more than 100 years ago, is still being studied by many researchers in the scientific community.

1.3 HUMAN NOISE PREDICTION IN EATING ESTABLISHMENTS

On an acoustic level, eating establishments can be particularly difficult environments for verbal communication. They are affected by high sources of noise due to both human activity, in which speech is the main factor, and the presence of machinery that generate sound (for example speakers for music). In particular, the presence of many people who speak at the same time causes the Lombard effect: the higher the noise level rises, the more people raise their tone of voice to be able to speak with another person. Because of these difficulties in being able to maintain a conversation, customers might leave the restaurant with a sense of discomfort and tiredness. These reasons have led to several initiatives and studies on acoustic comfort in these environments. According to a study investigating the appropriate acoustic environment for an enjoyable meal carried out on 5 eating establishments, typical noise levels in an empty environment range from 41 to 66 dBA while in an occupied one from 66 to 83 dBA with between 10 and 94 customers [13]. In 2004 Astolfi and Filippi evaluated the quality of verbal communication in 4 different pizzerias in Turin. Measurements made with people present, placing a microphone at the table where normally a person sits, show noise trends between the various locations ranging from 67.4 dBA to 76.8 dBA [14]. To study speech intelligibility, instead, they used a head-torso simulator as a voice source positioned at 1 m from a receiver, finding poor intelligibility with a "normal" tone level of voice and a noise level of 76.2 dBA, while better intelligibility at the expense of privacy can be achieved with a "raised" tone of voice. Of the two studies mentioned above, only the latter included the Lombard effect in its considerations by taking Gardner's studies into account. In fact, by studying changes in voice output in group situations he found that for every doubling of the number of people there was an increase of 6 dB in the total voice output once exceeded the 12-15 people present, assuming a third of them spoke at the same time [15]. In 1997 Tang *et al.* proposed a predictive model for sound pressure levels in an occupied enclosure. In their method they took into account the effect of rising voices in order to predict the variation of the A-weighted sound pressure levels in a closed environment where the number of occupants varies over time [16]. The model and an estimation of the Lombard effect was developed through measurements of noise levels in relation to the number of occupants in a university staff canteen. From the results obtained they deduced that people started to raise the level of their voice when the noise level exceeded 69 dBA. They also noted that the effect might not be seen with less than 50 occupants because of the difference in voice levels between people and the variability of the ways they could sit inside the canteen. In 2002, instead, Kang developed a radiosity-based computer model to predict sound pressure levels in dining spaces. Using this model, he conducted a parametric study to examine the basic features of intelligibility in a conversation in dining spaces and to investigate the effectiveness of acoustic treatments aimed at improving it [17]. The relationship between number of people and measured sound pressure levels was also reported by Navarro and Pimentel, in their 2006 publication [18]. They carried out a survey in twelve food courts and selected two case studies which presented poor quality of communication. The measured noise levels were 74 dBA in the presence of about 345 people in one case, while in the other 80 dBA with about 545 people. They also developed an analytical formula for the evaluation of SPL that provided good feedback on the measured data, although they suggested further studies in this regard. It was published in 2007, instead, the study conducted by Hodgson *et al.* on the measurement and prediction of noise and voice levels and the Lombard effect in eating establishments [19]. For this research they carried out measurements in 10 different eating establishments, measuring noise levels from 45 to 82 dBA. From these levels and from some questionnaires distributed to the diners they estimated the number of customers present and the level of noise to which a single diner was exposed. The processing of all the collected data led them to the development of an iterative model for predicting speech and noise levels, in which unknown parameters such as the definition of the Lombard effect, the number of speakers per client and the level of absorption per client, have been estimated by optimization techniques. The Lombard slope resulting from the proposed model was 0.69 dB/dB. Most of these studies formed the basis of the research published by Jens Holger Rindel in 2010: starting from the analyses made by Lazarus and using the data collected by Tang, Hodgson, Navarro and Pimentel proposed a new prediction model for average A-weighted noise levels noise level due to many people speaking in a room with assumed diffuse sound field [20]. The equation he suggested for noise prediction in an eating establishment is the following.

$$L_{N,A} = \frac{1}{1-c} \cdot \left(69 - c \cdot 45 - 10\log\left(g \cdot \left(\frac{0.16 \cdot V}{T_0 \cdot N} + A_p\right)\right)\right)$$

where the factors taken into account are:

- Lombard slope c $[dB/dB]$;
- Room volume V $[m^3]$;
- Reverberation time in empty state T_0 $[m^3]$;
- Number of people in the room N;
- Sound absorption per person A_p $[m^2]$;
- Group size $g = N/N_s$ which represents the average number of people per speaking person.

He also expressed that the precondition for his model is that the A-weighted ambient noise level is at least 45 dB, while the value found for the Lombard slope is 0.5 dB/dB in order to achieve good agreement with the measured data. The Rindel's model has been taken up in very recent years by D'Orazio et al. which first proposed a new formula for noise prediction applicable to a museum [21] and then one applicable to large food court that took into account the spatial decay (Dl2) [22]. In both cases the Lombard slope used was 0.4 dB/dB.

1.4 REFERENCE STANDARD FOR THE LOMBARD EFFECT: ISO 9921.

At the regulatory level, the Lombard effect is defined by the International Organization of Standard inside the ISO 9921 "Ergonomics - Assessment of speech communication" published in 2003 [23]. The aim was to advise the levels of speech communication quality required for conveying comprehensive messages in different applications and to define criteria applicable to the intelligibility of speech communications. The Lombard effect is defined as "Spontaneous increase of the vocal effort induced by the increase of the ambient noise level at the speaker's ear" and is a parameter to take into account in vocal effort measurement. Vocal effort is expressed by the equivalent A-weighted sound pressure level at a distance of 1 m in front of the mouth $L_{s,A,1m}$ and some of its typical values are given in Table 1.0 (male speaker). Figure 1.0 shows instead the relation between speech level and ambient noise level. The hatched area indicates the variability of the Lombard effect among speakers.

Table 1.0 Vocal effort at varius speech levels, ISO 9921

$L_{s,A,1m}$ [dB]	*Vocal effort*
54	Relaxed
60	Normal
66	Raised
72	Loud
78	Very Loud

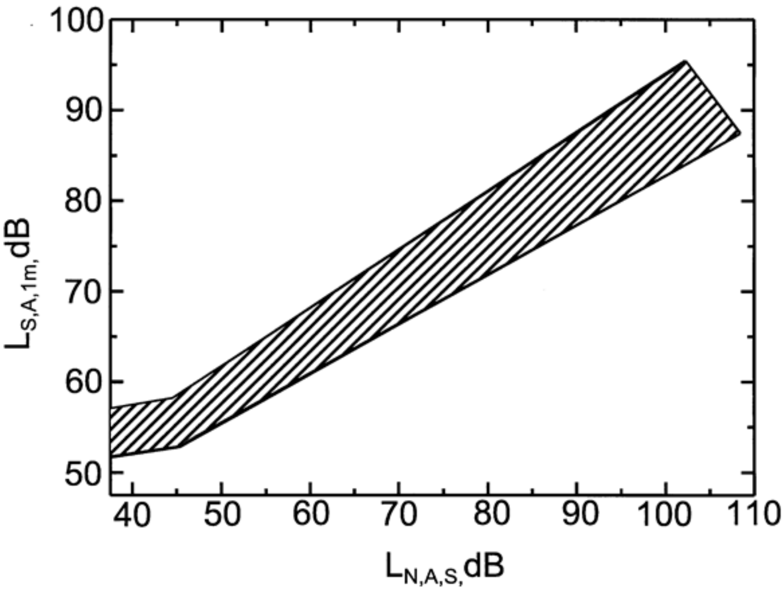


Figure 1.0 Relation between the range of vocal effort and the ambient noise level at the speaker's position (ISO 9921)

Chapter 2

MEASUREMENT PROCEDURES, LOCATION AND EQUIPMENT

2.1 OVERVIEW

Measurements took place in the staff canteen of Politecnico di Torino located in the main campus of the university, during four different days divided in two slots:

- First slot: November 25 and 26, 2020;
- Second slot: April 14 and 16, 2021.

On all days the time interval monitored was 12:00-14:00 p.m. in which the canteen serves lunch to employees of the polytechnic. This location was chosen for two main reasons:

1. It was one of the few structures opened during the second wave of coronavirus throughout the fall/winter of 2020 in the northern part of Italy;
2. Its architectural shape allows to control quite easily the entrances and the exits, a significant fact within the research.

Two types of devices were used to measure noise levels in order to monitor both the overall trend of noise levels and the individual vocal effort of some participants. These are respectively:

- XL2 Audio and Acoustic Analyzer (2 in total);
- ZOOM H4n Pro Audio Recorder + Shure Beta 54 Headworn Vocal Microphone (6 each).

Those who participated in the vocal effort measurements were asked to fill out a short questionnaire and to indicate on a map where they ate their meal. In addition, at the entrance and exits of the canteen devices were placed that counted the number of people present within the environment. All the equipment, the location and the procedures will be analyzed in particular in this chapter.

2.2 LOCATION: STAFF CANTEEN OF THE POLYTECHNIC OF TURIN

The staff canteen of the Polytechnic of Turin is located in the lower ground floor of the main campus situated in Corso Duca degli Abruzzi 26, Turin. Those who benefit of the canteen services are university employees such as teachers, researchers, contract workers etc. At a structural level, the space inside the canteen is organized as follows: a main entrance, accessible either by stairs or elevator, leads to a corridor where customers take the trays and choose among the foods present in that particular day and served by the canteen's staff; a main lounge with tables where clients can eat their meal and a smaller one near the bar with additional seats; the restroom near the two exits which lead one to the external courtyard and the other to the upper floor of the building (Fig. 1.1).

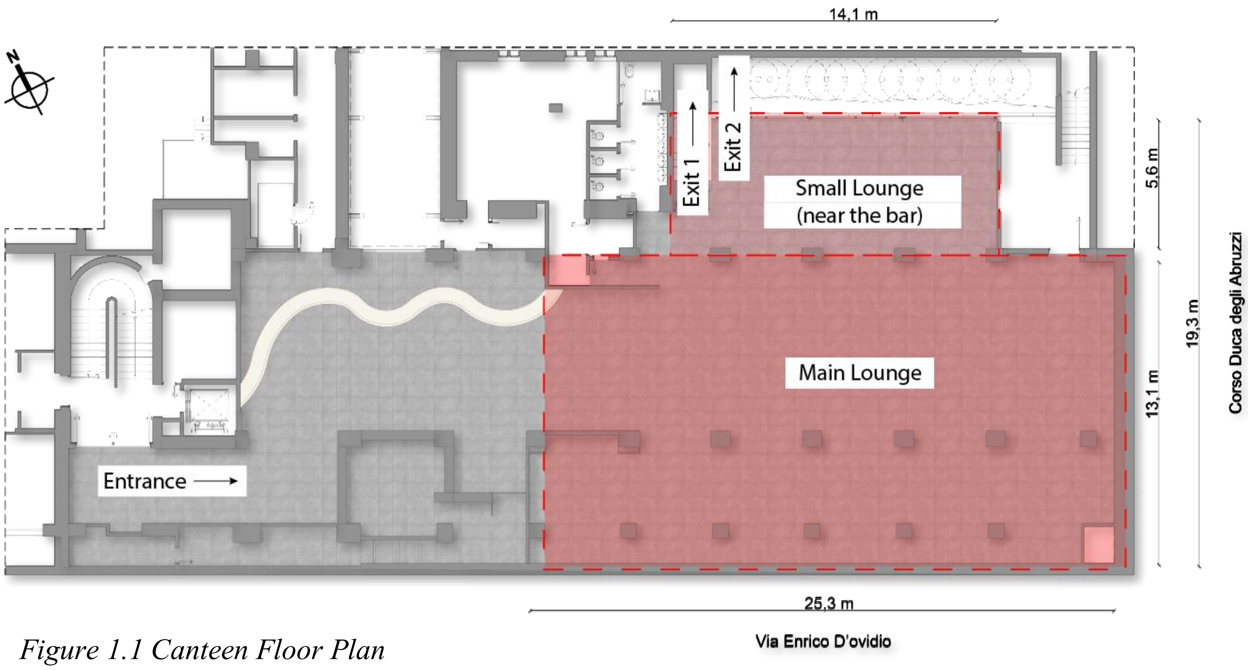


Figure 1.1 Canteen Floor Plan

The environment's total volume is 1115.0 m^3 and, in normal conditions, the maximum capacity of people present to eat at the same time is about 280 distributed in 82 tables, each of them long 1.2 m. This capacity was reduced at the time of the measurements according to the laws in force issued by the Italian Government to contrast Covid-19 pandemic in such a way as to ensure the social distance of at least 1 meter between the people present to decrease the number of infections [24]. Specifically, the number of tables in which customers could eat their meal has been reduced to 69 with 2-3 seats for each of them.

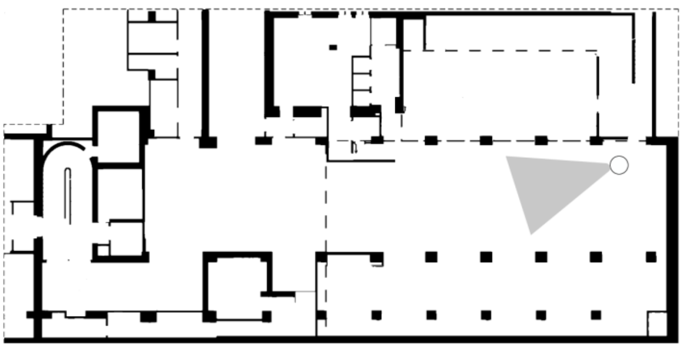


Figure 1.2 Image of the empty canteen (left) and respective point of view on the map (right)

The reverberation time in an empty environment has been measured with the result of T_0=0.82 s.

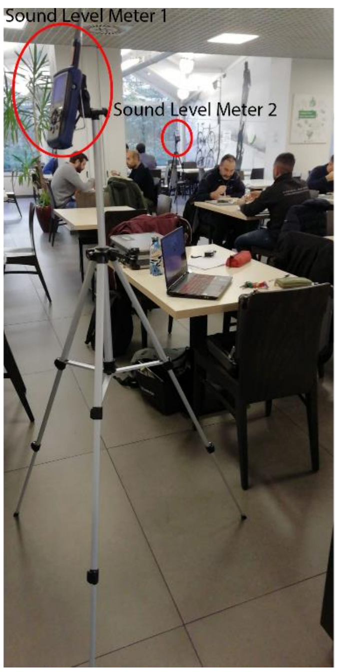
2.3 Equipment

2.3.1 Environmental noise monitoring: XL2 Audio and Acoustic Analyzer

The XL2 Audio and Acoustic Analyzer (Fig. 1.3) is a professional Sound Level Meter for environmental noise monitoring capable of measuring different parameters such as SPL actual, Lmin, Lmax, Lpeak, Leq [25].

Figure 1.3 XL2 Sound Level Meter

Two instruments of this type have been placed inside the canteen to be able to monitor the ambient noise in the larger hall and in the smaller hall near the bar (Fig. 1.4). The parameter taken into account and used in subsequent data processing is the one marked as Laeq_dt in the final report provided by the XL2, where the name indicates an equivalent sound pressure level with A frequency weighting calculated every second.



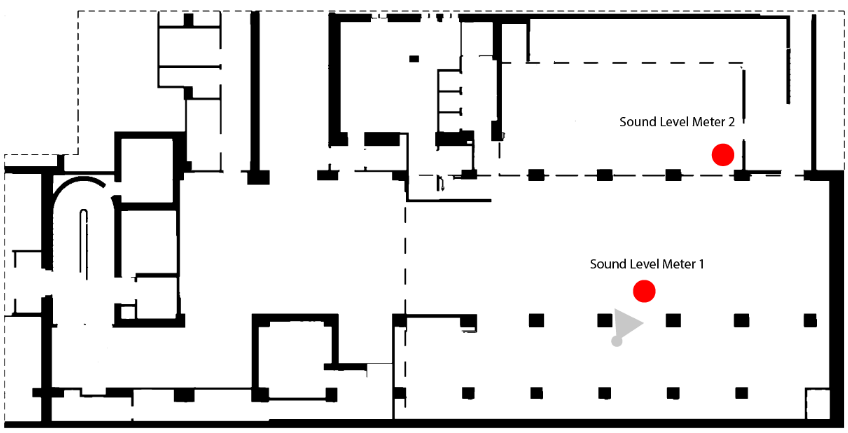


Figure 1.4 (a) Positions of the 2 XL2 (November 2020 measurements)

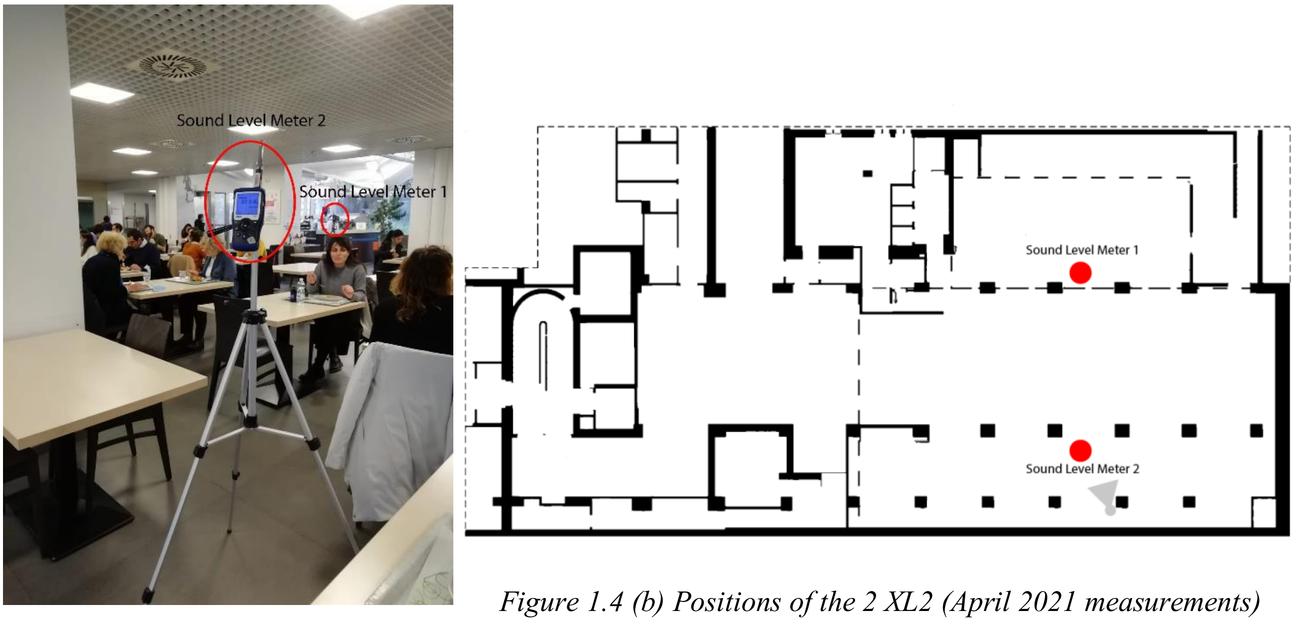


Figure 1.4 (b) Positions of the 2 XL2 (April 2021 measurements)

2.3.2 Voice monitoring: ZOOM H4n Pro Audio Recorder + Shure Beta 54 Headworn Vocal Microphone

In order to monitor the individual vocal effort, during the measurement has been asked to several people to wear a microphone throughout the meal inside the canteen. The device used is a Shure Beta 54 Subminiature Headworn Condenser Microphone (Fig. 1.5), which is characterized by a supercardioid polar pattern that provides maximum gain-first-feedback and environmental rejection and an extended frequency response designed for vocal performances [26].

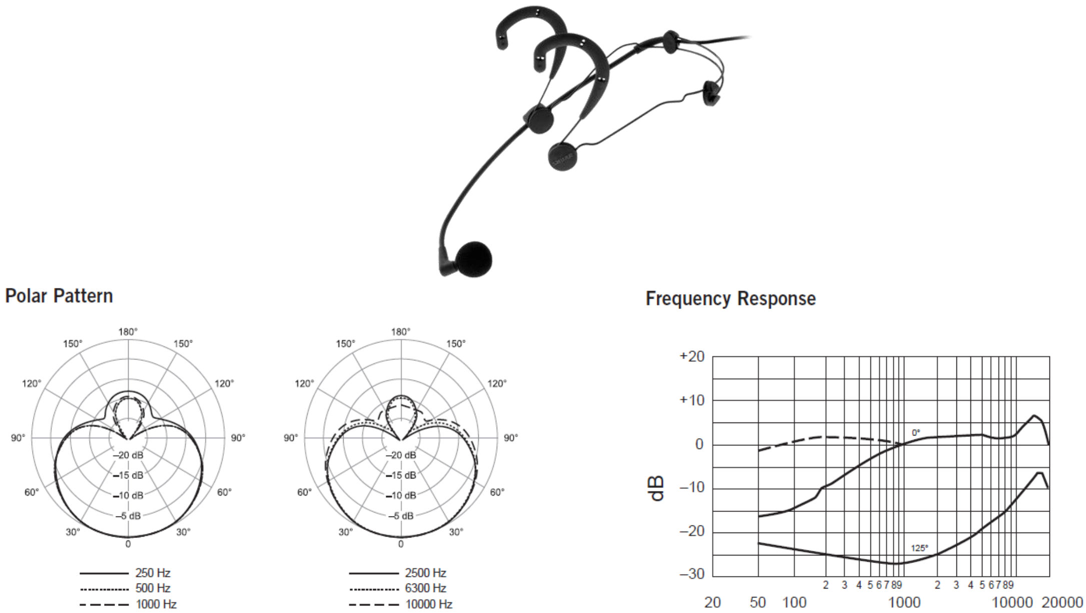


Figure 1.5 Shure Beta 54 Microphone (above) and its polar pattern and frequency response (below)

Every microphone has been connected via XLR balanced cable to a ZOOM H4n Pro Handy Recorder (Fig 1.6) [27] as external input to record voice tracks of the participants. H4n was used as a simple data logger, its built-in microphones were not utilized. For each audio recording the following settings has been used:

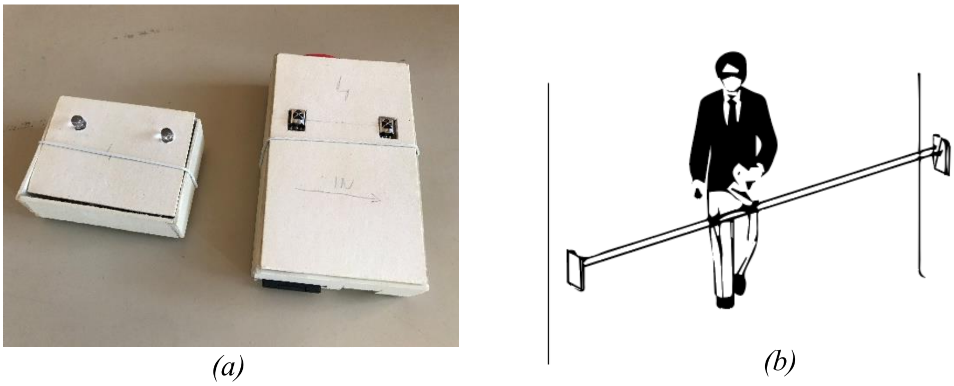
- Input REC LEVEL 80;
- Mode STEREO (external input on Left channel);
- Waveform Audio File Format (WAV);
- Sampling Frequency 44.1 kHz;
- Quantization 16 bit.

Figure 1.6 ZOOM H4n Pro Handy Audio Recorder

The coupling formed by the Beta 54 microphone and the H4n recorder will be called MIC/REC chain in the rest of this treatment. A total of 6 MIC/REC chains were used, distributed in turn to volunteers in both November and April measurements.

2.3.3 People Counter

Devices have been used to measure the flow of people inside the canteen. These instruments were designed and manufactured by two members of the research team belonging to the Department of Electronics and Telecommunications of the Politecnico di Torino. Such prototypes, hereinafter referred as "people-counters", are composed of two small boxes with an electronic circuit inside them: one of them acts as a transmitter (TX) while the other as a receiver (RX). An infrared beam is constantly emitted between them and every second a check is made to determine whether the flow is still available; if it is interrupted, for example by the passage of a person, in memory is added a unit to the local variable that stores the number of people. Since everyone of them operates in the same way (each adds up the passage of a person), the incoming and outgoing counters have been synchronized via software once the measures were finished subtracting output counter data with input counter data in the same time frame. The devices and an example of its operating mechanism are shown below (Fig. 1.7.).



(a) *(b)*

Figure 1.7 Image of the people counters (a) and example of operation (b)

In total, 3 people-counters were used to cover the main entrance and the two exits. At the entrance, a pair of transmitter/receiver was fixed on two chairs lined up and spaced enough to allow the passage of a person (Fig. 1.8). This choice was made primarily to try to preserve the stability of infrared flux, able to bounce on multiple surfaces. A member of the research remained present during all measurements to supervise the flow of incoming people.

Figure 1.8 Shot of people-counter at the entrance

In the first exit (the one leading to the upper floor inside the building) the counter was placed just outside the canteen attached to the walls of the staircase, while in the second exit (leading to the outdoor courtyard) it was placed in the railing near the door (Fig. 1.9).

Figure 1.9 Shot of people-counters at exit 1 (on the left) and exit 2 (on the right)

Figure 1.10 shows the positions of the people-counter on the map of the canteen with the respective points of view in which the photographs shown above were taken.

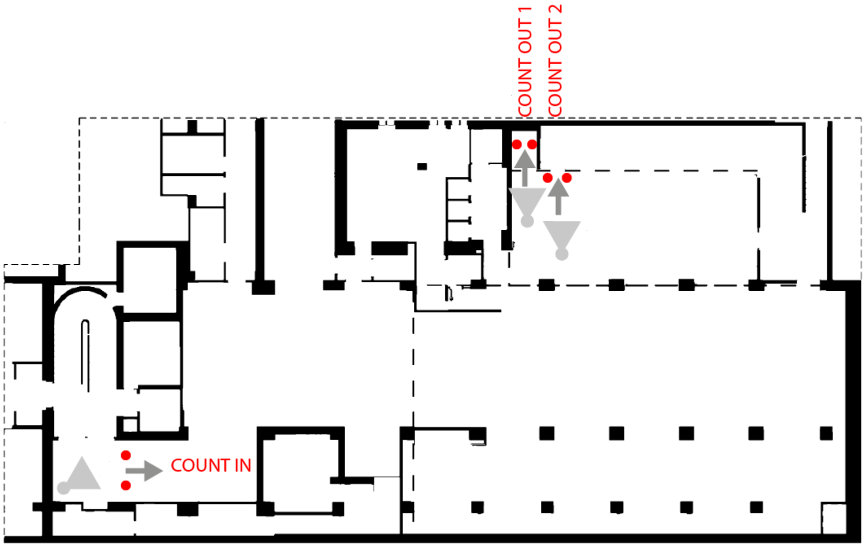


Figure 1.10 Positions of the people-counters on the canteen's map

2.4 MEASUREMENT PROTOCOL

Once the measuring instruments have been placed in the positions listed above, the following measurement procedure has been followed for all four measuring days (both for the November and April slots). Four members of the research were present: two inside the canteen to manage the two XL2 phonometers, one at the entrance to control the flow of people, one outside the cafeteria to distribute the individual monitoring tools (MIC/REC chain) to those who wanted to participate in the experiment. The measurements started at 12:00 pm and ended at 14:00 pm, following the opening and closing time of the canteen to the public; time was used for synchronization between the various instruments taking as reference the time indicated by https://www.oraesatta.co/. Each participant in the voice monitoring measures, whose involvement was only on a voluntary basis, was provided with a document (see Appendix A) which included:

- brief description of the research and the type of measures that it was intended to carry out;
- consent to the processing of personal data collected during measurements and privacy policy;
- short questionnaire of general character;
- map where to indicate the place where the meal was consumed.

In addition, the research team member carried out the following operations for each specific voice monitoring measure:

1. Note down the identification number of the MIC/REC chain used.
2. Place the microphone close to the speaker's mouth so that it does not bother the latter during the meal and remains in that position for the duration of the measurements.
3. Starting the recording on the data logger
4. Note down the start time of the measurement.

5. Deliver the informed consent document and questionnaire attached to each participant.
6. Upon return of the participant, stop the recording.
7. Note down the end time of the measurement.
8. Take back the MIC/REC chain from the participant.
9. Check that the document contains all the data and signatures of the participant.
10. Download data from memory to computer, annotating which document each recording refers to.

2.5 GENERAL TREND AND PEOPLE ATTENDANCE DURING MEASUREMENT DAYS

Despite the particular circumstances in which the measurements were carried out due to the COVID-19 pandemic, both in November and April a significant number of people took advantage of the service offered by the canteen employees of the Politecnico di Torino. In general, attendance was higher in April 2021 than in November 2020. This is due to the fact that in November the second wave of the pandemic had just broken out (with a greater impact than the first), the measures taken by the Italian Government were more stringent so that safety conditions had a higher standard to be respected. In addition, the vaccination campaign against the virus had not yet begun. This was ongoing in April, so some employees of the polytechnic had already received at least the first dose of the vaccine and the risk of contagion was significantly reduced.

2.5.1 November measurements

Measurements for November 2020 were made on 25 and 26. Most of the tables had three seats, as shown in the picture below (Fig. 1.11).

Figure 1.11 Shot of the staff canteen during November measurements

The total number of people present at the same time reached peaks of almost 120 for the 25th and more than 140 for the 26th. Specifically, analyzing the data measured with the counter-people is found (Fig. 1.12):

- for the 25th a trend ranging between 60 and 80 people in the time interval between 12:20 and 13:30 and a steep increase in the last half hour between 13:30 and 14:00;
- for the 26th a proximately increase, with small fluctuations between 12:15 and 13:45.

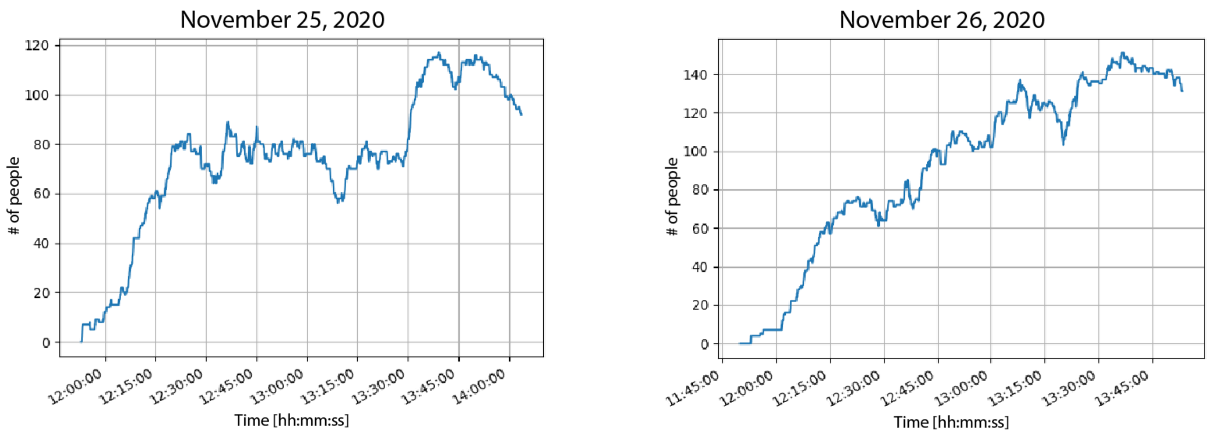


Figure 1.12 Global trend of the number of people present during November measurements on the 25th (left) and 26th (right)

2.5.2 April measurements

Measurements for April 2021 were made on 14 and 16. Most of the tables had two seats but, unlike November 2020, were closer to each other and greater in number.

Figure 1.13 Shot of the staff canteen during April measurements

The total number of people present at the same time reached peaks of more then 200 people in both days. Specifically, analyzing the data measured with the counter-people is found (Fig. 1.14):

- for the 14th, some oscillations between 150 and 190 in the time interval ranging from 12:35 to 13:25 and peaks over 200 from 13.30 onwards;
- for the 16th, an almost constant increase to 175 around 13:30 followed by a slight drop and a climb to the peak of 200 at 13:40.

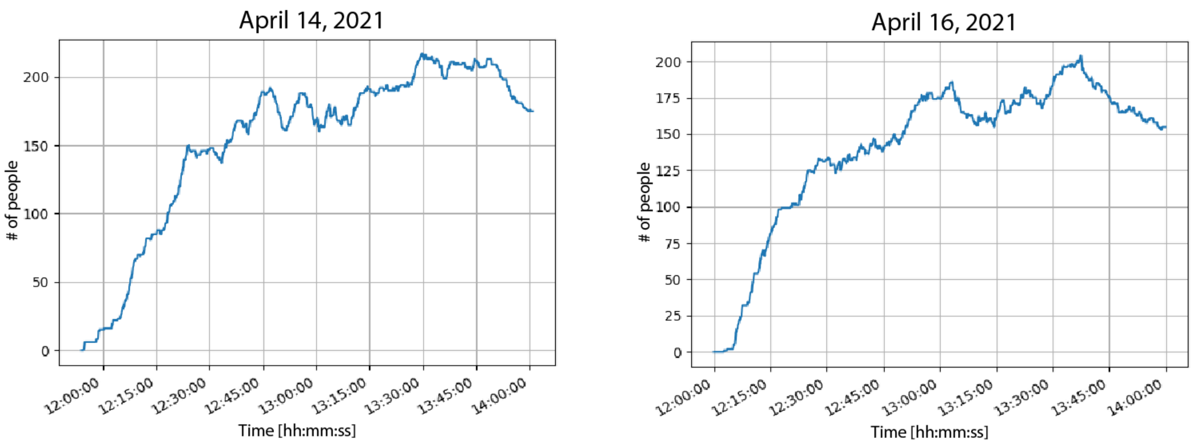


Figure 1.14 Global trend of the number of people present during April measurements on the 14th (left) and 16th (right)

2.5.3 Participants in voice effort measurements

In total, 35 people participated in vocal effort measurements divided almost equally between November 2020 (18 participants) and April 2021 (17 participants). In both cases more people joined during the second day of the respective slot. Of these 35 people, 22 were men and 13 women, with a population age ranging from 23 to 60 years; 10 identified themselves as smokers, 3 people reported having suffered from hypoacusia while none of them indicated having vocal diseases.

Chapter 3

SIGNAL PROCESSING: PRELIMINARY STUDY

3.1 OVERVIEW

This chapter describes the first part of signal processing derived from measurements, which will then form the basis of the automation and optimization of the choice of data. This procedure, in fact, constitutes a preliminary study on how to derive the parameters useful to determine the characteristic slope of the Lombard effect from the measurements described in the previous chapter. The operations described below were carried out for 34 of the 35 audio recordings related to participants who joined in the monitoring of their vocal effort; the first was discarded due to incorrect settings during recording. The software used for processing were:

- Adobe Audition 2020 for operations concerning the waveform of the signal;
- Matlab R2020b for numerical calculation and statistical operations;
- Microsoft Excel for storage and graphic representation of data.

The procedure is divided into:

- preliminary operations concerning signal cleaning, calculation of sound pressure levels and synchronization between all data;
- application of the Gaussian Mixture Model for the estimation of noise and speech levels;
- choice of parameters for the graphical representation of the Lombard slope.

The first results obtained led to the targeted selection of some signal's portions that constituted the idea behind the choice to carry out the analysis on some segments obtained through the optimization and automation algorithm described in Chapter 4. To each of the participants and therefore to the respective recorded audio tracks has been assigned a unique identifying name so indicated: ID + N with N = 1,2,...,35. This name will be used throughout the course of the discussion when referring to a specific measurement.

3.2 PRELIMINARY STEPS

3.2.1 Signal Cleaning

The 34 tracks were analyzed on Adobe Audition, a professional workstation for editing and manipulation of the audio signal. In particular, the goal was to search for and eliminate any points where the signal was distorted, by which we mean the typical problem of audio recordings called "Clipping". This type of distortion concerning the sound waveform occurs when an amplifier is driven too far beyond its maximum output capacity by a much higher voltage or signal. In order to identify it visually, it is necessary to examine the signal close to 0 dBFS (digital audio reference level equal to Full Scale) which refers to the maximum peak voltage level possible before “digital clipping” or digital overload (overs) of the data converter [28]. In case of clipping, the sine wave form is "cut" and the dynamic range of the signal is altered (Fig. 2.1).

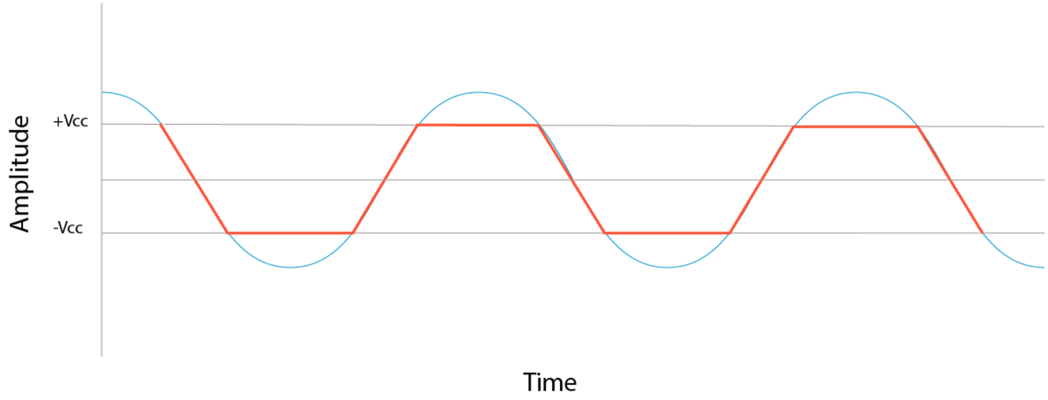
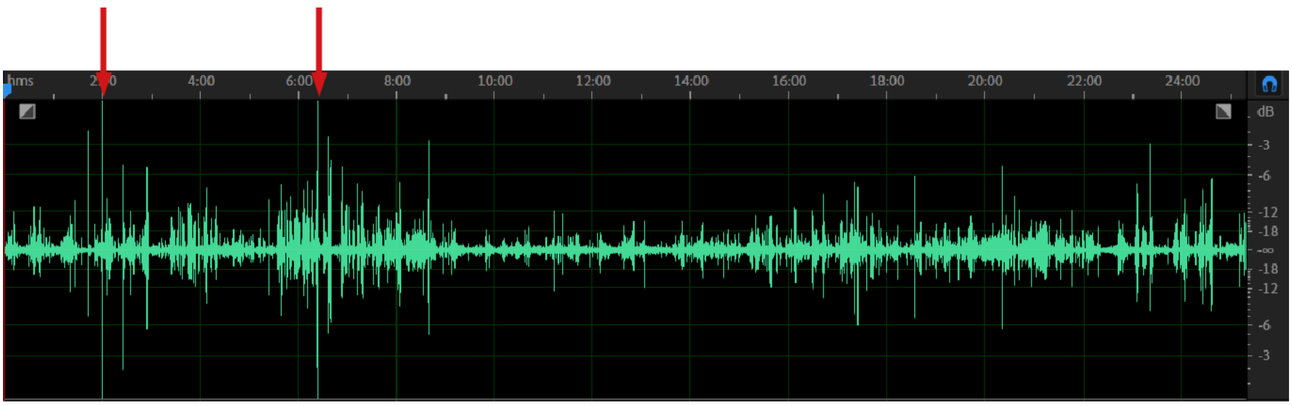
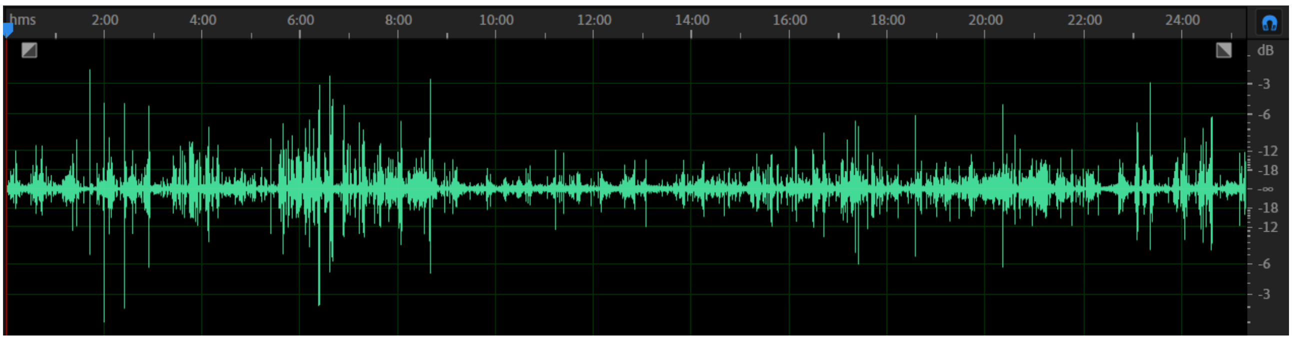


Figure 2.1 Clipped signal example

Figure 2.2 shows an example of a track before and after the removal of clipped signal points. This was done manually for each of the recordings by trying to remove as little signal as possible in order to preserve as much as possible the original audio. Only for ID8 it was not necessary to remove any point.



(a)



(b)

Figure 2.2 ID7 audio signal before (a) and after (b) removal of clipped points

3.2.2 Sound Pressure Levels calculation

According to the definition given by the International Electrotechnical Commission (IEC 801-22-07) the *sound pressure level* (SPL) represents the logarithm of the relationship between a given sound pressure and the reference sound pressure:

$$L_p = 10log\left(\frac{p}{p_{ref}}\right)^2$$

with $p_{ref} = 20\mu Pa$ corresponding to the conventional threshold of audibility at 1000 Hz. The sound pressure level is the most commonly used indicator of the acoustic wave strength and is measured in decibels. In order to extract the SPL levels from the recorded audio tracks these were processed on Matlab by using the *splMeter* System object™, dedicated to this type of operations. The object returns measurements for:

- frequency weighted sound levels;
- fast or slow time-weighted sound levels;
- equivalent-continuous sound levels;
- peak sound levels;
- maximum sound levels.

In this case the properties of interest were sect in this way:

- Time Weighting at ‘Fast’;
- Frequency Weighting at ‘A-weighting’;
- Sample Rate at ‘44100’ [kHz];
- Time Interval as ‘1’ [s].

In summary, the SPL values were calculated second by second using the isophonic curve A, which discriminates frequencies in a manner similar to the response of the ear. The unit of measurement is the dB(A). The following is the script that defines the *“Spl_dba*_th*”* function where the *splMeter* object was used.

```
function [SPL_mean,SPL_std,dB,windowTime] = SPL_dBA_TH(x,fs,durata)

SPL = splMeter('TimeWeighting','Fast', 'FrequencyWeighting','A-
weighting', 'SampleRate',fs, 'TimeInterval',durata);
[Lt,Leq,Lpeak,Lmax] = SPL(x);
dBtemp = downsample(Leq,durata*fs);
dB=dBtemp(2:end);
windowTime=durata:durata:(length(dB))*durata;
SPL_mean=mean(Leq);
SPL_std=std(Leq);
```

This function is used within a *for* loop to process all recorded audio tracks. For this purpose, a script called *"Analisi_Wav_SPL_dbA_raw"* was created in which the cycle was present, consisting of the following structure:

```
id_length=35;% number of participants
path='name of the path where audio files are present';
for j=2:id_length
    fname = strcat(path, 'ID',num2str(j), '.wav');
    [data,fs]=audioread(fname);
    dataclean=data(:,1);
    durata=1;
    [SPL_mean,SPL_std,dB,windowTime] = SPL_dBA_TH(dataclean,fs,durata);
    windowTime=windowTime';
    lim=length(dB);
    xlswrite ( strcat('ID',num2str(j),'_dBA.xlsx'),windowTime,
strcat('A1:A',num2str(lim)));
    xlswrite ( strcat('ID',num2str(j),'_dBA.xlsx'),dB,
strcat('B1:B',num2str(lim)));
end
```

The variable indicating the number of participants is external to the loop, as well as the name of the folder where the audio files are stored. The files are read through the Matlab audioread function, the left recording channel is examined (where the signal is present), the SPL values are calculated with the function shown above and these are written into an excel sheet. This process aims to build a time history of sound pressure levels for each recording.

3.2.3 Synchronization between all data

Once the sound pressure values have been calculated through the procedure described in the previous section, they have been synchronized with the data obtained from the XL2 phonometers and from the people-counter present in the canteen during the measurements. This has been possible thanks to the start and end monitoring times marked in the documents delivered to each research participant. In these documents, moreover, each participant was invited to mark on a map depicting the floor plan of the canteen the position where the meal was consumed. Thanks to these reports it was possible to associate each participant with the nearest sound level meter. Table 1.1 shows which of the two XL2s have been associated with each participant.

Table 1.1 Nearest sound level meter for each ID

ID	date	# Sound Level Meter
2	nov-20	2
3		1
4		1
5		1
6		2
7		1
8		1
9		2
10		2
11		1
12		1
13		2

14		2
15		1
16		2
17		2
18		1
19		2
20		1
21		1
22		2
23		2
24		1
25		1
6		1
27	apr-21	2
28		2
29		1
30		2
31		2
32		1
33		2
34		2
35		1

The Sound Level Meter number refers to the one shown on the maps in Fig. 1.4 in section 2.3.1. Once the nearest sound level meter has been identified, the equivalent noise levels of the same time slot indicated in the document associated with each participant have been selected. These levels are calculated second by second and are marked in the measurement report of the instrument under the variable *Laeq_dt*. The same procedure was done for the selection of data on the number of persons present measured by the people-counter. In summary, for each ID the following data were grouped in a Microsoft Excel sheet:

- measuring time (second per second);
- sound pressure levels measured with microphone beta 54;
- sound pressure levels measured with nearest XL2 phonometer;
- number of people inside the canteen at that moment.

3.3 ANALYSIS WITH THE GAUSSIAN MIXTURE MODEL

The Gaussian Mixture Model is a model-based clustering technique used in previous studies to measure speech levels [29]. With this procedure a probabilistic model is applied in which it is assumed that all the points constituting a data set are generated by a mixture of a finite number of Gaussian distributions. In this context, the number of components assumed is two: one is the background noise level of the audio signal while the other is the speaker's voice level.

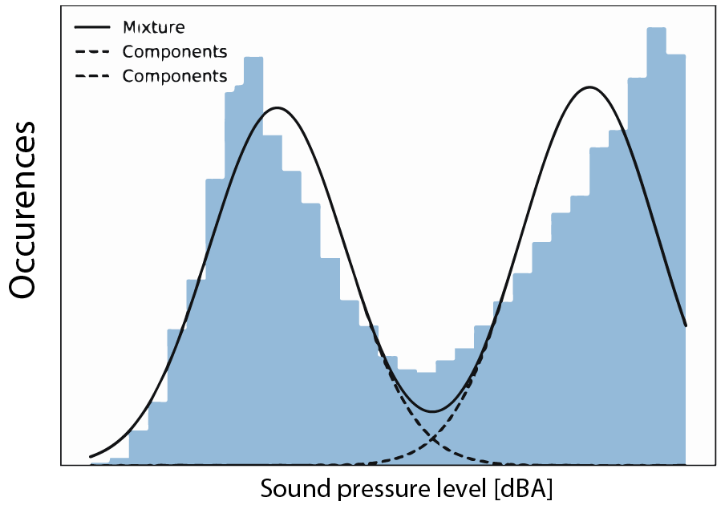


Figure 2.3 Gaussian Mixture Model (GMM) example

Figure 2.3 shows a general example of application of the GMM method: the left component indicates the samples identified as noise and the right one those identified as voice. The GMM technique has been applied to the sound pressure levels obtained for each ID (described in section 3.2.2) through the Matlab *fitgmdist* function, which taking in input a data set X returns a Gaussian mixture distribution model (Gmmodel) with k components.

$$GMM = fitgmdist(X,k)$$

This function has been used within a script called *"GMM_RAW"* in which for each ID the GMM models are calculated and its most significant parameters are saved.

```
id_length=35;
for j=2:id_length
    filename=strcat('ID',num2str(j),'.xlsx');
    sheet=1;
    xlRange='B:B';
    data_ID=xlsread(filename,sheet,xlRange);
    GMModel=fitgmdist(data_ID,2);
    mean=GMModel.mu;
    sigma=[GMModel.Sigma(:,:,1) GMModel.Sigma(:,:,2)]';
    dev_std=sqrt(sigma);
    xlswrite(filename,mean,sheet,'D1:D2');
    xlswrite(filename,sigma,sheet,'E1:E2');
    xlswrite(filename,dev_std,sheet,'D1:D2');
    h=histogram(data_ID,50);
    title(strcat('ID',num2str(j),'RAW'));
    xlabel('dB');
    ylabel('occurences');
    savefig(strcat('ID',num2str(j),'RAW.fig'))
    saveas(h,strcat('ID',num2str(j),'RAW.png'))
end
```

The sound pressure values are read by the excel sheets calculated previously through the script "Analisi_wav_spl_dba_raw" and the fitgmdist function is applied. The saved parameters are Gmmodel.mu and Gmmodel.sigma which represent the average and the variance of each component respectively; of the variance is then calculated the square root to obtain the standard deviation.

Finally, histograms are created and saved for the graphic representation of the two derived Gaussian distributions (an example is shown in Fig. 2.4).

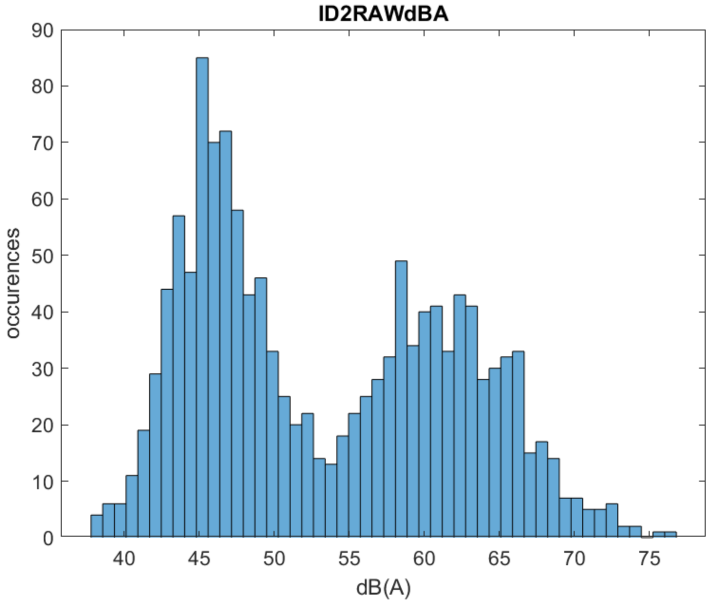


Figure 2.4 GMM distributions obtained for ID2

3.4 PARAMETERS FOR THE GRAPHICAL REPRESENTATION OF THE LOMBARD SLOPE

The parameter that best describes the Lombard effect is the rate ***c*** which represents the slope of the relationship between noise level and voice level. All the operations described in the previous sections of this chapter aim to put together all the data from which to choose what best represents noise level and voice level in order to calculate ***c***. In this methodology, the parameters taken into account are the following:

- the average levels of the two components of the Gaussian distributions obtained through the application of the GMM technique (referred to hereafter as L_{noise_fon} and L_{speech_mic}), which represent respectively the average noise level and the average voice level of the measurements made with beta 54 microphones;
- the average ambient noise level of the nearest phonometer in the same time frame (referred to hereafter L_{noise_fon}).

As the latter is a parameter measured through an instrument for monitoring ambient noise, it has been chosen as the general noise level (L_{noise}) to be used for the determination of the Lombard slope.

$$L_{noise} = L_{noise\_fon}$$

A calibration factor has been calculated according to the following equation (1) in order to relate the measurements made with beta 54 microphones and with XL2 sound level meters.

$$\Delta_{cal} = L_{noise\_fon} - L_{noise\_mic} \quad (1)$$

This factor has been applied to the average voice level calculated with the GMM technique in order to obtain the general voice level (L_{speech}) to be used for the determination of the Lombard slope.

$$L_{speech} = L_{speech\_mic} + \Delta_{cal}$$

By placing the values of L_{noise} and L_{speech} of a scatter graph and tracing the associated trend line it is possible to derive the Lombard slope from the angular coefficient of the indicated line (Fig. 2.5).

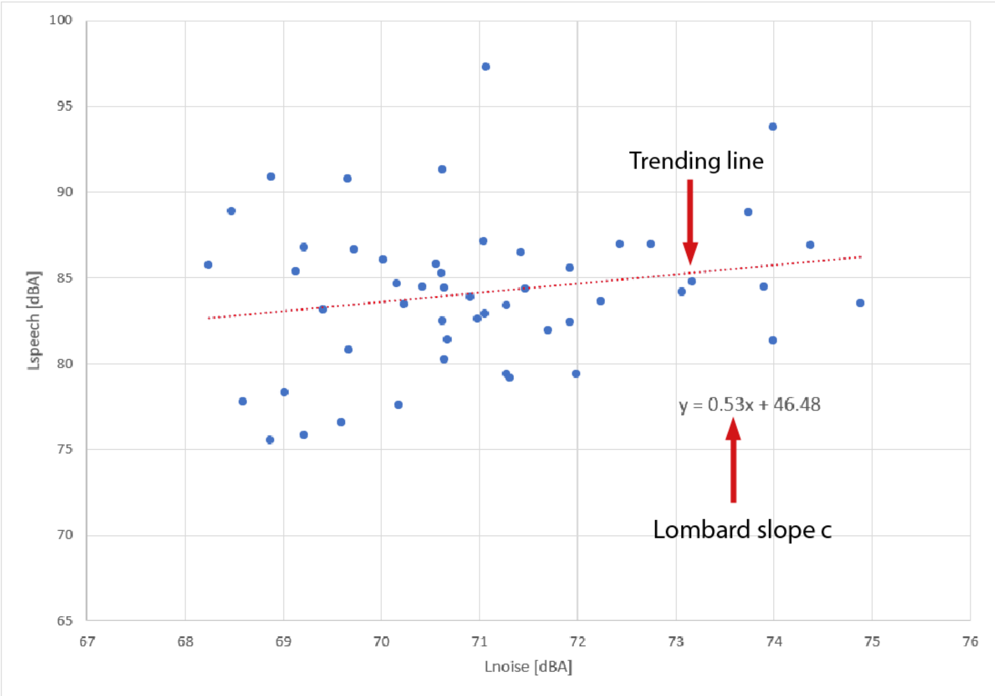
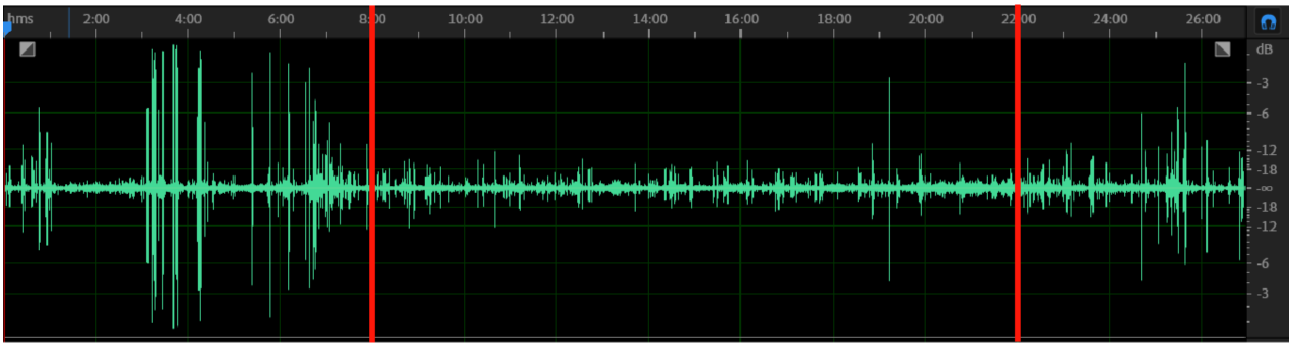


Figure 2.5 Example of graphic representation of the Lombard slope ***c***

3.5 PRELIMINARY RESULTS AND ADJUSTMENTS MADE

3.5.1 First considerations on preliminary results

The first results obtained by applying the operations (described in Section 3.2 to Section 3.4) to all 34 tracks recorded, taking into account the entire duration of each record, were not significant. For the November 2020 measurements a Lombard slope c = 0.01 dBA/dBA was obtained, while for the April 2021 measurements the result seems to be appropriate with c = 0.6 dBA/dBA. The latter, however, is heavily influenced by the values of L_{noise_mic} and L_{speech_mic} calculated for ID34 and ID35: for these, the difference between voice level and noise level is virtually non-existent being 2.8 dBA for the first and 0.1 dBA for the second. By removing these values from the analysis, in fact, the slope turns out to be negative with c = -0.1 dba/dba. It has been hypothesized that the main causes of the problem were due to two main factors: the participant spoke little during the meal or moved the microphone during the measurement. This assumption seems to have some foundation detectable by analyzing the waveform signal more in depth. Both factors can be observed visually in the variations in the amplitude of the signal but, being audio recordings, they are much more easily identifiable by listening to them. The following figure shows two indicative examples of: an audio track belonging to a participant who spoke little during the meal (Fig 2.6 a); an audio track belonging to a participant who moved the microphone during the measurement (Fig 2.6 b).

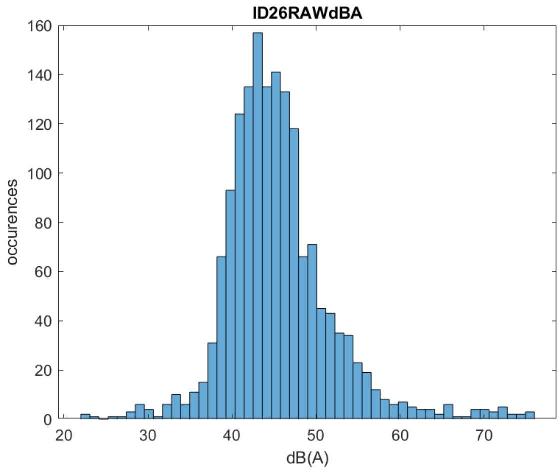


(a)

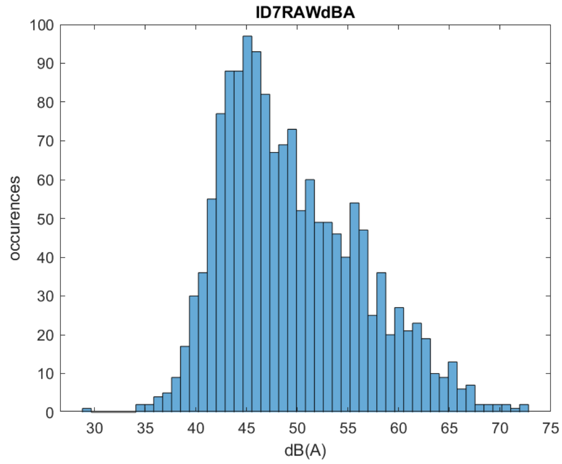
(b)

Figure 2.6 Waveforms of a recording in which the participant ID26 speaks little (a) and in which the microphone has been moved by participant ID7 (b)

Listening to both audio tracks it was verified that in the range indicated in red in Fig 2.6 (a) the participant has hardly ever spoken, while in the points indicated in Fig 2.6 (b) the waveform represents the voice of the same person with the same tone but with very different amplitude levels caused by the displacement of the measuring instrument. The latter factor is the result of the high sensitivity of Beta 54 microphones that, having a supercardioid polar diagram, are highly directional so even a small shift in the position can cause an alteration to the audio signal measured. The problems presented also have an effect on the Gaussian distributions obtained with the GMM technique. Figure 2.7 shows the distributions obtained for ID26 and ID7, the same participant to which belong the audio tracks of figure 2.6. In particular, it can be seen how in the case in which the participant spoke little (Fig. 2.6 a) the number of occurrences between 54 and 66 dba, range in which according to ISO 9921 [23] the tone of voice passes from *Relaxed* to *Raised*, is very reduced. In the second case (Fig. 2.6 b), however, samples in the desired range are present but the two Gaussian distributions of noise and voice levels cannot be distinguished. The latter situation was found in several other cases examined. These considerations have led to the choice of selecting, for each ID, portions of the signal with the more suitable characteristics for the analysis so far presented. The goal of this choice was to try to extract more samples in the range where the voice should be present, and to succeed in separating the Gaussian distributions as much as possible.



(a)



(b)

Figure 2.7 Gaussian distributions for ID26 (a) and ID7 (b)

3.5.2 First adjustments applied (15 minutes and 5 minutes cuts)

During the first phase of selection some cuts were made on audio tracks recorded for vocal effort measurements, taking portions of the signal in which there were present both moments of silence and speech and where the dynamic range of amplitude was adequate and consistent throughout the selection. These cuts were made manually and the search for the desired and listed features was carried out by listening to the recordings one by one. The first selections made for each track were of the duration of:

- a single cut of 15 minutes in total;
- three different cuts of 5 minutes each.

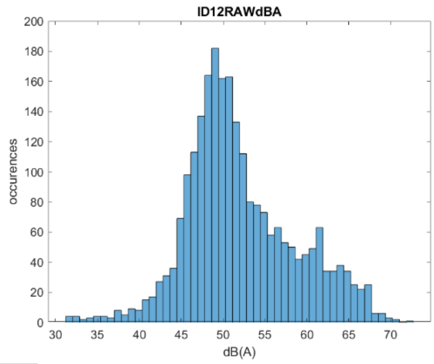
This choice was made by taking into consideration earlier studies on the subject [30]. It has been reported, in fact, that by carrying out long-term noise monitoring and randomly extracting short-term noise monitoring with different lengths of time (1, 5 and 15 minutes), the two types of monitoring can be considered compatible according to the confrontation made by applying the normalized error concept. On these selections the GMM technique has been applied again generating new values of L_{noise_mic} and L_{speech_mic} (Table 1.2) obtained from the average values of the components of the new Gaussian distributions.

Table 1.2 Lnoise_mic and Lspeech_mic values calculated for each full-length audio tracks (raw), single 15 minutes length cut and three 5 minutes length cuts

ID	raw		15 minutes cut		5 minutes cut					
	Lnoise_mic	*Lspeech_mic*	*Lnoise_mic*	*Lspeech_mic*	*Lnoise _mic*			*Lspeech _mic*		
					c1	*c2*	*c3*	*c1*	*c2*	*c3*
2	46.0	61.0	46.1	61.6	45.3	47.4	45.3	60.8	62.6	61.7
3	47.2	69.0	48.8	69.2	47.8	49.9	48.3	67.6	69.8	70.4
4	41.4	51.4	41.2	51.1	40.7	41.2	41.8	49.6	50.7	53.1
5	43.9	55.8	43.8	56.2	42.7	43.7	46.1	56.2	55.2	59.8
6	44.7	54.9	45.3	56.6	43.0	46.4	44.2	57.9	60.2	51.1
7	44.6	52.9	44.1	53.5	43.9	44.2	44.1	52.8	51.6	53.5
8	46.4	57.7	45.1	54.7	44.8	44.8	46.1	53.7	51.9	56.4

9	42.8	56.4	42.5	55.1	43.5	41.0	44.4	54.1	53.8	61.4
10	48.3	56.7	48.6	56.0	51.1	46.5	45.6	59.8	55.2	53.1
11	40.6	48.5	40.9	51.7	40.2	40.8	41.5	50.0	52.0	53.0
12	49.8	62.4	49.0	60.8	50.1	49.0	50.0	63.1	59.5	60.5
13	45.3	67.9	44.6	64.9	42.7	46.7	46.8	61.4	68.0	67.3
14	44.2	63.2	44.8	62.2	44.6	43.7	46.3	60.3	62.0	65.2
15	44.5	57.7	44.6	57.7	44.1	44.5	43.5	57.6	58.7	58.4
16	44.2	66.0	44.0	65.7	40.8	46.0	43.6	61.8	68.8	64.4
17	43.4	48.1	44.6	52.7	44.1	43.2	46.8	51.7	51.6	62.1
18	41.5	50.6	41.5	55.1	40.4	42.9	43.4	54.0	56.5	57.4
19	43.4	55.4	43.7	62.2	43.2	44.8	45.5	64.6	58.5	57.7
20	46.6	51.1	45.2	53.3	45.4	46.4	48.4	55.7	52.7	55.9
21	45.6	61.9	46.7	60.4	46.1	47.4	44.1	58.9	59.4	60.8
22	41.0	53.0	41.0	55.0	42.9	41.1	40.7	63.6	56.9	55.1
23	41.8	64.4	42.9	64.4	42.6	42.1	41.6	61.5	63.8	67.1
24	43.1	56.1	42.2	53.8	41.7	42.2	46.9	55.4	53.9	60.8
25	43.8	56.0	43.5	58.3	44.3	45.2	44.2	59.1	58.3	64.0
26	44.2	49.4	42.3	48.3	42.0	42.8	47.8	52.6	53.7	56.3
27	45.7	60.5	45.5	60.5	45.4	44.6	46.3	59.5	59.8	60.7
28	44.8	56.5	44.4	54.5	44.0	44.6	46.3	55.8	55.0	57.4
29	44.6	50.6	44.8	53.7	45.5	43.4	46.5	55.3	53.1	57.8
30	42.7	49.1	42.7	58.7	44.5	43.2	41.9	52.5	58.7	59.1
31	46.8	52.4	46.8	53.3	47.5	47.5	45.9	53.7	57.3	53.1
32	41.9	55.2	41.5	60.5	41.4	42.1	40.7	54.1	61.5	62.7
33	46.2	61.8	46.6	60.1	48.5	46.9	45.0	61.2	60.5	61.3
34	41.3	44.1	41.4	48.3	41.7	40.6	41.9	49.6	48.4	50.3
35	44.8	44.9	44.4	53.2	43.4	46.0	46.2	51.2	64.4	55.9

The new analyses carried out showed the first improvements in the results. First of all, already visually is noticeable an improvement by looking at the new derived histograms. Taking for example the case of ID12, it is possible to notice that the two Gaussian distributions are more distinguishable by reducing the analysis interval (Fig. 2.7).



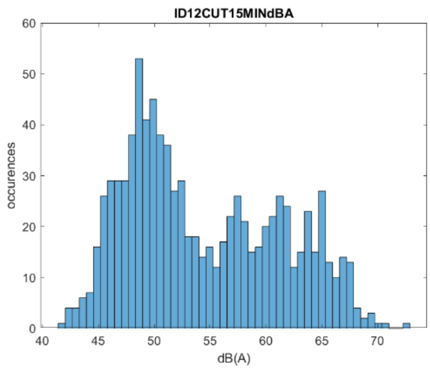
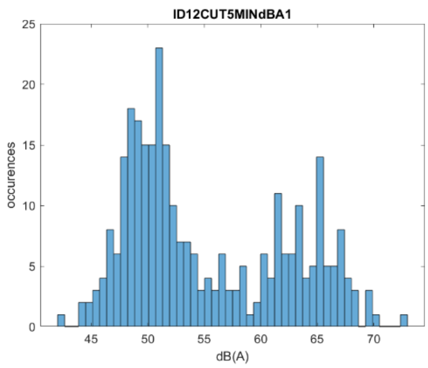


Figure 2.8 ID12 Gaussian distributions for the entire duration of the audio track (above), 15-minute cut (left) and first 5-minute cut (right)



Recalculating the parameters for the graphic representation of Lombard slope L_{noise} and L_{speech} with the same criterion presented in section 3.4, the Lombard slope begins to take values in accordance with those found in literature. This happens in the case of 5-minute cuts, with c = 0.5 dba/dba for November 2020 measurements and c = 0.4 dba/dba for April 2021 measurements (Fig. 2.8).

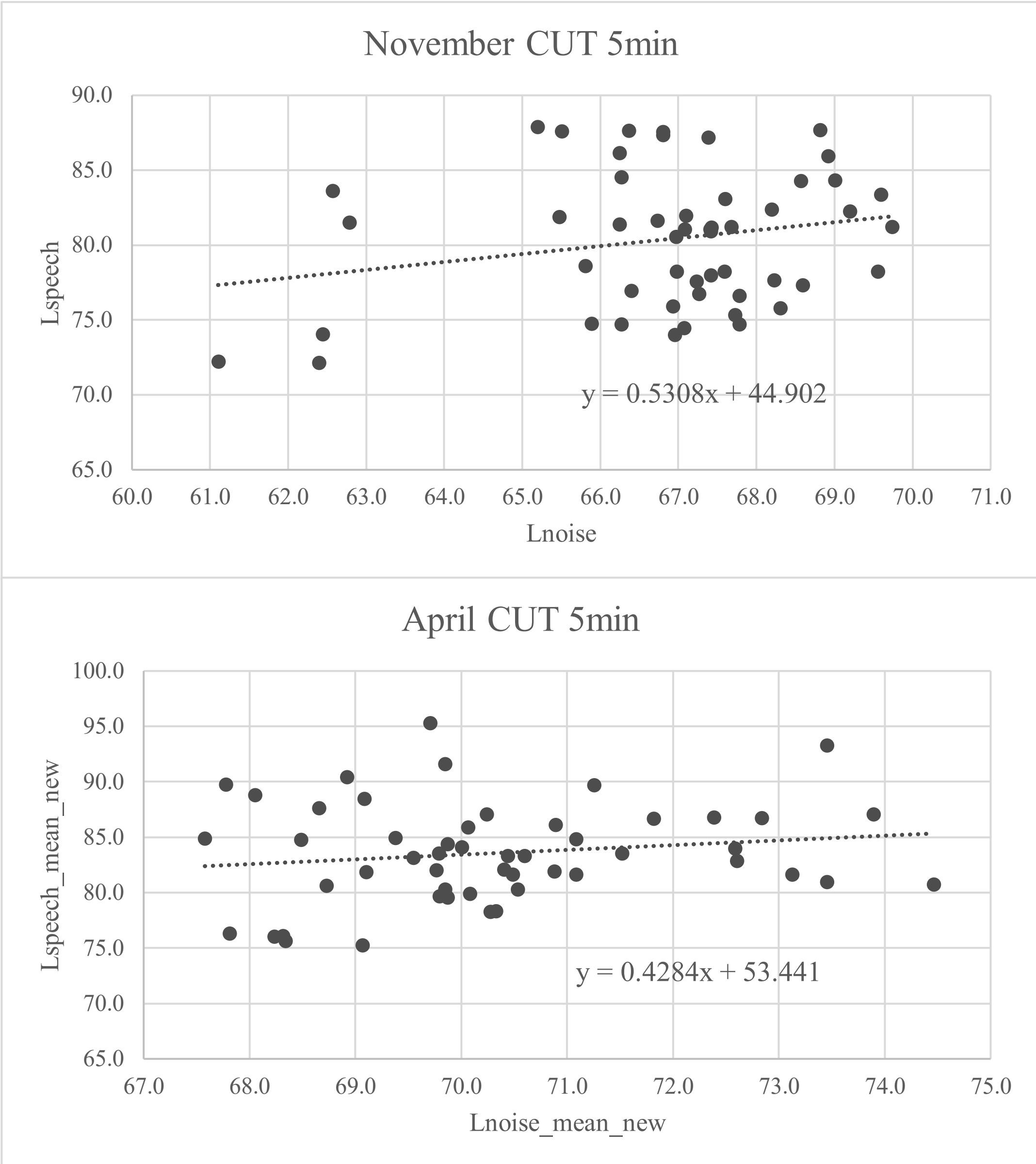


Figure 2.9 Lombard slope obtained by analyzing the 5-minute cuts on the recordings made in November 2020 (above) and April 2021 (below)

3.5.3 Final considerations on preliminary study

The analysis of the first results obtained and the consequent adjustments show how by reducing the length of the range in which to apply the sequence of operations listed in this chapter (from section 3.2 to section 3.4) and obtaining, consequently, a greater number of points, the slope tends to assume values that are consonant and representative of the Lombard effect. These motivations are at the base of the choice to realize a process that, starting from a predefined interval of analysis, can select independently without the supervision of an operator the portions of audio recordings with the best features for the calculation of the Lombard slope.

Chapter 4

AUTOMATION AND OPTIMIZATION PROCESS FOR DATA SELECTION

4.1 OVERVIEW

As discussed at the end of the previous chapter (section 3.5.3) the results obtained from the preliminary study on the application of the methodology presented so far indicate that the slope $\boldsymbol{c}$ could be obtained in a more consonant way by decreasing the analysis interval and by researching specific features in the measured data. At a preliminary level, these last two steps were carried out manually by the operator, so it became necessary to implement an algorithm that carried out such operations in an unsupervised manner. The first factor that was considered for the construction of this process was to decide which statistical criteria to use in order to discriminate the data during the selection. Thresholds have been chosen to meet this requirement by applying the statistical concept of interquartile range (IQR). These thresholds were then used as parameters to be respected within an algorithm that takes a data set as input, verify that the data comply with the established requirements and save them only if the verification has been successful; otherwise the data are discarded. This automation process has been implemented on the Matlab platform with the construction of a script that integrates the analysis operations with Gaussian Mixture Model described in the previous chapter with the data selection process just presented. All the steps are detailed in this chapter.

4.2 THRESHOLDS USED FOR SELECTION

4.2.1 Calculation of interquartiles

In statistics, quartiles are cut points dividing the range of a probability distribution in four intervals of more-or-less equal size. There are 3 quartiles:

- the first quartile Q_1 is the value that leaves 25% of the elements of the distribution on its left;
- the second quartile Q_2 coincides with the median since it is the one that leaves on its left 50% of the distribution;
- the third quartile Q_3 is the values that leaves 75% of the elements of the distribution on its left and 75% on its right.

The difference between third quartile and first quartile is a measure of statistical dispersion and is called interquartile range (IQR).

$$IQR = Q_3 - Q_1$$

The interquartile range concept was used in the calculation of the thresholds as follows.

L_{noise_mic} and L_{speech_mic} values and the respective standard deviations σ_{noise_mic} and σ_{speech_mic} derived from the 5-minute cut analyses (described in section 3.5.3) for both November 2020 and April 2021 measurements were examined. For each of these four variables, 204 values were calculated, resulting from the 3 distinct cuts made on the 34 recorded audio tracks. From these results two new factors were calculated:

- Signal-to-noise ratio (SNR) here calculated as
$$SNR = L_{speech\_mic} - L_{noise\_mic}$$
- difference of standard deviations of noise level and voice level with the respective general averages
$$\Delta_{noise\_std\_dev} = \left|\sigma_{noise\_mic} - \sigma_{noise\_gen}\right|$$
$$\Delta_{speech\_std\_dev} = \left|\sigma_{speech\_mic} - \sigma_{speec\_gen}\right|$$

where σ_{noise_gen} indicates the arithmetic mean of the 204 samples of σ_{noise_mic} and $\sigma_{speech_{gen}}$ indicates the arithmetic mean of the 204 samples of σ_{speech_mic}. All values are shown in Table 1.3.

Table 1.3 Values of SNR, Δ_noise_std_dev and Δnoise_std_dev for each ID (5-minutes cuts)

ID	**SNR**			**Δ_noise_std_dev**			**Δ_speech_std_dev**		
	c1	c2	c3	c1	c2	c3	c1	c2	c3
2	15.5	15.1	16.4	0.6	0.2	0.0	1.0	0.5	1.3
3	19.8	19.9	22.1	0.7	0.6	1.0	0.2	1.4	0.9
4	8.9	9.5	11.2	1.0	1.1	0.8	1.2	1.4	1.0
5	13.5	11.5	13.7	0.2	0.0	0.5	1.6	1.3	3.3
6	14.9	13.8	6.9	0.0	0.3	1.0	4.0	1.0	0.5
7	8.9	7.4	9.4	0.6	0.8	0.8	1.1	0.3	0.1
8	9.0	7.0	10.3	0.6	0.3	0.9	0.3	0.3	0.7
9	10.7	12.8	17.0	0.1	0.7	0.6	1.4	1.8	0.1
10	8.7	8.7	7.5	0.8	0.4	0.4	2.3	0.6	0.4
11	9.8	11.1	11.6	0.7	0.1	0.0	1.5	1.8	0.8
12	13.1	10.5	10.6	0.5	0.4	0.4	1.1	0.8	1.5
13	18.7	21.3	20.6	1.2	1.6	1.5	1.5	1.0	0.1
14	15.7	18.3	18.9	0.9	0.1	0.2	0.8	1.2	0.7
15	13.5	14.2	14.9	0.4	0.4	0.1	0.3	0.4	0.7
16	21.0	22.7	20.7	0.6	1.6	0.2	1.6	0.1	2.1
17	7.6	8.4	15.3	0.3	0.2	0.8	0.8	1.2	0.5
18	13.6	13.7	14.0	0.5	0.7	2.7	0.9	2.0	1.5
19	21.5	13.7	12.2	0.3	0.7	1.0	0.3	0.8	0.6
20	10.2	6.3	7.5	1.1	0.7	0.5	0.4	0.6	0.5
21	12.8	12.0	16.8	0.0	0.0	0.7	0.5	0.5	1.5
22	20.7	15.8	14.4	0.1	0.3	0.5	1.8	1.2	0.9
23	18.9	21.7	25.6	0.0	0.7	0.1	1.0	1.7	1.2
24	13.7	11.6	13.8	0.1	0.5	0.7	0.5	0.6	0.8
25	14.8	13.1	19.8	0.4	0.6	0.2	0.6	0.0	0.2
26	10.5	11.0	8.5	0.2	0.1	0.6	2.5	1.7	2.3
27	14.1	15.2	14.4	0.3	0.9	0.7	0.4	0.8	1.6

28	11.9	10.4	11.1	0.3	0.2	0.0	1.6	1.6	0.2
29	9.8	9.7	11.4	0.6	0.8	0.5	1.4	0.1	0.1
30	8.0	15.5	17.3	0.0	0.6	0.0	1.3	0.7	0.4
31	6.2	9.8	7.3	0.5	0.5	0.8	0.4	2.0	1.5
32	12.7	19.3	21.9	0.2	0.6	0.2	2.1	2.4	3.5
33	12.7	13.6	16.3	0.3	0.2	0.5	1.5	1.6	1.7
34	7.9	7.8	8.4	0.6	0.0	0.7	0.6	1.4	0.4
35	7.7	18.4	9.7	0.1	1.1	0.7	1.3	1.1	3.0

The first quartile Q_1 has been calculated on all the values of SNR, constituting the first threshold that will be used in the algorithm ($threshold_{SNR}$), while the third quartile Q_3 has been calculated on all the differences between the standard deviations constituting the second threshold that will be used in the algorithm ($threshold_{\Delta_{std\_dev}}$). The values obtained were 9.8 dBA for $threshold_{SNR}$ and 1.1 dBA for $threshold_{\Delta_{std\_dev}}$ (Fig. 3.1).

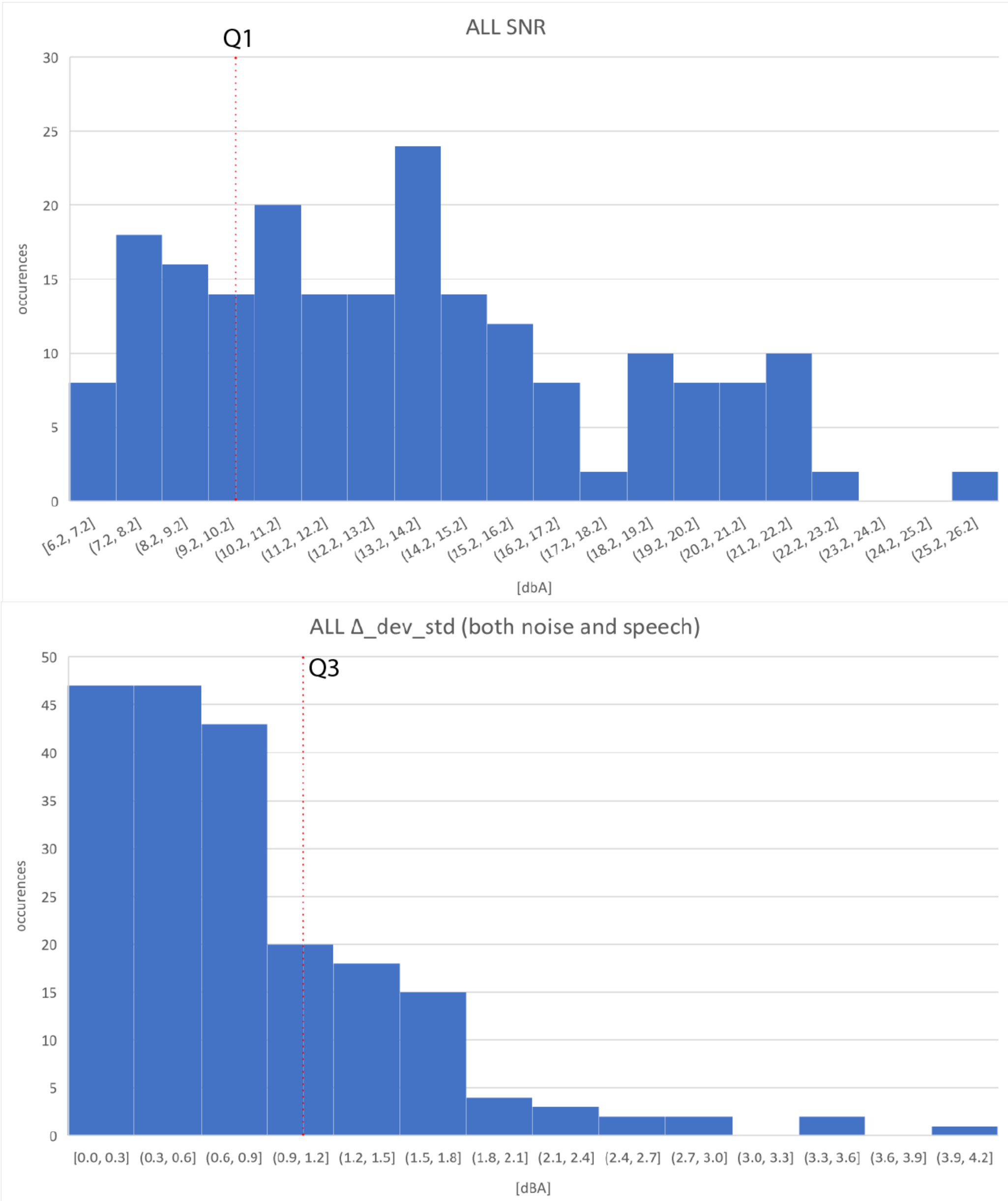


Figure 3.1 Calculated thresholds: first quartile on all SNR values (above) and third quartile on all Δ_dev_std values (below)

4.2.2 Selection criterion

After calculating the two thresholds regarding the signal to noise ratio and the difference between the standard deviations of noise and voice levels with their general averages, the selection criterion on the data was chosen as follows: given a segment on which the GMM method has been applied and from which the values of noise and voice levels (L_{noise_mic} and L_{speech_mic}) with the respective standard deviations (σ_{noise_mic} and σ_{speech_mic}) have been calculated, it is selected and used to derive the Lombard slope $\boldsymbol{c}$ if it has simultaneously

$$SNR > threshold_{SNR}$$

$$\Delta_{noise\_std\_dev} \leq threshold_{\Delta_{noise\_std\_dev}}$$

$$\Delta_{speech\_std\_dev} \leq threshold_{\Delta_{speech\_std\_dev}}$$

where all the quantities taken into consideration have been derived according to sections 3.2 and 4.2.1 present in this document. If a segment does not meet this selection criterion, it is discarded.

4.3 AUTOMATION PROCESS

4.3.1 Operating principle

In order to automate the process in which to integrate the analyses described in Chapter 3 with the selection of data presented in this chapter, an algorithm has been created to perform the following operations:

- a data set of sound pressure values is taken into the input;
- the data set is divided according to the length of the analysis interval established, deriving an N number of segments;
- GMM technique is applied to each segment and the values of interest are obtained;
- the selection criterion is checked: if the thresholds are met then all data relating to that segment are saved, otherwise it is discarded.

Applying this sequence of operations to the specific case examined in this thesis required a solution to two main problems:

1. The threshold for standard deviations of noise and voice levels has been calculated using their respective general averages; these averages update dynamically if a segment is discarded.
2. Discarding an entire segment is a significant waste of data. It is therefore necessary to apply a mechanism that shifts within the segment, discarding only a part of it.

This has led to the choice to structure the algorithm in such a way as to be able to best address both critical situations presented above. In particular, multiple loops have been created to solve the first problem, while for the second the concept of shift window has been implemented.

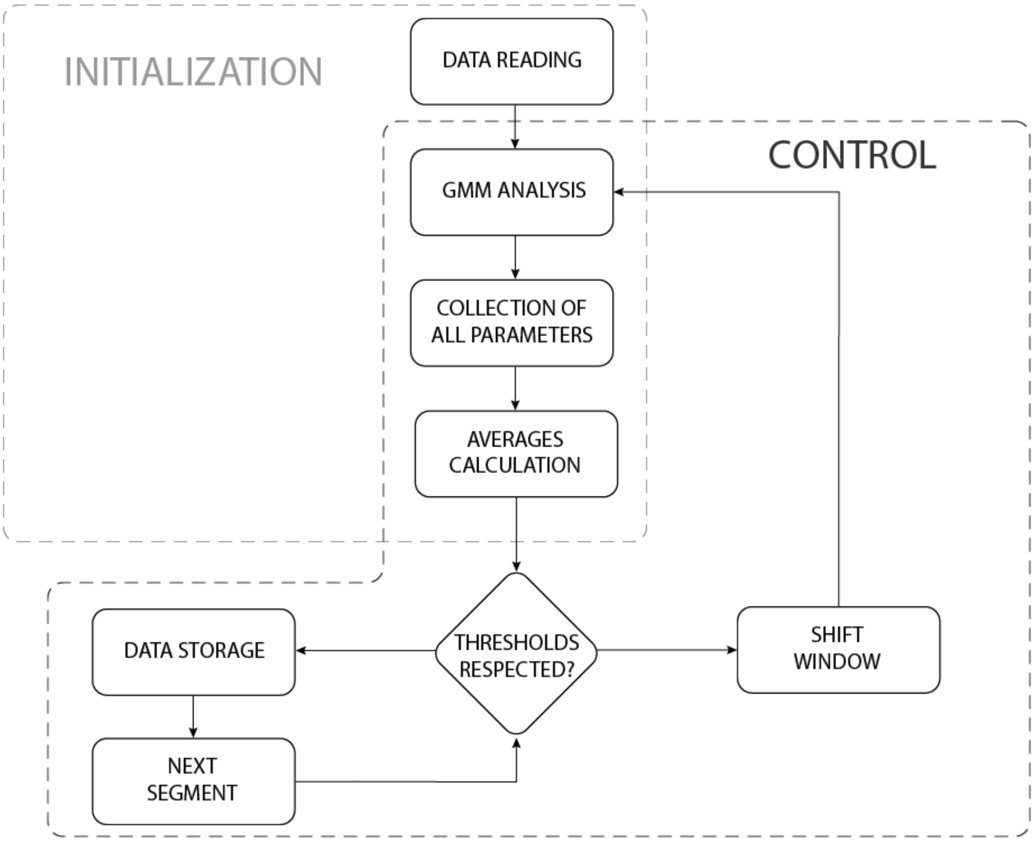


Figure 3.2 Flow chart of the entire algorithm

Figure 3.2 shows the operations performed in the execution of the algorithm. There are two stages:

- "Initialization", which includes a first phase of data reading, the application of the GMM technique, the creation of a structure in which all the parameters of interest are present and the calculation of averages of standard deviations (used together with SNR values for comparison with thresholds).
- "Control", in which the election criterion is checked: if it is respected, the corresponding data is saved and the next set of values is considered, otherwise only a portion of that set is discarded. The latter situation is addressed with a mechanism that operates a forward shift within the segment just examined (shift window); once this operation is completed, the new parameters are calculated.

The implementation and functioning of both stages are discussed in detail in the continuation of the paragraph.

4.3.2 Initialization phase

In this phase are collected and calculated all the parameters of interest for the next stage in which the actual selection on the data takes place. There are two nested cycles: the external one takes into account the number of participants for which vocal effort measurements have been made, while the inner one operates on the corresponding single recorded audio track. Out of both, a variable l is defined which indicates the length of the analysis interval. Taking into account a given recording of duration L, the number of possible segments n_s in which it can be broken down is given by the quotient from the ratio *L/l*. This concept is equivalent to transforming a column vector of length L into a matrix of dimensions $l \times n_s$ (Fig. 3.3).

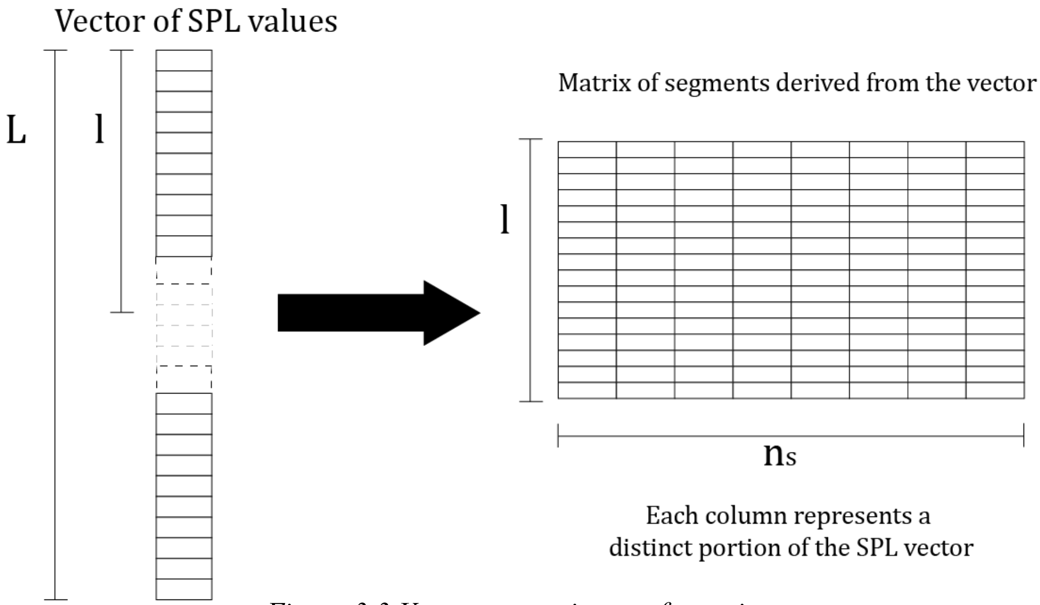


Figure 3.3 Vector to matrix transformation

The column vector represents the set of SPL values of a given audio recording: by transforming it into a matrix with a number of rows equal to the analysis interval and with a number of columns equal to the number of segments calculated as described above, GMM analysis can be applied to each segment independently. For example, if an audio recording lasts 30 minutes (L) and the length of the analysis interval is 5 minutes (*l*), it can be divided into 6 segments (n_s) of 5 minutes each. The average noise and voice levels with the respective standard deviations and the signal to noise ratio are calculated for each segment; all these parameters are then saved as local variables that can be used in the control phase. Taking the above example, 6 pairs of $L_{noise\_mic} \backslash L_{speech\_mic}$ values, 6 pairs of $\sigma_{noise\_mic} \backslash \sigma_{speech\_mic}$ and 6 SNR values are obtained. These operations are performed in the internal loop. In the external one, instead, a cell array data type is constructed where to contain all the values obtained for every participant (Fig. 3.4): every row corresponds, in fact, to a determined ID while every column to a category of parameters calculated with the GMM technique together with other data obtained during the first reading (time interval, sound pressure levels of the nearest sound level meter, number of people present).

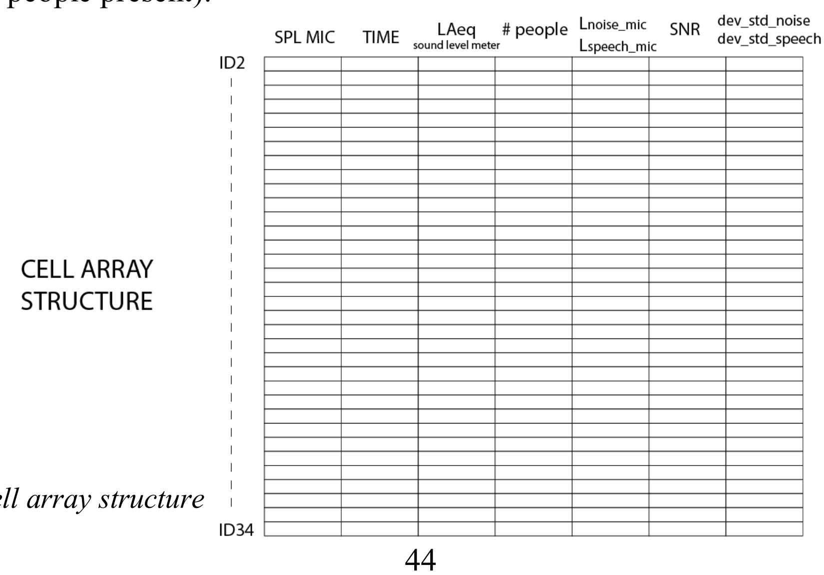


Figure 3.4 Cell array structure

Below is the portion of code used on Matlab for the implementation of the two cycles just described.

```
id_length=35;
length_interval=60;
filename_data='name of the path where audio files are present';
for j=2:id_length % EXTERNAL LOOP
sheet_ID=strcat('ID',num2str(j));
time=xlsread(filename_data,sheet_ID,'B:B');
dB=xlsread(filename_data,sheet_ID,'C:C');
Laeq_dt=xlsread(filename_data,sheet_ID,'D:D');
people=xlsread(filename_data,sheet_ID,'E:E');
lim=length(dB);
num_interval=fix(lim/length_interval);
      for i=1:num_interval% INTERNAL LOOP
      SPL_matrix(:,i)=dB(start_interval:(length_interval*i));         %VECTOR TO
      time_matrix(:,i)=time(start_interval:(length_interval*i));     %MATRIX
      Laeq_dt_matrix(:,i)=Laeq_dt(start_interval:(length_interval*i)); %AND GMM
      people_matrix(:,i)=people(start_interval:(length_interval*i)); %ANALYSIS
      GMModel=fitgmdist(SPL_matrix(:,i),2,'RegularizationValue',0.01);
      mean_val(:,i)=sort(GMModel.mu);
      sigma_val(:,i)=sort([GMModel.Sigma(:,:,1) GMModel.Sigma(:,:,2)]');
      dev_std=sort(sqrt(sigma_val));
      SNR=max(mean_val)-min(mean_val);
      start_interval=(i*length_interval)+1;
      end
ALL_IDs{j,1}=SPL_matrix;             %CELL ARRAY CONSTRUCTION
ALL_IDs{j,2}=time_matrix;
ALL_IDs{j,3}=Laeq_dt_matrix;
ALL_IDs{j,4}=people_matrix;
ALL_IDs{j,5}=mean_val;
ALL_IDs{j,6}=SNR;
ALL_IDs{j,7}=sigma_val;
ALL_IDs{j,8}=dev_std;
SPL_matrix=[];
time_matrix=[];
Laeq_dt_matrix=[];
people_matrix=[];
mean_val=[];
SNR=[];
sigma_val=[];
dev_std=[];
start_interval=1;
dB=[];
time=[];
Laeq_dt=[];
people=[];
end
dev_std_mean=mean([ALL_IDs{:,8}],2);
```

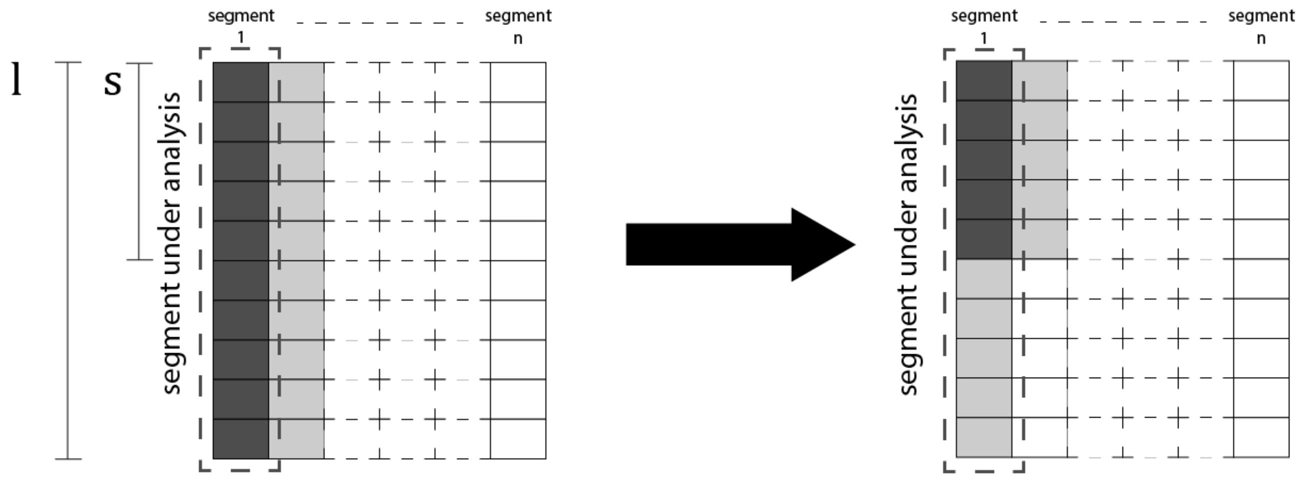
The mean values of the standard deviations of both components (noise and voice) are calculated with the last row of the code; they will be used later for check compliance with the thresholds.

4.3.3 Control phase

In this phase the actual selection between the segments takes place: these can be saved and then used for the determination of the Lombard slope or be discarded. The processing is structured as follows:

- A data set calculated for a given ID is examined; this is represented by an entire row of the cell array produced during initialization.
- The values of SNR, σ_{noise_mic}, σ_{noise_gen}, σ_{speech_mic} and σ_{speech_gen} are read.
- The selection criterion (section 4.2.2) is checked; this can lead to two cases:
    1. the values fall within the range bounded by the thresholds, for which the segment is selected; this involves saving on an excel sheet all the data related to that segment;
    2. the criterion is not met, so that segment is not selected; this involves the actuation of the shifting mechanism.

This last operation called "shifting" has been thought in order to be able to discard the least possible number of data in case of negative result if the thresholds were not respected. It operates within the segment itself in this way: given a segment formed by l elements and defined a shifting window of length s, if the selection criterion is not met then the elements from one to s of the segment under consideration are discarded while the elements from s+1 to l are analyzed again together with the other s elements of the next segment. For example, if a segment lasts 60 seconds and the window is 30 seconds, the first 30 seconds of the current segment are discarded while the remaining 30 seconds are analyzed together with the first 30 seconds of the next segment.



only the first s elements of the segment are discarded, while the next ones are shifted

Figure 3.5 Shifting example

The shifting operation requires a recalculation of all segments for a given ID. For that reason there is also at this stage a reading of the data from the original file, a new analysis with the GMM technique and the determination of all the new parameters to be used for the next control. All these operations have been implemented in a manner similar to that done in the initialization's phase, by reconstructing the single matrix for the ID under analysis and updating the data present in the global cell array. In this way the general averages of the standard deviations are dynamically recalculated and used for the next control cycle. A special attention has been paid to the case in which shifts are no longer available. Below is the full code used for the implementation of the control stage.

```
threshold_delta_dev_std_gen=1.1;
threshold_SNR=9.8;
length_interval=60;
window=30;
for j=2:id_length
```

```
        SPL_matrix=[ALL_IDs{j,1}];
        time_matrix=[ALL_IDs{j,2}];
        Laeq_dt_matrix=[ALL_IDs{j,3}];
        people_matrix=[ALL_IDs{j,4}];
        mean_val=[ALL_IDs{j,5}];
        SNR=[ALL_IDs{j,6}];
        sigma_val=[ALL_IDs{j,7}];
        dev_std=[ALL_IDs{j,8}];
        Laeq_dt_mean=mean(Laeq_dt_matrix);
        Laeq_dt_mode=mode(Laeq_dt_matrix);
        num_interval=length(mean_val);
        start_write=1;
        end_write=length_interval;
        start_shift=window+1;
        end_shift=window+length_interval;
        k=1;
        while k<=num_interval
            if(SNR(k)>threshold_SNR) && (abs((dev_std(1,k)-
dev_std_mean(1,1)))<=threshold_delta_dev_std_gen) && (abs((dev_std(2,k)-
dev_std_mean(2,1)))<=threshold_delta_dev_std_gen)
                X = sprintf('Il segmento %d di ID %d va bene.',k,j);
                disp(X)
                count_segment=count_segment+1;
                end_write_ALL=count_segment;
                sheet_ID=strcat('ID',num2str(j));
                sheet_ALL='ALL_IDs';
                filename=strcat('ID19-
ID35_analysis_dBA_',num2str(length_interval),'_seconds_shift_window_',num2str(wi
ndow),'_seconds_SNR_6dB.xlsx');

xlswrite(filename,j,sheet_ALL,strcat('A',num2str(start_write_ALL),':A',num2str(e
nd_write_ALL)));

xlswrite(filename,k,sheet_ALL,strcat('B',num2str(start_write_ALL),':B',num2str(e
nd_write_ALL)));

xlswrite(filename,mean_val(1,k),sheet_ALL,strcat('C',num2str(start_write_ALL),':
C',num2str(end_write_ALL)));

xlswrite(filename,mean_val(2,k),sheet_ALL,strcat('D',num2str(start_write_ALL),':
D',num2str(end_write_ALL)));

xlswrite(filename,Laeq_dt_mean(1,k),sheet_ALL,strcat('H',num2str(start_write_ALL
),':H',num2str(end_write_ALL)));

xlswrite(filename,Laeq_dt_mode(1,k),sheet_ALL,strcat('I',num2str(start_write_ALL
),':I',num2str(end_write_ALL)));

xlswrite(filename,k,sheet_ID,strcat('A',num2str(start_write),':A',num2str(end_wr
ite)));

xlswrite(filename,time_matrix(:,k),sheet_ID,strcat('B',num2str(start_write),':B'
,num2str(end_write)));

xlswrite(filename,SPL_matrix(:,k),sheet_ID,strcat('C',num2str(start_write),':C',
num2str(end_write)));

xlswrite(filename,Laeq_dt_matrix(:,k),sheet_ID,strcat('D',num2str(start_write),'
:D',num2str(end_write)));

xlswrite(filename,people_matrix(:,k),sheet_ID,strcat('E',num2str(start_write),':
E',num2str(end_write)));
```

```
xlswrite(filename,mean_val(:,k),sheet_ID,strcat('F',num2str(start_write),':F',nu
m2str(end_write)));

xlswrite(filename,sigma_val(:,k),sheet_ID,strcat('G',num2str(start_write),':G',n
um2str(end_write)));

xlswrite(filename,dev_std(:,k),sheet_ID,strcat('H',num2str(start_write),':H',num
2str(end_write)));

xlswrite(filename,Laeq_dt_mean(:,k),sheet_ID,strcat('I',num2str(start_write),':I
',num2str(end_write)));

xlswrite(filename,Laeq_dt_mode(:,k),sheet_ID,strcat('J',num2str(start_write),':J
',num2str(end_write)));
                h=histogram(SPL_matrix(:,k),50);

title(strcat('ID',num2str(j),'INTERVAL',num2str(k),'LENGTH',num2str(length_inter
val),'SECSNR6dB'));

savefig(strcat('ID',num2str(j),'_INTERVAL',num2str(k),'_LENGTH',num2str(length_i
nterval),'_SEC_SNR_6dB.fig'))

saveas(h,strcat('ID',num2str(j),'_INTERVAL',num2str(k),'_LENGTH',num2str(length_
interval),'_SEC_SNR_6dB.png'))
                start_write=end_write+1;
                end_write=end_write+length_interval;
                start_write_ALL=end_write_ALL+1;
                start_shift=start_shift+length_interval;
                end_shift=end_shift+length_interval;
                k=k+1;
                if (num_interval==2) && (size(SPL_matrix,2)==1)
                    num_interval=num_interval-1;
                    W=sprintf('Non ci sono più intervalli disponibili');
                    disp(W)
                end
            else
                Y = sprintf('Il segmento %d di ID %d NON va bene. Aggiungo %d
secondi.',k,j,window);
                disp(Y)
                sheet_ID=strcat('ID',num2str(j));
                time=xlsread(filename_data,sheet_ID,'B:B');
                dB=xlsread(filename_data,sheet_ID,'C:C');
                Laeq_dt=xlsread(filename_data,sheet_ID,'D:D');
                people=xlsread(filename_data,sheet_ID,'E:E');
                lim=length(dB);
                lim_new=lim-start_shift+1;
                num_interval_new=fix(lim_new/length_interval);
                start_interval=start_shift;
                end_interval=end_shift;
                if (num_interval_new==0)
                    Z=sprintf('Non ci sono più intervalli disponibili');
                    disp(Z)
                    SPL_matrix=[SPL_matrix(:,1:k-1)];
                    num_interval=num_interval-1;
                else
                    for z=1:num_interval_new
                        SPL_matrix_new(:,z)=dB(start_interval:end_interval);
                        time_matrix_new(:,z)=time(start_interval:end_interval);

Laeq_dt_matrix_new(:,z)=Laeq_dt(start_interval:end_interval);
```

```
people_matrix_new(:,z)=people(start_interval:end_interval);
                        % disp(SPL_matrix_new)

GMModel=fitgmdist(SPL_matrix_new(:,z),2,'RegularizationValue',0.01);
                        mean_val_new(:,z)=sort(GMModel.mu);
                        sigma_val_new(:,z)=sort([GMModel.Sigma(:,:,1)
GMModel.Sigma(:,:,2)]');
                        dev_std_new=sort(sqrt(sigma_val_new));
                        SNR_new=max(mean_val_new)-min(mean_val_new);
                        Laeq_dt_mean_new=mean(Laeq_dt_matrix_new);
                        Laeq_dt_mode_new=mode(Laeq_dt_matrix_new);
                        start_interval=start_interval+length_interval;
                        end_interval=end_interval+length_interval;
                    end
                    if (k==1)
                        SPL_matrix=[];
                        time_matrix=[];
                        Laeq_dt_matrix=[];
                        people_matrix=[];
                        mean_val=[];
                        SNR=[];
                        sigma_val=[]
                        dev_std=[];
                        Laeq_dt_mean=[];
                        Laeq_dt_mode=[];
                        SPL_matrix=SPL_matrix_new;
                        time_matrix=time_matrix_new;
                        Laeq_dt_matrix=Laeq_dt_matrix_new;
                        people_matrix=people_matrix_new;
                        mean_val=mean_val_new;
                        SNR=SNR_new;
                        sigma_val=sigma_val_new;
                        dev_std=dev_std_new;
                        Laeq_dt_mean=Laeq_dt_mean_new;
                        Laeq_dt_mode=Laeq_dt_mode_new;
                    else
                        SPL_matrix=[SPL_matrix(:,1:(k-1)) SPL_matrix_new];
                        time_matrix=[time_matrix(:,1:(k-1)) time_matrix_new];
                        Laeq_dt_matrix=[Laeq_dt_matrix(:,1:(k-1))
Laeq_dt_matrix_new];
                        people_matrix=[people_matrix(:,1:(k-1))
people_matrix_new];
                        mean_val=[mean_val(:,1:(k-1)) mean_val_new];
                        SNR=[SNR(:,1:(k-1)) SNR_new];
                        sigma_val=[sigma_val(:,1:(k-1)) sigma_val_new];
                        dev_std=[dev_std(:,1:(k-1)) dev_std_new];
                        Laeq_dt_mean=[Laeq_dt_mean(:,1:(k-1)) Laeq_dt_mean_new];
                        Laeq_dt_mode=[Laeq_dt_mode(:,1:(k-1)) Laeq_dt_mode_new];
                    end
                    ALL_IDs{j,1}=SPL_matrix;
                    ALL_IDs{j,2}=time_matrix;
                    ALL_IDs{j,3}=Laeq_dt_matrix;
                    ALL_IDs{j,4}=people_matrix;
                    ALL_IDs{j,5}=mean_val;
                    ALL_IDs{j,6}=SNR;
                    ALL_IDs{j,7}=sigma_val;
                    ALL_IDs{j,8}=dev_std;
                    dev_std_mean=mean([ALL_IDs{:,8}],2);
                    num_interval=length(mean_val);
                    SPL_matrix_new=[];
                    time_matrix_new=[];
```

```
                    Laeq_dt_matrix_new=[];
                    people_matrix_new=[];
                    mean_val_new=[];
                    SNR_new=[];
                    sigma_val_new=[];
                    dev_std_new=[];
                    Laeq_dt_mean_new=[];
                    Laeq_dt_mode_new=[];
                    start_shift=start_shift+window;
                    end_shift=end_shift+window;
                end
            end
        end
    end
end
```

4.3.4 Final adjustment

After some tests carried out using the algorithm described in this chapter, a further adjustment has been chosen. In particular, it was noted that the first threshold obtained by calculating the first quartile on the SNR values influenced the choice of segments too significantly. In the drawn dispersion graphs there was, in fact, a sharp cut in the selected points: this trend clearly traced the slope. For this reason it was decided to lower this threshold from a value of 9.8 dBA to a value of 6 dBA. This in acoustics represents a significant value: a doubling of the sound pressure produced by a source implies an increase of +6dB. So taking the two levels L_{speech_mic} and L_{noise_mic} found with GMM analysis, if there is a difference of at least 6 dB between them then the contribution produced by L_{noise_mic} could be considered not significant compared to L_{speech_mic}.

Chapter 5

RESULTS

5.1 OVERVIEW

In this chapter the results obtained through the application of the methodology developed to the case study of the canteen employees of the Politecnico di Torino are reported. The data collected and analyzed as described in Chapter 3 have been processed by the selection algorithm presented in Chapter 4. Analysis intervals of 1,2,3,4 and 5 minutes were used with a 30-second shift window, for both November 2020 and April 2021 data. The thresholds chosen are:

$$threshold_{SNR} = 6\ dBA$$

$$threshold_{\Delta_{std\_dev}} = 1.1\ dBA$$

For the determination of the Lombard slope the L_{noise} and L_{speech} parameters were used following the definitions presented in section 3.4. Slopes have been obtained for each analysis interval by examining separately the measurements of November 2020 and April 2021 and finally putting together all the data divided into the ranges indicated.

5.2 LOMBARD SLOPE FOR NOVEMBER 2020

In the November 2020 measurements, 18 participants were monitored, 7 during the first day and 11 during the second. The audio recording of the first participant was discarded for technical reasons.

5.2.1 60-second segments

Choosing 60 seconds as the length of the analysis interval, the algorithm selected 234 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 10 useful segments were selected for 15 of them; 8 out of 17 have at least 15 segments. For the two not included in this group, 7 (for ID4) and 9 (for ID11) segments were selected respectively. The maximum number of segments selected is 25 per ID15, the only one exceeding 20 units. The values of L_{noise_mic} and L_{speech_mic} vary between 30.8 dBA and 53 dBA for the first and 43.7 dBA and 74.3 dBA for the second. For general noise and voice levels, instead, the values vary between 60.1 dBA and 71.3 dBA for L_{noise}, and between 68.5 dBA and 95.1 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.57$ (Fig. 4.1).

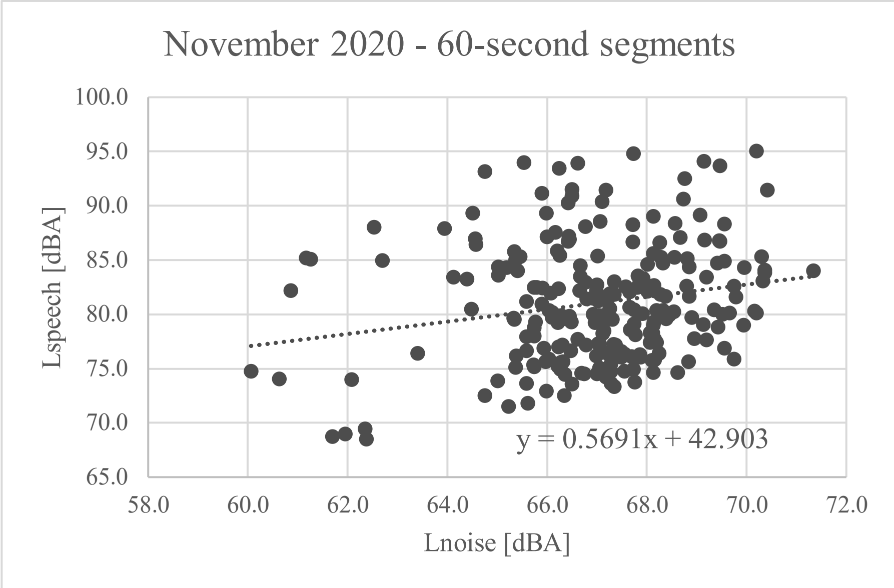


Figure 4.1 Lombard slope for November 2020, 60-second segments

5.2.2 120-second segments

Choosing 120 seconds as the length of the analysis interval, the algorithm selected 153 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 10 useful segments were selected for 7 of them; only two exceed the 10 units, that is ID12 with 13 and ID15 with 16. The latter represents the maximum number of selected segments, while the minimum is 3 for ID4. The values of L_{noise_mic} and L_{speech_mic} vary between 37.8 dBA and 51.4 dBA for the first and 45.7 dBA and 73.1 dBA for the second. For general noise and voice levels, instead, the values vary between 61.4 dBA and 70.0 dBA for L_{noise}, and between 68.3 dBA and 99.4 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.76$ (Fig. 4.2).

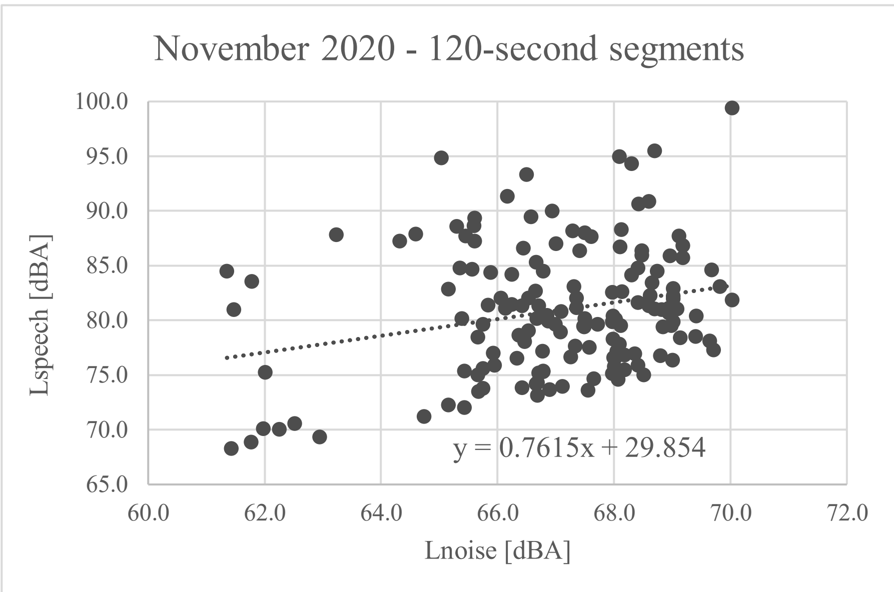


Figure 4.2 Lombard slope for November 2020, 120-second segments

5.2.3 180-second segments

Choosing 180 seconds as the length of the analysis interval, the algorithm selected 107 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 7 useful segments were selected for 9 of them; three of which reach a maximum of 9 segments each (ID6, ID12 and ID15). The minimum number of segments selected, instead, is 2 for ID4. The values of L_{noise_mic} and L_{speech_mic} vary between 39.0 dBA and 52.0 dBA for the first and 46.5 dBA and 71.3 dBA for the second. For general noise and voice levels, instead, the values vary between 60.8 dBA and 69.7 dBA for L_{noise}, and between 67.0 dBA and 97.7 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.77$ (Fig. 4.3).

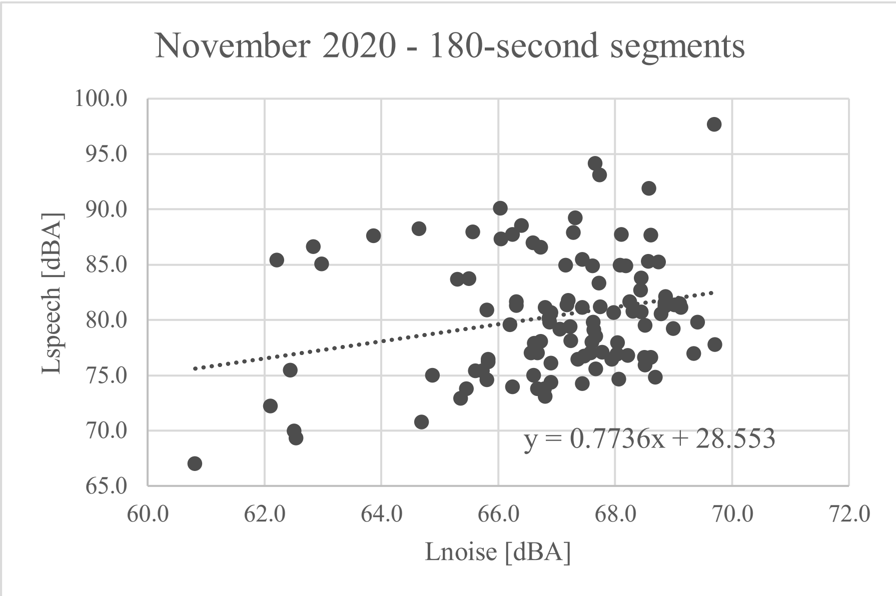


Figure 4.3 Lombard slope for November 2020, 180-second segments

5.2.4 240-second segments

Choosing 240 seconds as the length of the analysis interval, the algorithm selected 80 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 5 useful segments were selected for 10 of them; three of which reach a maximum of 7 segments each (ID6, ID12 and ID15). The minimum number of segments selected, instead, is 2 for ID4. The values of L_{noise_mic} and L_{speech_mic} vary between 39.4 dBA and 51.6 dBA for the first and 45.8 dBA and 71.4 dBA for the second. For general noise and voice levels, instead, the values vary between 62.3 dBA and 70.0 dBA for L_{noise}, and between 68.7 dBA and 97.6 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.81$ (Fig. 4.4).

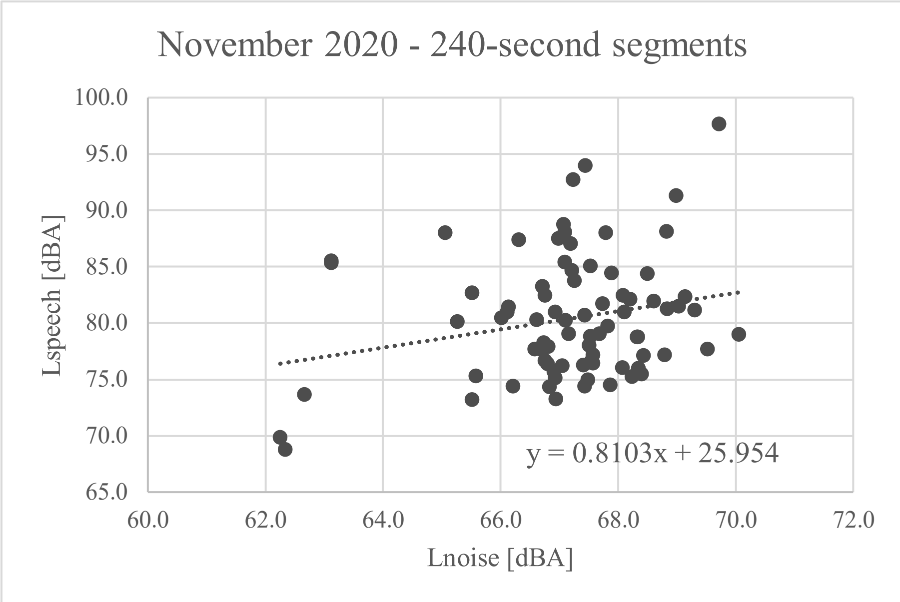


Figure 4.4 Lombard slope for November 2020, 240-second segments

5.2.5 300-second segments

Choosing 300 seconds as the length of the analysis interval, the algorithm selected 67 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 5 useful segments have been selected for only 5 of them. The minimum and maximum number of selected segments were 4 for ID4 and 6 for ID15. The values of L_{noise_mic} and L_{speech_mic} vary between 40.5 dBA and 51.3 dBA for the first and 47.6 dBA and 70.9 dBA for the second. For general noise and voice levels, instead, the values vary between 61.6 dBA and 69.5 dBA for L_{noise}, and between 69.4 dBA and 96.8 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.57$ (Fig. 4.5).

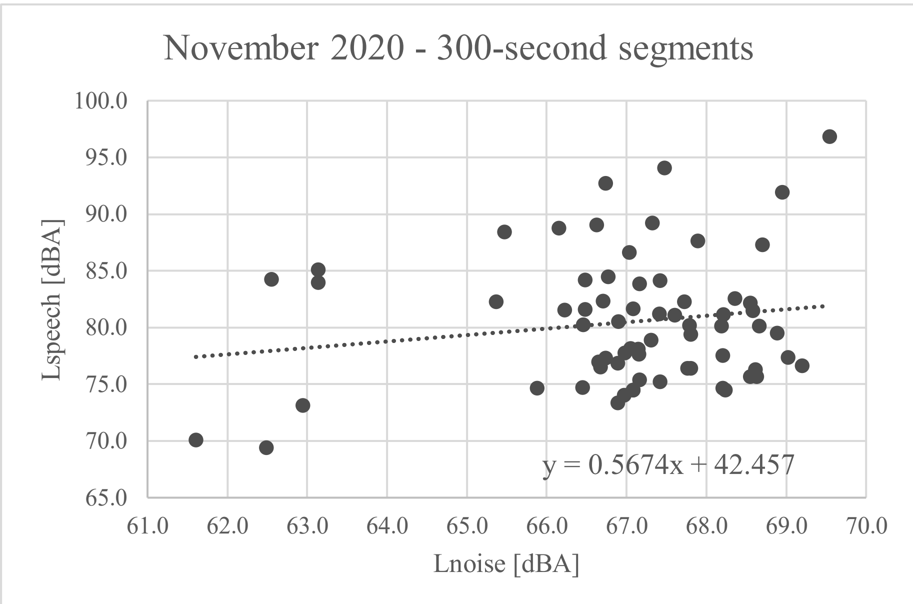


Figure 4.5 Lombard slope for November 2020, 300-second segments

5.3 LOMBARD SLOPE FOR APRIL 2021

In the April 2021 measurements, 17 participants were monitored, 11 during the first day and 6 during the second. All audio tracks have been analyzed.

5.3.1 60-second segments

Choosing 60 seconds as the length of the analysis interval, the algorithm selected 301 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 20 useful segments have been selected for 6 of them; only for one of them (ID26) it goes below the 10 units with 8 segments. The maximum number of selected segments were 44 for ID19. The values of L_{noise_mic} and L_{speech_mic} vary between 28.2 dBA and 54.8 dBA for the first and 37.6 dBA and 73.2 dBA for the second. For general noise and voice levels, instead, the values vary between 62.1 dBA and 74.7 dBA for L_{noise}, and between 73.7 dBA and 101.7 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.62$ (Fig. 4.6).

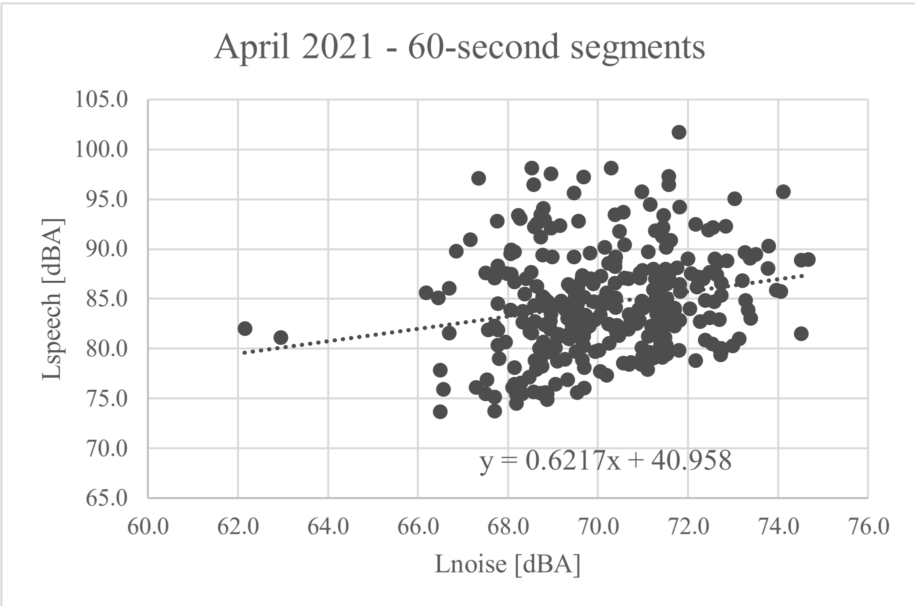


Figure 4.6 Lombard slope for April 2021, 60-second segments

5.3.2 120-second segments

Choosing 120 seconds as the length of the analysis interval, the algorithm selected 176 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 10 useful segments have been selected for 9 of them. The minimum and maximum number of selected segments were 4 for ID35 and 24 for ID19. The values of L_{noise_mic} and L_{speech_mic} vary between 33.5 dBA and 51.2 dBA for the first and 41.4 dBA and 71.6 dBA for the second. For general noise and voice levels, instead, the values vary between 63.0 dBA and 75.0 dBA for L_{noise}, and between 74.2 dBA and 99.0 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.45$ (Fig. 4.7).

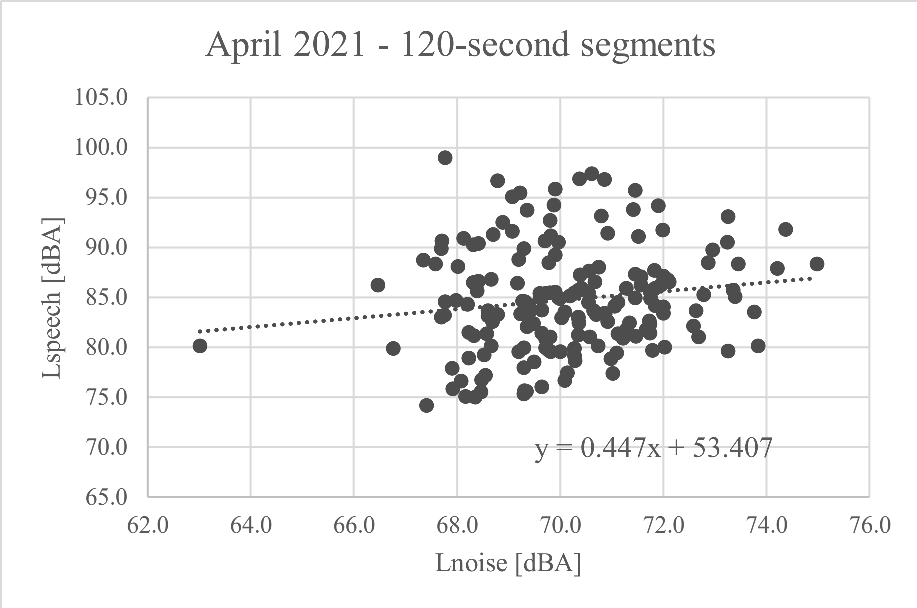


Figure 4.7 Lombard slope for April 2021, 120-second segments

5.3.3 180-second segments

Choosing 180 seconds as the length of the analysis interval, the algorithm selected 127 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 5 useful segments have been selected for all of them but two (ID32 and ID33 with 4). At least 10 segments were selected for four of them, with a maximum of 16 for ID19. The values of L_{noise_mic} and L_{speech_mic} vary between 36.9 dBA and 50.6 dBA for the first and 46.9 dBA and 72.2 dBA for the second. For general noise and voice levels, instead, the values vary between 63.3 dBA and 74.6 dBA for L_{noise}, and between 74.4 dBA and 100.5 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.24$ (Fig. 4.8).

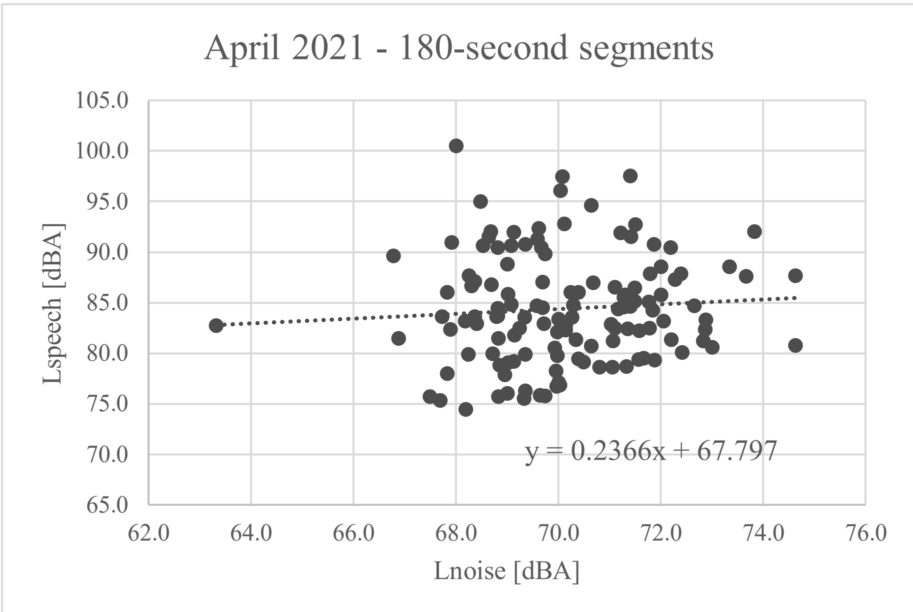


Figure 4.8 Lombard slope for April 2021, 180-second segments

5.3.4 240-second segments

Choosing 240 seconds as the length of the analysis interval, the algorithm selected 96 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 5 useful segments have been selected for 9 of them. The minimum and maximum number of selected segments were 2 for ID35 and 12 for ID19. The values of L_{noise_mic} and L_{speech_mic} vary between 38.5 dBA and 49.4 dBA for the first and 47.5 dBA and 71.3 dBA for the second. For general noise and voice levels, instead, the values vary between 63.6 dBA and 74.9 dBA for L_{noise}, and between 75.0 dBA and 97.5 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.67$ (Fig. 4.9).

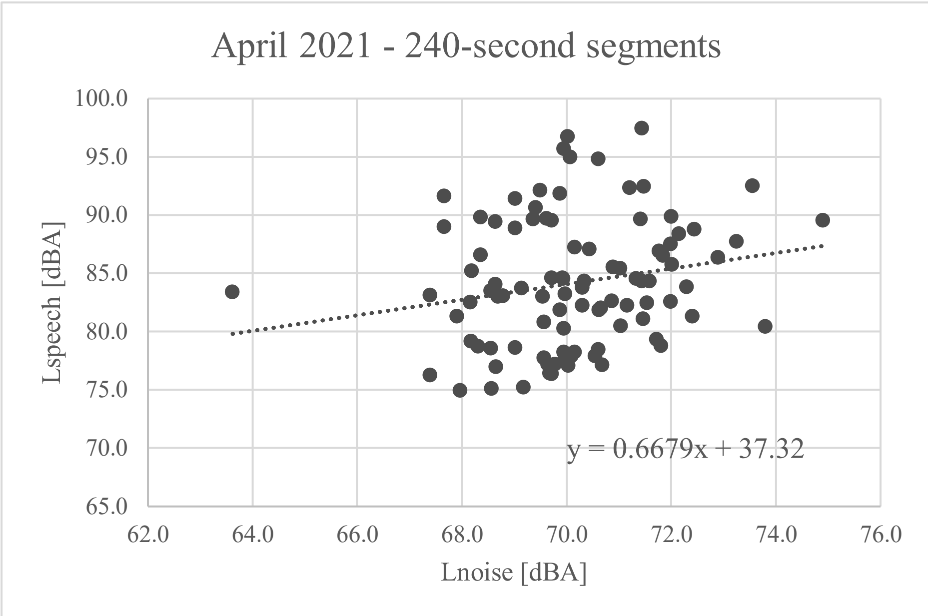


Figure 4.9 Lombard slope for April 2021, 240-second segments

5.3.5 300-second segments

Choosing 300 seconds as the length of the analysis interval, the algorithm selected 83 useful points for the determination of the Lombard slope that respected the indicated thresholds. Of the 17 audio tracks examined, at least 3 useful segments have been selected for all of them but one (ID35 with 2). At least 5 useful segments were selected for 8 of them, with a maximum of 11 for ID19. The values of L_{noise_mic} and L_{speech_mic} vary between 38.7 dBA and 49.4 dBA for the first and 48.0 dBA and 74.2 dBA for the second. For general noise and voice levels, instead, the values vary between 64.0 dBA and 74.9 dBA for L_{noise}, and between 74.6 dBA and 98.4 dBA for L_{speech}. The derived Lombard slope is $\boldsymbol{c} = 0.71$ (Fig. 4.10).

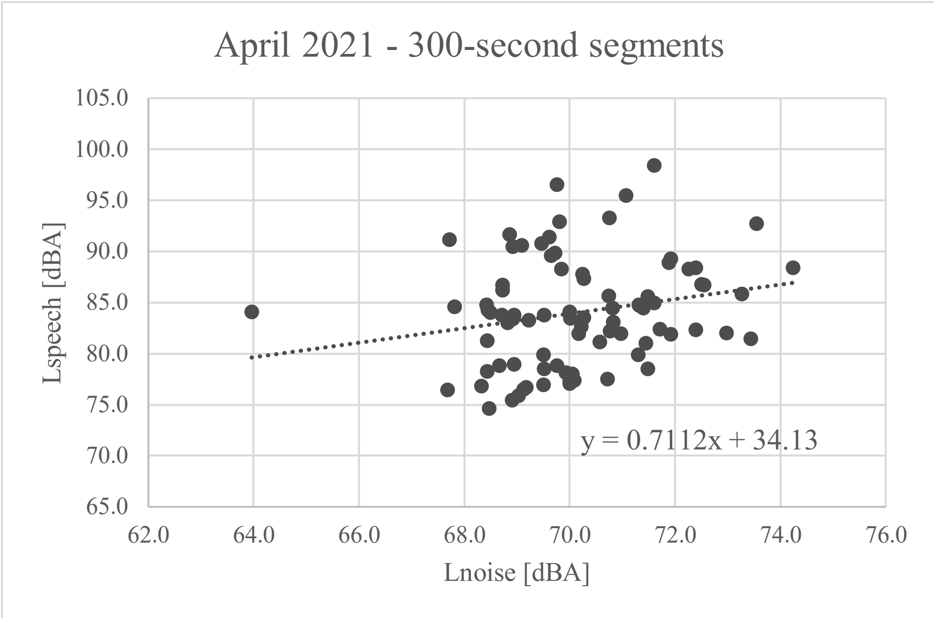


Figure 2.10 Lombard slope for April 2021, 300-second segments

5.4 LOMBARD SLOPE FOR ALL MEASUREMENTS

Below are the results obtained by combining all the data obtained from all the measurements, divided according to the length of the analysis interval.

5.4.1 60-second segments

With an analysis interval of 60 seconds in total 535 segments were selected. L_{noise} varies between 60.1 dBA and 74.7 dBA, while L_{speech} between 68.5 dBA and 101.7 dBA. The derived Lombard slope is $\boldsymbol{c} = 0.81$ (Fig. 4.11).

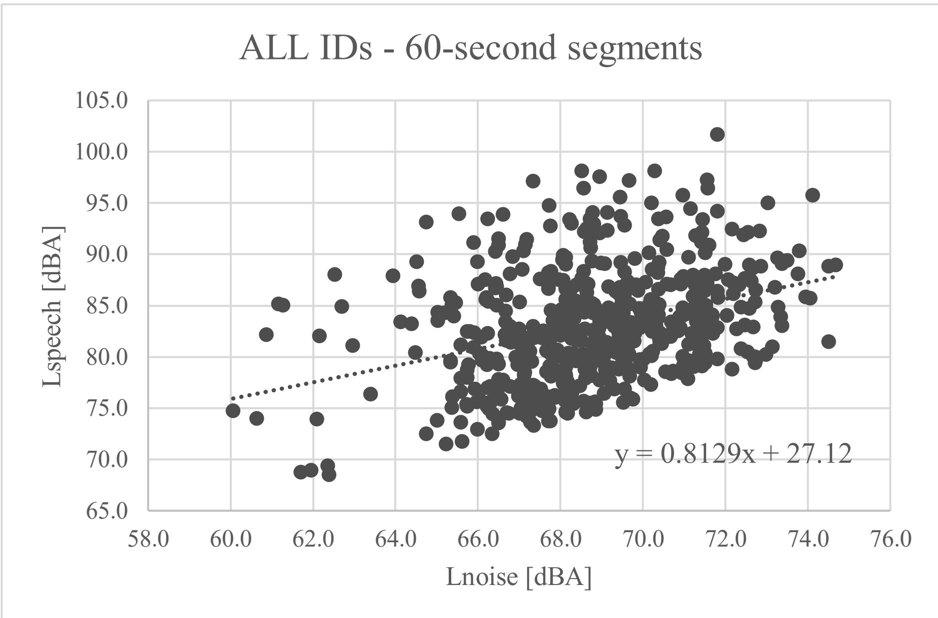


Figure 4.11 Lombard slope for all IDs, 60-second segments

5.4.2 120-second segments

With an analysis interval of 120 seconds in total 329 segments were selected. L_{noise} varies between 61.4 dBA and 75.0 dBA, while L_{speech} between 68.3 dBA and 99.4 dBA. The derived Lombard slope is $\boldsymbol{c} = 0.88$ (Fig. 4.12).

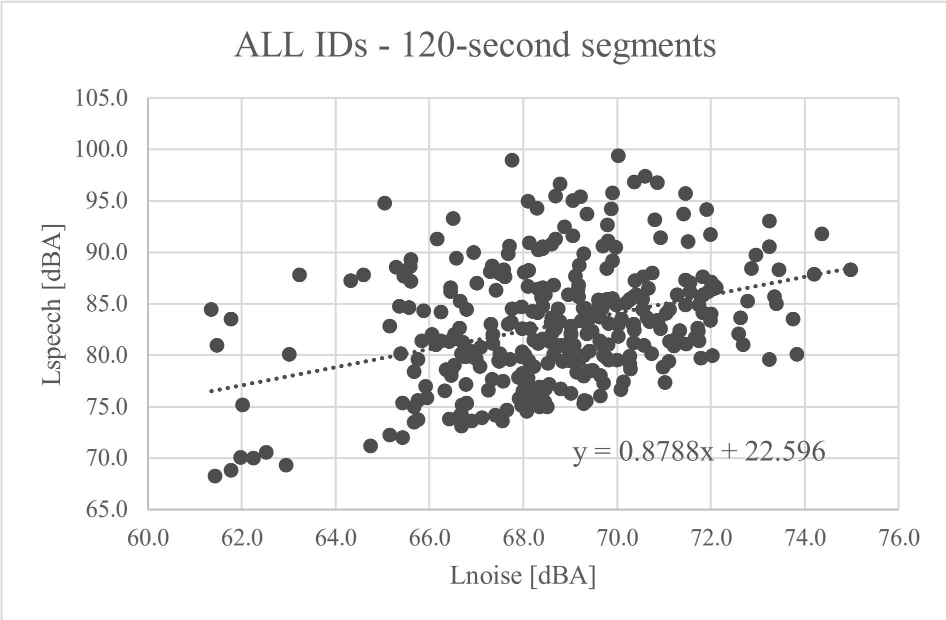


Figure 4.12 Lombard slope for all IDs, 120-second segments

5.4.3 180-second segments

With an analysis interval of 180 seconds in total 234 segments were selected. L_{noise} varies between 60.8 dBA and 74.6 dBA, while L_{speech} between 67.0 dBA and 100.5 dBA. The derived Lombard slope is $\boldsymbol{c} = 0.84$ (Fig. 4.13).

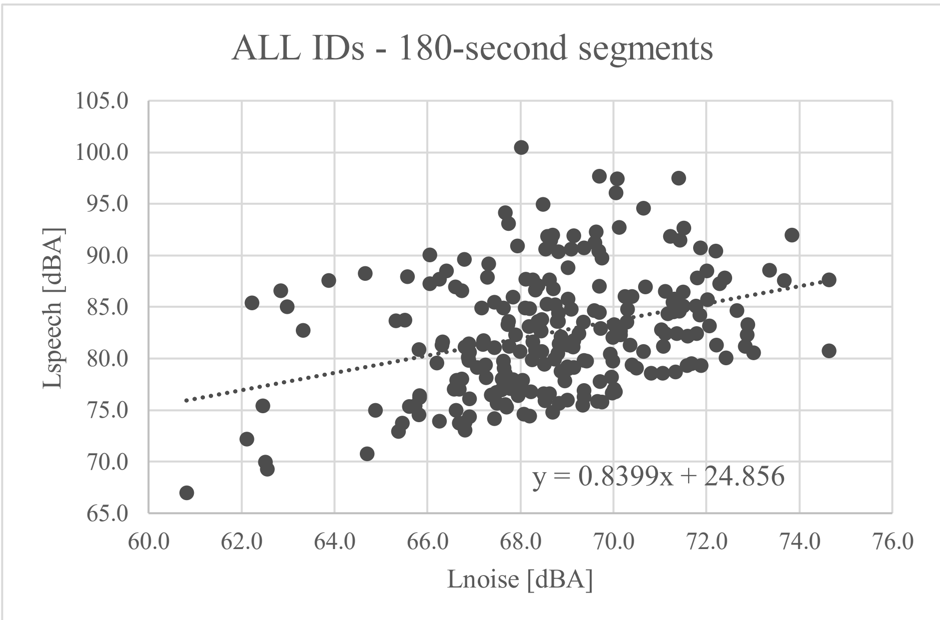


Figure 4.13 Lombard slope for all IDs, 180-second segments

5.4.4 240-second segments

With an analysis interval of 240 seconds in total 176 segments were selected. L_{noise} varies between 62.3 dBA and 74.9 dBA, while L_{speech} between 68.7 dBA and 97.6 dBA. The derived Lombard slope is $\boldsymbol{c} = 0.98$ (Fig. 4.14).

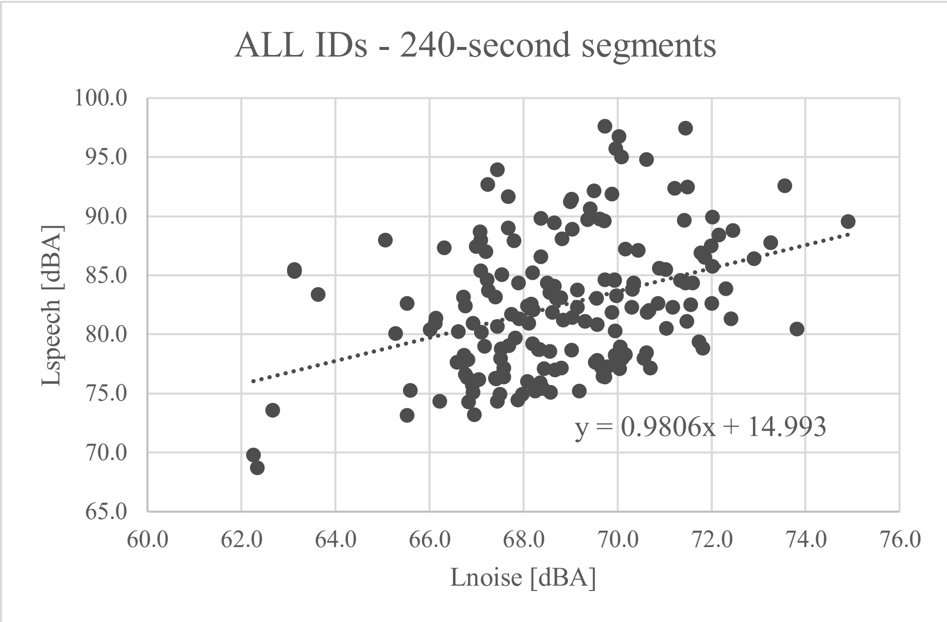


Figure 3.14 Lombard slope for all IDs, 240-second segments

5.4.5 300-second segments

With an analysis interval of 300 seconds in total 150 segments were selected. L_{noise} varies between 61.6 dBA and 74.2 dBA, while L_{speech} between 69.4 dBA and 98.4 dBA. The derived Lombard slope is $\boldsymbol{c} = 0.87$ (Fig. 4.15).

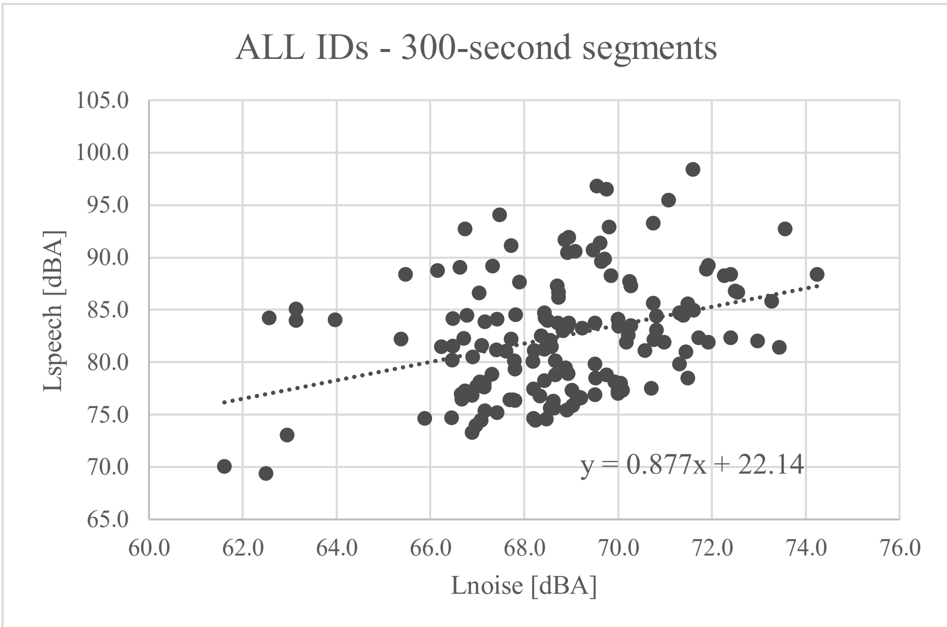


Figure 4.15 Lombard slope for all IDs, 300-second segments

Chapter 6

CONCLUSIONS

The aim of this work was to develop and implement a methodology for measuring and assessing the Lombard effect in an eating establishment. This phenomenon is described in the literature by rate **c**, which represents the slope of the relationship between noise level and voice level. Although it has been studied in many ways since its discovery in 1911, in most research its estimation has been done under controlled conditions such as laboratories, anechoic rooms, soundproof cabins etc. For this reason it was necessary to develop a methodology for on-field measurement, which obviously presents more difficult and challenging conditions. This process required essentially 3 stages:

1. Stage 1: Measurements on-field;
2. Stage 2: Method's general definition and preliminary study;
3. Stage 3: Automation and optimization process.

In the first stage, the actual measurements were made taking as a case study the canteen employees of the Polytechnic of Turin. Measurements were divided into two slots consisting of two days each: 25 and 26 November 2020 and 14 and 16 April 2021. Environmental noise monitoring devices and individual devices for voice monitoring of some participants in the experiment were used; the flow of people in and out of the canteen was also measured. Conditions were heavily affected by the presence of the COVID-19 pandemic, especially in the first slot in which the second wave of the virus was in full progress. Despite this, the presence of people in the cafeteria was significant with peaks of 140 people in November and 200 in April. 35 people participated in the vocal effort measurements. The second stage formed the basis of the methodology, identifying the operations to be carried out for the actual processing of data. The procedure was divided into:

- preliminary operations concerning signal cleaning, calculation of sound pressure levels and synchronization between all data;
- application of the Gaussian Mixture Model for the estimation of noise and speech levels;
- choice of parameters for the graphical representation of the Lombard slope;
- considerations on preliminary results.

This last point became necessary given the low relevance of the first results obtained in which the Lombard slope did not have values consonant with those found in literature. A selection of the most significant signal portions was then carried out, which led to an improvement in the results. The final step was therefore to define the statistical criteria to be respected in the selection and to automate the process, a procedure carried out in the third stage. By setting certain thresholds and implementing an algorithm that performed the selection, acceptable results have been obtained (Table 1.4).

Table 1.4 Final results 1

Interval Length	*November 2020*	*April 2021*	*ALL data*
60 s	0.57	0.62	0.81
120 s	0.76	0.45	0.88
180 s	0.77	0.24	0.84
240 s	0.81	0.67	0.98
300 s	0.57	0.71	0.87

The analysis were done by taking portions of the signal of 1, 2, 3, 4 and 5 minutes and putting together the selected segments first divided into the two slots and then all together. In all cases the values obtained are compatible with those found in the literature, although slightly higher than average. In particular, putting all the data together the Lombard slope seems to settle on a value of about 0.8 dBA/dBA: this would imply that a speaker increases his vocal effort by 0.8 dBA for each 1 dBA increase in the surrounding environmental noise. The value is quite high, but it could be due to the exceptional circumstance in which it was measured. The regulations implemented to combat the COVID-19 pandemic provide not only a reduction in the capacity of the room, but also the need to maintain a social distance of at least 1 m with others. This increase of the distances could have provoked a meaningful increase of the medium voice level of the people present during measurements: the presence of only 2-3 seats per table and the greater spacing did not, but on the contrary communications took place between on tables. Another reason for these results can also be given by the significant uncertainty with which the monitoring of vocal effort was conducted. The microphones Beta 54 used have a supercardioid polar pattern so they are strongly directional. Their high sensitivity becomes a double-edged weapon: the best performance can be achieved only if the microphone is not moved during measurement; even a small deviation can be significant. It was found that several participants moved the microphone during the meal, so a more accurate processing work was required. For these reasons, a comparison should be made both by taking measurements and reusing this procedure under normal conditions at the end of the pandemic, and by trying to change instrumentation by replacing supercardioid microphones with piezoelectric contact microphones.

Bibliography

[1] E. Lombard, “Le signe de l’élévation de la voix _The characteristics of the elevation of the voice_,” Annales des Maladies de l’Oreille, du Larynx, du Nez et du Pharynx, **37**, 101–119 _1911_.

[2] Brumm, Henrik, and Sue Anne Zollinger. "The evolution of the Lombard effect: 100 years of psychoacoustic research." Behaviour 148.11-13 (2011): 1173-1198.

[3] Lane, Harlan, and Bernard Tranel. "The Lombard sign and the role of hearing in speech." Journal of Speech and Hearing Research 14.4 (1971): 677-709.

[4] Hotchkin, Cara, and Susan Parks. "The Lombard effect and other noise-induced vocal modifications: insight from mammalian communication systems." Biological Reviews 88.4 (2013): 809-824.

[5] Korn, T. S. "Effect of psychological feedback on conversational noise reduction in rooms." The Journal of the Acoustical Society of America 26.5 (1954): 793-794.

[6] Pickett, J. M. "Limits of direct speech communication in noise." The Journal of the Acoustical Society of America 30.4 (1958): 278-281.

[7] Lazarus, H. "Prediction of verbal communication is noise—A review: Part 1." Applied Acoustics 19.6 (1986): 439-464.

[8] Dodd, George, and James Whitlock. "Auditory and behavioural mechanism influencing speech intelligibility in primary school children." Proceedings of the 18th International Congress on Acoustics. The Acoustical Society of Japan, 2004.

[9] Sato, Hiroshi, and John S. Bradley. "Evaluation of acoustical conditions for speech communication in active elementary school classrooms." Proceedings of ICA, Kyoto (2004).

[10] Sutherland, Louis, David Lubman, and Karl Pearsons. "Acoustic environment challenges for the unique communication conditions in group learning classes in elementary school classrooms." The Journal of the Acoustical Society of America 117.4 (2005): 2366-2366.

[11] Bottalico, Pasquale, et al. "Evaluation of the starting point of the Lombard effect." Acta Acustica United With Acustica 103.1 (2017): 169-172.

[12] Bottalico, Pasquale. "Lombard effect, ambient noise, and willingness to spend time and money in a restaurant." The Journal of the Acoustical Society of America 144.3 (2018): EL209-EL214.

[13] A. White, “The effect of the building environment on occupants: the acoustics of dining spaces,” M.Phil. dissertation, University of Cambridge, _1999_.

[14] Astolfi, Arianna, and Marco Filippi. "Good acoustical quality in restaurants: a compromise between speech intelligibility and privacy." Proc. of ICA. 2004.

[15] Gardner, Mark B. "Factors affecting individual and group levels in verbal communication." Journal of the Audio Engineering Society 19.7 (1971): 560-569.

[16] Tang, Shiu Keung, Daniel WT Chan, and K. C. Chan. "Prediction of sound-pressure level in an occupied enclosure." The Journal of the Acoustical Society of America 101.5 (1997): 2990-2993.

[17] Kang, Jian. "Numerical modelling of the speech intelligibility in dining spaces." Applied Acoustics 63.12 (2002): 1315-1333.

[18] Navarro, Manuelina Porto Nunes, and Roberto Leal Pimentel. "Speech interference in food courts of shopping centres." Applied acoustics 68.3 (2007): 364-375.

[19] Hodgson, Murray, Gavin Steininger, and Zohreh Razavi. "Measurement and prediction of speech and noise levels and the Lombard effect in eating establishments." The Journal of the Acoustical Society of America 121.4 (2007): 2023-2033.

[20] Rindel, Jens Holger. "Verbal communication and noise in eating establishments." Applied Acoustics 71.12 (2010): 1156-1161.

[21] D'Orazio, Dario, Federico Montoschi, and Massimo Garai. "Acoustic comfort in highly attended museums: A dynamical model." Building and Environment 183 (2020): 107176.

[22] Rossi, E., D'Orazio, D., Fratoni, G., & Garai, M. (2020, December). Improving the acoustic quality of a food court by using rate spatial decay. In Forum Acusticum (pp. 3053-3054).

[23] ISO 9921. Ergonomics – assessment of speech communication. Geneva; 2003.

[24] http://www.sitiarcheologici.palazzochigi.it/www.governo.it/febbraio%202021/it/coronavirus-misure-del-governo.html

[25] https://www.nti-audio.com/Portals/0/data/en/XL2-Specifications.pdf

[26] https://www.shure.com/en-MEA/products/microphones/beta_54

[27] https://zoomcorp.com/media/documents/E_H4n_Pro_REV3.pdf

[28] AES-6id-2006 (s2017): AES information document for digital audio - Personal computer audio quality measurements

[29] D'Orazio, D., De Salvio, D., Anderlucci, L., and Garai, M. (2020). "Measuring the speech level and the student activity in lecture halls: Visual- vs blind-segmentation methods," Appl. Acoust. 169, 107448.

[30] Puglisi, G. E., Astolfi, A., Cantor Cutiva, L. C., & Carullo, A. (2015, May). Assessment of indoor ambient noise level in school classrooms. In Proceedings of the Conference on Noise Control–EuroNoise2015.

Appendix A

FOGLIO INFORMATIVO

Misura dello sforzo vocale in presenza di rumore antropico

Gentile interessato/a,
vogliamo proporti di partecipare ad una ricerca e, al fine di informarti circa lo scopo e le caratteristiche dello studio affinché tu possa decidere in modo consapevole e libero se partecipare, ti invitiamo a leggere attentamente quanto riportato di seguito. I ricercatori coinvolti in questo progetto sono a disposizione per rispondere alle sue domande.

Titolare del trattamento dei dati	
Politecnico di Torino	Dati di contatto: PEC: politecnicoditorino@pec.polito.it Per informazioni e chiarimenti: privacy@polito.it
Responsabile del trattamento dei dati (*eventuale*) ex art.28 GDPR	
Data Protection Officer d'Ateneo	Dati di contatto: avv. Nicoletta Roz Gastaldi PEC: dpo@pec.polito.it Mail: dpo@polito.it
Responsabile scientifico dello studio:	Prof.ssa Arianna Astolfi

<u>Qual è lo scopo di questo studio?</u>

Lo scopo del presente studio è di valutare il livello di sforzo vocale e l'effetto Lombard in presenza di rumore antropico da chiacchiericcio all'interno della mensa del Politecnico.

Come si svolgerà lo studio?

Indosserai un microfono a guancia e un registratore DAT durante il pasto nel quale parlerai con altri commensali. Dal file .wav registrato otterremo dei livelli sonori dai quali calcoleremo dei parametri statistici connessi al tempo di misura.

Per quale ragione ti proponiamo di partecipare?

Per capire se in base al numero di persone che parlano e al rumore antropico registrato il tuo sforzo vocale, rappresentato dal livello della tua voce ad una distanza fissa dalla bocca, aumenterà e di quanto.

Sei obbligato/a partecipare allo studio?

La tua partecipazione è completamente libera, il rifiuto di partecipare non comporterà alcuna conseguenza negativa. Inoltre, se dovessi cambiare idea e volessi ritirarti, in qualsiasi momento sei libero/a di farlo senza dover fornire alcuna spiegazione.

In caso di ritiro i dati precedentemente acquisiti saranno comunque impiegati.

Quali sono i passaggi necessari per partecipare allo studio?

La partecipazione allo studio avviene previa dettagliata informazione sulle caratteristiche, sui rischi e benefici dello stesso. Al termine della fase informativa potrai acconsentire alla partecipazione allo studio firmando il modulo di consenso informato. Solo dopo che avrai espresso per iscritto il tuo consenso, potrai attivamente partecipare allo studio proposto.

Che cosa ti verrà chiesto di fare?

La procedura sperimentale/il progetto di ricerca prevede l'ingresso nella mensa del Politecnico di Torino, dove ti verrà fatto indossare un microfono a guancia. Di seguito potrai prendere il cibo al banco e potrai sederti al tavolo parlando con i tuoi commensali. La durata complessiva dell'esperimento corrisponde al tempo necessario a pranzare in mensa. Restituirai il microfono a guancia al termine del pranzo.

Quali sono i possibili rischi ed i disagi dello studio?

Non vi sono rischi noti rispetto allo studio.

Quali sono i possibili benefici derivanti dallo studio?

Lo studio non comporta diretti benefici per il partecipante. Tuttavia, lo studio consentirà di incrementare le conoscenze nell'ambito dello sforzo vocale del parlato in ambiente competitivo.

Cosa accadrà se nel corso dello studio emergessero informazioni che riguardano la tua salute?

Qualora emergessero dallo studio informazioni potenzialmente utili per la tua salute, potrai esprimere la scelta di essere informato/a o meno, nella sezione “Espressione di consenso informato”.

<u>Come viene garantita la riservatezza delle informazioni/dati/campioni?</u>

Lo sperimentatore ti chiederà di fornire alcuni dati personali, quali nome, cognome, data di nascita, età, genere, disturbi uditivi.

Queste informazioni così come i dati che emergeranno nel corso della ricerca sono importanti per il corretto svolgimento dello studio. La riservatezza di tutte le informazioni sarà garantita secondo la normativa vigente (Regolamento europeo UE 2016/679 concernente la tutela delle persone fisiche con riguardo al trattamento dei dati personali e la libertà di circolazione di tali dati - https://www.garanteprivacy.it/regolamentoue)

I dati verranno utilizzati per la ricerca in oggetto o per altre affini.

<u>Come saranno usati i tuoi dati personali?</u>

I dati verranno raccolti e conservati in forma anonima. Ad ogni soggetto verrà assegnato un codice alfanumerico e la corrispondenza codice nome utente sarà in un file conservato nel computer del responsabile della ricerca.
I risultati verranno presentati in una relazione a fine progetto in forma anonima e/o aggregata.

<u>Che ne sarà dei tuoi dati alla fine della ricerca?</u>
I dati verranno conservati alla fine della ricerca per poterli riutilizzare in ricerche successive. Il file con la corrispondenza nome codice invece verrà distrutta.

<u>I tuoi dati potranno essere ceduti a terzi?</u>
I dati relativi alla ricerca potranno essere comunicati in forma anonima e/o aggregata ad altri soggetti, quali, ad esempio, altre Università, istituzioni e organismi pubblici e privati aventi finalità di ricerca, limitatamente ad informazioni prive di dati identificativi e per scopi storici o scientifici.

<u>Altre informazioni importanti</u>
Ti informiamo anche che questo studio sarà stato sottoposto alla valutazione del Comitato Etico del Politecnico di Torino, di prossimo insediamento.
L’originale del Consenso informato scritto da te firmato verrà conservato dal responsabile del presente studio, mentre tu hai diritto a riceverne una copia.
Durante lo studio, potrai contattare il responsabile dello studio per qualsiasi informazione.

LA RINGRAZIAMO PER LA DISPONIBILITÀ

DICHIARAZIONE DEL RESPONSABILE DELLO STUDIO

Dichiaro di aver fornito alla/al partecipante informazioni complete e spiegazioni dettagliate circa la natura, le finalità, le procedure e la durata di questo progetto di ricerca. Dichiaro, inoltre, di aver fornito al/alla partecipante il foglio informativo.

Arianna Astolfi

Data

25 novembre 2020

ARIANNA ASTOLFI

FIRMA INFORMATIVA

Dichiaro di aver ricevuto informazioni che mi hanno permesso di comprendere il progetto di ricerca, anche alla luce degli ulteriori chiarimenti da me richiesti. Confermo che mi è stata consegnata copia del presente foglio informativo.

FIRMA Data

ESPRESSIONE DI CONSENSO INFORMATO

Io sottoscritto/a ____

DICHIARO

- di aver ricevuto spiegazioni esaurienti in merito alla richiesta di partecipazione allo studio sperimentale in oggetto e sufficienti informazioni riguardo ai rischi e ai benefici implicati nello studio, secondo quanto riportato nel foglio informativo qui allegato.
- di aver potuto discutere tali spiegazioni, di aver potuto porre tutte le domande che ho ritenuto necessarie e di aver ricevuto in merito risposte soddisfacenti;
- di manlevare il Politecnico di Torino da qualsiasi tipo di responsabilità non direttamente imputabile allo stesso per qualunque danno o lesione possa derivare a me stesso/a in occasione della mia partecipazione;
- di essere stato, inoltre, informato del mio diritto di ritirarmi in qualsiasi momento dalla ricerca stessa.

Alla luce delle informazioni che mi sono state fornite, pertanto:

☐	ACCONSENTO	☐	NON ACCONSENTO	a partecipare allo studio
☐	ACCONSENTO	☐	NON ACCONSENTO	ad essere informata/o su eventuali risultati utili alla mia salute derivanti dallo studio stesso *(se pertinente)*

LUOGO DATA

FIRMA DEL PARTECIPANTE

LUOGO DATA

FIRMA DEL RESPONSABILE DELLO STUDIO

INFORMATIVA AI SENSI DELL'ART. 13 DEL REGOLAMENTO GENERALE SULLA PROTEZIONE DEI DATI UE 679/2016 PER LA PARTECIPAZIONE ALLO STUDIO

Gentile interessato/a, come previsto dal Regolamento Generale sulla protezione dei dati (Regolamento EU 2016/679 – noto anche come "GDPR") ti forniamo le seguenti informazioni che riguardano il trattamento dei tuoi dati personali.

DATI DI CONTATTO

Titolare del trattamento dei dati è il Politecnico di Torino, con sede in Corso Duca degli Abruzzi, n. 24, 10129 – Torino, nella persona del Rettore.

Il dato di contatto del Titolare è PEC: politecnicoditorino@pec.polito.it

Per ulteriori informazioni e chiarimenti: privacy@polito.it

Il **Responsabile della protezione dati** ("**DPO**") del Politecnico di Torino, al quale gli interessati possono rivolgersi per questioni relative al trattamento dei loro dati personali e all'esercizio dei loro diritti, è contattabile ai seguenti indirizzi: dpo@polito.it ; PEC: dpo@pec.polito.it

PRINCIPI, BASE GIURIDICA E FINALITA' DEL TRATTAMENTO

Nel rispetto dei principi di liceità, correttezza, trasparenza, adeguatezza, pertinenza e necessità di cui all'art. 5, paragrafo 1, del GDPR, il Politecnico di Torino, in qualità di Titolare, provvederà al trattamento dei tuoi dati personali ai sensi dell'art. 6, paragrafo 1, lettere a), ("*l'interessato ha espresso il consenso al trattamento dei propri dati personali per una o più specifiche finalità*") ed e) ("*il trattamento è necessario per l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri di cui è investito il titolare*") nel perseguimento delle finalità istituzionali connesse al progresso nella ricerca scientifica come previsto dallo Statuto di Ateneo e, con riferimento ai dati personali appartenenti alle categorie particolari di cui all'art. 9 del GDPR, previo tuo consenso ai sensi dell'art. 9, paragrafo 2, lettera a) ("*l'interessato ha prestato il proprio consenso esplicito al trattamento di tali dati personali per una o più finalità specifiche* […]").

In particolare, i tuoi dati personali saranno trattati dal Politecnico di Torino, con modalità cartacea e/o informatizzata, per condurre uno studio in ambito del monitoraggio dello sforzo vocale nelle mense.

TRASFERIMENTO DEI DATI

I dati trattati per le finalità di cui sopra verranno comunicati, o saranno comunque accessibili, ai dipendenti e collaboratori assegnati ai competenti uffici del Politecnico di Torino, che saranno adeguatamente istruiti dal Titolare.

I dati relativi alla ricerca potranno essere comunicati in forma anonima e/o aggregata ad altri soggetti, quali, ad esempio, altre Università, istituzioni e organismi pubblici e privati aventi finalità di ricerca, limitatamente ad informazioni prive di dati identificativi e per scopi storici o scientifici.

La divulgazione dei risultati (ad esempio mediante pubblicazione di articoli scientifici, partecipazione a convegni, redazione di tesi ecc.) avverrà soltanto con modalità che rendano difficile la tua identificazione.

La gestione e la conservazione dei dati personali raccolti dal Politecnico di Torino avviene su sistemi ubicati all'interno dell'Ateneo e/o esterni di fornitori di alcuni servizi necessari alla gestione tecnico – amministrativa che, ai soli fini della prestazione richiesta, potrebbero venire a conoscenza dei dati personali degli interessati e che saranno debitamente nominati come Responsabili del trattamento a norma dell'art. 28 del GDPR.

I dati raccolti non saranno oggetto di trasferimento verso un Paese non appartenente all'Unione Europea (c.d. Paese terzo).

PERIODO DI CONSERVAZIONE DEI DATI

I dati personali inerenti il trattamento, compresi quelli appartenenti alle categorie particolari, saranno conservati per il tempo necessario allo Studio.

CONFERIMENTO DEI DATI

Il conferimento dei dati personali è obbligatorio, l'eventuale rifiuto comporta l'impossibilità di partecipare allo Studio.

DIRITTI DELL'INTERESSATO

In qualità di interessato hai diritto di chiedere al Titolare del trattamento, conformemente agli artt. 15 e ss. del GDPR,

- l'accesso ai propri dati personali ed a tutte le informazioni di cui all'art. 15 del GDPR;
- la rettifica dei propri dati personali inesatti e l'integrazione di quelli incompleti;
- la cancellazione dei propri dati, fatta eccezione per quelli contenuti in atti che devono essere obbligatoriamente conservati dall'Ateneo, e salvo che sussista un motivo legittimo prevalente per procedere al trattamento;
- la limitazione del trattamento nelle ipotesi di cui all'art. 18 del GDPR.

Hai, altresì, il diritto:

- di opporti al trattamento dei dati personali, fermo quanto previsto con riguardo alla necessità ed obbligatorietà del trattamento dati per poter fruire dei servizi offerti;
- di revocare il consenso eventualmente prestato per i trattamenti non obbligatori dei dati, senza con ciò pregiudicare la liceità del trattamento basata sul consenso prestato prima della revoca;
- alla portabilità dei dati.

Se desideri esercitare qualsiasi dei tuoi diritti, puoi rivolgerti al Titolare del trattamento.

RECLAMO

Hai il diritto di rivolgerti al Garante per la protezione dei dati personali secondo le modalità indicate al seguente link: https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/4535524

La presente informativa è aggiornata al 05.03.2020

ESPRESSIONE DEL CONSENSO PER IL TRATTAMENTO DEI DATI APPARTENENTI ALLE CATEGORIE PARTICOLARI DI DATI PERSONALI e UTILIZZO IMMAGINI

Alla luce di quanto premesso e delle informazioni che mi sono state fornite:

Io sottoscritta/o ____

DICHIARO

- di aver preso visione dell'informativa sul trattamento dati personali resa ai sensi dell'art. 13 del Regolamento EU 679/2016 (noto come "GDPR");
- di:

☐	ACCONSENTIRE	☐	NON ACCONSENTIRE	al trattamento dei propri dati personali
☐	ACCONSENTIRE	☐	NON ACCONSENTIRE	al trattamento dei propri dati personali appartenenti alle categorie particolari di cui all'art. 9 del GDPR (*se pertinente*)

☐	AUTORIZZARE	☐	NON AUTORIZZARE	il Politecnico di Torino, ai sensi degli artt. 96 e 97 della Legge in materia di protezione del diritto d'autore e di altri diritti connessi al suo esercizio n. 633 del 22 aprile 1941 nonché dell'art. 10 codice civile, a: - esercitare i diritti previsti dagli artt. 12 e ss. della Legge n. 633/1941; - riprodurre, a titolo gratuito, unicamente per le finalità connesse alla ricerca, anche di carattere divulgativo, le immagini e le videoriprese che lo ritraggono acquisite durante la partecipazione all'attività organizzata; - di rinunciare a qualunque corrispettivo per la posa, l'utilizzo, la riproduzione e la diffusione delle immagini. ***per poter partecipare è necessario fornire l'autorizzazione***

Il/la sottoscritto/a vieta altresì l'uso delle immagini in contesti che ne pregiudichino la dignità personale ed il decoro.

LUOGO DATA

FIRMA DEL PARTECIPANTE

LUOGO DATA

FIRMA DEL RESPONSABILE DELLO STUDIO

IL RUMORE NEGLI AMBIENTI RISTORATIVI

Data: __________ Ora: ____________

Area mensa: ____

Codice soggetto: ______________

Finalità della ricerca è la determinazione della qualità acustica degli ambienti ristorativi, in particolare la misura in cui il livello di rumore e la riverberazione sonora interferiscano sullo sforzo vocale.

Le chiediamo gentilmente di rispondere a queste domande. I dati derivanti dal questionario e dalle misurazioni di sforzo vocale saranno aggregati anonimamente utilizzando codici univoci.

Grazie per la gentile collaborazione.

INFORMAZIONI DI CARATTERE GENERALE

(1) Sesso:

- o F
- o M

(2) Età: ________

(3) È fumatrice/fumatore:

- o Sì
- o No

(4) Ha mai sofferto o soffre di ipoacusia (diminuzione dell'udito causata da differenti fattori quali età, ereditarietà, rumori, infezioni etc.):

- o Sì
- o No

(5) Ha mai sofferto o soffre di patologie vocali:

- o Sì
- o No

(6) Numero di commensali presenti al tavolo: ________