Politecnico di Torino

Master’s degree in Biomedical Engineering
A.y. 2020/2021
Graduation session of December 2021

Speech intelligibility of speakers with masks

How facial masks affect communication in children with cochlear implant

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

Candidate:
Michele Di Iulio

ABSTRACT

The thesis work focuses on the assessment of speech intelligibility of children with cochlear implants in real complex acoustic scenarios, which include both competitive sound environments (i.e., with reverberation and noise) and the use of face masks that modify speech intelligibility towards a listener. The research is inspired by the recent events related to the global pandemic, as well as by the need of ensuring optimal acoustic scenarios for the support of auditory abilities especially for hearing impaired listeners. On the one side, the dramatic spread of the Coronavirus disease made it necessary face masks in public and private spaces. Although the use of face masks is needed to prevent the spreading of Coronavirus, several studies have already shown their negative effects on speech intelligibility (SI) and comprehension due to the coverings' acoustic attenuation. On the other side, research is now focusing on the understanding of the extent to which realistic acoustic scenarios, that account for the combined presence of noise with informative content and reverberation time, affect speech intelligibility. Besides behavioural (i.e., the use of face masks) and environmental (i.e., complex acoustic scenarios) issues, there are only few studies that consider them in relation to children with hearing impairments, and particularly wearing cochlear implants (CI). These subjects have special acoustic needs that are difficult to be met in everyday life and would need the built environment to support their hearing abilities. This research thus followed the open questions that arouse from the abovementioned baseline. Children between 7 and 15 years of age have been involved, who were divided into an experimental group of 14 children with CI, and a control group of 6 normal hearing children and were administrated via web-platform a SI test. Different acoustic conditions were evaluated in terms of noise (i.e., no noise, signal-to noise ratio of 0 dB, +5 dB, +10 dB) and considering three types of masks that differed for intrinsic characteristics (i.e., acoustic attenuation, filtration efficiency and breathability). SI was evaluated in terms of percentage of items correctly understood in relation to the speech material used in the test (which is the one of the Simplified Matrix Sentence Test for the Italian language), and the listening difficulty was also evaluated based on the subjective rating given by the subjects on a 5-points scale. Results showed that SI was higher in the condition without a mask and varied depending on the type of mask used. More challenging conditions, e.g., with a lower thus more competitive signal-to-noise ratio corresponded to lower SI as expected. Results on SI related to the case of no-mask were similar to those obtained with the mask that had the lowest acoustic attenuation, although characterized by a low filtration efficiency. The optimal match between intrinsic features of masks thus needs to be investigated with respect to the consequences on SI and listening difficulty. The preliminary results discussed in this work suggest the need of guaranteeing optimal acoustic

conditions, especially in terms of high signal-to-noise ratios, and of selecting the most appropriate face masks for teachers to be understood by children wearing CIs. These outcomes can be transferred to improve speech communication (that is, speech production and speech understanding) in classrooms for everyday design and usage practices to be implemented to support optimal listening conditions.

CONTENENTS

FIGURE INDEX

TABLE INDEX

INTRODUCTION

With the diffusion of SARS-CoV-2, better known as COVID-19, the population's habits and the lifestyle changed dramatically. The virus appeared for the first time at the beginning of 2020 and the medic science classified it as a whole new type of disease. It is a viral form belonging to the strain of coronavirus which, until now, had not yet been identified in the human organism. As medical researchers have been affected by the lack of knowledge and specific information about this phenomenon, it appears to be quite difficult to define a good therapy. Furthermore, as the first events have demonstrated, this disease can cause serious health problems or even death. Therefore, prevention is vital in order to restrain the contagion and further diffusion of Covid. Virus transmission accouris through airborne transmission or contact, therefore the precautionary measures require social distancing, limitation of travel and the use of Individual Protection Devices (DPI). In short time, face masks have become a common and indispensable part of our everyday life. Several studies have already shown its efficiency in slowing down the infection or in protecting hospital staff. There are several models of masks currently available on the market that are different for composition, attenuation, and filtering capacity. However, facial mask not only reduce the transmission of aeresol particles through a filtering operation, but it also affects the communication itself. In fact, the mask represents an obvious obstacle to verbal understanding, and it has a considerable impact on body language, facial expression and lip reading as well, all useful strategies in everyday communication, for people affected by hearing loss. It is therefore very important to analyse how and to what extend the use of the mask affects communication, social relations and the learning process [1].

To evaluate speech intelligibility, several authors have implemented different audiometric tests (matrix test) based on the combination of different words and sentences. The first speech test, proposed by Hagerman in 1982 on the Swedish language, immediately highlighted some critical issues related to this method. In general, any audiometric test is optimized for the language for which it was carried out and therefore its applicability is limited, or in some cases unsuitable, for languages that are not the original one. This aspect is the main topic of the studies conducted by Kollmeier in 2015, through which a new version of the matrix-test was developed; a version that allows to obtain reliable information on intelligibility efficiently and also to make a comparison with other languages [2]. Later, Kollmeier's results have been adapted and processed to build a matrix-test for the Italian language (ITAMatrix) to study the speech intelligibility with noise [3]. The structure of the ITAMatrix test is made by 50 words from which randomized, unpredictable and phonetically balanced, five-word sentences are built. Experimentation and validation on have demonstrated the accuracy, reliability and validity of this method adult subjects. However, this procedure is not always suitable, five-word sentences are too long and can be difficult for children or adult with reduced memory skills. For this reason, Puglisi et al. developed a simplified model with three-word sentences to ease the test and make it suitable for every situation. Speech material, which is needed for the test, is collect according to psychometric theory in order to obtain an evaluation as objective as possible and independent of the age, knowledge and language skills of the subject [4].

This audiometric test is going to be used in this study for the assessment of speech intelligibility in children hearing impaired and with cochlear implant (CI). This research project aims to understand how facial masks affects the working of the CI system and, consequently, the communication. Different acoustic conditions will be tested, considering 3 types of masks, in the absence or in the presence of noise. Children are able to run the test anonymously and remotely at home using a specific platform under the supervision of a parent. The results related to this survey are particularly useful for the development of possible future applications and to understand how to reduce the negative effects related to the mask, especially in environments, such as schools or universities, where communication is the basis of the learning process.

1. FUNCTIONAL ANATOMY OF THE EAR

1.1 AUDITORY SYSTEM

The ear is the main component of the auditory system which, together with other sensory organs, plays an essential role in sound perception and in maintaining the static and dynamic balance of the body. The human ear is located in the temporal region, and it is topographically divided into three compartments: **middle ear**, **outer ear**, and **inner ear** (Fig.1).

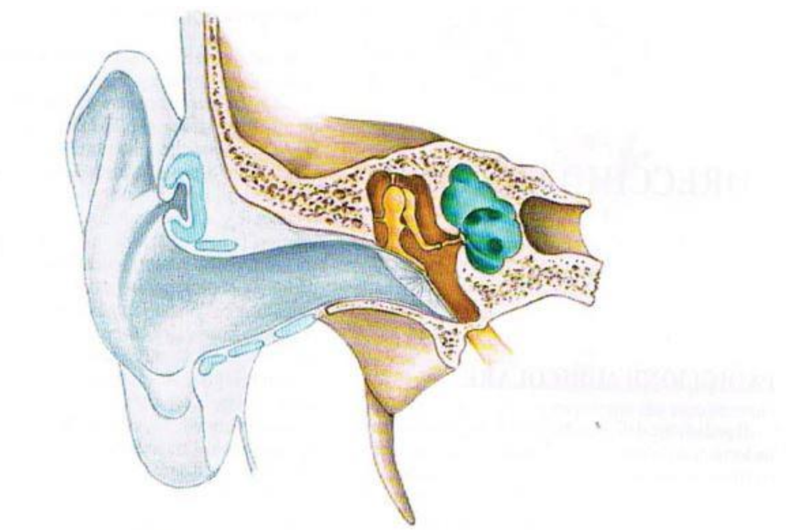


Figure 1. Organization of the auditory system: outer ear (blue), middle ear (yellow), middle ear (green) [5].

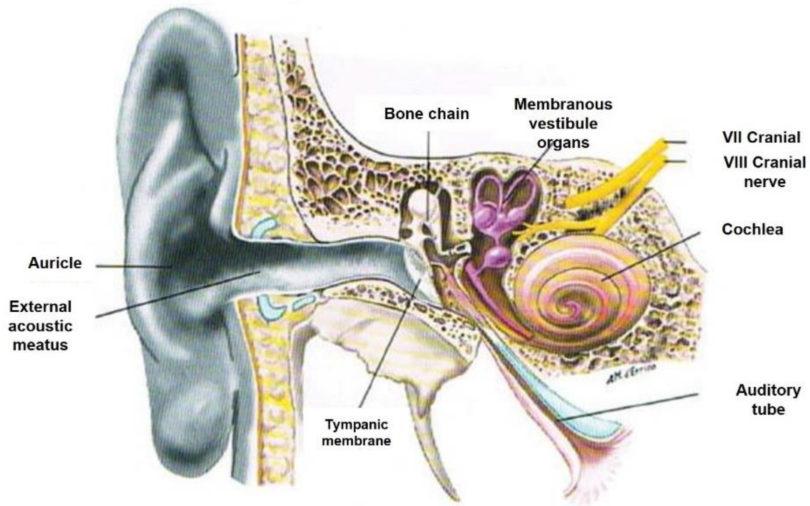


Figure 2. Main components of the human ear [5].

The outer and middle ears manage auditory sensitivity through specific acoustic receptors. The inner ear, on the other hand, has acoustic receptors in the trochlear duc and statokinetic receptors for the control of balance in the vestibular organs and semimembranous canals (Fig.2) [5].

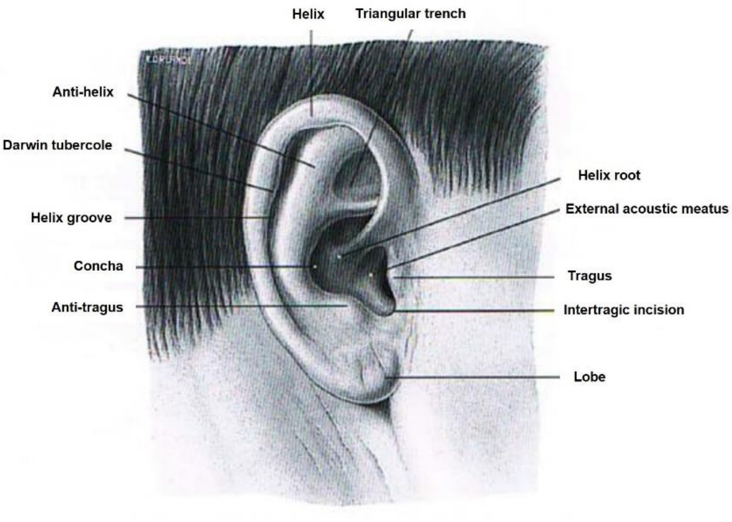


Figure 3. Lateral view of the outer ear [5]

The outer ear (Fig. 3), consisting of the *auricle* and the *acoustic meatus*, collects sound waves and directs them toward the eardrum membrane. From a functional point of view, the musculature of the outer ear is useful for spatial localization of the sound source, but in humans, unlike many other animals, it is of considerably reduced importance [5].

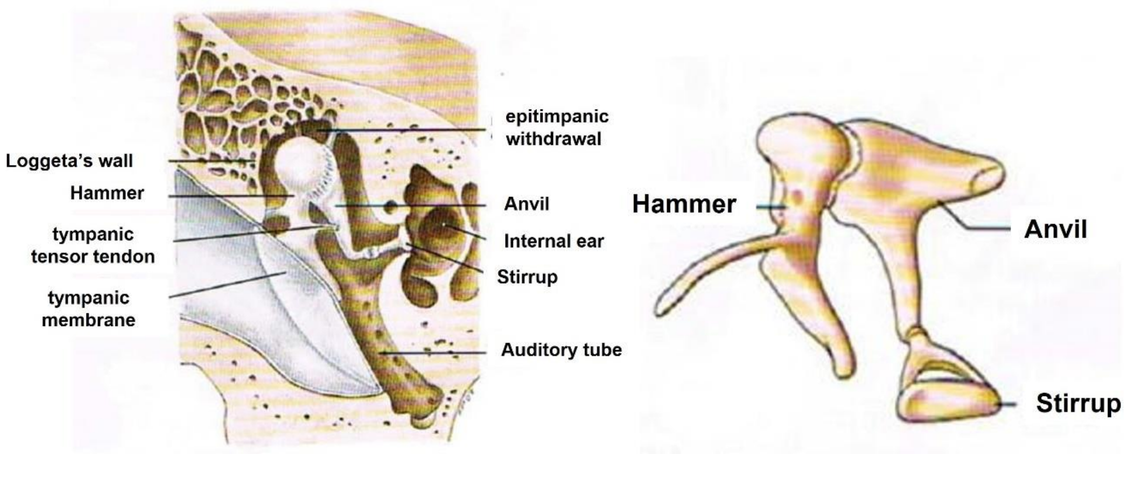


Figure 4. Frontal section of the middle ear (left); auditory ossicle chain (right) [5].

The middle ear (Fig. 4) is formed by the *tympanic cord*, *tympanic membrane*, *chain of auditory ossicles*, *mastoid apparatus*, and *auditory tube*. It allows acoustic pressure waves to be converted into mechanical vibrations. In particular, the auditory ossicles, represented by *hammer*, *anvil* and *stirrup*, are essential for the transmission of sounds from the outer ear to the inner ear; the vibration of the tympanic membrane is transmitted in sequence to the hammer, the anvil and finally to the stirrup. However, sound stimuli can also be due to a simple movement of the air contained in the tympanic cavity or the bone structure surrounding the cochlea.

The inner ear, which consists primarily of the cochlea and the vestibular apparatus, has the task of transducing sounds into nerve impulses. From a structural point of view, it is possible to distinguish two components: *bony labyrinth* and *membranous labyrinth*. The first one, represented in Figure 5, is characterized by a series of bony channels with semicircular shape, oriented according to the three directions of space and, therefore, very important for balance and walking; in fact, thanks to specific receptors, these structures are sensitive to angular accelerations due to rotational movements of the head.

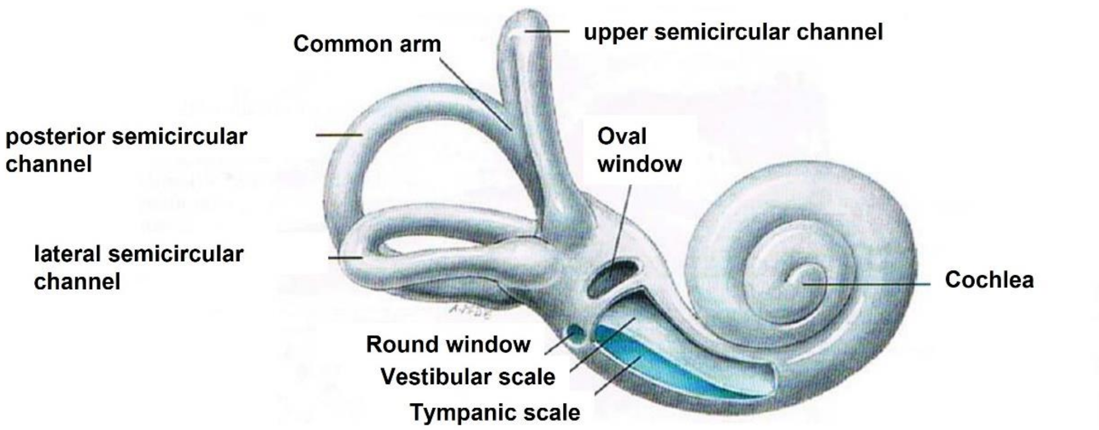


Figure 5. Bony labyrinth observed externally [5].

The cochlea belongs to the membranous labyrinth, and it is a channel spirally wound around a bony axis with of conical shape called the *modiole*. In the internal part there is a membrane, called *basilar membrane*, covered by some cylindrical cells that form the so-called *Organ of the Courts* (Fig. 6), responsible for nerve transmission.

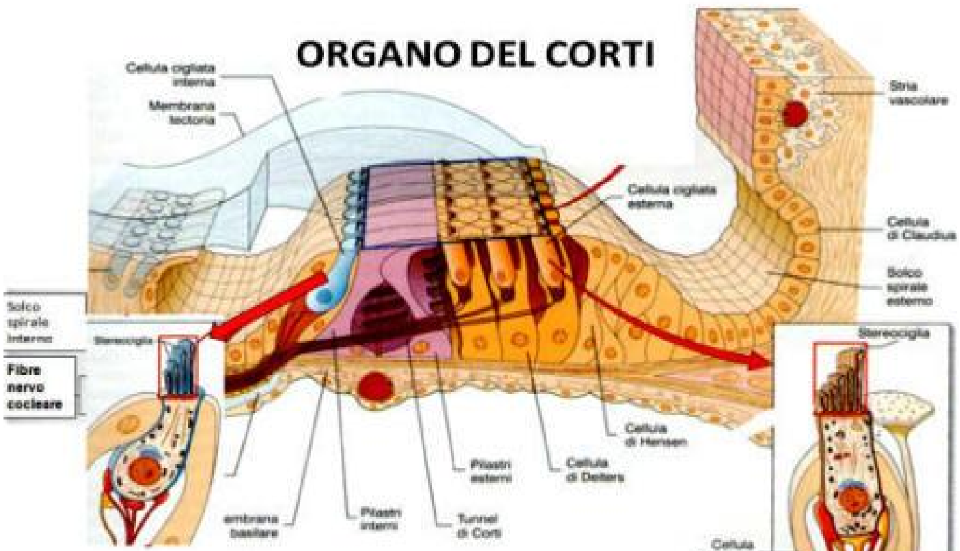


Figure 6. Organ *of the* Courts *[7].*

It has two types of cells:

- *supportive* which give support to the structure.
- *sensory* or *acoustic* wich are hair cells in direct contact with the nerve cells of the vestibulocochlear nerve; they have different mechanical characteristics, and they are positioned in different regions of the organ. The mechanism of movement of these cells, at the bottom of the phenomenon of "*Tonotipicity of the cochlea*", depends on the frequency of the sound stimulus; each fiber of the acoustic nerve, in fact, has its own resonant frequency corresponding to the frequency of response of the hair cell associated with it [6].

1.2 THEORIES OF HEARING

The structure of the human ear is designed to receive airborne acoustic energy in a frequency range of about 16 Hz to about 20,000 Hz and for a sound pressure range of more than 120 dB. Currently there are two theories related to the principle of operation of the ear, "The Place Theory" and "The Temporal Theory", but neither of these is able to describe all the possible stimuli and physical phenomena associated with the auditory mechanism.

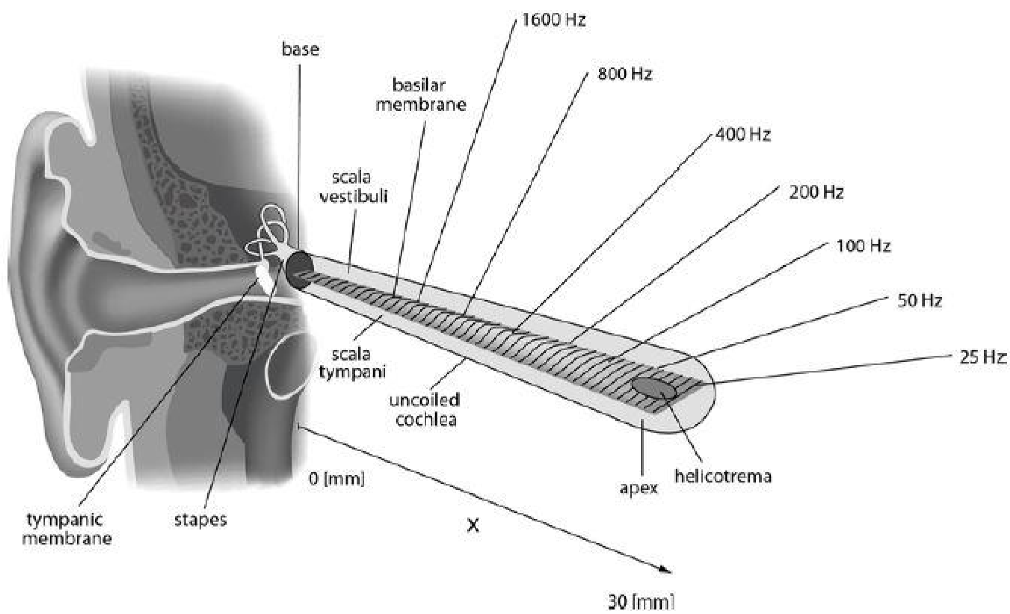


Figure 7. Resonance frequency of the cochlea [8].

The Place Theory represents one of the most important theories in the study of sound perception; it develops from the concept that each region of the cochlea corresponds to a specific frequency. In fact, as it can be deduced from the tonotypical structure of the cochlea, each zone vibrates at a different frequency: the part closest to the stirrup has a higher resonance frequency, while the part farthest away from the ossicle system responds to lower frequency stimuli (Fig. 7) [8].

In figure 8 there is a mechanical model spring-damper representative of the phenomenon just described: every mass simulates a different region of the cochlea and is characterized by own frequency and a specific amplitude of oscillation.

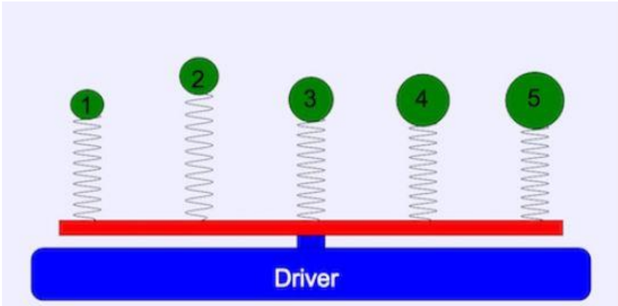


Figure 8. Mechanical model for the cochlea [8].

The resonance curves of the cochlea, as shown in figure 9, are very wide and often overlap each other; in fact, the different areas act as resonators with low quality factor and consequently the ear can hardly distinguish stimuli with very similar frequency. Audiometric studies have shown that, for most people, the perceivable difference for frequencies below 1000 Hz is about 1 Hz. If "The Place Theory" were correct, it should be possible that, for example, between the frequencies from 250 Hz to 240 Hz different regions are excited and therefore different nerve cells respond to the stimuli; however, this does not happen because the frequency responses overlap each other, and the resulting nerve stimulus is almost identical in both cases. In addition, according to what is defined as the "Uncertainty Principle", a strong resonance limits the information about the duration of the sound: if resonance peaks were the only possible elements for the detection of very close frequencies, it is not possible to detect sudden changes in frequency; however, the human ear, can easily perceive variations up to a tenth of a second. Therefore, the mechanism at the bottom of the auditory process is actually much more complex and articulated. A valid reason, to justify the validity of this theory, can be related to the fact that the activity of neighboring nerve cells is inhibited by the firing rate of central cells of the excited area, as it happens also for the sight and the touch [8].

According to **The Temporal Theory**, however, is the nerve impulse, in response to the stimulus which provides information about the type of sound perceived. The propagation occurs as a sine wave at a frequency of 500 Hz and a period of 1/500 Hz = 0.002 s: an excitation determines a vibration of nerve fibers every 0.002 s, which is then transmitted to the central nervous system and intepreted as a sound at a frequency of 500 Hz. The same mechanism is also valid for more complex sounds in which the different sections of the waveform are repeated periodically. The hair cells follow the same periodicity and, in the most complicated cases, in correspondence of the part of the wave with greater amplitude, several nerve fibers are activated in order to store the information on the shape of the wave and send it to the brain. However, the nerve fibres in communication with the cochlea do not always respect this theory and, moreover, they are not sensitive to stimuli around 20,000 Hz. This is because nerve cells can react in a disordered way and combine several signals, or even filter them, giving rise to messages that are difficult to decode. However, in some cases, although the fibers vibrate with periods other than 0.002 s, the information is still transmitted correctly, and the Tempolar Theory is still valid; some studies are investigating this last aspect, but the experimental results obtained until now are still insufficient [9].

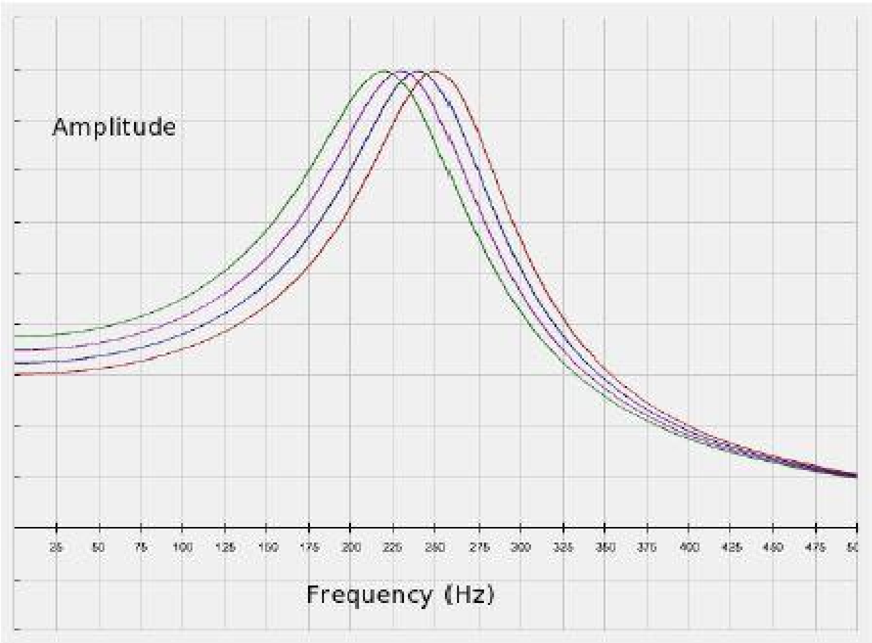


Figure 9. Overlapping phenomenon for four different regions of the cochlea [8].

2. THE HUMAN VOICE

The human voice, like many musical instruments, can be likened to a vibration amplified by resonance: the vocal cords are the vibrating element, while the throat, mouth, nasal cavities and bronchial canals act as resonant cavities that transform vibrations into sound [10].

2.1 THE VOCAL DUCT

The major components of the voice duct are depicted in Figure 10. The nasal cavities (1), represented by a series of interconnected chambers, are located in the frontal part of the face, above the oral cavity (2). The hard palate (3) separates the mouth from the nasal cavity and extends to form the soft palate (4) connected to the uvula (6). The teeth (5), lips (7) and different parts of the tongue (9, 11, 13, 15) also contribute to modifying the sound emitted. The pharynx (8) connects the mouth to the trachea (16), the epiglottis, on the other hand, (10) closes the trachea during swallowing and thus pushes food towards the esophagus. Finally, the vocal cords (12), essential for phonation, are located in the glottis (14); the outer part of the glottis forms the larynx (17).

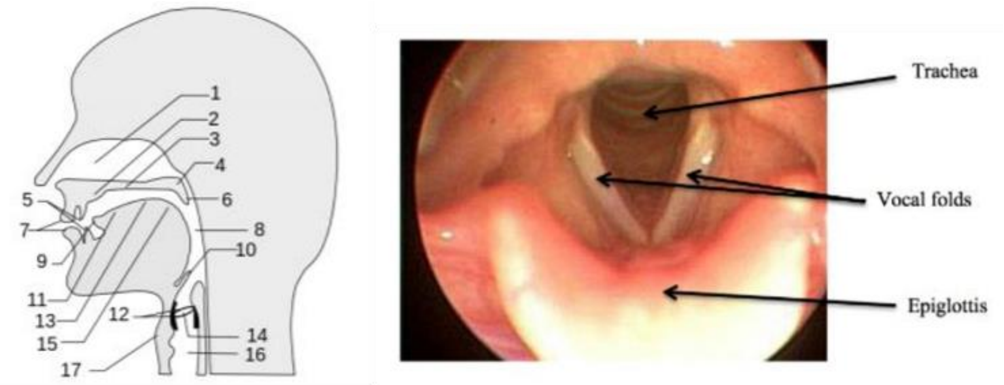


Figure 10. The vocal duct (left) and vocal cords form an internal view of the mouth, in relaxed and open conditions (right) [10].

The vocal cords (Fig. 10, right), are similar to "pockets" made of skin whose movement regulates the flow of air along the trachea; when they close, the Bernulli effect causes them to snap shut, blocking the passage of air, until the flow is strong enough to open them again. Then, they act as vibrating elements generating a sort of "hum" whose frequency depends, in part, also on the muscular contraction exerted on the strings. The following picture shows the typycal movemets of vocal chord vibration that leads to the vocal emission [10].

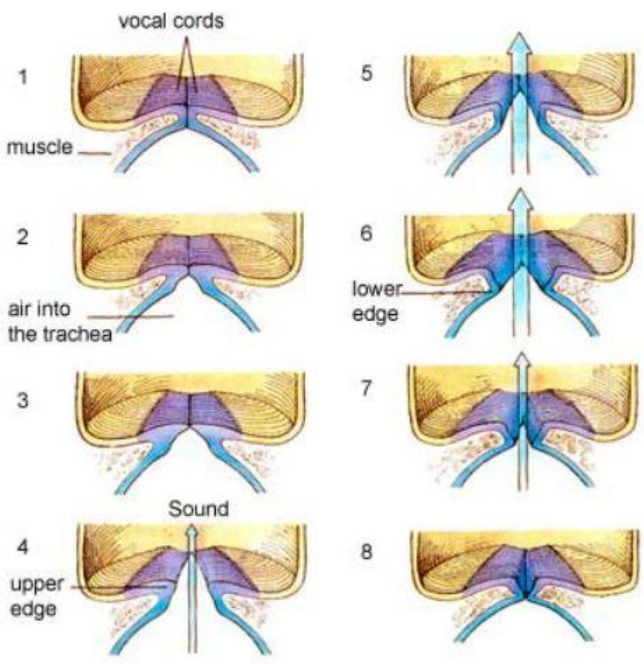


Figure 11. Diagram of the vocal cord vibration process [10].

2.2 THE FORMANTS OF THE VOICE

The characteristics of the voice sound depend on the size, the distribution and the shape of the resonant cavities within the vocal duct that affect the vocal cords vibration in some specific frequencies. The sound resulting from the vibration of the vocal cords covers a wide range of frequencies, from about 60 Hz up to 7000 Hz, considering the harmonics associated with the so-called "pink noise" (Fig. 12) [11].

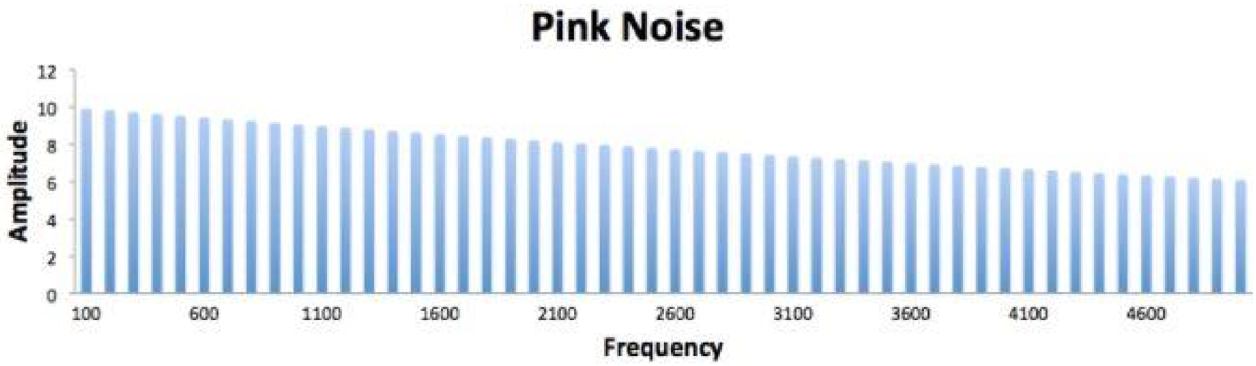


Figure 12. Spectrogram of pink-noise: the amplitude decreases with the frequency.
"Pink- noise" is a particular acoustic signal with a frequency spectrum such that the power spectral density is inversely proportional to the frequency of the signal. In pink noise, each octave carries an equal amount of noise energy [11].

It can be useful to observe what happens within a resonance phenomenon in order to better understand how vocal emission works. Consider, for example, three Helmholtz resonators of different sizes and connected to each other; each of these is characterized by its own resonant frequency (Fig. 13). In general, the acoustic response of a resonator depends on its physical prorieties, particularly on the material and the size; as it can be deduced from the figure 13 the larger the resonator is, the lower the resonant frequency is.

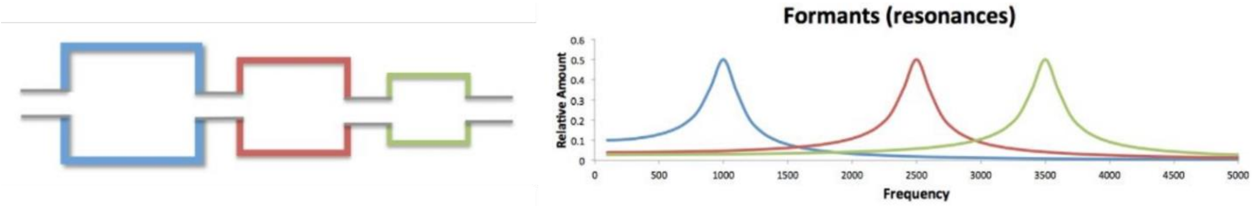


Figure 13. Interconnected Helmholtz acoustic resonators and corresponding resonance peaks [11].

The frequency response of the three resonators, in the case of pink noise, is shown in Figure 14. As can be easily observed, the amplitude of the spectrum is limited by the envelope given by the resonant frequencies of each resonator defined as **formants**. Therefore, the formant represents the timbre component most present in the sound and its nature depends on and is controlled by the type and shape of the resonant cavity [11].

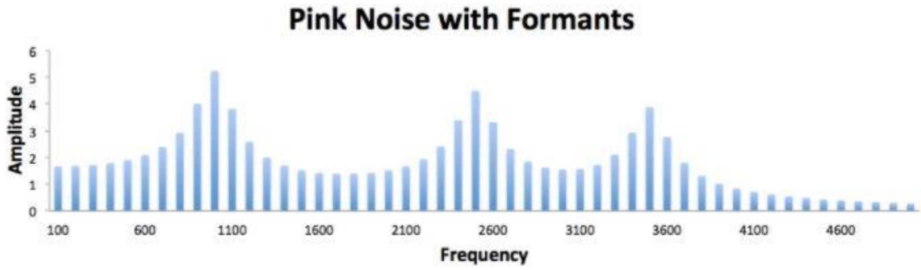


Figure 14. Pink-noise and the formats due to the three Helmholtz resonators [11].

The most influential cavities in the vocal emission are the mouth and the larynx; both change shape while speaking or singing and regulate phonation. The following figure shows, in fact, how the mouth assumes a very narrow position during the reproduction of the "ee" sound, compared to the "ah" sound in which, instead, the mouth is much more open [11].

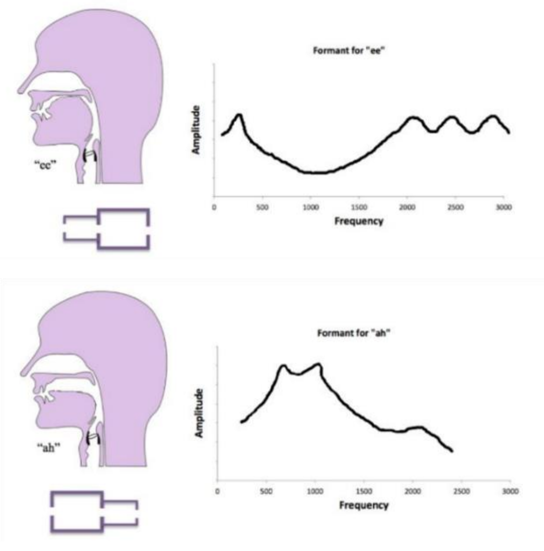


Figure 15. Resonant cavities of the mouth and larynx and formants during the emission of the "ee" and "ah" sounds [11].

2.3 PHONETIC VOICE ANALYSIS

The different sounds that, combined, can form the words and meanings of a language, are called phonemes. Phonetics deals precisely with this topic and it studies the phonemes of a language and reconstructs their origin; This science is divided in *articulatory phonetics*, *acoustic phonetics* and *perceptual phonetics* [12].

3.3.1 Articulatory phonetics

Articulatory phonetics studies the production of sound through the articulatory apparatus. As before stated, the emission of sound depends on an egressive flow of air which passes through the vocal canal. This flow can determine the vibration of the vocal cords and generate a "***sonorous sound***", or alternatively it can give rise to a "***deaf sound***". This aspect is very interesting in the Italian language because it is useful to distinguish, for example, a deaf "f" from a sound "v" and also to analyze the phonetic difference between vowels and consonants. The vowels are sonorous and generated by a free and unobstructed passage of air; between one vowel and another, the only variable elements are the positions of the tongue and jaw and also the opening of the mouth. In contrast, the production of consonants necessarily involves some form of obstruction. The following table details, in terms of place and mode, the articulation of consonants and vowels in the Italian language [12].

Table 1. Articulation for the Italian consonants and vowels [12].

Punto → ↓Modo	bilabiali	labiodentali	dentali	alveolari	palato-alveolari	palatali	velari
occlusive	p b		t d				k g
nasali	m	ɱ		n		ɲ	ɳ
laterali				l		ʎ	
polivibranti				r			
fricative		f v		s z	ʃ ʒ		
affricate				ts dz	tʃ dʒ		
semiconsonanti						j	w

2.3.2 Acoustic phonetics

Acoustic phonetics focuses on the instrumental analysis of the voice. In spite of the progressive development and the diffusion of new representation techniques, the **spectrogram** is the most widely used method for the acoustic interpretation of the human voice; it is a *sonogram*, a graph in wich shows the distribution of the spectral energy of a given sound recording, at various frequencies and for the entire temporal duration of the recording. Spectrograms can be *narrow-band* or *wide-band*, but the latter are more advantageous because they allow a more precise evaluation of the evolution of formants over time [13]. Figure 16 compares the two types of graphs for the pronunciation of the word "yi" or rather the verb "lose" in Mandarin Chinese.

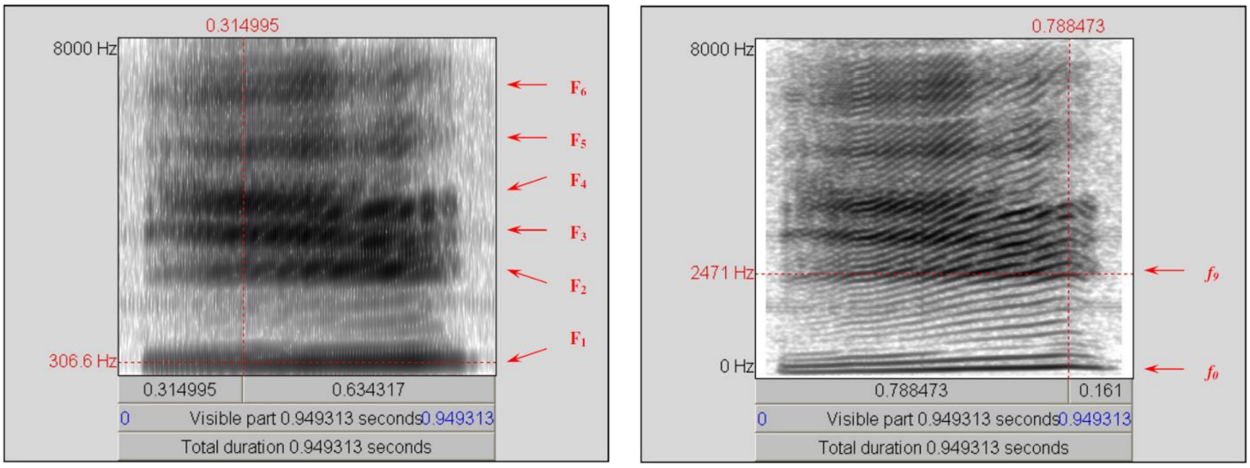


Figure 16. Wide-band (left) and narrow-band (right) spectrograms related to the pronunciation of the verb "lose" in Mandarin Chinese [13].

In the wide-band spectrogram the vertical dotted line indicates the boundary between the two syllables of the word; it can be observed that there is a gradual increase in energy passing from one syllable to another and also the first six formants (F1, F2...etc.) of the word, represented by a horizontal distribution, can be easily identified. On the other hand, from the narrow-band spectrogram it's easy to distinguish not only the formants, often segmented into a series of parallel striations, but also the harmonics (f0, f1...etc.) which, in fact, depend on the characteristics of the sound produced and not on how it is modified by resonance phenomena [13]. It is also possible to identify the value of the fundamental frequency, indicated by f0; the distance between each harmonic corresponds precisely to the value of the fundamental frequency that can be calculated simply by dividing the frequency value associated with the n-th harmonic (fn) by "n", which is the value that indicates the number of the harmonic in succession (Eq. 1).

$$f_0 = \mathrm{n} \cdot fn \tag{1}$$

Spectrographic analysis is flanked, then, by other representation techniques such as:

- ***Oscillogram*** related to the trend of the instantaneous amplitude of the waveform on a scale of different levels, usually shown above the spectrogram (Fig. 17)
- ***Melodic profile*****,** also called ***f0 curve*** or ***pitch***, it follows the fundamental frequency trend on a scale in Hz, usually superimposed on the spectrogram (Fig. 18.A).
- ***Sound intensity*** or ***energy curve*** which refers to the instantaneous volume (loudness or intensity) expressed in dB and superimposed on the spectrogram (Fig. 18.B).
- ***Formantics plot***, is an estimate of formantics changes during sound reproduction, superimposed on the spectrogram (Fig. 18.C). (Fig. 18.C).

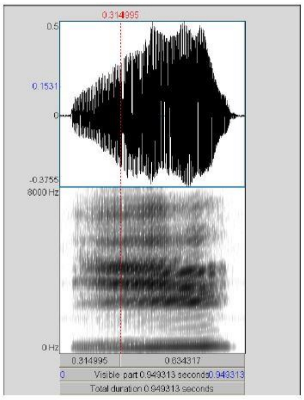


Figure 17. Oscillogram (top) and spectrogram (bottom) related to the verb "lose" in Mandarin Chinese [13].

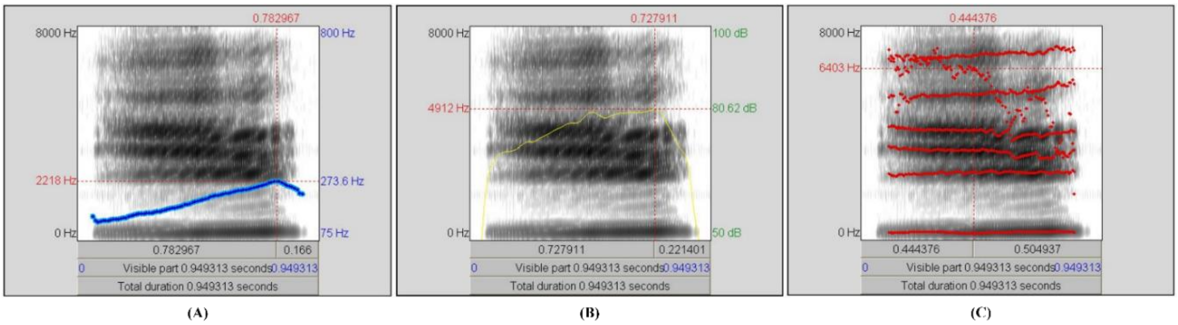


Figure 18. Melodic Profile (A), Sound Intensity (B) and Formant Trace (C) during the pronunciation of the verb "lose" in Mandarin Chinese [13].

2.4 MODELLING OF SPEECH PRODUCTION

Figure 19. Movements of speech organs during natural speaking, extracted from the Niebergall's study about Real-time magnetic resonance imaging for speech analysis [14].

The emission of speech signal, as described before, is characterized by a complex phonetic process. In this case, mathematic models can be very useful to understand how the articulatory mechanism works and to build speech processing algorithm for speech signal or coding too. These models derive from some studies focused on movements that occur during phonation and on the analysis of the vocal organs involved [14]. In the following picture there is an “engineering-oriented” acoustic model for the speech production.

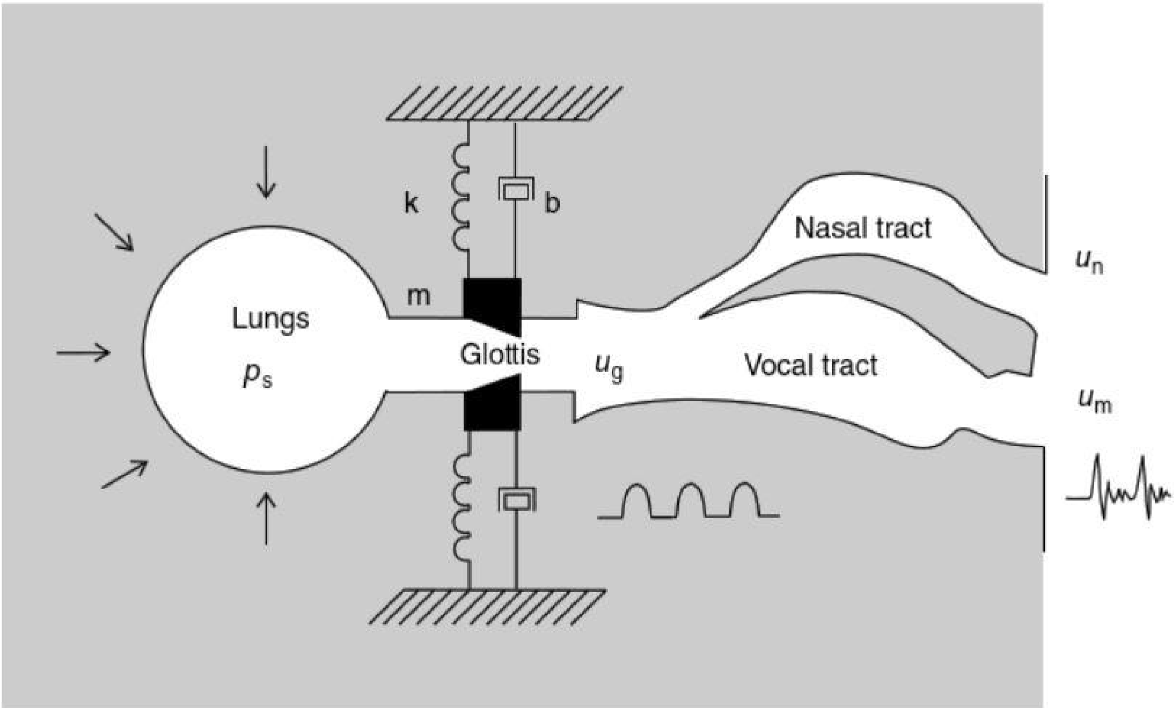


Figure 20. Mechanical model for the speech production [15].

The lungs can be considered as the driving force of the system, they work as a sort of pump that pushes air out and leads it to the epiglottis. The vibrational movement and mechanical action performed by the epiglottis are modeled with a mass- spring oscillator on which the characteristics of the sound emission depend. The vocal and the nasal tract can act as acoustic resonators or filters and modify, like a modulator or an encoder, the speech sound that propagates from the mouth or from the nose [15].

Starting from this model a circuit analogy can be built (Fig. 21).

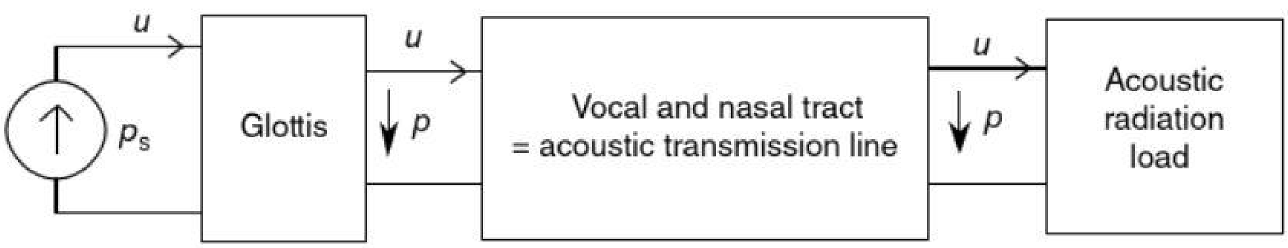


Figure 21. analogy circuit modelling the vocal emission process [15].

It is represented by a *three cascaded system*:

- The generator which represents the excitation of the system given by the pumping work of the lungs
- The transmission line modelling the glottis
- The acoustic radiation load which considers the radiation impedance.

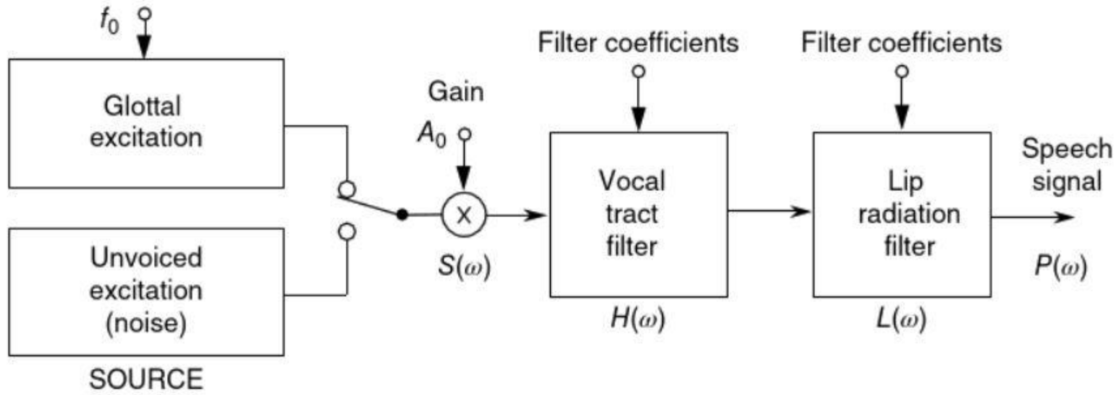


Figure 22. Signal model for the photation analysis [15].

The quantities, involved in this circuit, which regulate the functioning of the model are the *sound pressure* (p) and the *volume velocity* (u) in the system.

The figure 22 illustrates another model based on the consideration of the voice as a signal. During the phonation and in voiced signal the source is given by the oscillation of the epiglottis, in unvoiced signal there is not a specific source, but the excitation is due to the frication noise generated by pressure changes following a sudden closure of the vocal tract [15].

The variables that describe this model, in relation to pressure and the volume velocity of the mechanical model, should be the voltage and the current. However, in this case the system depends only on a single transfer function derived from the following Fourier Transform formulation:

$$P(\omega) = S(\omega) \cdot H(\omega) \cdot L(\omega) \tag{2}$$

Upstream of the structure there is a switch system, thanks to which this solution is applicable for both voiced and unvoiced signal. This diagram considers not only the source but also how the sound changes before reaching the outside world and radiating. In fact, it is also called "***source- filter model***":

- $P(\omega)$ is the function that describes the waveform of the speech
- $S(\omega)$ represents the excitation signal

- $H(\omega)$ transfer function for the filtering function of vocal/nose tract
- $L(\omega)$ transfer function for the filtering action of the lip

In the graph below (Fig. 23), for example, it is possible to observe how the sound undergoes alterations in the vocal tract.

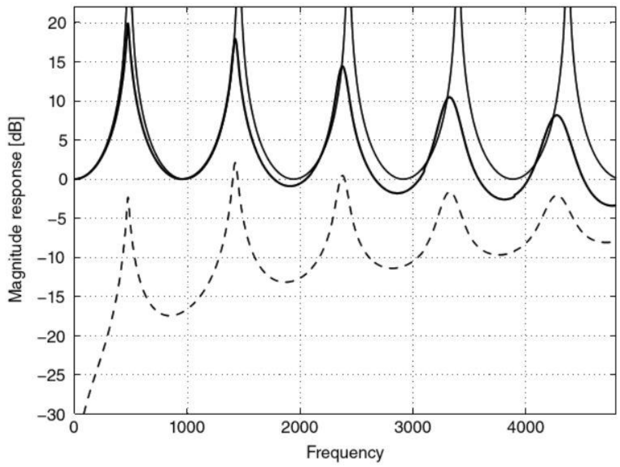


Figure 23. Magnitude response of the volume velocity transfer along the vocal duct [15].

The dotted line is related to an ideal condition and without losses in which the trend of the curve follows the resonant frequency, the upper one is referred to a simulation with losses due to the phenomenon of friction and thermal loss [15].

3. HEARING IMPAIRMENT

3.1 THE HUMAN AUDITORY FIELD

The human ear can perceive stimuli from 20 Hz to 20 kHz. This range of frequencies represents the ***audible sounds***, the frequencies below 20 Hz are classified as ***infrasounds*** and the sounds above 20 kHz belong to the ***ultrasounds***, which can be detected only by some animals such as dogs, bats, or dolphins (Fig. 24) [16].

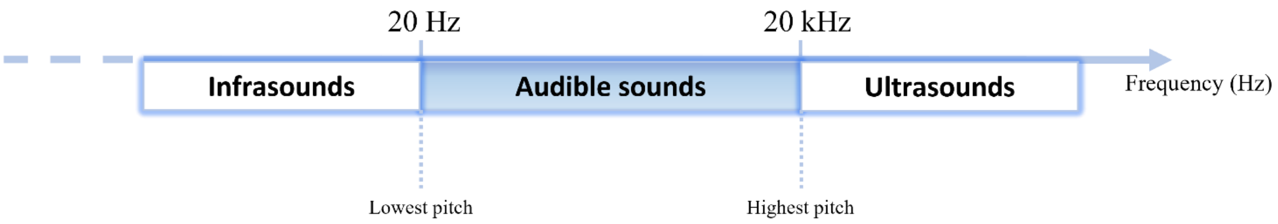


Figure 24. Audible frequencies for the human ear.

The dynamic range of the intensities lies about between 0 dB and 140 dB, however, for higher and lower audible frequencies, it is narrowed. The lower limit is called **threshold of audibility,** it is the minimum intensity value to perceive acoustic stimuli at different frequencies. The upper limit is the **threshold of pain** whose value depends on the subject. In general, all sounds above 90 dB could damage the inner ear or create irreversible damage over 120 dB (Fig. 25).

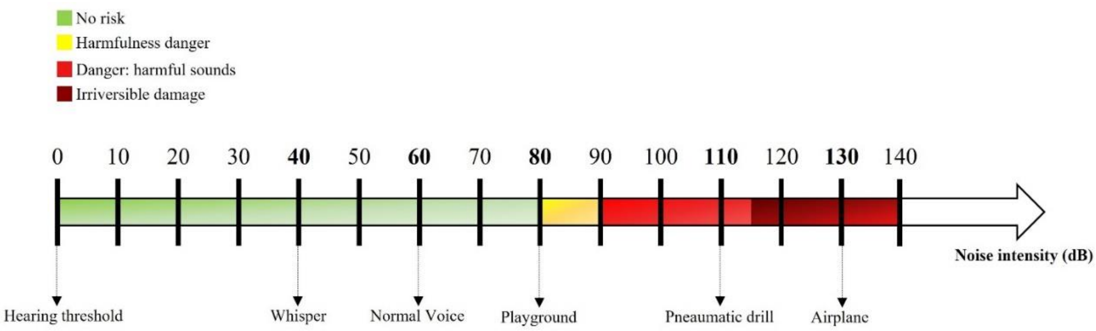


Figure 25. Dynamic range for human ear.

The graph illustrated in figure 26 is the ***human auditory field*** which is represented by the set of sounds perceivable on average by a normal hearing human [17]. It is obtained testing acoustic response of different subjects by sending them stimuli with several intensity and frequency: each point is uniquely identified by a frequency value (ordinate) and an intensity value (abscissa). It is possible to individuate a region, between 2000 Hz e 5000 Hz, in which the energy/frequency ratio is optimal, and the ear is able to perceive faint sounds near the threshold of audibility. Even the frequency interval from 200 Hz to 7 kHz is very sensitive, an energy of about 10 dB is not enough to detect a sound [15]. The conversation area (dark green) shows the range of sounds most commonly used in human voice perception; when hearing loss affects this area, communication is impaired [17].

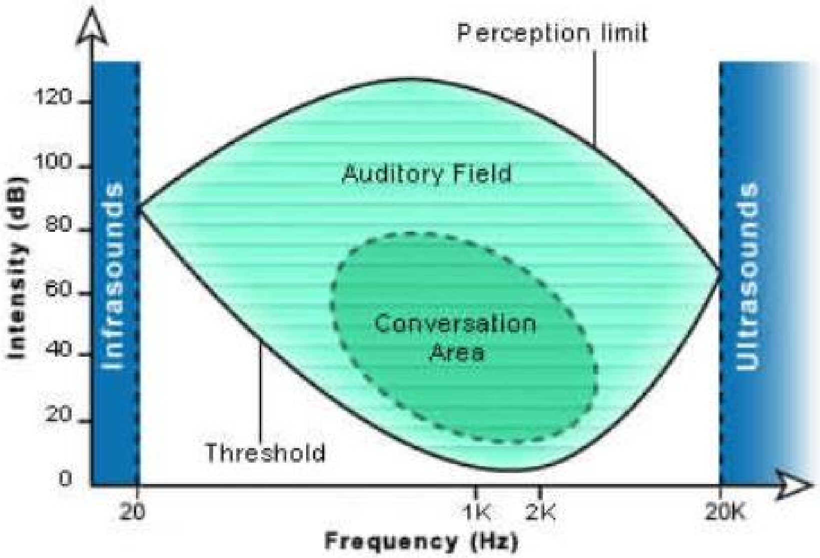


Figure 26. Human auditory field [17].

3.2 BASIC PSYCHOACOUSTIC QUANTITIES

The most important physical quantities that describe the characteristics of a sound are frequency, level, magnitude spectrum and time. The ***auditory psychophysics***, better known as ***psychoacoustic***, is the science which studies acoustic aspect related to the human perception of the sound. It's based on some experiments that allow to evaluate complex and physical auditory system safely and in a non-invasive way. Psychoacoustic uses specific quantities which are expressed as function of physical variables; below will be introduced and analysed **pitch**, **loudness**, **timbre,** and **duration**.

3.1.1 Pitch

Pitch is defined by the American National Standards Institute as "*That auditory attribute of sound according to which sounds can be ordered on a scale from low to high*" [18]. From a physical point of view, pitch is mostly related to the frequency of a repetition in a signal; it is a property attributable to any type of sound such as instruments, voice, or noise. However, pitch's individuation is not a simple task because often a sound may not have an easily identifiable pitch or, in some cases, even more than one, but is speech analysis the study is restricted to sound with a predominant pitch. Pitch is a an extremely important variable in communication, it allows for the distinction between a male, female, or child's voice; pattern and changes in pitch influence the prosody of speech and give information about the geometry, size and physical proprieties of the source. [15]

The following picture (Fig. 27) is referred to *pitch strength* intended as the clarity and salience of pitch perception. This parameter has been measured testing psychoacoustic conditions for different signals at 250 Hz. A 250 Hz sinusoid is considered as reference in the scale; in general, periodic signals have higher pitch strength than aperiodic ones, in fact, sinusoids and low-pass filtered harmonic have the best pitch perception [15].

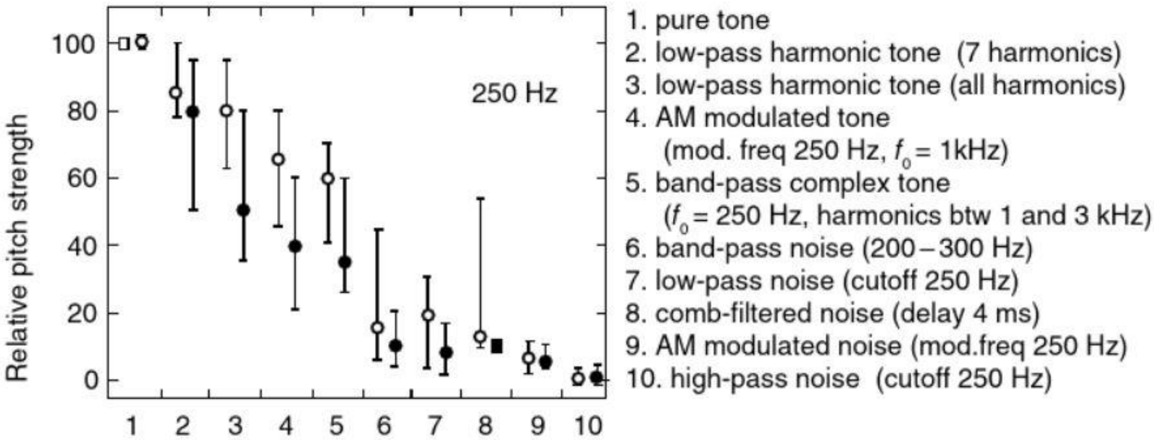


Figure 27. . Pitch strength measurement for some different signals [15].

Loudness is "*that attribute of auditory sensation in terms of which sounds can be ordered on a scale from quiet to loud*" [18], therefore, it is a sort of subjective measure of perceived sound intensity. It is strictly connected with the concept of sound pressure and related to the *sound intensity* by a logarithmic relation. In fact, a low value of intensity sets a not audible sound, on the contrary, high intensity sounds are ascribable to a painful and dangerous condition for the hearing. Usually, the human ear can tolerate sound intensity which is about 10^{12} greater than the perceptible one. Obviously, this range varies with frequency and depends on the subject [19]. Loudness is based on the use of ***phon***, a specific unit which permits to have a subjective measurement of the sound intensity; the intensity of a sound, expressed in phon, is the same intensity of a pure 1 kHz tone, expressed in decibels, judged by the listener to be equally loud. In this way, the objective decibel scale is converted into a partially subjective scale whose, however, characteristics depend on the judgment of the listener involved in the measurement [19].

Loudness is measured in Sound Pressure Level (SPL) on a logarithmic scale in sound decibel (dB SPL):

$$L_{SPL} = 10 \log \left(\frac{P}{P_0} \right)^2 = 20 \log \frac{P}{P_0} \qquad [dB\ SPL] \tag{3}$$

Where P is the effective value of the pressure of the sound and P_0 is the reference value refers to value of the minimum sound pressure required to perceive a sound of 1 kHz and 20 μPa. However, in audiology, audiometric decibel (dB HL) is used to indicate the subjective hearing ability respect to the standards. In this case, the minimum pressure value, audible for all the frequencies, is defined in dB SPL. This means that dB HL represents, for a given frequency, the hearing threshold of a normal hearing person.

Even for Loudness, like the Pitch, some experimental studies were ran. The figure 28 illustrates the ***equal loudness contours***, built on the response of several people to a specific audiometric test. This graph shows curves of averaged sensitivity for which a listener, subjected to pure tone acoustic stimuli, perceives a sound with equal and constant intensity, for different frequencies [21].

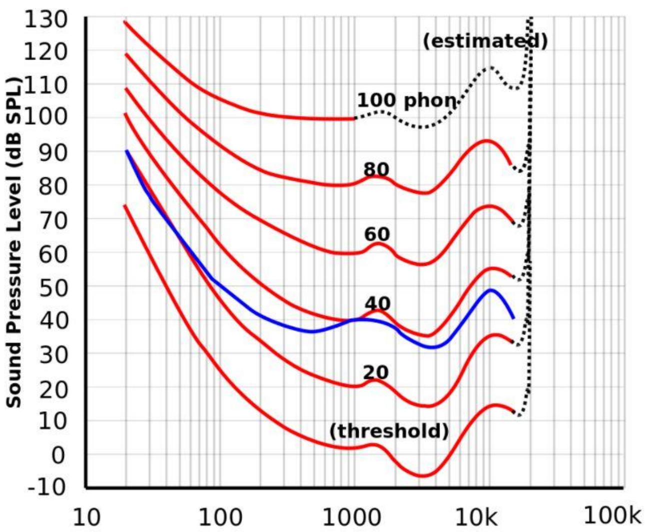


Figure 28. Equal-loudness contours [20].

Subjective quantities were defined because they've resulted being more useful to study the human ear's proprieties. In general, engineer applications use more objective scale, and medical or biological fields tend to use more-subjective scale [19].

3.1.3 Timbre

Timbre, or tonal colour, is a multidimensional psychoacoustic measure. It is the only quantity thanks to which is possible to distinguish two, or more, sounds with same value of pitch, loudness and duration. This is probably due to a physical concept for which the acoustic spectrum changes along time. The number of possible timbres is very big. Considering that the human ear has a frequency resolution of 1 ERB, the auditory field is composed mostly by 42 bands, and a level resolution of 1 dB for a dynamic range of 100 dB, there are about 100^{42} timbres. Obviously, if the dynamic range is not as large the former one, because of some masking phenomena, the number of timbres appears to be less but still relevant [15].

In general, in *steady-state conditions* the loudness of the sounds represents quite faithfully the corresponding timbre. Otherwise, if the sound is characterized by *modulations* in amplitude or in frequency the perception of the timbre can be different, because of the changes on the auditory band. Also, the *onset* of the sound is important for the timbre; some vocal or instrument emissions have specific onset that, acting on the component of the sound, lead to a different detection of timbre [15].

3.1.4 Duration of the sound

Duration is the attribute for which is possible to order a sound from "short" to long"; as well as fo pitch and loudness, the duration is a subjective quantity, and some tests are needed to define it. For example, in the figure 29 the subjective duration of a burst tone is shown; This graph was built considering the result of a particular task in which the subjects have to find sounds with half or twice duration of a reference sound. Subjective duration is measured in dura, a specific unit related to the perception of 1 s of 1 kHz tone. From the curve it can be understood that the effective and perceived durations are quite similar for sounds longer than 200 ms, on the contrary, if the physical sound is shorter than 200 ms, the subjective duration is received longer than the real one [15].

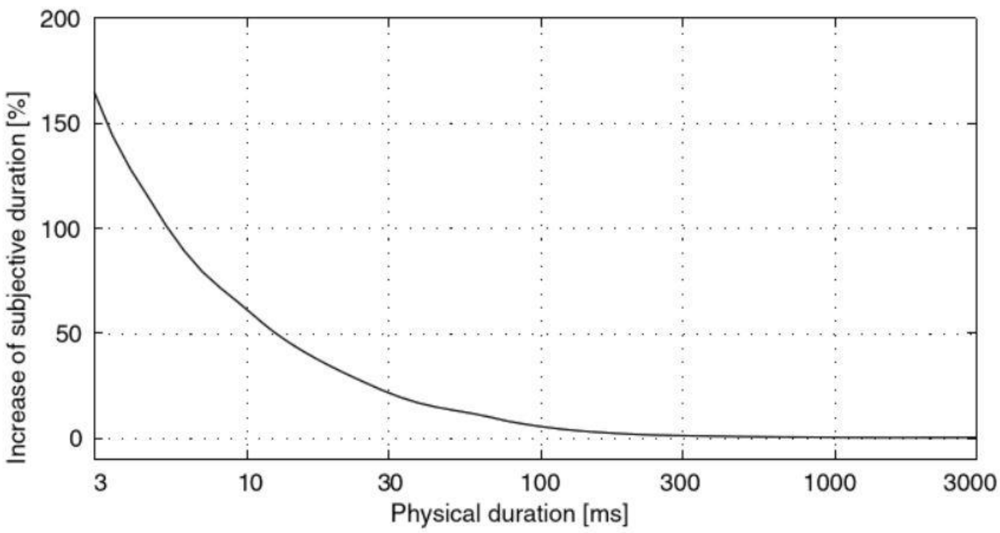


Figure 29. Subjective duration of a burst tone [20]

3.3 DEAFNESS AND HEARING LOSS IN CHILDREN

3.3.1 Types of hearing loss

The hearing loss is referred to reduced functioning of the ear that can affect the loudness and/or timbre of sound [22]; according to the ***disability level***, it can be classified in different types [24]:

- **Conductive hearing loss** (CHL) it is due to an interference in the mechanical transmission of sound trough the external and the middle ear. It affects mainly children, because of inner hear infection, a very diffused problem during childhood [23]. It can cause permanent impairment; however, it is usually temporary and can be treated with medicine or surgery.
- **Sensorineural hearing loss** (SNHL) reflects on the problem in transducing vibration into neural impulses within the cochlea or transmitting them down the auditory vestibule-cochlear nerve, this may be due to damage to nerve itself or even to the hair cells [25]. Generally, it is permanent and very difficult to treat; it is the most common type of hearing loss and the most severe one [23]. The auditory nerve has the task of carrying fundamental information about pitch, loudness and meaning of sound to the central nervous system; SNHL alters its functioning and, therefore, this condition can often lead to difficulty understanding sound or speech, even if it is loud enough to perceive, which result distorted as well as softer.
- **Mixed hearing loss** is given by the combination of CHL and SNHL and it is usually linked to a damage of the middle or the inner hear.
- **Central hearing loss** is related to some brain disfunction. It is due to auditory neuropathy spectrum disorder, it means that sounds and acoustic stimuli enter the ear normally but, because of a damage of the inner ear or a malfunctioning of auditory nerve, the information cannot be organized in such a manner to be understood correctly by the brain.

However, hearing impairment can also be classified according to the severity of the damage and the resulting ***degree of deafness***; The table 2 gives information about the variety of hearing disorder and highlights their range. According to the severity of the loss, there may be a variety of problems with cognitive development and language; high degrees of deafness require early intervention and, in some cases, the use of hearing aids [26, 27].

Deafness can also be described as [28]:

- **Unilateral or Bilateral**

Hearing disorders affect only one ear (unilateral) or both (bilateral).

- **Symmetrical or Asymmetrical**

  If the characteristics of hearing loss are the same in each ear the configuration is symmetrical, on the contrary, it is asymmetrical.

- **Progressive or Sudden**

  Hearing loss worsens over time (progressive) or happens quickly (sudden).

- **Pre-lingual or Post-lingual**

  Hearing loss happened before a person learned to talk (pre-lingual) or after (post-lingual).

- **Fluctuating or Stable**

  This depends on the changes of hearing loss along time, it can be stable or not. Fluctuating condition is typically due infection or problems with middle ear fluid.

Table 2. Degree of deafness: dynamic range and its consequences [26,27].

Hearing Loss label	Hearing Threshold (dB)	Description
Normal Hearing	-10 – 15	No significative hearing loss symptoms
Slight Hearing Loss	16 – 25	• Compaired to the hearing ability when the index fingers are placed in ears; • Difficulty in hearing faint or distant speech (when the speaker is at a distance greater than three feet); • Problems in hearing in noise background (classroom, restaurant); for a child could be hard focusing on class.
Mild Hearing Loss	26 – 40	• Hearing problems greater than “plugged ears”; • Child can hear but misses fragments of speech which leads to misunderstandings. • At **30 dB** hearing loss child can miss up to 25-40% of speech signal, at **40 dB** child may miss 50% of classroom discussions. • Hitch in first learning skills (reading, letter/sound association)
Moderate Hearing Loss	41 – 55	• At **50 dB** hearing loss child may miss up to 80% of speech signal. • For child early amplification is needed to limit problems of delayed or disordered syntax, limited vocabulary, imperfect speech production, and flat voice quality. • Hearing aids can only limit the problem but they are not useful if the noise level is too high
Modarately Hearing Loss	56 – 70	• This condition could lead the child to miss up to 100% of speech content; • Early intervention is essential for learning development; it is very important to identify this deficit before 1 year of age to avoid learning disabilities.
Severe Hearing Loss	71 – 90	• Whitout amplification, child can hear only close and loud noise; • It’s crucial that to the use of hearing amplification be added intensive activities focused on language; • Early intervention, before 6 months, is critical to functionalise the brain to specific input; • This condition often requires the use of **cochlear implants.**
Profound Hearing Loss	> 90	Child affected by this impairment cannot hear with traditional hearing aids.

Only after having established the nature of the disease, in terms of its degree and type, is it possible to define an appropriate treatment. Less serious cases can be solved through the adoption of hearing aids, aural rehabilitation, and cochlear implants; although, significant conditions require more invasive medical intervention or surgery [28].

3.3.1 Etiology

The hearing loss may be caused by several factors which can be **congenital** or **acquired**.

CONGENITAL HEARING LOSS

According to the etiology, ***congenital*** causes can have *genetic* or *non-genetic* nature; the former type, which is the most diffused, is usually related to genetic mutation and it is often further subdivided into *syndromic* versus *non-syndromic* categories [25]. So far, more than 500 forms of syndromic hearing loss have been identified; some of the most important are [24]:

- Waardenburg syndrome can generate hearing loss of varying degrees.
- Usher syndrome affects hearing and sight; it is often associated to SNHL, retinal damage and vestibular dysfunction.
- Pendred syndrome leads to SNHL.
- Alport syndrome is inherited disease which elicits progressive SNHL and nephrite.

Congenital and non-genetic hearing reduction is due to several situations referable to conditions like trauma or drugs during the prenatal period, preterm birth, or viruses. Cytomegalovirus (CMV), for example, has been identified as one of the most diffused infective agents of non-genetic deafness [24].

ACQUIRED HEARING LOSS

Obviously, children can be affected by hearing loss not only after birth but also in later years. When the proprieties of deafness are not related to congenital factors, hearing impairment can be due to ***acquired*** causes. Several studies have demonstrated that *otitis* can expose children to dangerous hearing conditions; otitis media with effusion (OME), a particular type of otitis characterized by an accumulation of fluid in the middle ear due to altered eustachian tube functioning, in fact, is the first cause of pediatric CHL. Typically is a non-severe pathological condition that resolves on its own or with the insertion of an artificial ventilation system in the middle ear [24]. However, otitis can also damage the ear without generating infections, therefore, it is recommended a combination between visual inspection and tympanometry [25]. CHL can also have *tumoral* nature, Cholesteatoma, a benign tumour of skin which can affect the middle ear, for example, may be asymptomatic or results in hearing loss that increases as the metastasis progresses [25]. In addition, trauma is also very important; accidents or *traumatic events* involving the temporal bone are potentially dangerous because they risk damaging the cochlea and lead to CHL or SNHL; children who present significant hearing problems after a *physical trauma* should be seen promptly by a competent physician to identify the type and severity of damage sustained and to understand how it can be treated. However, *acoustic trauma* can create hearing disorders too, in fact, they are considered as the main culprit in hearing damage over 4000 Hz. In general, severity of impairment depends on the intensity of the sound and the type of exposure; As illustrated (in Chapter 4.1 "*The Human Auditory Field*"), the threshold of pain has not a specific value because it depends on the person, but usually a sound over 80-90 dB should damage the ear; anyway , if not only the intensity but also the exposure is considered, sounds greater than 85 dB with a continuous exposure or sudden noise, greater than 140 dB, are defined as the limit of damaging noise levels. Then there are some *bacterial diseases* which can damage auditory abilities; meningitis, for example, is a common source of pediatric SNHL which, in severe cases, may even result in profound hearing loss; therefore, it's important that a child affected by acute meningitis undergoes a hearing screening as soon as possible [24].

Other possible risk factor for neonates can be [25]:

- Family history
- In utero infections
- Craniofacial anomalies
- Birthweight lower than 1,5 Kg
- Need to prolonged intubation

Table 3. Congenital and acquired factors for CHL and SNHL [24].

Conductive Hearing Loss	Sensorineural Hearing Loss
Congenital • Microtia/atresia • Tympanic membrane abnormalities • Ossicular malformations	• Genetic disorders (syndromic, connexin 26, mitochondrial) • In utero infections (cytomegalovirus, measles, mumps, rubella, varicella, syphilis) • Anatomic abnormalities of the cochlea or temporal bone • Exposure to ototoxic drugs during pregnancy (alcohol, isotretinoin, cisplatinum) • Hyperbilirubinemia
Acquired • Infection (acute otitis media, otitis externa, ossicular erosion) • Otitis media with effusion • Foreign body (including cerumen) • Cholesteatoma • Trauma (ossicular disruption, tympanic membrane perforation)	• Infections (bacterial meningitis, measles, mumps, rubella, Lyme disease) • Trauma (physical or acoustic) • Radiation therapy for head and neck tumors • Neurodegenerative or demyelinating disorders (Alport, Cogan syndromes)

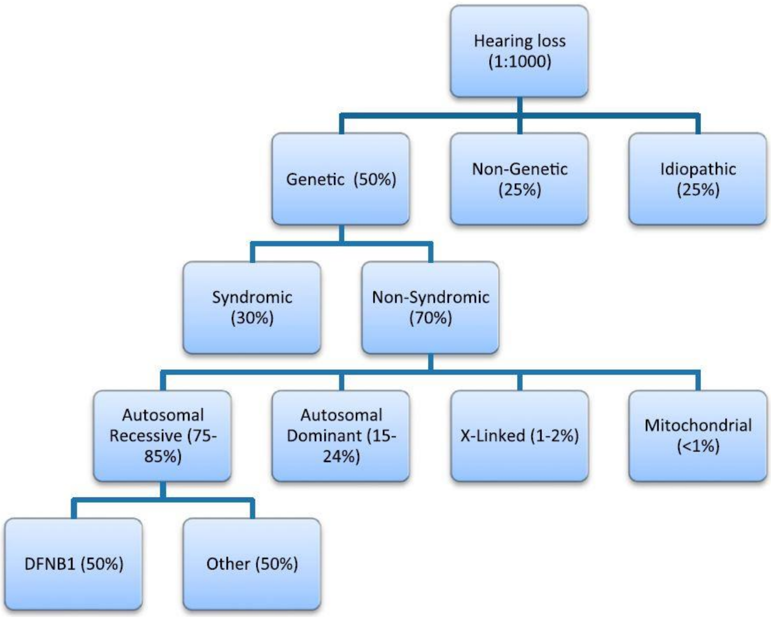


Figure 30. Etiology of prelingual hearing loss in developed countries [29]

3.3.2 Evaluation

It is known that children can deal with the auditory world from early age and the language acquisition lasts until 36 months of age, therefore, it's very important to expose the child to spoken language as soon as possible. Hearing loss can be a hindrance in linguistic and cognitive development, however, early intervention on children with hearing problems can offset these difficulties. In fact, some studies have demonstrated that children, treated with early hearing detection and intervention (EHDI) protocol, have better performance in test of vocabulary skills and intellectual development or even similar abilities to children whose auditory capacity is unimpaired. According to this idea, it's clear that children's hearing abilities must be constantly monitored since birth to detect all hearing problems and to intervene as soon as possible [24].

However, hearing problems in infants, differently from adults, cannot be diagnosed through simple observation because hearing-impaired infants are obviously not able to speak, therefore, some objectives analyses are needed. Here are the tests, divided by age group, that are usually carried out on children [25]:

- **Neonates**
  - Otoacoustic Emissions are based on the use of specific acoustic stimuli that cause the contraction of cochlear outer hair cells detectable from the external ear canal using a probe and transducer [29]. The stimulation consists of a low intensity sound energy [30], which means that a general anaesthesia is not necessary, in fact, the test is easy to perform and immediate.

- Automated Auditory Brainstem Response is an electrophysiological technique which permits to monitor brain activity thanks to electrodes placed on the scalp. According to the type of stimuli sent to the subject, the brain responses in a different manner. To evaluate hearing proprieties, obviously, triggering acoustic signal are used; whit this method is possible to assess the integrity of the entire auditory canal, from the outer ear to the brainstem [31].

- **6-8 months**

  Distraction techniques can be descripted as a type of test in which the doctor or the tester try to engage the child's attention. After that some external sounds are emitted around the child, the evaluation consists in observing if the child turns towards the sound source.

- **9 – 36 months**

  Visual Reinforcement Audiometry: this test requires a specific spatial configuration. The child is sitting in front of a table which some toys, beside the child two sound sources are placed. The child is asked to play and focus on that, if his/she is distracted because of a sound, a flashlight is emitted.

- **24-60 months**

  Conditioned Play Audiometry: the child is asked to perform a task after listening a specific acoustic stimulus. After the child has understood the task, the intensity of the sound is reduced in order to assess the value of hearing threshold.

- **Over 2 years**

  Pure Tone Audiometry subjectively determines how the individual processes auditory information; it is a challenging test which requires a high degree of attention; therefore, it is usually run with children older than 5 years old. The audiometric test is based on finding the hearing thresholds using sounds with different frequencies and intensities. The aim is to detect the quietest sound that children can hear at least 50% of the time.

Logically, depending on the symptoms shown by the child and the characteristics of the individual case, different measures and screening techniques can be adopted. For example, if the hearing loss shows syndromic behaviour, genetic test is advised. Also medical imaging plays an essential role in diagnostics, computed tomography (CT) or magnetic resonance imaging (MRI) are the most used for investigate cochlear disorders. Electrocardiography (ECG) can be useful to analyse QT interval which is an index of hearing functioning, if it's too long, it can lead to deafness. Moreover, some new genetic methods have been proposed; in fact, in some cases, just measuring the renal function and monitoring the presence of a specific protein (connexin-26), is possible to detect SNHL condition [24].

3.3.4 Treatment and management

After that, thanks to the possible diagnostic programs and the various screening tests, hearing problems are identified, it is crucial to define an intervention plan adaptable to the specific case and more detailed as possible, to limit hearing loss and maximize the functioning of the damaged auditory system. In addition, early intervention is important to prevent the problem from worsening and affecting other sensory systems. Typically, the more severe the ear damage, the greater is the likelihood of becoming completely deaf without proper intervention [32].

Hearing deficits often affect language and cognitive development, particularly in children, and, therefore, it influences everyday life. In fact, they can often cause problems in relationships or in social interactions [31]. In this regard, several authors have pointed out how essential is to make the child feel comfortable and at ease, not only acoustically but also socially; then they accordingly recommend **conservative management** focused on the child's needs. In this type of therapy, child has to be supported and advised and family is a key element in that. Various behavioural measures can help the child during everyday action, which consist in some cautions like limiting background noise, talking face-on, and clear intonation; this aspect becomes relevant in environments, such as schools, where hearing process is the basis of the communication [25].

There are also particular conditions related to phenomena like otitis, bacterial infections, or viral diseases, for which it is necessary to intervene by administering *antibiotics* or subjecting the child to some imaging methods that can help to understand the cause of the problem. For example, for severe childhood SNHL, ECG can be useful to measure QT interval or also to evaluate hereditary disease [31]. In some cases, *surgical intervention* is needed; OME often requires the insertion of ventilation tube to flue ear and cholesteatoma has to be removed surgically [25].

For children with mild to severe hearing loss, ***artificial amplification*** could be a great solution. The modern technology is suitable for the first months of age, and it allows to regulate the amplification according to the characteristics of the hearing profile of the child. This type of treatment is very effective because it gives the child the opportunity to have access to the entire sound spectrum of the speech signal, therefore, it permits to obtain clear and satisfactory improvements in learning skills in long term [32].

3.3.5 Hearing aids

The **hearing aid** (HA) is a simple signal processing system with an integrated sound reproduction technology. It focuses on the collection of external sounds, processes and amplifies them so as to compensate the lack of information caused by sound comprehension difficulties in individuals with hearing loss.The improvement of the users' linguistic and auditive skills is therefore the major objective; the intent of modern bioengineering is to design systems as versatile as possible that are able to ensure a good signal-to-noise ratio (SNR) by changing the value of amplification depending on the circumstances and the external environment. The simplest configuration consists of only 3 parts [15]:

- a *microphone* that detects sound from outside
- An *amplifier* that amplifies the signal with a constant gain over the entire range of audible frequencies
- a *speaker* that transmits the amplified sound into the ear canal.

This set up, however, is obviously not suitable for more complicated cases, which require selective amplification and more complex instrumentation [15]. Here, it will be briefly described the most used models [25]:

- **Binaural air conduction** is typically used in subject affected by partially hearing loss; in fact, it exploits the still functioning inner ear and central auditory system. The working principle is based on a microphone which collects the sound which is converted in a digital signal and then amplified. After that it is converted again in analogic signal which is sent to the ear. The entire processing module can be placed externally near the ear, in the ear or in the auditory canal (Fig. 31).

Figure 31. Binaural air conduction hearing aid models (from left to right): external, in the ear, in the auditory canal [15]

- **Bone conduction hearing** is advised for CHL treatment and in particular conditions in which hearing aid is difficult to place, because of physical deformities or chronical infections. This solution, in fact, requires a surgical intervention to anchor the system to temporal bone. In this way, the sound is transmitted directly to the inner ear, bypassing the middle ear. However, the installation process in invasive, therefore, this model is usually used for almost 4 years old children, even if they are some although there are some variants that allow the application even in the first months of life.
- **Contralateral routing of sound (CROS)** is recommended for unilateral sensorineural hearing loss. In this case, no amplification is needed, the sound is detected and immediately sent to the ear with better hearing. BiCROS can also be used for the treatment of both ears.

Hearing aids are the most used artificial amplification system used to improve hearing abilities in children with no serious problems, but sometimes a more complex solution is needed to recover and compensate hearing deficits, like ***Cochlear implants***. In the following chapter some information about this medical device and its functioning are given.

4. COCHLEAR IMPLANT

4.1 GENERAL ASPECTS

Cochlear implant and its functioning are the main topic of this research. Children involved in this study are treated using cochlear implant. In fact, the aim is to assess the speech intelligibility and to test the auditory skills of the subjects in particular acoustic conditions. The cochlear implant's configuration is a parameter that greatly influences the execution of the test; therefore, participants are given specific information on how to correctly set the device in order to ensure a good quality of results and the success of the test.

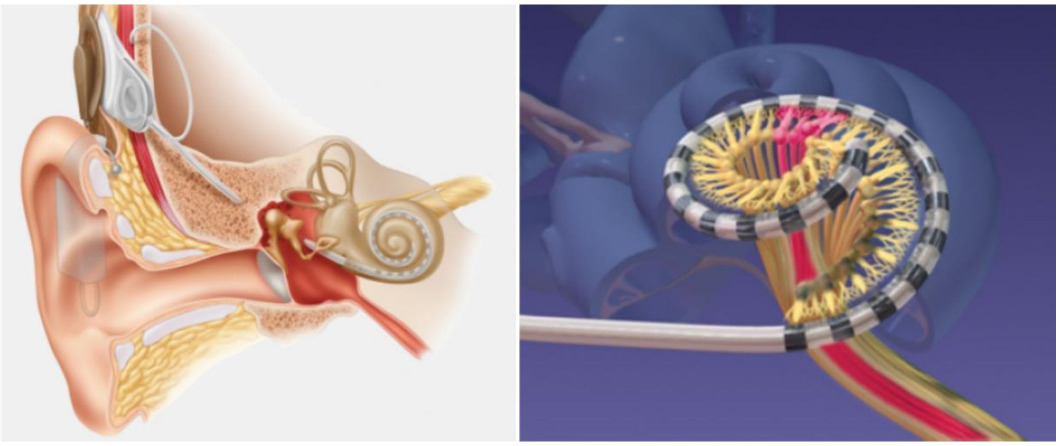


Figure 32. Cochlear implant system (from Cochlear Company) [33].

According to the science the cochlear implant is the first artificial organ sense in the bioengineering history therefore it can be considered as one of the most important success of the modern medicine [20]. Nowadays, this is a safe and effective solution adopted for childhood hearing loss which helps to improve language and communication. It is a medical electronic device (Fig. 32) properly designed for the treatment of severe to profound SNHL which is caused, as it has been formerly described, by a disfunction of the hair cells in the inner ear or by a damage of the auditory nerve. The working principle is similar to that of traditional hearing aids, in fact, it is characterized by a system which collects external sound and converts them in electric signal; however, in this case, the signals are directly transmitted to the nerves around the cochlea and the brain interprets them as a sound. Therefore, while hearing aids, used for less severe cases of deafness, seek to promote normal ear function by amplifying the signal and exploiting the residual cochlear function, cochlear implantations aim to completely replace the cochlea [33].

4.2 STRUCTURE AND FUNCTIONING OF THE DEVICE

4.2.1 Main component of a cochlear implant

The structure of the device can be separated in two parts: **internal** and **outer**. The former (Fig. 33) is surgically implanted and consists in [34]:

- *Receiving unit* made by a receiver coil with a magnet
- *Stimulation unit* represented by some current generators
- *Array of electrodes*

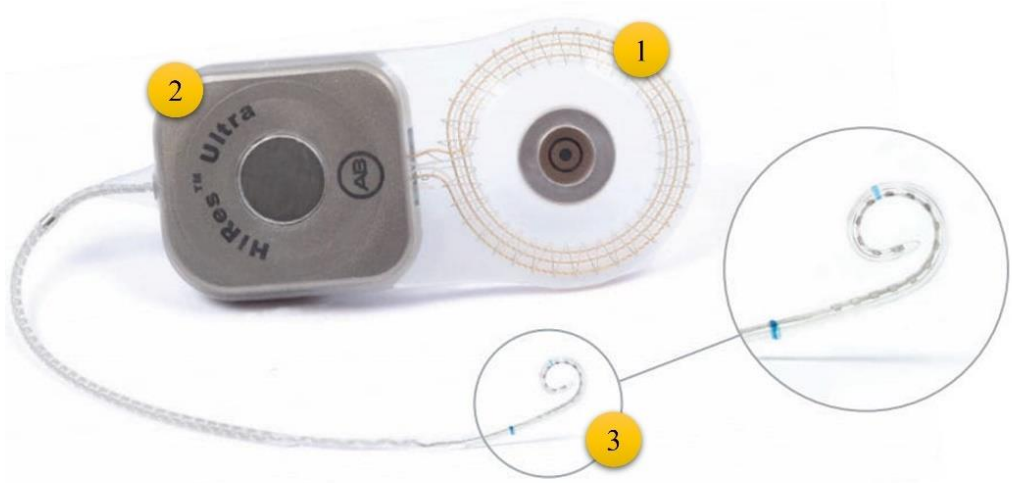


Figure 33. Internal module of a cochlear implant: receiving unit (1), stimulation unit (2), array of electrodes (3) [35].

The stimulation unit, surgically inserted subcutaneously, acts as a receiver/stimulator module which receives information about the sounds from the external antenna, then it sends, through some platinum wires, to the micro-electrodes for the stimulation of the cochlea. The size and the dimension of this circuit depend on the type of the implant and the physical characteristics of the subject [36]. The connection between the internal receiver coil and the external antenna is transcutaneous and based on radiofrequency; it is bidirectional; this means that it is suitable for the electrical stimulation of the cochlear nerve, but it can be also used to control what happens on the interface between the acoustic nerve and the electrodes. Furthermore, this configuration allows the recording of action potentials which are essential for the regulation of the implant [34].

The external part is typically composed by [34]:

- *Processing unit*: microphone and processor
- *Control unit:* battery compartment and controls
- *Transmission unit:* external antenna and magnet

The following picture shows more in detail the most important elements of the external module.

Figure 34. External module of a cochlear implant [37].

Here, there will be described the characteristics of each component [38]:

1. ***Battery*** supplies energy to the processor unit. It can be a standard version that uses disposable batteries or a rechargeable one; this module consists of a battery container and a battery cover.
2. ***Coil cord*** which interconnects the coil and the processor.
3. ***Microphone*** captures sound, converts it into an electrical signal and sends it to the processor.
4. ***Coil*** is directly in contact with the scalp, it is equipped with a specific magnetic circuit that ensures the correct positioning respect to the implanted subcutaneous part. This configuration is essential for the transmission and the receiving processes between external and internal modules.
5. ***Volume button*** thanks to which is possible to regulate the intensity of the sound and so the amount of amplification of the acoustic input.
6. ***Earhook*** has the task of supporting the entire external device maintaining the processor in right position.

7. ***Processor*** is the functional unit of the device; it elaborates the sound received from the external antenna, decodes, converts in an electrical signal, and then brings it to the stimulation unit.

4.2.3 Working principle

The figure 35 schematize the typical functioning of a CI; in order the steps are [39]:

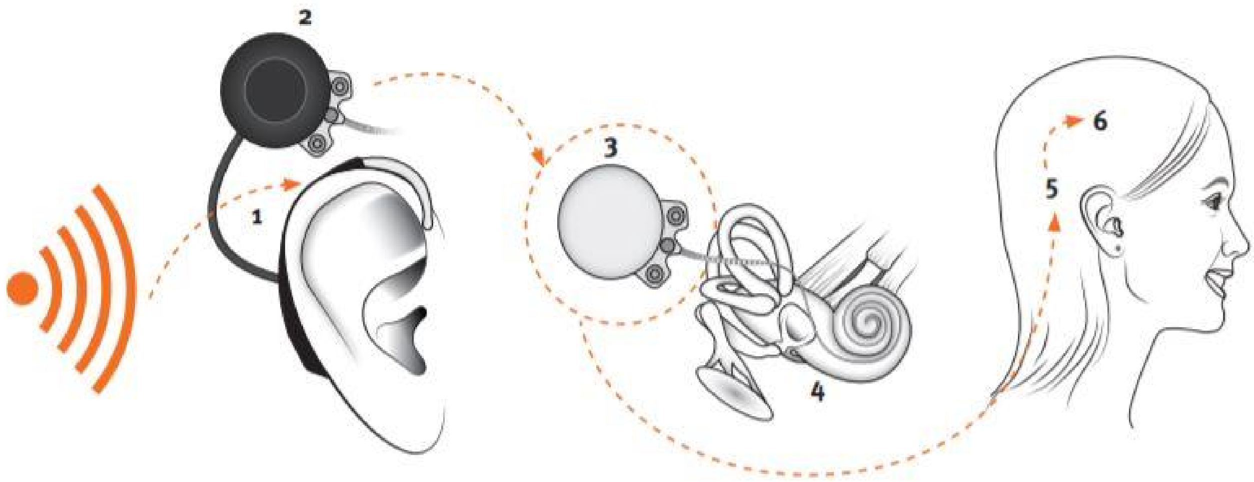


Figure 35. Working principle for a cochlear implant [39]

1. The processor collects the sound and converts it in a digital signal
2. The coil receives the information from the processor and sends them to the subcutaneous receiver unit.
3. The internal magnetic receiver, paired with the external coil, decodes the information in an electrical signal which is used as stimulus to activate the hearing functioning of the cochlea.
4. The micro-electrodes system is contained in an appropriate holder inside the cochlea. It is possible to transmit different forms of impulses because each electrode is associated with a specific frequency.
5. The codified signal is transmitted to the microelectrode array, then he vestibule-cochlear nerve the nerve is stressed and responds accordingly to the nature of the stimulus.
6. The brain receives the information about the sound and interpreters it. The resulting perception depends on number and type of neurons activate [20].

The data processing and the type of conversion performed by the processor is regulated by software thanks to which is possible to define custom programs for each user. In fact, the installation of the device requires a very detailed procedure; the audiologist must set up the audio-processor to adapt it to the patient's physiology. Using the software, a sort of univocal *sound map* is created on the basis of which the limits of each electrode are established to guarantee a cleaner and more comfortable listening as possible [39].

4.3 TECHNICAL FEATURES

According to the points that have been named, each implant is unique and specific to its user, however, the main elements, that influence the technical features of the device, are related to ***electrode, type of stimulation, connection*** between the implanted part and the external part, and ***signal processing*** [20]. Here, there will be analysed the most relevant aspects about them.

4.3.1 Electrode

Electrodes are probably the most influent parameter in a CI design. The **number** of electrodes defines the type of transmission and the structure of the system. According to this, a cochlear implant can have *single channel* (Fig 36a) or *multi-channel* (Fig. 36b) configuration.

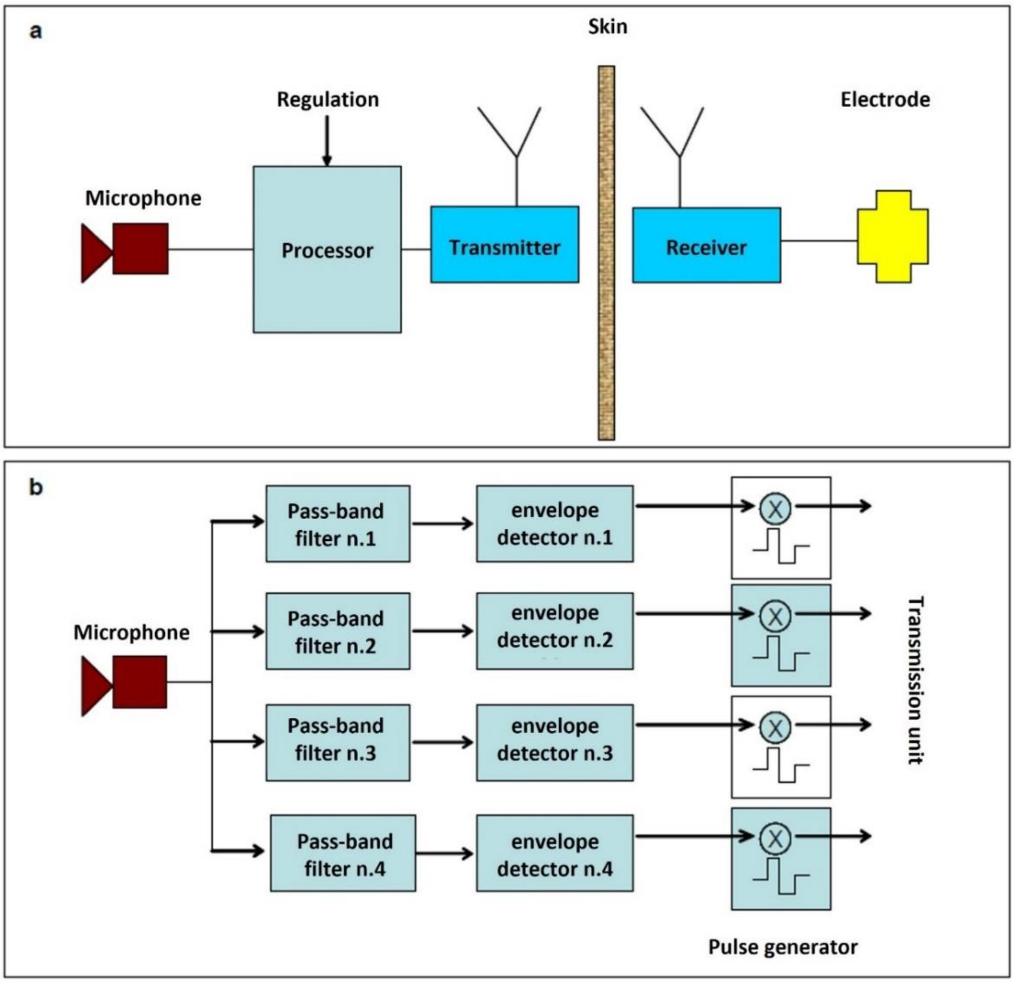


Figure 36. Cochlear implant: single channel (38a) and multi-channel (38b) configurations [20].

The former uses only one channel to collect information, transmit and decode them for the stimulation. The multichannel design, instead, is characterized by an array of electrodes which is insert in the cochlea in order to respect the tonotipicity of this structure. In fact, according to "*The Place Theory*", described in chapter 2.2, each region of the cochlea has its resonant frequency and leads to a different hearing perception. Consequently, the electrodes near the base of the cochlea are stimulated at high frequency, while those near the apex are stimulated at low frequency. In figure 36 is shown an example for a 4-channel structure; The sound is conditioned by 4 bandpass filters, the current pulses generated are proportional to the energy of each channel and transmitted to the 4 electrodes using radiofrequency. Therefore, multi-channel structure is typically more complex than the single channel, and it requires a compound network for the signal processing, but it is also more versatile, and it allows to obtain a more precise stimulation [20].

Positioning is essential to have a correct stimulation. There are three possibilities, electrodes can be placed *extracochlearly* around the round window, *intracochlearly* on the tympanic ramp, or *on the surface* of the cochlear nucleus. Nowadays, the intracochlear solution is the most used strategy because it permits to respect and preserve the mechanism of the tonotipicity, the practice recommends the insertion of an electrode array at a depth of about 20-30 mm in the cochlea [20].

The *frequency resolution* depends on the number of electrodes but especially on the **distance** between them. It is important to consider that the resulting perception is influenced by the number of survived neurons in the cochlea but also by the extension of the cochlear zones related to a specific stimulation. The electrical signal, which is transmitted by the device, tends to distribute symmetrically respect to the point of excitation. As a result, the current stimulus does not involve only a single, isolated site of auditory neurons, but several. This phenomenon is more present in the monopolar configuration. because the active electrode is distant from the reference electrode, which acts as a ground for all electrodes (Fig. 37). On the contrary, in the bipolar configuration, the active and reference electrodes are very close, this produces a more localized stimulation area [20,40].

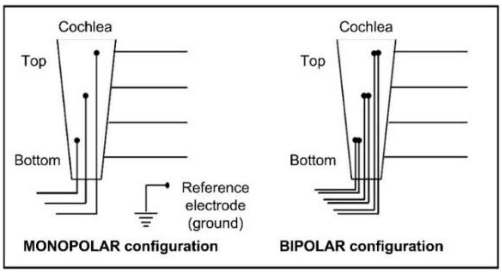


Figure 37. Monopolar and Bipolar configurations for the electrode [20].

4.3.2 Stimulation

Two types of stimulation are possible: **analogic** and **pulsed**. In the former an analogic and electric signal with the same characteristics of the incident acoustic stimulus is sent to the electrode. In multi-channel configuration the signal is efficiently filtered in each channel and then leads to the matching electrode. However, the simultaneous stimulation creates an interaction between the channels which may modify the waveform of the original signal. This problem is not present in the pulsed stimulation in which the information is transmitted in the form of a train of impulses, which can be sent simultaneously to the electrodes but without overlapping, minimizing the interaction between adjacent channels [20].

4.3.3 Connection between the implanted and the external parts

There are two methods of transmitting signals from the processor to the electrodes:

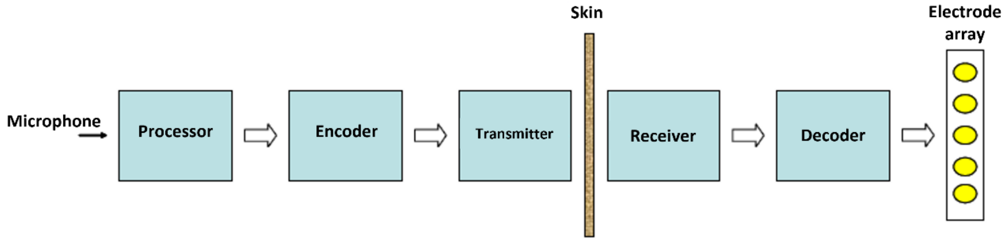


Figure 38. Transcutaneous connection [20].

- **transcutaneous transmission** (Fig. 38) is based on a radiofrequency connection between the external coil and the internal coil which sends stimuli to the electrodes. It is a very safely solution because the positioning between the transmitter and the receiver is guaranteed by a magnetic coupling; however, if there is a problem with the implant it's necessary to intervene surgically modifying the system, more over it creates problem in magnetic resonance imaging.
- **percutaneous transmission** (Fig 39) in which the internal and external components are directly connected; This defines a more versatile and flexible system, in fact, electrical problems can be solved without surgical intervention, and it is also possible to update the processing software in a very short time.

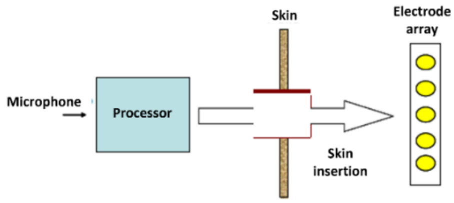


Figure 39. Percutaneous connection [20].

4.3.4 Signal processing strategies

The purpose of the cochlear implant is to help the user hearing all sounds coming from the surrounding environment but, obviously, speech signal is the element on which this system is focused because it's essential for everyday life, social relationship and learning process. From an acoustic point of view, as it has been described in detail before, the speech signal has three relevant parameters: *loudness* which represents the power of the signal, *pitch* corresponding to the key frequency in the signal, and the timbre indicating the number of harmonics greater than the key frequency, and which carries the psychoacoustic information.

To extract the significant elements about the message and the content, the signal is not analysed in terms of energy or mean value, but it is usually divided in window of 10-20 ms and then processed to retrieve temporal and spectral information [20]. The types of algorithms used to process the signal differ according to the structure (single or multi channels), model and manufacturing company. Being this aspect complex and wide, it results difficult to elaborate further on within this thesis.

4.4 COCHLEAR IMPLANT MANUFACTURING

The current manufacturing companies of cochlear implants are **Medel**, **Cochlear** e **Advanced Bionics.** Each industry proposes different models and uses specific strategies to produce CI system. The main parameters that influence the performance of the device can be:

- speech processing method
- number of stimulation channels
- technical features
- material
- design

However, all the solutions, adopted and marketed so far, have good adaptability and biocompatibility, and they are very useful to improve the communication and people's experience with hearing impairments. Developers and manufacturers always work with the aim of updating and perfecting their products, but currently the most used models are based on ***Cochlear Nucleus 7***, ***Advanced Bionics Naida CI Q90***, and ***Medel Sonnet*** processors. The main characteristics about them are respectively shown in figures. 40, 41, 42. The children involved in this study all have this type of devices.

Cochlear Nucleus® 7

- **Weight and Dimensions:**
  - 9.8 g NUCLEUS® 7 audio processor
  - Coil: 3.9 g, 30.3 mm x 5.8 mm (diameter x depth)
  - Volume: 3.9 cm^3
  - Durable spiral cables available in three lengths:6.8 cm (2.67 in.), 11 cm (4.33 in.), 25 cm (9.84 in.)
- **Power supply:** 2 x type 675 zinc-air hearing aid batteries
- Hardware:
  - Digital signal processing
  - Various programmable parameters
  - Up to 17 bandpasses
  - Frequency range from 250 to 7,000 Hz
  - Automatic Gain Control
  - BAHA FITTING system software
  - Cochlear Remote Control

- **Audio Input**
  - Sensitivity -3.2 dBV
  - Impedance 80% measured according to EN 45502-2-3
- **Temperature and humidity range**
  - Operating temperature range: 5°C to 40 °C
  - Storage temperature range: -10°C to 55°C
  - Relative humidity range: 0% to 90%
- **Radio frequency (RF) connection**
  - Receive frequency band 2.4 kHz
- **Connectivity**
  - FM cables
  - TV Streamer
  - Nucleus smart App
  - Roger™ 20 system-
  - Telecoil
  - Mini Microphone

- **Product features**
  - Dual microphone
  - Data Recording
  - Omni-directional microphone-
  - Two-way wireless communication capability
- **Cost:** 20.466,40 €
  - Internal part: 11.100 €
  - External part: 8.560 €

Figure 40 Main features for Cochlear Nucleus [37].

Advanced Bionics Naida CI Q90

- **Weight and Dimensions:**
  - 11 g NAIDA CI Q90 audio processor
  - Coil: 31 mm x 3.7 mm (diameter x depth)
  - Permanent magnet 1.5 mm x 2 mm (diameter x length)
  - Volume 4.8 cm^3
- **Power supply:** 4 hearing aid batteries type 675 zinc-air
- **Hardware:**
  - Fully digital signal processing
  - Various programmable parameters
  - Up to 120 bandpasses
  - Frequency range from 150 Hz up to 10,000 Hz
  - Configurable Automatic Gain Control
  - CLARION system software
  - AB myPilot remote control

- **Audio Input**
  - Programmable sensitivity up to +/- 10 dBm (corresponds to 96 dB)
- **Temperature and humidity range**
  - emperature range during use 0°C to 45°C
  - Storage temperature range -20°C (-4°F) to 55°C (131°F)
  - Relative humidity range 0% to 95%
- **Radio frequency (RF) connection**
  - Receive frequency band 17.4 kHz

- **Connectivity**
  - Roger System
  - Phone coil
  - Phonek DECT Phone-Easy Call
    (send cordless and mobile phone calls directly to the processor)
  - Phonak TVLink II
  - Listening Check (portable diagnostic device for checking acoustici nput)
  - FM Systems Mini Microphone
- **Product features**
  - Dual microphones (front and rear of processor)
  - Omni-directional microphone
  - Wind Noise Reduction
  - Automatic Gain Control with Cambridge Dual Loop AGC
  - Full range of wirelles technology for audio streaming
  - Soud-relax (developed to attenuate loud and impulsive noises)
  - Data recording
- **Cost:** 21.840 €
  - Internal part: 12.700 €
  - External part: 8.300 €

Figure 41. Main features for Advanced Bionics Naida CI Q90 [41].

Medel Sonnet

- **Weight and Dimensions:**
  - 8.1 g SONNET audio processor
  - DL reel: 4.6 g, 32.8 mm x 5.8 mm
  - DL coil magnets available in five strengths
  - Durable spiral cables available in three lengths:6.5 cm (2.6 inches), 9 cm (3.5 inches), 28 cm (11.0 inches)
- **Power supply:** 3 x type 675 zinc-air hearing aid batteries
- **Hardware:**
  - Fully digital signal processing
  - Various programmable parameters
  - Programmable non-linear amplification
  - Up to 12 bandpasses
  - Frequency range: up to 10,000 Hz
  - Configurable Automatic Gain Control
  - Audio processor self-diagnosis
  - MAX Programming Interface
  - MAESTRO System Software
  - Remote Control FineTuner

- **Audio Input**
  - Sensitivity -61.14 dBV (corresponds to 70 dB SPL at 1 kHz)
  - Impedance 2.9 kΩProgrammable sensitivity up to +/- 10 dBm (corresponds to 96 dB)
- **Temperature and humidity range**
  - Temperature range during use 10 °C to 45 °C
  - Storage temperature range -20°C to 60°C
  - Relative humidity range 10% to 90%**Radio frequency (RF) connection**
  - Receive frequency band 17.4 kHz
- **Connectivity**
  - 2.4 GHz wireless chip for future applications
  - Roger System
  - Telecoil
  - Direct audio input
  - FM systems

- **Product features**
  - - Dual microphone
  - - Data recording
  - - Automatic Sound Management 2.0 (ASM 2.0)
  - - Microphone directionality
  - - Wind noise reduction
  - - Automatic volume control with AGC Dual-Loop
  - - Tamper-proof design with integrated child safety features
  - - Coil connection check function with LED indicator
- **Cost:** 21.320 €
  - Internal part: 12.859 €
  - External part: 7.941 €

Figure 42. Main features for Medel Sonnet [42].

5. FACIAL MASKS

5.1 MAIN FEATURES OF MASKS

The facial masks can be classified in two different categories: *surgical masks* and *personal protective equipment* (PPE). The formers are recommended for healthcare workers, doctors and nurses and serve to protect patients within a hospital facility. On the contrary, the PPE (such as FFP1, FFP2, FFP, etc.) have been designed to protect hospital staff from external contaminations. Usually, the surgical mask is blue or white and disposable; PPE can protect the user and people in the surrounding environment (FFP2) but it can also have a breathing valve (FFP1) and in this case it protects only the wearer. The figure 41 illustrates most common models of facial masks in the market.

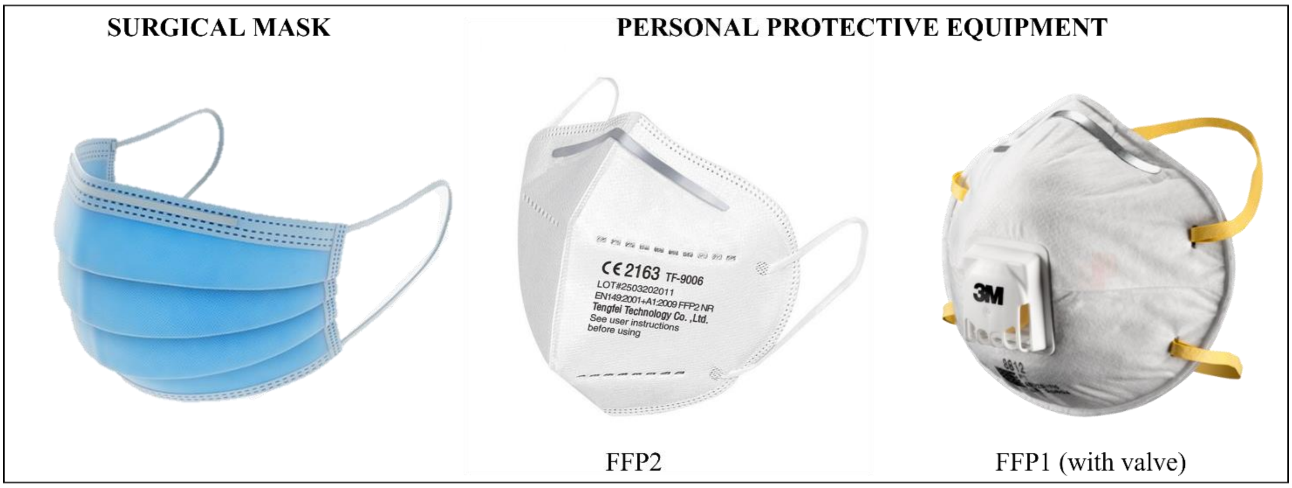


Figure 43. From left to right: Surgical mask, PPE without valve, PPE with valve.

The traditional composition of a facial mask is a combination of different types of fabric, suitably arranged to form a multilayer structure. The materials used may be natural, such as cotton, or synthetic, mainly of polymeric origin, polypropylene, polyester, and polyurethane are the most used. Typically, the outside layer of a surgical mask exposed to the air is made by of a spun bond type material, often used in the industrial or automotive fields, which gives mechanical resistance and hydrophobic propriety to the mask. The intermediate layer, characterized by a nonwoven fabric structure made of microfibres, performs the filtering function. The filtering capacity (FC) of these masks is almost total towards the outside (over 95% for bacteria), while they have a reduced filtering capacity from the outside towards the wearer of about 20%, mainly due to the poor adherence to the face. The PPE masks have a great FC and are often adopted in infectious diseases wards in hospitals. In these models, the outer layer of the mask protects against larger particles, the middle layer filters out smaller particles. The inner layer, in contact with the face, has the dual function of maintaining the shape of the mask and preserving the mask from moisture produced by breathing, coughing or sneezing. Furthermore, the filtering process is also facilitated by an electrostatic mechanism whereby charged fibres attract and capture particles. These masks have a filtering capacity of 73% for FFP1, 92% for FFP2 and 98% for FFP3.

5.2 PROBLEMS ASSOCIATED WITH THE USE OF MASKS

An important aspect, from an acoustic point of view, is that fabrics are characterised by a highly porous internal structure; the presence of these pores leads to a decrease and damping of the sound wave due to friction phenomena; this is one of the main reasons why fabric is often used for sound absorption. However, the porous material has a higher absorption coefficient for high frequencies range, therefore, masks act as **low-pass filters**, attenuating the mid-high frequencies of the voice, which are the most important in verbal communication and comprehension. The range frequency, in which the most important element of speech content lies, is about 0.5-4 .5 kHz; on the contrary, the range, for which the filtering action of the mask is relevant, is 2-8 kHz. It is therefore inevitable for some information to get lost and the SI to decreases. On top of that, classrooms normally don't have particularly favourable acoustic conditions, in terms of noise or reverberation, therefore the SI is generally already suboptimal [43].

A further issue is represented by the fact that traditional masks cover the bottom part of the face, mostly nose and mouth, which are the principal element involved in the contagion due to air transmission. Several speech-language studies have shown that, in silent or noisy conditions, language perception is better when accompanied by audio-visual inputs. In particular, the lack of visual stimuli is detrimental to communication, especially for people with hearing loss. Often, one of the most frequently used techniques in cases of poor hearing is **speechreading** or **lipreading**. This is a therapeutic strategy to improve communication by using body language and observing the speaker's facial expressions. A study conducted by Preminger et al. in 1988 [44] showed that when the nose and mouth are covered, the interpretation of visas can become complex. A possible solution to this problem could be the use of transparent masks (Fig. 42) that can facilitate conversation. The work " *The effect of conventional and transparent surgical masks on speech understanding in individuals with and without hearing loss*", published in the Journal of the American Academy of Audiology, highlights that a subject with hearing disorders has more difficulty understanding in the absence of visual stimuli, and at the same time, showed that the use of a transparent mask does not negatively affect the linguistic perception [45].

Figure 44. Transparent facial mask used by deafness or hearing impairment people.

Moreover, the attenuation introduced by the masks is compounded by the problems associated with the noise condition. Some studies have shown that noise has a greater negative impact on intelligibility when the speaker wears a mask. However, experimental results have shown that the attenuation by the mask may be negligible compared to the influence of noise on speech comprehension [46].

6. RESEARCH PROJECT DESCRIPTION

6.1 PRESENTATION AND PURPOSE OF THE RESEARCH

This project is the result of a synergic collaboration between doctors, researchers, and professionals in the field of acoustics; it has been elaborated by the Energy Department of the **Polytechnic of Turin**, in collaboration with the **University of Illinois**, the **Martini Hospital** and the non-profit association "**Ciao Ci Sentiamo**" in Turin. Furthermore, it can also be considered as one of the most innovative works to test acoustic perception remotely and not in direct contact with the participants.

The study focuses on a very topical issue of worldwide interest; it aims to assess how the recent pandemic, due to the spread of the Covid-19, has inevitably affected the population's lifestyle, not only in terms of habits, but also in education. Health regulations prescribe the use of personal protective equipment (PPE), often commonly called *facial masks*, to reduce infenctions and limit the spread of the disease. Face masks are indeed extremely useful in preventing transmission by air, but, at the same time, they can make communication more difficult. In fact, the face masks cover the mouth, preventing the listener from lip reading, and therefore hinder vocal emission and act on the speech signal by attenuating it [47]. The attenuation degree, which is usually around 3-4 % [48], depends on the physical characteristics of the mask and its materials; obviously the filtering capacity and breathability rely on these aspects as well [1].

In this type of investigation, the ***speech intelligibility*** (SI), defined as the percentage of words correctly understood by the listener, is the main parameter to be measured. According to different studies, in fact, SI should be lower if the speaker wears a mask [47]. Consequently, a specific audiometric test in Italian language, was implemented to assess intelligibility and subjective listening difficulty for different acoustic conditions, in terms of noise and types of masks. In this case, the participants involved in this study are hearing impaired ***children with cochlear implant***, and therefore, the test must necessarily meet specific requirements to be suitable, reliable and provide an objective assessment of children's speech perception abilities [4].

In conclusion, the aim of this study is to improve the perception and acoustic sensitivity of children with hearing disabilities, particularly in environments where communication plays an essential role. Therefore, depending on the results of this work, it is possible to make changes to the processing system of the cochlear implant, for example by varying the gain as to compensate the loss of information, due to the listening difficulty. Moreover, from a more architectural point of view, the design choices for the classroom's constructions can be functionalized in order to make their acoustic characteristics closer to the needs of children, and thus to ensure a sufficiently high intelligibility of speech.

In addition, the work may prove useful not only to make strictly engineering considerations, related to the IC or the design of the environments, but also to propose some strategic solutions to improve the comprehension in the classroom during the lessons. For instance, the research may help to understand which type of mask has the least impact on speech signal and is therefore advisable for teachers to wear. Furthermore, it could be useful to evaluate the organization of spaces, the arrangement of the desks, the child-teacher distance and other aspects that may seem less influential but that, if analysed carefully, could help children, and simplify the school education program.

6.2 PARTICIPANTS

6.2.1 Recruitment method

Twenty children between 7 and 15 years of age (M = 10.05, SD = 2.46) were enrolled in this study. They were divided into two groups: an **experimental group** (EG) of children with hearing problems and cochlear implants, and a **control group** (CG) of normal hearing children. The recruitment process of potential listeners was entrusted to the hospital staff of the Otolaryngology Department of the Martini Hospital in Turin. Patrizia Consolino, MD, head of the audiology outpatients' department, and a medical specialist with professional skills in the diagnosis and early treatment of childhood sensorineural and transmissive deafness, was responsible for selecting the participants. In fact, with the collaboration of the non-profit association "C.I.A.O Ci Sentiamo", she dealt with the identification, among her patients, of the children whose profile was deemed compatible with what required by the project. After having asked the parents for their consent, the data of each participant were provided to the research team in a completely *anonymous* way. 14 members of the experimental group were recruited using this method, the remaining 6 forming the control group were selected from the siblings of EG. The listeners were selected without distinction of sex or age. The experimental group consisted of 8 males and 6 females, the control group of 2 males and 4 females.

The criteria for the right choice of the participants for EG are here described below.

INCLUSION CRITERIA

1. Ability to speak and understand the Italian language
2. Binaural Stimulation:
    - Bilateral cochlear implant

- Bimodal stimulation with cochlear implant and contralateral hearing aid.

3. Accelerated diagnosis and treatment pathway.
4. Full phonetic and phonological levels.
5. Auditory perception test in binaural mode with words, phrases, and phonemic confusions (in order to verify optimal perceptual competence).
6. Vocal audiometric examination with achievement of the 50 dB HL threshold of cognition.

<u>EXCLUSION CRITERIA</u>

1. Attention, concentration, and speech disorders.
2. Cognitive impairment.
3. Poor functional gain

As for the EG, also for the control group it is of the outmost importance that the child be able to speak and understand the Italian language, has normal hearing abilities and does not suffer from any language problems, attention deficit or learning difficulties. However, their abilities have not been certified in any way, as they proceeded to self-assess themselves.

6.2.2 Information data about the children

During the recruitment phase, the participant's data about the physiological condition and biographical information are collected.

ASL
CITTÀ DI TORINO

SERVIZIO SANITARIO NAZIONALE
REGIONE PIEMONTE
Azienda Sanitaria Locale "Città di Torino"
[redacted]

S.C. Direzione Sanitaria Presidio Ospedaliero Martini
[redacted]

Centro Infantile di Audiologia, ed Otologia
Centro Impianti Cocleari
[redacted]

CODICE IDENTIFICATIVO PZ: ____

DATA DI NASCITA: ___/___/_____
ETA': _____
CLASSE FREQUENTATA: ____________________

DIAGNOSI:

ETA' DEL PZ. ALLA DIAGNOSI: ____________
EZIOPATOGENESI:

TIPO DI STIMOLAZIONE BINAURALE:
☐ IC BILATERALE
☐ IC+ P.A.
Raggiungimento della soglia di intellezione a ___ dBHL
SPECIFICHE IC/P.A. DX:

SPECIFICHE IC/P.A. SN:

ETA' DEL PRIMO INTERVENTO DI IC: ____________________
ETA' DEL SECONDO INTERVENTO DI IC:

UTILIZZO DEL SISTEMA FM:

REGIONE PIEMONTE

☐ SI ____________________
☐ NO

LIVELLO FONOLOGICO:
☐ COMPLETO
☐ INCOMPLETO

RITARDO COGNITIVO:
☐ SI
☐ NO
ALTRI DISTURBI ASSOCIATI: -

DATA DELLA SOMMINISTRAZIONE: ___/___/_____

Dott.ssa Patrizia Consolino
Audiologo Foniatra

REGIONE PIEMONTE

Figure 45. Anamnestic sheet containing the patient's biographical and medical information.

Data acquisition is based on the compilation of an **anamnestic form**, carried out in Italian, specific for each patient. The figure 43 shows a facsimile of this form in which sensitive data have been appropriately obscured. The most advantageous aspect of this data collection method is the possibility of anonymising the test. This is achieved by assigning each patient a numerical identification code. The combination code-identity of the participant is confidential and accessible only to the project manager and some members of the research team to supervise the test. To simplify the data organisation during the project, this numerical code is paired with a letter identifying the type of participant: the letter "**p**" indicates a subject belonging to the experimental group; the letter "**c**" is used for the control group.

In addition to the information collected through this form, other parameters describing the subject's medical profile in more detail are added (regarding the *type of treatment*, *CI model*, *installation date*, *device transmission* etc.). The following table summarises available information of the tested sample. For each of them, a brief description is given, in order to facilitate the interpretation and subsequent processing of the data.

Table 4. Main parameters and information data about the participants.

PARAMETERS	DESCRIPTION
Identification code	It's a random **alpha-numeric code** which identifies a subject; it is assigned by the Martini Hospital.
Type of group	It is represented by a **letter** and gives information about the group. If the subject belongs to the experimental group, is indicated with **p** (it reminds to patient), if he/she belongs to the control group, the letter **c** is used.
Age	It's a **numerical value** wich expresses the age of the listener in ***years.***
Diagnosis	It's a **numerical value** that expresses, in terms of ***months***, when the problem was diagnosed.
Prothesis	It's a **numerical value** that expresses, in terms of ***months***, when the child was prosthetised.
Installation date 1° IC	Date, expressed as ***dd/mm/yy***, for the first cochlear implant.
Installation date 2° IC	Date, expressed as ***dd/mm/yy***, for the second cochlear implant.
Type of treatment for the right ear	It indicates the nature of treatment for the **right** ear in terms of *type of device* (CI or HA), *brand* of the device (Cochlear, MED-EL, Phonak…) and the *model*.
Type of treatment for the left ear	It indicates the nature of treatment for the **left** ear in terms of *type* of device (CI or HA), *brand* of the device (Cochlear, MED-EL, Phonak…) and the *model*.
Pure tone audiometry for the right/left ear	It's a number with one decimal place referred to the value of the *tonal gain* resulting from the pure tone audiometric test (for the right and for the left ears).
Transmission device	It is related to the transmission device used with the CI or HA.

It is crucial that the most significant parameters for the control group be only the following: *Identification code*, *Type of group*, and *Age*. Prior to the test's start, the listeners are asked to indicate in detail the *model of headphones* used. This aspect is important because different headphones have different frequency responses, introducing a confounding variable [45].

All participants were made aware of the purpose of the procedure and provided with a document containing instructions on how to run the test, however, this aspect will be discussed in more depth further on.

6.3 DESCRIPTION OF THE MASKS

Speech intelligibility is analysed by considering **three** different types of masks. Before choosing them, a detailed analysis of the physical and acoustic properties for more than 30 facial masks was carried out. The main parameters of interest, on which this preliminary investigation focused, are the degree of **acoustic attenuation**, **filtering capacity** and **breathability**. The data were collected at the polytechnic of Turin several months before the official start of the project. All measurements were run in order to minimise measurement error and limit any source of disturbance that could introduce variability into the results. The experimental set-up and the specific instrumentation vary, of course, depending on the type of quantity to examinate. Below, the procedure required for the acoustic attenuation, which is more relevant for this study, is briefly described.

ACOUSTIC ATTENUATION MEASUREMENT

The instrumentation includes a **Head And Torso Simulator** (HATS, model 4128-C by Brüel & Kjær) (Fig. 44), a half-bust dummy specific for this type of measurement, and a cheek microphone (NTi Audio MC230A + preamplifier MA220, integrated with NTi Audio XL2 Sound Level Meter). It is important, in this case, that the acoustic conditions of the environment are known and much controllable as possible. For this reason, the project manager and the research team decided to carry out this task in an **anechoic chamber** at the Polytechnic of Turin, a room with very particular structural characteristics; the walls are designed in such a way as to obtain the maximum reduction of sound reflections through the adoption of sound-absorbing panels with different shapes (wedge, pyramid, etc.).

Figure 46. Head and Torso simulator type 4128 by Brüel & Kjær [49].

The measurement process requires a precise arrangement of the instrumentation. HATS is placed inside the chamber, and the microphone is positioned at a horizontal distance of **one meter** from the mouth of the HATS (Fig. 44). After the instrumentation is correctly positioned, a **white noise** is emitted through the mouth of the HATS, the microphone collects the outgoing sound and SPL is measured using the sound level meter of the microphone. For each type of mask, **3 samples** were tested, and so the complete procedure requiring a minimum of 4 measurements: 3 for each sample and one for the unmasked condition. After the samples were acquired, they have been analysed and w frequency range as this is the most important for SI and voice production.

Figure 47. Experimental setup for acoustic attenuation measurement without mask (right) and with mask (left).

Based on the results obtained, three masks were chosen, each corresponds to a different acoustic condition. From table 5, it is evident that, unlike filtering capacity and breathability, sound attenuation has almost similar values between the masks. Therefore, it is expected that the SI are comparable and similar for all masks, however, it is important to remember that the type of mask is not the only variable, but also noise, fatigue, and other subjective factors must be considered.

Additionally, to facilitate the organisation and interpretation of data, each mask is assigned an abbreviation.

Table 5. Information data about physical and acoustic features of the selected mask.

	Abbreviation	**Breathability** (< 40 o 60 Pa/cm^2)	**Filtering capacity** (> 80 %)	**Acoustic attenuation** Δ 0.5-4 kHz oct (dB)
Mask 1	M1	96	98,5	3,33
Mask 2	M2	25	93,0	1,02
Mask 3	M3	NA	NA	1,15

The figure 46 shows the experimental setup in the anechoic chamber and the HATS wearing the three masks involved in this project. It is important that the mask is correctly positioned on the manikin's face. HATS is designed to simulate the vocal emission of the human body, with a speaker inside the mouth to reproduce the speech signal. Therefore, the mask has to cover the whole of the mouth and nose and the elastic bands must be firmly anchored over the ears to ensure that the mask adheres correctly.

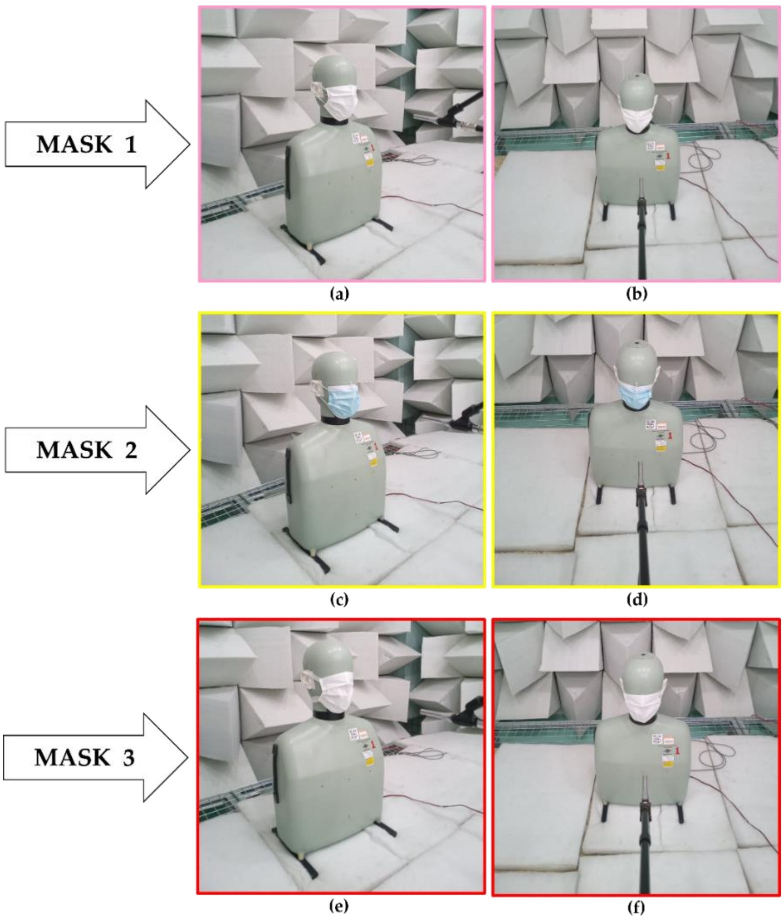


Figure 48. HATS wearing: mask 1, lateral (a) and frontal (b) views; mask 2, lateral (c) and frontal (d) views; mask 3, lateral (e) and frontal (f) views.

6.4 SPEECH MATERIAL

To the aim of the project, a specific audiometric test to assess speech intelligibility in an accurate and reliable manner is needed. It is therefore crucial for the method of administering the acoustic stimuli to be compatible with the Italian language and, on the other hand, still be suitable for more complex applications, such as this one, in which the subjects being tested are children suffering from hearing problems.

6.4.1 Matrix sentence test

The structure of the test proposed in this work is inspired by the audiometric investigation, published by Puglisi et al. in 2021 [4] and developed in cooperation with the Cluster of Excellence “Hearing 4all” of the University of Oldenburg, concerning the use of the **simplified Italian matrix test** (SiIMax) for SI measurement in cases which are difficult to analyse, such as children or adults with hearing impairment. In fact, SiIMax can be considered a simplified version of the speech *Italian matrix sentence test* (ItaMatrix) [3], in which the length of the sentences is reduced from 5 to 3 words. It was developed therefore as a response to the result of some studies, which have shown that sentences of 5 words may be too long and therefore effortful for assessing auditory abilities of children with cochlear implants or with hearing perception problems in general. The speech material, extracted from the original 50-word base matrix of the ITAMatrix test, was built considering the psychometric theories to make it is suitable for children of different ages and hearing abilities. In this way, the results are comparable with each other and independent of the subject's level of language and knowledge [4]. SiIMax has already proven to be valid for SI measurements on younger children [50-52]; moreover, some authors, including Neumann et al. [50] and Weißgerber et al. [51] have suggested using the simplified matrix not only for diagnostic purposes but also to assess the medical effectiveness of hearing aids or cochlear implants.

The speech material for this experimental study was taken from the SiIMax speech material. Thus, the **7x3 matrix of three-word sentences** was used and randomly built sentences being semantically unpredictable, phonetically balanced and with a grammatical structure fixed in the Italian form **number-object-adjective** were used. The words of the sentences were simple, of common use and easy understandable even for young children, who are subjects not having a huge vocabulary knowledge. An example of a sentence is "*Quattro matite rosse*" meaning “Four red pencils”. For reading clarity, table 6 shows all the elements (number-object-adjective) used in this work.

Table 6. Matrix word with number-objective-adjective used in this study (taken from Puglisi et al. 2021).

Number	Object	Adjective
Due	Macchine	Azzurre
Quattro	Sedie	Piccole
Cinque	Pietre	Bianche
Sette	Porte	Nere
Otto	Palle	Rosse
Nove	Tavole	Grandi
Dieci	Matite	Nuove

These words have been randomly combined to form different but semantically similar sentences. Obviously, numbers and adjectives have been declined respecting singular/plural and masculine/feminine of the objects.

6.4.2 Speech stimuli

Each sentence was read and repeated aloud by **a native Italian female** speaker with **standard Italian pronunciation**; therefore, they were recorded and stored according to the method used by Puglisi et al. in 2015 [3].

MASK CONDITIONS

Four different conditions were considered: **mask 1**, **mask 2**, **mask 3** and **un-masked** (Figure 49).

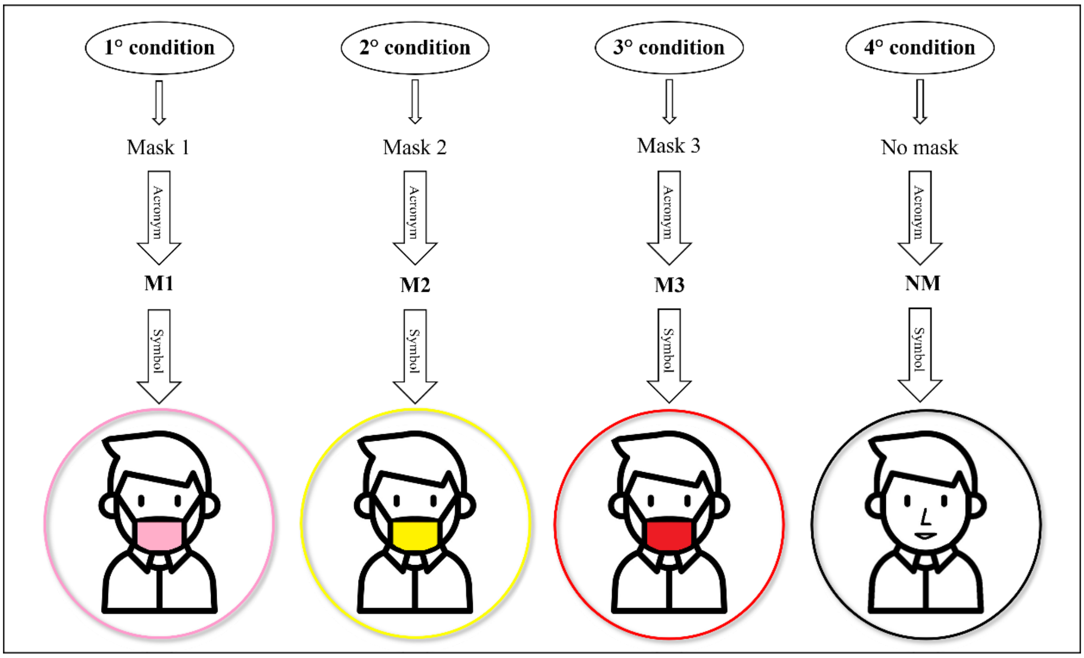


Figure 49. Scheme for the 4 mask conditions: symbols and acronyms.

The signals were processed to introduce the factor related to face masks in the way that follows: the HATS was equipped with the three masks (i.e., mask 1, mask 2, mask 3) and was set to play the sentences of SiIMax test. Recordings were perfomed in the anechoic room of Department of Energy of Politecnico di Torino in order to avoid any effects of the environment, such as reverberation and external noises, on the speech stimuli. In particular, sentences were recorded positioning the microphone at 1 meter from the mouth of the HATS, then used for convolution with the noise stimuli. To simplify the notation and the graphical representation of the results, each condition was associated with an acronym and a symbol.

NOISE CONDITIONS

The signal-to-noise ratio (SNR) is the most important parameter to analyse an acoustic signal affected by noise; this expresses the ratio between the power of the useful signal and the noise. For this application, “signal" indicates the useful information of the speech content, while "noise" means any disturbance or acoustic interference that overlaps and degrades the useful signal. The following formula represents the mathematic definition of SNR:

$$SNR = 10\log\left(\frac{P_{Speech\ signal}}{P_{noise}}\right) \tag{4}$$

Where $P_{Speech\ signal}$ and P_{noise} are respectively the power of the speech signal and the noise, usually expressed in *watt*. In this formulation SNR is expressed in dB because it is useful for practical applications. Typically, 0 dB means that useful signal and noise are equivalent, in terms of acoustic power, while an SNR greater than zero indicates that the signal is "louder" than the noise and hence the interfering signals.

Four different noise conditions were considered, represented respectively by the following abbreviations: **0**, **5**, **10**, **QUIET**. Table 7 describes in detail the main characteristics of each noise them focusing on the relation between the signal and the noise.

Table 7 Description of different noise conditions.

Noise condition	Abbreviation	Description
1° condition	**0**	It corresponds to SNR = 0 dB → the speech signal and the noise are equivalent.
2° condition	**5**	It corresponds to SNR = 5 dB → the speech signal is 5 times louder than the noise.
3° condition	**10**	It corresponds to SNR = 5 dB → the speech signal is 5 times louder than the noise.
4° condition	**QUIET**	The acoustic stimulus contains only speech contenent → no noise is mixed with the useful signal.

Mixing noise → The noise used in this work was extracted and recorded from a classroom while the children were talking freely among themselves. This type of noise, similar to a typical everyday chattering, was chosen firstly due to it being suitable for this study's main aim, secondly because it is a perfect simulation of everyday activities. The noise was mixed with the voice recordings in an additive manner using special software (*Praat*) and in such a way as to respect the desired SNR values. For each sentence the noise was digitally mixed by an operator belonging to the research team, who thanks to a simple graphic interface can choose the speech stimuli, the noise, and the SNR value (Figure 50) Therefore, the material was stored according to the test specifications.

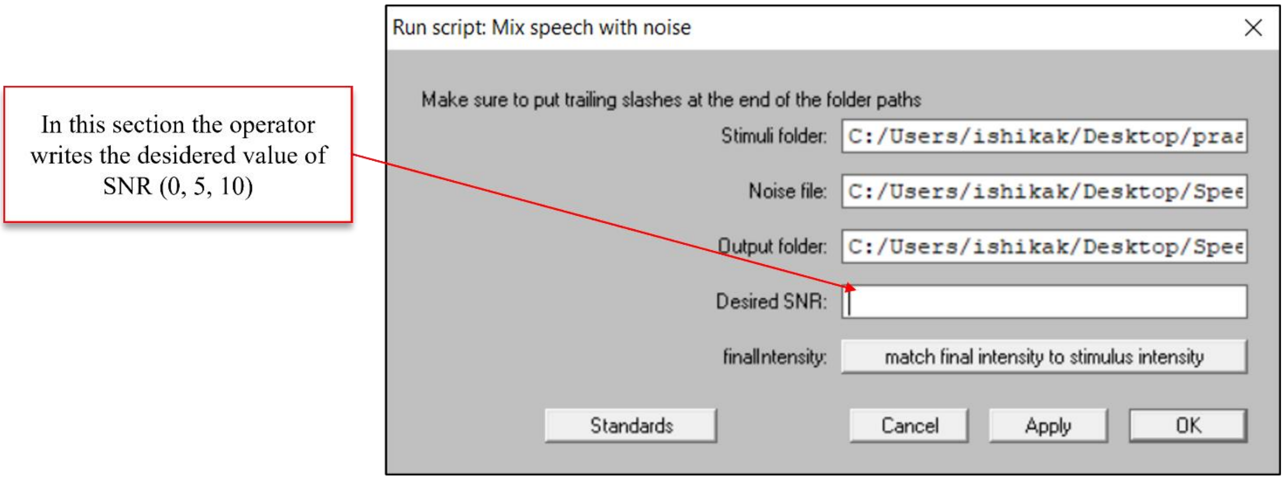


Figure 50. Praat interface for mix speech with noise

6.5 TEST STRUCTURE

The sentences were initially organised to form **14-sentences lists**. Some authors, such as Puglisi et al. in 2021 [4], have already used 14-sentences lists and demonstrated that this test structure is suitable for the assessment of SI in children with hearing problems. However, the test proposed in this project requires listeners to run the test at home in remote mode. This method of administering the test greatly influences the results, in fact, the performances of the first 5 participants showed that 14-sentences lists were too long, in terms of duration, and exhausting. The children used to lose their attention and complain with the responsible who supervised the execution of the test. Therefore, the number of sentences were reduced from 14 to 10, not considering the last 4 sentences of each list, for measurement of SI of the other listeners. The **10-sentences** lists proved to be equally reliable, in terms of the validity and stability of the results. This aspect will be further investigated and appropriately documented in the "Results" chapter.

The whole test was structured with **16 lists** each one corresponding to a different noise condition (0, 5, 10, QUIET). These were then divided into 4 subgroups relating to the 4 mask conditions (M1, M2, M3, NM). Considering all possible combinations then, each survey consisted of a total of $14 \times 16 = 224$ sentences, for the full 14-sentences version of the test, or $14 \times 10 = 140$ for the reduced 10-sentences version (Tables 8, 9, 10,11). This structure was the same for each participant but the order with which the lists are presented to the listener was different.

Table 8. Sentences of lists n. 1, 2, 3, 4 associated to M1 condition.

	0			5			10			QUIET	
	LIST n.1			**LIST n.2**			**LIST n.3**			**LIST n.4**	
Number	**Object**	**Adjective**	**Number**	**Object**	**Adjective**	**Number**	**Object**	**Adjective**	**Number**	**Object**	**Adjective**
quattro	macchine	azzurre	due	macchine	rosse	otto	palle	azzurre	nove	matite	azzurre
quattro	macchine	azzurre	quattro	pietre	nere	due	porte	piccole	cinque	matite	rosse
dieci	sedie	piccole	dieci	matite	nuove	quattro	matite	nere	sette	tavole	nuove
sette	pietre	bianche	due	matite	bianche	sette	porte	azzurre	quattro	tavole	bianche
cinque	pietre	nere	nove	tavole	bianche	dieci	macchine	rosse	due	macchine	bianche
otto	porte	rosse	cinque	palle	grandi	dieci	palle	rosse	quattro	sedie	azzurre
nove	palle	grandi	dieci	tavole	rosse	sette	tavole	bianche	dieci	palle	piccole
due	macchine	piccole	quattro	porte	piccole	nove	tavole	grandi	cinque	porte	piccole
sette	porte	nuove	otto	pietre	azzurre	quattro	sedie	grandi	otto	pietre	grandi
quattro	tavole	nuove	cinque	porte	azzurre	due	pietre	nuove	otto	palle	grandi
otto	matite	rosse	otto	sedie	nuove	cinque	sedie	bianche	nove	porte	nere
dieci	matite	nere	nove	palle	piccole	nove	pietre	nere	sette	macchine	rosse
due	tavole	bianche	sette	macchine	grandi	cinque	matite	nuove	dieci	pietre	nuove
nove	sedie	grandi	sette	sedie	nere	otto	macchine	piccole	due	sedie	nere

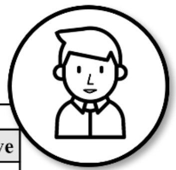
Table 9. Sentences of lists n. 5, 6, 7, 8 associated to M2 condition.

	0			5			10			QUIET	
	LIST n.5			**LIST n.6**			**LIST n.7**			**LIST n.8**	
Number	**Object**	**Adjective**	**Number**	**Object**	**Adjective**	**Number**	**Object**	**Adjective**	**Number**	**Object**	**Adjective**
dieci	sedie	grandi	cinque	porte	bianche	sette	macchine	rosse	dieci	tavole	piccole
otto	tavole	piccole	nove	porte	rosse	due	macchine	bianche	otto	pietre	nere
quattro	matite	azzurre	otto	pietre	nere	otto	pietre	grandi	nove	sedie	azzurre
due	porte	bianche	due	macchine	nuove	otto	palle	grandi	sette	pietre	grandi
quattro	porte	nuove	sette	palle	grandi	nove	matite	azzurre	quattro	matite	rosse
sette	palle	nuove	cinque	macchine	bianche	cinque	porte	piccole	quattro	sedie	nuove
cinque	pietre	rosse	quattro	sedie	nuove	cinque	matite	rosse	dieci	matite	azzurre
nove	tavole	nere	due	palle	piccole	due	sedie	nere	sette	palle	grandi
sette	matite	bianche	quattro	matite	rosse	dieci	pietre	nuove	otto	tavole	nere
nove	pietre	azzurre	sette	pietre	grandi	quattro	sedie	azzurre	due	palle	piccole
cinque	palle	rosse	dieci	matite	azzurre	quattro	tavole	bianche	cinque	macchine	bianche
otto	sedie	piccole	otto	tavole	nere	dieci	palle	piccole	nove	porte	rosse
due	macchine	nere	dieci	tavole	piccole	nove	porte	nere	due	macchine	nuove
dieci	macchine	grandi	nove	sedie	azzurre	sette	tavole	nuove	cinque	porte	bianche

Table 10. Sentences of lists n. 9,10, 11, 12 associated to M3 condition.

0			5			10			QUIET		
LIST n.9			LIST n.10			LIST n.11			LIST n.12		
Number	Object	Adjective	Number	Object	Adjective	Number	Object	Adjective	Number	Object	Adjective
sette	sedie	nere	due	macchine	nuove	otto	palle	grandi	due	tavole	grandi
due	macchine	rosse	quattro	matite	rosse	nove	porte	nere	cinque	porte	rosse
nove	palle	piccole	cinque	macchine	bianche	due	macchine	bianche	nove	matite	nere
otto	pietre	azzurre	sette	pietre	grandi	otto	pietre	grandi	sette	macchine	bianche
dieci	tavole	rosse	dieci	tavole	piccole	sette	macchine	rosse	dieci	macchine	azzurre
cinque	porte	azzurre	cinque	porte	bianche	dieci	pietre	nuove	due	sedie	bianche
dieci	matite	nuove	nove	sedie	azzurre	quattro	tavole	bianche	cinque	pietre	azzurre
quattro	pietre	nere	dieci	matite	azzurre	cinque	porte	piccole	nove	palle	nere
sette	macchine	grandi	quattro	sedie	nuove	quattro	sedie	azzurre	otto	tavole	nuove
cinque	palle	grandi	sette	palle	grandi	due	sedie	nere	dieci	porte	piccole
due	matite	bianche	nove	porte	rosse	sette	tavole	nuove	quattro	sedie	rosse
nove	tavole	bianche	due	palle	piccole	cinque	matite	rosse	otto	palle	nuove
otto	sedie	nuove	otto	tavole	nere	dieci	palle	piccole	quattro	pietre	piccole
quattro	porte	piccole	otto	pietre	nere	nove	matite	azzurre	sette	matite	grandi

Table 11. Sentences of lists n. 13, 14, 15, 16 associated to NM condition.



0			5			10			QUIET		
LIST n.13			LIST n.14			LIST n.15			LIST n.16		
Number	Object	Adjective	Number	Object	Adjective	Number	Object	Adjective	Number	Object	Adjective
otto	palle	azzurre	otto	tavole	azzurre	due	macchine	nere	dieci	macchine	grandi
dieci	macchine	rosse	nove	pietre	nuove	dieci	macchine	grandi	nove	tavole	grandi
sette	porte	azzurre	otto	porte	nuove	otto	sedie	piccole	sette	palle	nuove
due	porte	piccole	sette	tavole	piccole	nove	pietre	azzurre	otto	sedie	piccole
otto	pietre	nere	due	matite	rosse	quattro	matite	azzurre	dieci	sedie	grandi
due	pietre	nuove	quattro	matite	grandi	sette	palle	nuove	due	porte	bianche
quattro	matite	nere	dieci	palle	nere	otto	tavole	piccole	quattro	matite	azzurre
quattro	sedie	grandi	due	palle	bianche	dieci	sedie	grandi	cinque	pietre	rosse
sette	tavole	bianche	dieci	porte	nere	cinque	palle	rosse	otto	tavole	piccole
cinque	sedie	bianche	sette	macchine	piccole	nove	tavole	nere	due	macchine	nere
dieci	palle	rosse	quattro	pietre	azzurre	quattro	porte	nuove	nove	pietre	azzurre
otto	macchine	piccole	cinque	macchine	grandi	sette	matite	bianche	sette	matite	bianche
cinque	matite	nuove	cinque	sedie	rosse	due	porte	bianche	quattro	porte	nuove
nove	tavole	grandi	nove	sedie	bianche	cinque	pietre	rosse	cinque	palle	rosse

6.6 METHODOLOGY

6.6.1 Survey editing

The survey is based on the use of **Qualtrics**, a free, online survey creator software. This platform has already proven to be suitable for similar applications and for remote SI measurements. However, this survey was edited specifically for the purposes of this study and is not derived from previous versions. Some aspects have been deepened and improved to make the test as suitable as possible for children (as it can be deducible from the figure 51). In fact, the graphical interface is very user-friendly and easy to understand, therefore, the child is able to run the test independently.

Figure 51. Introductive page of the Qualtrics survey.

Each sentence corresponds to a request which the child must fulfil by performing a task. The figure 52 shows a typical Qualtrics interface for a sentence. The child, or a parent, has to click on a player to play the sentence, then the child has to listen to the sentence and repeat aloud what he/she has understood. In addition, to collect information on subjective listening difficulty (LD), a discrete 5-points scale from 1 to 5 coded with smiley faces of different colours is used. At the beginning, an operator of the research team explains the meaning of this scale and asks the child to click on the face that best suits his/her perception. After listening, repeating and choosing the face, the child can skip to the next sentence by scrolling down the page. The same procedure is repeated for each sentence until the corresponding list is finished. At the end of the page there is a blue arrow, simply clicking on it, the user is directed to the web page of the next list.

Figure 52. Qualtrics interface for a sentence.

Before the actual audiometric test, each participant must answer to some preliminary questions useful for the future data organization or for training process. The flowchart below (Figure 53) gives information on the sequence in which these questions are presented to the child.

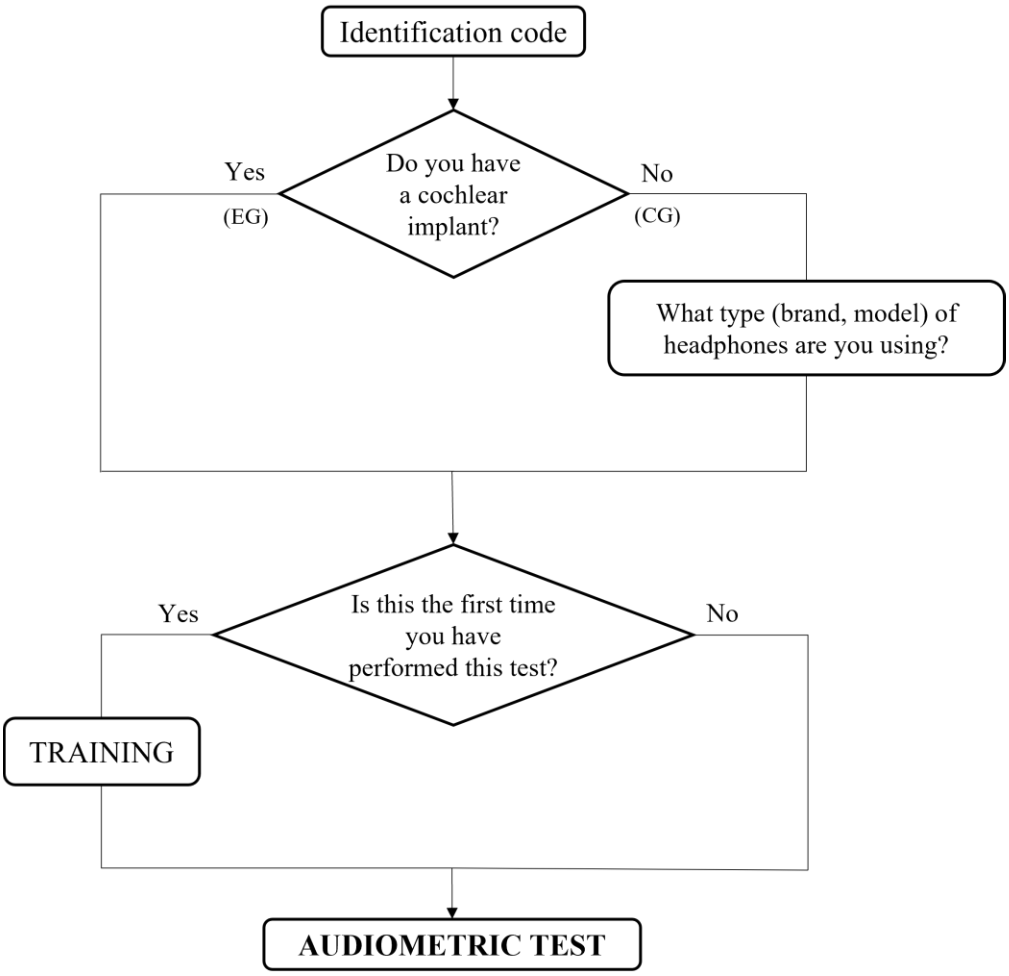


Figure 53. Flowchart for the execution of the test.

- **Identification code**: the child is asked to enter the identification code on the keyboard. In this first step, the operator indicates the code to be written (e.g."2p").
- **"Do you have a cochlear implant"**: the listener can choose between two options can choose between two options (yes or no) and consequently be classified in the EG or CG group. For the CG group additional information about the headphones is needed.
- **"Is this the first time you have performed this test?"**: if the answer to this this question is "No", the child has to run a TRAINING section through which is possible to familiarise with the test and understand how it works.
- **Audiometric test**: the participant performs the test following the guidelines indicated in the research project.

RANDOMIZATION

In addition, the Quatrics software has a very useful functionality for the purpose of this study. It allows to select a "**randomization mode**" wich can be customised directly by the test designer. This function consists of using a *randomizer block* in which only the elements to be randomised are inserted. For this specific application, it has been chosen to randomise the lists, but not the sentences within a list. This means that the order in which the lists are presented to the listener is unique and different for each subject, but the position of the sentences within a list is fixed and remains unchanged. Furthermore, Qualtrics records the randomisation order for each test, therefore it is easier for the operator to organise the data about SI values.

6.6.2 Training

As mentioned earlier, there is a preliminary **training session** before the actual start of the survey. This section, introduced by a specific image, makes clear to the subject that what he/she is going to do is a simulation (Figure 54). It becomes, therefore, fundamental to understand the task in its integrity and be able to familiarise with the test itself. During the training session, a member of the research team will explain the task to the subject and make sure he/she understands how to carry it out correctly. He will also describe the function and meaning of the "smiley faces" through practical examples, advising the subject to repeat the sentences out loud and in a comprehensible manner, as to help the child to carry out the task without efforts or incomprehension whatsoever.

Figure 54. Introductive image of the training session.

The training session has a similar structure to the rest of the test and consists of an 8-sentences list, related to the *NM mask condition*. The sentences are further on divided into 4 subgroups of two sentences each, corresponding to the noise conditions 0, 5 ,10, QUIET (Table 12).

Table 12. List of sentences for the training session.

	TRAINING LIST		
	Number	**Objective**	**Adjective**
QUIET	due	macchine	nere
	dieci	macchine	grandi
10	otto	sedie	piccole
	nove	pietre	azzure
5	quattro	matite	azzurre
	sette	palle	nuove
0	otto	tavole	piccole
	dieci	sedie	grandi

6.6.3 Procedure

The families of the participants give their consent for their children to take part of the project only after having been thoroughly informed on its main characteristics. The appointments were then made by contacting the families directly via email, which did not only contain basic knowledge on the topic, but they also included precise information on guidelines and materials which were required in order to perform the test correctly. Some of the key points communicated to the parents follow:

- A computer (tablet, mobile phone) and a stable internet connection are needed.
- The test is carried out at home in a remote mode using the "ZOOM meeting" videoconferencing platform. Participants received an email with the link to the meeting few hours before the test.
- The call is recorded but the cameras are switched off during the survey respecting the privacy of the participants.
- The environment must be as quiet as possible and free of other sources of noise (television, radio, chatting, household appliances, etc.).
- The volume of the device (computer, tablet, mobile phone) must be set at a comfortable level and never changed for the entire duration of the test.

- The cochlear implant must be used in “Roger mode”: the microlink FM device set to the mode defined and suggested by the Martini Hospital Audiology Centre for performing the test. This condition is essential to isolate the listener from the external environment.

By clicking on the ZOOM link, the listener joins an online meeting in direct contact with a researcher who supervises the test or intervenes in the event technical problems should occur. The experimenters first step is to verify that all conditions are met, making sure that the devices function correctly and are set as above indicated. The child is therefore instructed on the procedure as well. In more detail, it is fundamental for the subject to listen to the 3-word sentences and follow these instructions:

1. Repeat the 3 words out loud.
2. Click on the emoticon that corresponds to the listening difficulty (green > no difficulty; red > very difficult).
3. Skip to the next sentence

The survey starts with the training session, then the audiometric test follows. The latter is made of two parts, each composed by 8 lists of a duration of 20 minutes ca. Once this first step is completed, there is a compulsory break, which is crucial as to prevent the subject to lose focus. However, it is possible to interrupt the experiment at any time in case the child is too tired to continue. In the event the subject were too young and/or not quite familiar with the computer, it is possible for a parent to manage the test or alternatively, for the operator to request remote control of the device.

At the end of each test, a member of the research team has the task to listen to all recordings, extract and enter the SI data into an Excel file. SI values must be organized using **0/1 binary code**, giving 1 if the word was correctly repeated, considering singular/ plural and masculine/feminine, otherwise assigning 0. The working sheet will also contain data on the subjective LD, encountered by the subject, and the order of the lists, which are automatically recorded by Qualtrics. The remaining information concerning the recruitment phase are then integrated.

7. STATISTICAL ANALYSIS

7.1 EXTENDED vs. REDUCED VERSION

As already mentioned in the previous chapter, the test was initially implemented based on the work of Puglisi et al. in 2021 [4]. Thus, the first version of the audiometric test was made by lists of 14 sentences each. However, the tests on the first five EG subjects emphasised that lists of 14 sentences were too long and tiring. Therefore, being the subjects of the test young children, it ends being quite hard for them to concentrate completely on the task for a long amount of time. Moreover, it is important to consider also that the remote mode already requires a greater effort than the traditional test in presence. This issue was already foreseen in the preliminary set up of the test and, in fact, in order to limit it, the survey was divided into two parts with a compulsory break in between. However, this strategy proved to be insufficient, as the children lost focus, complaining that they were tired and often asked for a break. Consequently, the time necessary for data acquisition increased, making the situation more stressful than usual not only for the subjects, but also for the experimenter. Therefore, the team responsible for this research chose to create a reduced version of the test with 10 sentences instead of 14. The choice was based on a statistic confrontation between the extended and the reduced version of the test.

The results of a brief analysis, carried out using the Excel software, made clear that the intelligibility values obtained by the full version are comparable with those of the latter version and that there is, in fact, no statistically significant difference between the two. Table 13 shows SI values for the EG groups who ran the extended and reduced version. The mean values and standard deviation for each patient (SI_avg_P, SI_std_P) and test version (SI_avg_version, SI_std_version) are reported in percentages.

Table 13. SI values for extended and reduced test versions.

Test version	ID	SI_avg_P(%)	SI_std_P(%)	SI_avg_version (%)	SI_std_version (%)
Extended	14p	71.45	25.20	71.66	9.05
	13p	67.99	32.14		
	5p	58.84	21.69		
	26p	82.33	23.83		
	25p	77.68	17.19		
Reduced	11p	83.75	23.12	84.45	10.77
	29p	82.92	25.41		
	23p	88.54	11.02		
	34p	86.25	17.88		
	43p	94.79	9.73		
	7p	86.04	15.60		
	16p	89.68	10.21		
	8p	90.63	11.69		
	38p	57.43	38.38		

As can easily be deduced from the table and from the graph below (Figure 55), the SI of the reduced version is on average higher than those of the extended one. The experimenters who supervised the survey can testify that this phenomenon is probably linked to fatigue and to children losing focus, in the extended version, in order to answer quickly and finish first.

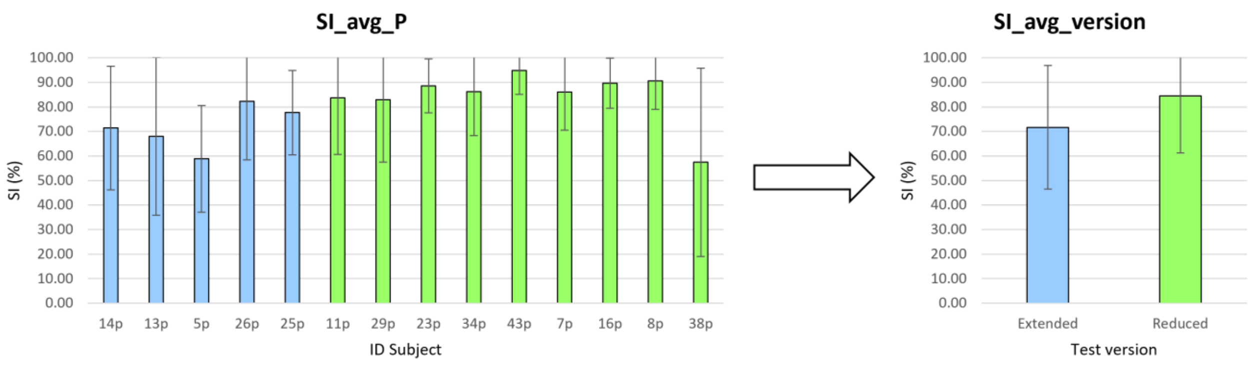


Figure 55. Bar diagrams of SI for each patient (left) and test version (right).

In conclusion, the reduced version proved to be reliable and suitable to fulfil the purpose of this work. In fact, in this case children were able to complete the test in a much shorter amount of time and some without even having to ask for a break.

7.1 STATISTICAL ANALYSIS METHODS

After acquiring the data and completing the tests for all participants, a statistical analysis was carried out. Both SI values and subjective LD scores were investigated, taking into account the different noise levels, type of mask and also factors such as sex, age or fatigue. In this chapter all methods and tools used to interpretate the data, are described.

The strategy used for processing the data is inspired from the Linear mixed-effects models using Eigen and S4 (lme4) proposed by Bates in 2015 [53]. In fact, **generalized linear mixed models** (GLMM) were implemented, using the software R3.6.0 and the lme4 package, to obtain a solid statistical evaluation. During the data acquisition phase, the information about the SI was recorded using a 0/1 binary code. As a consequence, for the SI assessment, a GLMM with a binomial distribution, according to Laplace approximation, was used to represent the binary outcome variable. In this analysis four independent variables were considered: ***Patients Vs Control Group***; ***Mask***; ***SNR***; and ***Type of word***. The term “patients” was used to indicate CI users and so member of the experimental group and “SNR” was related to the noise conditions. Table 14 indicates and describes the levels associated to each variable

Table 14. Independent variables in the GLMM model.

Independent variables	Levels
Patient Vs Control Group	Patient
	Control Group
Mask	M1
	M2
	M3
	NM
SNR	0 dB
	5 dB
	10 dB
	QUIET
Type of word	Number
	Object
	Adjective

According to the different acoustic conditions named earlier, the Mask predictor consisted of four factors (NM, M1, M2, M3), as the SNR predictor (QUIET, 0, 5, 10); the Type of word predictor consisted of three levels: Number, Object, Adjective. In addition, random factors are best defined as noise in the data. These are effects that arise from uncontrollable variability within the sample. Subject level variability is a random effect, therefore, in the proposed models, the listener was used as a random factor. On the other hand, the LD was evaluated, as mentioned before, with a discrete 5- points scale. It was noticed that the LD data is characterized by a gamma distribution, thus a GLMM with gamma distribution was used to model this quantity. As the SI models, the independent variables are the same of the table 14, and the listener was considered as a random factor.

For both quantities, SI and LD, **Tukey's posthoc pair-wise comparisons** were performed to examine the differences between all levels of the mask, SNR and, only for SI, type of word factors. These are pair-wise z tests, where the *z* statistic represents the difference between an observed statistic and its hypothesized population parameter in units of the standard deviation. The *p* values for these tests were adjusted using the default single-step method [54]. The *GLMM outputs* include the **estimates** of the fixed effects coefficients, the **standard error** associated with the estimate, the **test statistic *z*,** and the ***p*-value**. From the estimates, it is possible to calculate the **odds ratio** (OR) as the exponential function of the estimate. An OR is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.

8. RESULTS

Information deriving from the statistical analysis have been correctly organised in tables and charts facilitating its main meaning. The interested quantities, which are subjects of the mentioned tables, are speech intelligibility and listening difficulty, with the former being the most significative for the purpose of this study.

Therefore, starting from the first table (Table 15), results extracted from the GLMM for SI are illustrated; meanwhile its mean values and standard errors are reported and represented in the table 16 and figure 56. In this regard, the reference levels were EG for group, No Mask for the Mask predictor, QUIET for SNR, and the object for the type of word. Furthermore, listeners were used as a random factor.

*Table 15. GLMM (binomial family) for response variables Speech Intelligibility considering as predictors (1) Patients Vs Control group; (2) Mask (4 levels); (3) SNR (4 levels); and (4) Type of word (3 levels). The reference levels were Patients for group, NM for the Mask predictor, QUIET for SNR, and the object for the Type of word. Listeners were used as a random factor. The significance codes for the p-values ' *** '<0.001 '**'<0.01 '*'<0.05.*

Speech Intelligibility (-)	Estimate	Std. Error	z value	p-value	
(Intercept)	3.63	0.32	11.34	<0.001	***
ControlControl	2.62	0.18	14.56	<0.001	***
Mask M1	-0.52	0.09	-5.91	<0.001	***
Mask M2	-0.46	0.09	-5.10	<0.001	***
Mask M3	-0.08	0.09	-0.84	0.402	
SNR 10 dB	-1.22	0.13	-9.18	<0.001	***
SNR 5 dB	-1.89	0.13	-14.68	<0.001	***
SNR 0 dB	-3.21	0.12	-25.76	<0.001	***
Type number	-0.18	0.08	-2.33	0.020	*
Type adjective	-0.37	0.08	-4.86	<0.001	***

Table 16. Mean values and standard error of speech intelligibility grouped by patient and control, masks, and SNR.

Speech Intelligibility									
Patient	NM		M1		M2		M3		
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.95	0.01	0.97	0.01	0.98	0.01	0.96	0.01
10 dB	0.91	0.01	0.83	0.02	0.87	0.02	0.90	0.01
5 dB	0.81	0.02	0.76	0.02	0.77	0.02	0.87	0.02
0 dB	0.64	0.02	0.51	0.02	0.47	0.02	0.55	0.02
Control	NM		M1		M2		M3	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.98	0.01	1.00	0.00	0.95	0.02	0.98	0.01
10 dB	0.96	0.02	0.91	0.02	0.97	0.01	0.95	0.02
5 dB	0.97	0.01	0.88	0.02	0.91	0.02	0.98	0.01
0 dB	0.89	0.02	0.86	0.03	0.85	0.03	0.87	0.02

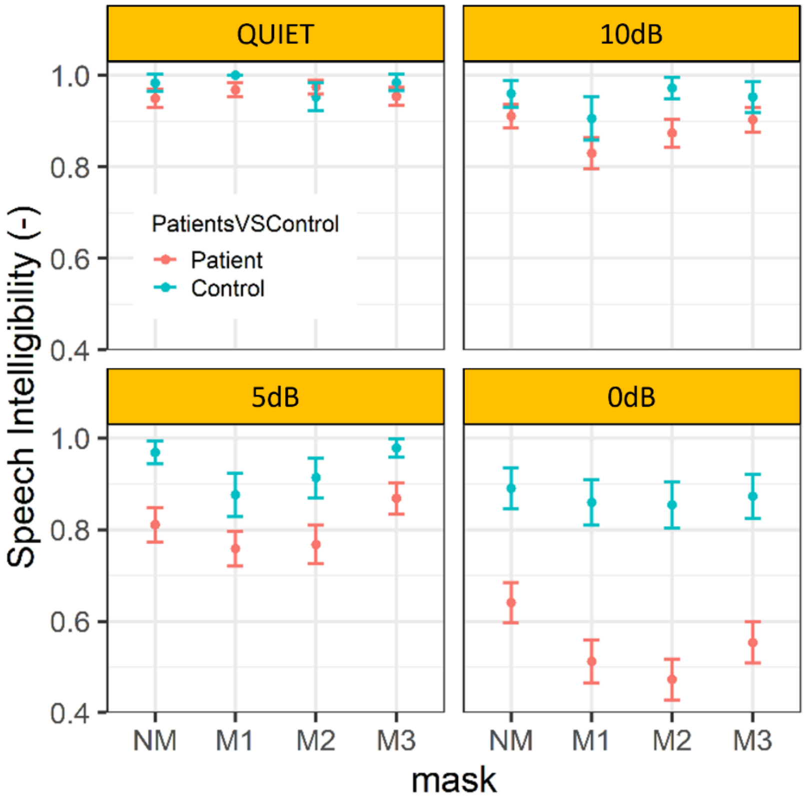


Figure 56. Mean Speech Intelligibility across the two groups of participants (CI users and control group) in the conditions with and without masks, under four SNR levels. Error bands indicate ± 95% confidence intervals.

As earlier mentioned, these outcomes are characterised by a high variability related to the listener. In this particular case, the estimate of standard deviation for random effects was 1.15 for listener. In addition, a crucial difference between the experimental and the control group can be easily deduced: the **probability of correctly recognising a word** (PCR) among the participants of EG was 86% less than the one in CG (X2 = 2.62, OR = 0.14, $p < 0.001$). The performances of mask conditions, M1 and M2, are lower than in NM condition (Fig. 57). PCRs were respectively 40% (X2 = -0.52, OR = 0.60, $p < 0.001$) and 37% (X2 = -0.46, OR = 0.63, $p < 0.001$) less compared to the unmasked condition. On the contrary, the use of mask 3 had a lower impact on speech comprehension and, in fact, no statistically significant difference from NM was found.

For what concerns the different noise conditions, QUIET recorded higher SI values, as predicted. Speech stimuli mixed with noise determine a foreseeable reduction of the number of understood words. Notably, PCRs for SNR of 10 dB, 5 dB, 0 dB, were respectively 70% (X2 = -1.22, OR = 0.30, $p < 0.001$), 85% (X2 = -1.89, OR = 0.15, $p < 0.001$) and 96% (X2 =-3.21, OR = 0.04, $p < 0.001$) less than in the quiet condition. In relation to the Type of word, PCR of numbers was 17% less than the one for objects (X2 = -0.18, OR = 0.83, p =0.02), PCR for adjectives was 31% less than the one for objects (X2 = -0.37, OR = 0.69, $p < 0.001$).

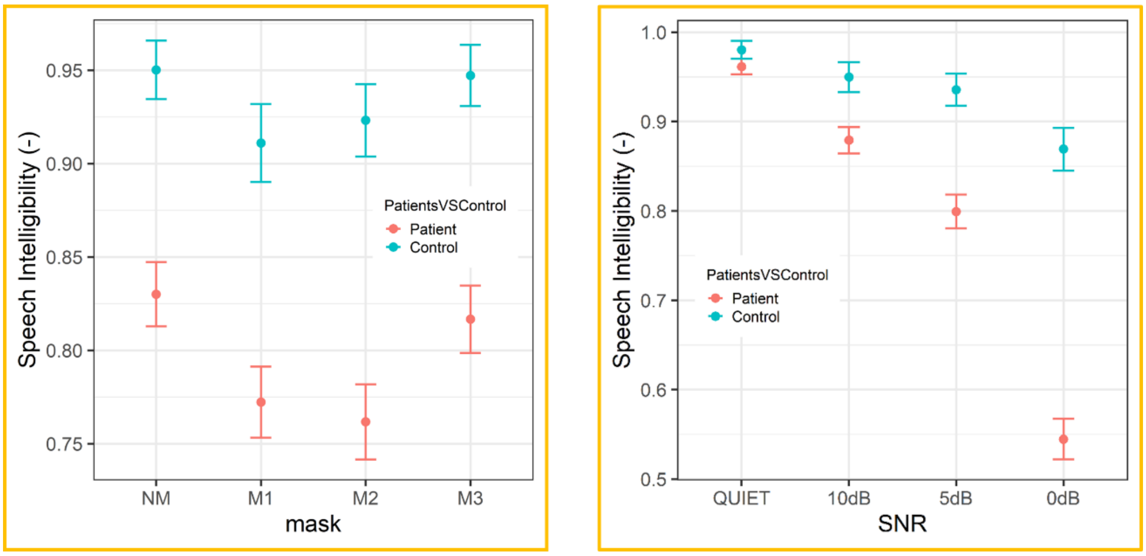


Figure 57. Mean Speech Intelligibility across the two groups of participants (CI users and control group) considering mask contribution (left) and noise (right). Error bands indicate ± 95% confidence interval.

Table 17. Mean values and standard error of SI for number grouped by patient and control, masks, and SNR

SI for number

Patient								
	NM		**M1**		**M2**		**M3**	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.96	0.02	0.96	0.02	0.99	0.01	0.95	0.02
10 dB	0.95	0.02	0.83	0.03	0.86	0.03	0.91	0.02
5 dB	0.82	0.03	0.76	0.03	0.78	0.04	0.90	0.03
0 dB	0.64	0.04	0.43	0.04	0.43	0.04	0.55	0.04
Control								
	NM		**M1**		**M2**		**M3**	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.98	0.02	1.00	0.00	0.97	0.02	0.98	0.02
10 dB	0.97	0.02	0.94	0.03	0.98	0.02	0.96	0.03
5 dB	0.97	0.02	0.85	0.05	0.93	0.04	0.97	0.02
0 dB	0.94	0.03	0.86	0.04	0.89	0.04	0.89	0.04

Table 18. Mean values and standard error of SI for object grouped by patient and control, masks, and SNR

SI for object

Patient								
	NM		**M1**		**M2**		**M3**	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.94	0.02	0.97	0.01	0.97	0.01	0.97	0.01
10 dB	0.90	0.02	0.84	0.03	0.90	0.02	0.93	0.02
5 dB	0.84	0.03	0.76	0.03	0.83	0.03	0.89	0.03
0 dB	0.68	0.04	0.61	0.04	0.53	0.04	0.59	0.04
Control								
	NM		**M1**		**M2**		**M3**	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.98	0.02	1.00	0.00	0.95	0.03	0.98	0.02
10 dB	0.97	0.02	0.90	0.04	0.97	0.02	0.96	0.03
5 dB	0.97	0.02	0.90	0.04	0.94	0.03	0.98	0.02
0 dB	0.91	0.04	0.86	0.04	0.86	0.04	0.90	0.04

Table 19. Mean values and standard error of SI for adjective grouped by patient and control, masks, and SNR

SI for adjective

Patient	NM		M1		M2		M3	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.96	0.02	0.97	0.01	0.96	0.02	0.95	0.02
10 dB	0.88	0.03	0.82	0.03	0.86	0.03	0.87	0.03
5 dB	0.77	0.04	0.76	0.03	0.69	0.04	0.83	0.03
0 dB	0.61	0.04	0.50	0.04	0.46	0.04	0.52	0.04
Control	**NM**		**M1**		**M2**		**M3**	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	0.98	0.02	1.00	0.00	0.94	0.03	0.98	0.02
10 dB	0.95	0.03	0.88	0.05	0.97	0.02	0.94	0.03
5 dB	0.97	0.02	0.87	0.04	0.87	0.05	0.98	0.02
0 dB	0.83	0.05	0.86	0.04	0.81	0.05	0.83	0.05

Table 20 and figure 58 show that SI for number is higher than object and adjective, for CI users and for normal hearing too. Furthermore, for patients, the mean values of SI for number in the QUIET condition are at least 0.95 regardless of the type of mask (Table 17).

Table 20. Mean values and standard error of SI for number, object and adjective grouped by patient and control.

	Patient		Control	
	Mean	*Std. Error*	*Mean*	*Std. Error*
Number	0.98	0.02	1.00	0.00
Object	0.97	0.02	0.90	0.04
Adjective	0.91	0.04	0.86	0.04

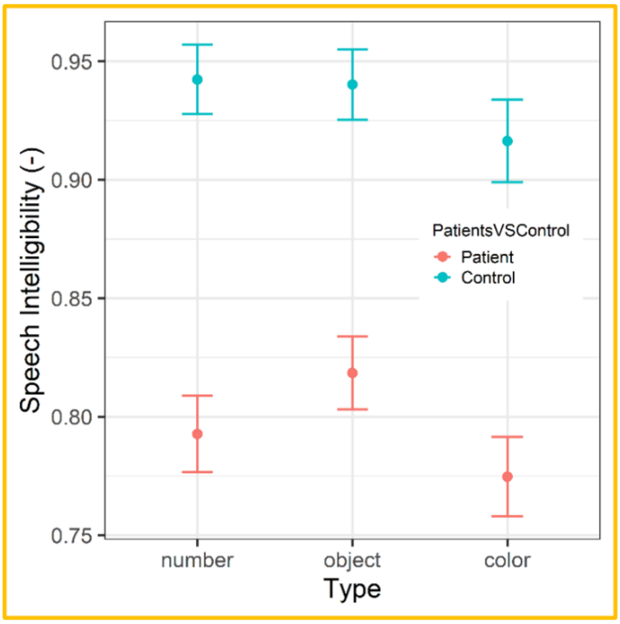


Figure 58. Mean Speech Intelligibility across the two groups of participants (CI users and control group) according the type of word. Error bands indicate ± 95% confidence interval.

*Table 21. Multiple comparisons of means using Tukey contrasts for the predictive variables SI for the factor mask, SNR, and Type of word. The significance codes for the p-values ' *** '<0.001 '**'<0.01 '*'<0.05.*

Speech Intelligibility (-) vs Mask	***Estimate***	***Std. Error***	***z value***	***p-value***	
M1-NM	-0.52	0.09	-5.91	<0.001	***
M2-NM	-0.46	0.09	-5.11	<0.001	***
M3-NM	-0.08	0.09	-0.84	0.836	
M2-M1	0.06	0.08	0.75	0.879	
M3-M1	0.44	0.09	5.01	<0.001	***
M3-M2	0.38	0.09	4.23	<0.001	***
Speech Intelligibility (-) vs SNR	***Estimate***	***Std. Error***	***z value***	***p-value***	
10 dB - QUIET	-1.22	0.13	-9.18	<0.001	***
5 dB - QUIET	-1.89	0.13	-14.68	<0.001	***
0 dB - QUIET	-3.21	0.12	-25.75	<0.001	***
5 dB - 10 dB	-0.67	0.09	-7.27	<0.001	***
0 dB - 10 dB	-1.99	0.09	-23.28	<0.001	***
0 dB - 5 dB	-1.32	0.08	-17.15	<0.001	***
Speech Intelligibility (-) vs Type of word	***Estimate***	***Std. Error***	***z value***	***p-value***	
Number - Object	-0.18	0.08	-2.33	0.05	
Adjective - Object	-0.37	0.08	-4.86	<0.001	***
Adjective - Number	-0.19	0.08	-2.55	0.03	*

In addition, information extracted from the posthoc comparisons (Table 21) confirmed that overall, the SI differences among the four types of masks were statistically significant except for the difference between the M3 and NM and between M2 and M1. The same aspect was indagated considering the four noise conditions and, also in this case, differences in SI were all statistically significant when controlled for multiple comparisons. In the same way, also for the Type of word there were statistical differences in SI values except for the difference between numbers and objects.

Using the same strategy, data on listening difficulty were analysed; in the table 22 are reported the results derived from the models, while its mean values and standard errors are illustrated in table 23 and figure 60. These values demonstrate that listeners belonging to the EG experienced more LD compared to the one of the CG (X2=0.09, $p < 0.001$). Furthermore, in relation to the type of masks, M1 (X2= -0.06, $p < 0.001$) and M2 (X2= -0.05, $p < 0.001$) were the most competitive mask conditions with higher LD values than NM condition. On the contrary, an evident statistical difference between M3 and NM was not identified (X2= -0.02, $p < 197$). This can also be easily noticed from Figure 1, where the mean LD values are shown considering only the mask condition. Moreover, the QUIET condition resulted to be easier compared to SNR of 10 dB (X2= -0.21, $p < 0.001$), SNR of 5 dB (X2= -0.34, $p < 0.001$) and SNR at 0 dB (X2=-0.47, $p < 0.001$). Furthermore, LD values in QUIET condition proved to be very similar between patients and normal hearing subjects (Fig 59).

*Table 22. GLMM (binomial family) for response variables Speech Intelligibility considering as predictors (1) Patients Vs Control group; (2) Mask (4 levels); (3) SNR (4 levels); and (4) Type of word (3 levels). The reference levels were Patients for group, NM for the Mask predictor, QUIET for SNR, and the object for the Type of word. Listeners were used as a random factor. The significance codes for the p-values ' *** '<0.001 '**'<0.01 '*'<0.05.*

Listening Difficulty (-)	***Estimate***	***Std. Error***	***z value***	***p-value***	
(Intercept)	0.96	0.04	21.51	<0.001	***
control group	0.09	0.02	5.83	<0.001	***
mask M1	-0.06	0.01	-4.91	<0.001	***
mask M2	-0.05	0.01	-4.04	<0.001	***
mask M3	-0.02	0.01	-1.29	0.197	
SNR 10 dB	-0.21	0.02	-11.96	<0.001	***
SNR 5 dB	-0.34	0.02	-20.60	<0.001	***
SNR 0 dB	-0.47	0.02	-31.02	<0.001	***
(Intercept)	0.96	0.04	21.51	<0.001	***
control group	0.09	0.02	5.83	<0.001	***

Table 23. Mean values and standard error of Listening Difficulty grouped by patient and control, masks, and SNR.

Listening difficulty	**Patient**							
	NM		**M1**		**M2**		**M3**	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	1.11	0.03	1.09	0.02	1.10	0.02	1.06	0.02
10 dB	1.18	0.03	1.56	0.05	1.43	0.04	1.41	0.04
5 dB	1.81	0.06	1.89	0.06	1.95	0.07	1.49	0.05
0 dB	2.25	0.07	2.63	0.08	2.59	0.08	2.60	0.08
	Control							
	NM		**M1**		**M2**		**M3**	
	Mean	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*	*Mean*	*Std. Error*
QUIET	1.06	0.04	1.06	0.02	1.13	0.04	1.08	0.04
10 dB	1.41	0.06	1.60	0.08	1.52	0.05	1.58	0.07
5 dB	1.60	0.07	1.88	0.08	1.67	0.08	1.42	0.06
0 dB	1.78	0.09	1.89	0.09	2.00	0.09	1.91	0.09

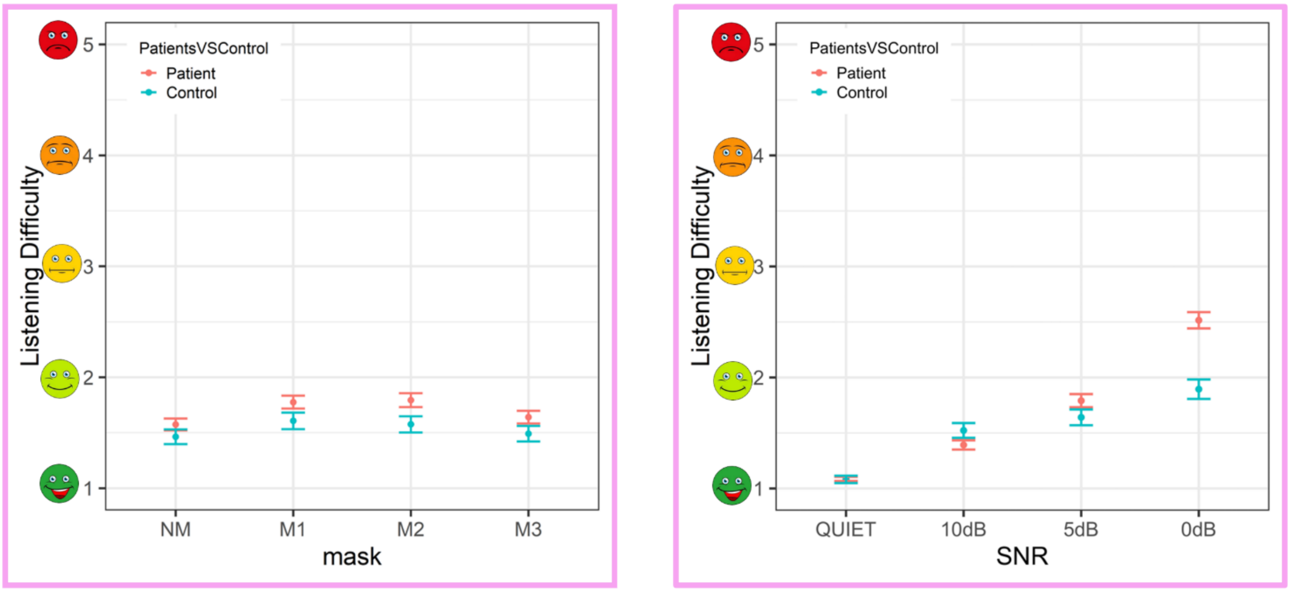


Figure 59. Mean Listening Difficulty across the two groups of participants (CI users and control group) considering mask contribution (left) and noise (right). Error bands indicate ± 95% confidence interval.

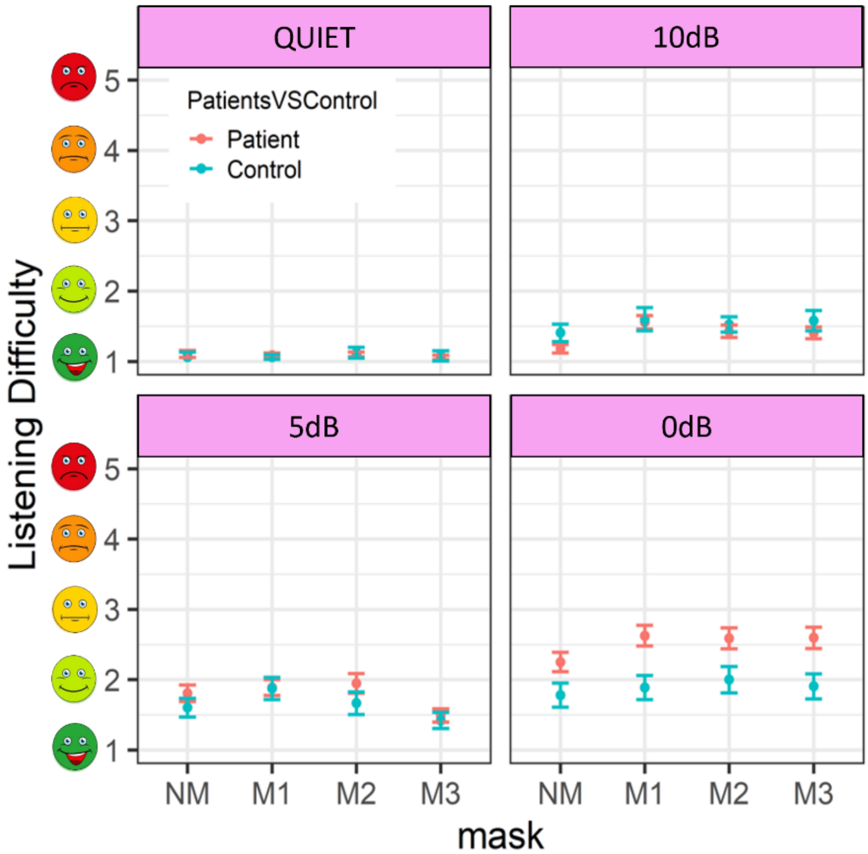


Figure 60. Mean Listening Difficulty across the two groups of participants (CI users and control group) in the conditions with and without masks, under four SNR levels. Error bands indicate ± 95% confidence intervals.

As for the SI, posthoc comparisons proved that there was a significant statistical difference in LD values among the four mask conditions except for the difference between the M3 and NM and between M2 and M1. On the contrary, the differences in LD among noise conditions were all statistically significant for multiple comparisons (Table 24).

*Table 24. Multiple comparisons of means using Tukey contrasts for the predictive variables SI for the factor mask and SNR. The significance codes for the p-values ' *** '<0.001 '**'<0.01 '*'<0.05.*

Listening Difficulty (-) vs Mask	***Estimate***	***Std. Error***	***z value***	***p-value***	
M1-NM	-0.06	0.01	-4.91	<0.001	***
M2-NM	-0.05	0.01	-4.04	<0.001	***
M3-NM	-0.02	0.01	-1.29	0.569	
M2-M1	0.01	0.01	0.88	0.816	
M3-M1	0.04	0.01	3.60	0.002	**
M3-M2	0.03	0.01	2.75	0.031	*
Listening Difficulty (-) vs SNR	***Estimate***	***Std. Error***	***z value***	***p-value***	
10 dB - QUIET	-0.21	0.02	-11.96	<0.001	***
5 dB - QUIET	-0.34	0.02	-20.60	<0.001	***
0 dB - QUIET	-0.47	0.02	-31.02	<0.001	***
5 dB - 10 dB	-0.13	0.01	-9.43	<0.001	***
0 dB - 10 dB	-0.26	0.01	-21.51	<0.001	***
0 dB - 5 dB	-0.13	0.01	-12.51	<0.001	***

9. LIMITATIONS OF THE STUDY

The audiometric test used in this project proved to be an excellent method for the assessment of speech comprehension and the remote mode was a valid alternative to the traditional performing execution in presence. However, the duration of the survey must be suitable to avoid the child becoming tired and exhausted. This aspect is crucial to collect information which are significative and representative of listener's hearing abilities. In this case, the 14-sentence lists, already used in previous studies, proved to be too demanding, unlike the 10-sentence lists which resulted more suitable for our purpose.

The data on listening difficulties were rather low, and never higher than 3. This is probably due to the fact that after the first lists of sentences, the listener paid less attention to the choice of the smiley and continued in an almost automatic way. In fact, this problem was confirmed by the examiner who was responsible for the data acquisition and who supervised the child during the test. In this application, the survey was structured in order to record an LD value for each sentence, which means that for each list the child was asked, 14 or 10 times, to click on the smiley depending on the extended or reduced version. This design choice aimed to collect as much data as possible, however, an idea for future modification could be to reduce the frequency with which the child was asked to click on the smiley face. For example, considering the reduced version, LD could be evaluated every 5 sentences. In this way we will have 2 LD values for each list instead of 10 and therefore less information, but it could probably be a good strategy to obtain more reliable and representative data.

The number of participants involved in this study was relatively low. In fact, the statistical analysis revealed a clear variability between the different subjects and the listener was found to be a random factor. In order to reduce this phenomenon, it is necessary to increase the sample's numerosity, especially for the control group. In this way it is possible to have more stable and significant results.

10. CONCLUSION

The aim of this study was to assess speech intelligibility in hearing impaired and cochlear implanted children. Twenty subjects, divided into a control group of 14 children with CI and a control group of 6 normal hearing children, were recruited. A specific audio-metric test according to the requirements of this project was built. The participants were able to carry out the test remotely using a device (PC, tablet...) and internet network. This strategy allows to overcome the limitations imposed by the pandemic rules and thus to collect data even without the possibility of performing the test in presence. In addition, the difficulty of listening was measured in order to have a subjective evaluation about speech comprehension in competitive acoustic scenarios.

The results extracted from this experiment showed that speech intelligibility is influenced by many factors and varies depending on the acoustic conditions in which the auditory stimulus is presented to the listener. The effect of noise and masks on speech intelligibility are discussed below.

<u>EFFECT OF NOISE</u>

It was observed that in absence of noise (QUIET condition) SI was always above 90% and results from EG and CG were significantly different, on the contrary it decreased when speech was affected by noise. In fact, acoustic stimuli mixed with noise had lower intelligibility. Additionally, the negative influence of noise on speech comprehension has also been confirmed by an increase of listening difficulty. The stimuli with a signal-to-noise ratio of 10 dB and 5 dB had good intelligibility and were higher than 0 dB condition. In particular, 5 dB condition was characterized by a mean SI of about 70%, a fairly satisfactory value and still acceptable within a classroom. A possible strategy could be to acoustically design the room and organize spaces in such a way as to guarantee a minimum SNR of 5 dB. It is clear, however, that the best context would be reaching optimal acoustic scenarios without any noise, but this condition is quite difficult to obtain, or rather impossible in school environments characterised by extreme acoustic variability and reverberation. Although having lower SI values than 10 dB, SNR of 5 dB is easier to achieve in terms of architecture and materials, and therefore represents a hypothetically valid acoustic treatment.

<u>EFFECT OF MASK</u>

Considering the mask condition, the data confirmed that masks play an essential role in communication. The un-masked condition proved to be the one with higher SI values, as expected. A correlation between the type of masks and their intrinsic characteristics was found, considering attenuation of the speech signal and filtering capacity. For acoustic stimuli in presence of M1 and

M2, lower SI values than NM were recorded. Furthermore, for these two conditions, LD data highlighted that child experienced more difficulty in sentence comprehension. On the contrary, the mask 3, which was characterized by a low acoustic attenuation, had values comparable to the no mask condition. In fact, it apparently had the least impact on speech comprehension and, therefore, this can be considered as the most suitable for the teacher to wear. Thus, it is advisable for teachers to use masks with the same or similar features of M3.

EFFECT OF TYPE OF WORD

The type of word also turned out to be a discriminating factor in speech intelligibility. The statistics demonstrated that objects were the most correctly understood words, followed by numbers and adjectives. From SI value, can be deduced a cognitive flexion for which the adjective is the least correctly understood word type. Probably this phenomenon can be related to tiredness or also to the fact that the children paid more attention to the first words of the sentence. Moreover, it might be interesting to investigate this aspect in a future study.

EFFECT OF COCHLEAR IMPLANT

Overall, the data showed that children with cochlear implants had lower intelligibility than the normal hearing subject. This situation is more evident when the stimulus is characterised by a high level of noise. In cases of SNR of 5 and 10 dB, there was always a statistically significant difference between patients and controls. Therefore, this means that noise is the most influential and discriminating element for comprehension in children with hearing problems. A possible way to improve hearing abilities of children with CI, might be to directly modify the signal processing system of the implant in order to adapt it to competitive acoustic situations, even considering speakers wearing masks. However, considering the narrowness of the outcomes of this experiment, it is logical that deeper and specific studies are required. A more heterogenic sample could reduce variability among the listeners, leading to a greater reliability of SI assessment. Therefore, a viable idea would be a future investigation based on the same audiometric test, involving participants with similar hearing characteristics, but considering a wider range of subjects dividing them in different groups according to the type of CI device.

FINAL CONSIDERATION

Finally, the preliminary results discussed in this work suggest the need of guaranteeing optimal acoustic conditions, especially in terms of high signal-to-noise ratios, and of selecting the most appropriate face masks for speakers. These outcomes can be transferred to improve speech production and understanding in classrooms for everyday design and usage to implement optimal listening conditions.

The content and the results of this thesis will be summarised in a paper and published in the "*International Journal of Environmental Research and Public Health (IJERPH)*" in the special issue "*Speech Communication in Complex Auditory Scenes and Effects on Voice Behaviour and Health, Listening Comfort, Well-being, and Learning*". This could prove essential to support audiological research and future studies focused on improving CI users' life.

BILBIORAFIA

[1] Caniato M., Marzi A., Gasparella A.; "How much COVID-19 face protections influence speech intelligibility in classrooms?" in Elsevier, v. 178, pp. 1-14 (2021).

[2] Kollmeier B., Warzybok A., Hochmuth S., Zokoll M. A., Uslar V., Brand T., Wagener K. C.; "The multilingual matrix test: Principles, applications, and comparison across languages: A review" in International Journal of Audiology v.54, pp. 3–16 (2015).

[3] Puglisi G. E., Warzybok A., Hochmuth S., Visentin C., Astolfi A., Prodi N., Kollmeier B.; "An Italian matrix sentence test for the evaluation of speech intelligibility in noise" in International Journal of Audiology v.54, pp. 44–50 (2015).

[4] Puglisi G. E., Di Berardino F., Montuschi C., Sellami F., Albera A., Zanetti D., Albera R., Astolfi A., Kollmeier B., Warzybok A.; "Evaluation of Italian Simplified Matrix Test for Speech-Recognition Measurements in Noise" in Audiology Research, v.11, pp 73-78 (2021).

[5] Anastasi G., Balboni G.; *Trattato di Anatomia umana*, v. IV, 3 ed., Milano, Edi-Ermes, 2006.

[6] "Cenni di anatomia dell'orecchio", from www.audiofit.it/orecchio.pdf.

[7] https://www.howtomed.it/epitelio-sensoriale/.

[8] Forinash, K., & Christian, W.; "The Place Theory of Hearing", Sound-An Intercative Book, Libretexts, from from https://phys.libretexts.org/@go/page/26224 (2021).

[9] Forinash, K., & Christian, W.; "The Temporal Theory of Hearing", Sound-An Intercative Book, Libretexts, from https://phys.libretexts.org/@go/page/26225 (2021).

[10] Forinash, K., & Christian, W.; "The Vocal Tract", Sound-An Intercative Book, Libretexts, from https://phys.libretexts.org/@go/page/26269 (2021).

[11] Forinash, K., & Christian, W.; "Vocal Formants", Sound-An Intercative Book, Libretexts, from https://phys.libretexts.org/@go/page/26270 (2021).

[12] "Fonetica e fonologia: dalla teoria alla riabilitazione nel bambino e nell'audulto (parte 1: la teoria)", from https://www.trainingcognitivo.it/fonetica-fonologia-dalla-teoria-alla-riabilitazione-nel-bambino-nelladulto/.

[13] Romano A., Rivoira M., Minissi N.; *Manuale di fonetica,* v. IV "L'analisi fonetica strumentale", Alessandria, Dell'Orso, in prep.

[14] Niebergall A., Uecker M., Zhang S. Voit D., Merboldt K.S., Frahm J.; "Real time MRI-speaking", Biomedizinische NMR Forschungs GmbH, https://commons.wikimedia.org/wiki/File:Real-time_MRI_-_Speaking_(English).ogv (2011).

[15] Pullki V., Karjalainen M.; *Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics,* 1 ed., John Wiley & Sons, Incorporated, 2015.

[16] Leoni F. A.; "Il ruolo dell'udito nella comunicazione linguistica. Il caso della prosodia" in Rivista di Linguistica, vol. 13.1, pp. 45-68 (2001).

[17] Rémi Pujol; "Journey into the world of hearing: Human auditory range", from http://www.cochlea.org/en/hear/human-auditory-range (2018).

[18] Standards Secretariat, Acoustical Society of America, "American National Standards: Acoustical Terminology" New York, (2013).

[19] T. Editors of Encyclopaedia; *Encyclopedia Britannica*: *"Loudness"*, Britannica, from https://www.britannica.com/science/loudness (2008).

[20] Giansanti D., Morelli S., Maccioni G., Grigioni M.; "Impianti cocleari: indagine su stato dell'arte, problematiche riscontrate, aspetti clinici e normativi.", Rapporti ISTISAN, Istituto Superiore di Sanità (2009).

[21] Lindosland; "Equal loudness countour", http://en.wikipedia.org/wiki/Image:Lindos1.svg (2005).

[22] Fastl H., Zwicker E.; *Psychoacoustic: Fact and models*, v. XXII, 3 ed., Berlin, Springer, 2007.

[23] Katz, Laurie, Schery, Teris K.; "Including Children with Hearing Loss in Early Childhood Programs" in Young Children, v. 61, pp 86-95 (2006).

[24] Gifford K.A., Holmes M. G., Bernstein H. H.; "Hearing Loss in Children" in Pediatrics Review, v. 30, pp 207-216 (2009).

[25] Dimitrov L., Gossman.W. G.; *Pediatric Hearing Loss,* 1 ed., Treasure Island (FL), StatPearls Publishing, 2021.

[26] Clark J. G.; "Uses and abuses of hearing loss classification" in Asha: a journal of the American Speech-Language-Hearing Association, v. 23, pp. 493-500 (1981).

[27] Tharpe A. M., Seewald R.; *Comprehensive Handbook of Pediatric Audiology*, 2 ed., San Diego, Plural Publishing, 2017.

[28] Alshuaib W. B., Al-Kandari J. M. Hasan S. M.; *Update in Hearing Loss,* ch. 2 "Classification of Hearing Loss", IntechOpen, 2015.

[29] Adam MP, Ardinger HH, Pagon RA, et al..; *GeneReview: "*Deafness and hereditary hearing loss overview", University of Washington, Seattle, 2012.

[30] Vasconcelos R. M., Serra L. M., De Farias Aragão V. M.; "Transient evoked otoacustic emissions and distortion product in school children" in Bras Otorrinolaringol, v. 74, pp 503-507 (2008).

[31] Delgado R. E.; "Automated auditory brainstem response interpretation" in IEEE Engineering in Medicine and Biology Magazine, v. 3, pp. 227-237 (1994).

[32] Grindle C. R.; "Pediatric Hearing Loss" in Pediatrics in Review, v. 35, pp. 456-464 (2014).

[33] Lenarz T:; "Cochlear implant – state of the art" in GMS Current Topics in Otorhinolaryngology - Head and Neck Surgery, v. 16 (2017).

[34] Cuda D.; *Impianti cocleari,* Galatina, TorGraf, 2008 from https://www.aooi.it/contents/attachment/c4/impiantiCocleari.pdf.

[35] Medical Expo, prodotti: Impianto cocleare componente interno HiRes Ultra from https://www.medicalexpo.it/prod/advanced-bionics/product-70818-823680.html.

[36] De Filippis Cippone A.; *L'impianto cocleare in età pediatrica*, 1 ed., Masson, Milano, 1997.

[37] Medel, Estimulación eléctrico-acústica: processador de audio SONNET EAS from https://www.medel.com/latam/hearing-solutions/electric-acoustic-stimulation/sonnet-eas/sonnet-eas---colour-configurator.

[38] Cochlear Limited, "Guida dell'utente del processore Nucleus CP810", 2010 from https://docplayer.it/5919315-Guida-dell-utente-del-processore-nucleus-cp810.html.

[39] Oticon Medical: because sounds matter, "Descrizione dettagliata di un impianto cocleare" from www.oticalmedical.com.

[40] Van den Honert C, Stypulkowsi P.; "Single fiber mapping of spatial excitation patterns in the electrically stimulated nerve" in Hearing research, v. 29, pp. 195-206 (1987).

[41] Cochlear, products: nucleus sound processor 7 from https://www.cochlear.com/it/it/home/products-and-accessories/cochlear-nucleus-system/nucleus-sound-processors/nucleus-7.

[42] Advanced Bionics, solutions ,processors: Naida CI Q90 from https://advancedbionics.com/us/en/home/solutions/processors/naida-ci/naida-ci-q90.html.

[43] Bottalico P., Murgia S., Puglisi G. E., Astolfi A.; "Effect of masks on speech intelligibility in auralized classrooms" in Journal of the Acoustical Society of America, v.148, pp. 2878- 2884 (2020).

[44] Preminger J. E., Lin H., Payen M., Levitt H.; "Selective Visual Masking in Speechreading" in Journal of Speech, Language and Hearing Research, v. 41, pp. 564-575 (1988).

[45] Atcherson S., Mendel L., Baltimore W., Patro C.; "The effect of conventional and transparent surgical masks on speech understanding in individuals with and without hearing loss" in Journal of the American Academy of Audiology, v. 28, pp. 58-67 (2016).

[46] Mendel L.L., Gardino J.A., Atcherson S.R., "Speech Understanding Using Surgical Masks: A Problem in Health Care?" in J Am Acad Audiol, v.19, pp.686–695 (2008).

[47] Cohn M., Pycha A., Zellou G.; "Intelligibility of face-masked speech depends on speaking style: Comparing casual, clear, and emotional speech" in Elsevier, v. 210, pp. 1-5 (2021).

[48] Palmiero A. J., Symons D., Morgan J. W., Shaffer R. E.; "Speech intelligibility assessment of protective facemasks and air-purifying respirators" in Journal of Occupational and Environmental Hygiene, v. 13, pp. 960-968 (2016).

[49] Brüel & Kjær an HBK company/TRANSDUCER/SIMULATORS/ HEAD AND TORSO SIMULATOR (HATS)/ TYPE 4128-C from https://www.bksv.com/en/transducers/simulators/head-and-torso/hats-type-4128c

[50] Neumann K., Baumeister N., Baumann U., Sick U., Euler H., Weißgerber T.; "Speech audiometry in quiet with the OldenburgSentence Test for Children" in International Journal of Audiology, v.51, pp. 157–163 (2012).

[51] Weißgerber T., Baumann U., Brand T., Neumann K.; "German Oldenburg Sentence Test for Children: A Useful Speech Audiometry Tool for Hearing-Impaired Children at Kindergarten and School Age" in Folia Phoniatrica et Logopaedica, v. 64, pp. 228–233 (2012).

[52] Wagener K., Kollmeier B.; "Evaluation des Oldenburger Satztests mit Kindern und Oldenburger Kinder-Satztest (Evaluation of Oldenburg sentence test with children and Oldenburg children sentence test)" in Z. Audiology, v. 44, pp. 134–143 (2005).

[53] Bates D., Maechler M., Bolker B., Walker S., Haubo Bojesen Christensen R., "lme4: Linear mixed-effects models using Eigen and S4," R package version 1.1–7 (2015).

[54] Hothorn T., Bretz F., Westfall P., Heiberger R. M., Schuetzenmeister A., Scheibe S., Hothorn M. T.; "R Package 'multcomp'. Simultaneous inference in general parametric models," Project for Statistical Computing, Vienna, Austria (2016).

ACKNOWLEDGMENTS

Ringrazio tutti coloro che hanno deciso di prendere parte a questo progetto di ricerca ed in particolar modo i miei relatori, Arianna Astolfi, Giuseppina Emma Puglisi, Pasquale Bottalico e Marco Masoero che mi hanno seguito e mi hanno offerto un aiuto prezioso nella stesura di questo elaborato. Grazie per le call improvvise e ad orari improponibili, grazie per avermi accolto in questo progetto e grazie per la fiducia che avete riposto in me.

Ringrazio poi la mia famiglia, mia mamma, mio padre e mia sorella che hanno sempre sostenuto nelle mie scelte e mi sono stati accanto nei momenti più difficili, questo traguardo è soprattutto loro. I loro sacrifici mi hanno insegnato cosa vuol dire per dei genitori avere un figlio, e spero un giorno di poter fare lo stesso con i miei. Mamma, Papà…vostro figlio ha finito l'Università ed è diventato grande! Probabilmente rimpiangerete questi tempi più di me! Vi voglio bene.

Ringrazio Veronica per aver dovuto sopportare tutti i miei scleri, le mie paure, le mie lacrime (seguite da improvvise risate) ed i miei silenzi. La ringrazio perché quando hai un peso sulle spalle è sempre bello sapere di avere qualcuno con cui condividerlo. Grazie per le chiamate interminabili in cui avevo sempre da lamentarmi e tu hai sempre avuto la pazienza di ascoltarmi e capirmi. Prometto di fare lo stesso con te….anche perché ho le spalle molto più forti delle tue! Ti voglio super bene.

Ringrazio le mie amiche, Marica e Costanza che sono state ancora una volta al mio fianco e mi hanno accompagnato in questo viaggio. Spero che il mondo vi riservi il meglio, siete delle persone meravigliose e sono sicuro che ancora una volta (così come nella triennale) non ce l'avrei fatta così facilmente. Con il senno di poi mi viene quasi da dire "Viva i gruppi studio!" (scherzo). Però devo ammettere di essere un pochino triste che questa "laurea a ciclo unico", come la chiamavamo noi, è invece giunta al termine! Asciugatevi le lacrime e rimboccatevi le maniche amiche, che il bello arriva adesso!

Ringrazio i miei coinquilini, Antonio, Juseppe ed Andrea che hanno saputo rendere meno amari anche le giornate più intense. Non scordatevi che abbiamo brevettato "Dark ball", sono sicuro che un giorno spopolerà e diventeremo ricchi…….O forse no.

Infine, e non per ordine di importanza, voglio anche ringraziare il mio gatto Francesco (anche detto Frank/Franco/Pampo) per essere stato sempre con me seduto di fianco al camino e per avermi fatto compagnia negli ultimi mesi della mia carriera universitaria. TVB FRENK!

Ah, dimenticavo, ringrazio anche me stesso per tutta la vita vissuta finora. Così, perché mi sembrava giusto farlo.