

POLITECNICO DI TORINO

**Corso di Laurea Magistrale
in Ingegneria Biomedica**

Tesi di Laurea Magistrale

Caratterizzazione di una coclea artificiale bio-ispirata



Relatori

prof. Antonio Gliozzi

prof. Federico Bosia

prof. Alberto Botter

Candidato

Simone Cicala

Anno Accademico 2024-2025

Sommario

L'impianto cocleare è una neuroprotesi in grado di ripristinare l'udito in persone affette da perdita uditiva. Essa è una delle neuroprotesi più diffuse al mondo e come tale ha subito un notevole sviluppo durante gli anni, al fine di affrontare sempre più efficacemente il problema clinico. L'obiettivo di tale impianto è andare a stimolare la parte più interna dell'orecchio: la coclea. La coclea è costituita da una struttura a spirale che presenta una diversa sensibilità in frequenza in base alla posizione. Questo effetto viene indicato con il termine "tonotopia".

Il primo obiettivo di questa tesi è di passare in rassegna la tecnologia esistente nel settore degli impianti cocleari, mettendo in evidenza vantaggi e svantaggi, e relativa adeguatezza a diversi casi clinici.

Il secondo obiettivo della tesi è di valutare l'efficacia di un nuovo tipo di sensore tonotopico passivo bioispirato che possa contribuire alla tecnologia degli impianti cocleari.

Gli impianti cocleari tradizionali sfruttano la corrente elettrica per stimolare la coclea in punti ben specifici al fine di sfruttare la tonotopia di quest'ultima. Il sensore artificiale proposto, invece, è passivo, e sfrutta le onde acustiche/meccaniche per effettuare la stimolazione, simulando il comportamento della coclea. Per come è realizzato, questo dispositivo ha una tonotopia invertita rispetto a quella naturale.

Per caratterizzare il sensore e valutarne la tonotopia, è stata svolta una caratterizzazione sperimentale della sua risposta in frequenza. Successivamente, per valutare se esso è in grado di effettuare una discriminazione del parlato in modo simile agli impianti cocleari presenti in commercio, è stato effettuato un test di riconoscimento vocale basato sulla sua risposta in frequenza.

Le prove sperimentali hanno confermato il comportamento tonotopico del sensore e mostrato un buon livello di affidabilità nella discriminazione del parlato, nonostante la semplicità della struttura e il suo funzionamento passivo. Questi risultati sono promettenti per la realizzazione di impianti cocleari futuri funzionanti a bassa potenza, con un alto livello di affidabilità.

Ringraziamenti

Lo svolgimento di questa tesi è stato un viaggio impegnativo e ricco di soddisfazioni in cui ho imparato tantissime nuove cose. Non sarei arrivato fin qui senza il supporto di tantissime persone alle quali devo esprimere la mia più sincera gratitudine.

Il primo ringraziamento va alla mia famiglia per aver sempre creduto in me e senza i quali non avrei potuto far nulla nella vita.

Un pensiero speciale va ai miei amici sia di Torino che del mio paese. In particolare ringrazio i miei coinquilini per le bellissime esperienze e le belle canzoni scritte insieme e i miei amici del Poli con cui ho passato più tempo che con la mia famiglia.

Un ringraziamento speciale, ovviamente, va anche ai professori che mi hanno seguito durante tutto il lavoro svolto.

Indice

1	Introduzione	1
	Anatomia dell'udito	2
	Fisiologia dell'udito	7
	Perdita uditiva	10
2	Stato dell'arte.....	12
	Dispositivo	12
	Aziende produttrici e costi	16
	Processore del suono (microfoni e strategie di elaborazione).....	20
	Impianto interno (array di elettrodi).....	28
3	Materiali e metodi.....	30
	Descrizione coclea artificiale	30
	Strumenti utilizzati	33
	Descrizione delle misure (primo e secondo dataset).....	36
4	Risultati ed analisi dati	40
5	Conclusioni	57
6	Bibliografia	61

1 Introduzione

Più di 1,5 miliardi di persone in tutto il mondo, cioè il 20% della popolazione mondiale, soffrono di ipoacusia. Indipendentemente dal suo grado di severità, la perdita uditiva è la terza causa di disabilità nel mondo. Come si può osservare nella figura 1, il grado di ipoacusia più popolare è l'ipoacusia di tipo lieve; ciononostante i restanti gradi di severità non sono trascurabili: circa 400 milioni di persone convivono con un'ipoacusia compresa tra moderata e grave e ben 30 milioni con una perdita uditiva di tipo profonda o completa in entrambe le orecchie [1]. In quest'ultima parte di popolazione molto spesso il trattamento più adatto è l'impianto cocleare; infatti, quando gli apparecchi acustici non sono sufficienti a risolvere il problema si rende necessario intervenire in maniera più drastica ed invasiva con l'impianto cocleare.

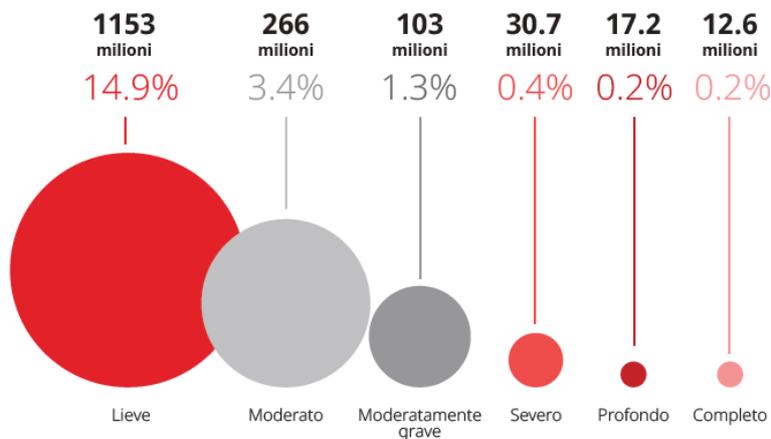


Figura 1: Percentuale di persone affette da ipoacusia in base al grado di severità [1].

L'impianto cocleare è una neuroprotesi composta da due parti: una parte esterna, posta dietro l'orecchio, e una parte interna, impiantata chirurgicamente all'interno dell'osso temporale del cranio. La parte esterna si occupa di analizzare il suono invece la parte interna si occupa di stimolare la coclea, l'organo di senso dell'udito umano. La coclea è un organo sensoriale in grado di comunicare con il sistema nervoso così che il suono possa essere percepito ed interpretato dal cervello.

Nel presente elaborato di tesi viene presentata la base concettuale per lo sviluppo di un nuovo impianto, basato su un modello di coclea artificiale, che sfrutta la vibrazione di un risonatore elastico con caratteristiche appositamente sviluppate per ottimizzare la risposta in frequenza nel range dell'udibile e consentire una migliore stimolazione del nervo uditivo. Questo dispositivo è bio-ispirato, cioè la sua geometria ricorda, per alcune caratteristiche, quella della coclea biologica. Quando il modello artificiale è posto in vibrazione da un'onda acustica, questa viene trasdotta in un'onda meccanica che si propaga lungo la coclea artificiale riproducendo la tonotopia della coclea biologica. Per tonotopia si intende la capacità della coclea biologica di analizzare l'onda sonora secondo una specifica mappa posizione-frequenza, cioè discriminare l'onda nel dominio della frequenza attraverso una specifica posizione lungo l'organo sensoriale.

Si ottiene così un risonatore elettro-meccanico che, in base alla posizione lungo la spirale, enfatizza determinate frequenze rispetto ad altre.

Nella progettazione di un dispositivo medico, il punto di partenza è sempre il problema clinico specifico che il dispositivo intende risolvere. Nel caso dell'impianto cocleare, è essenziale comprendere il sistema uditivo, sia dal punto di vista anatomico che fisiologico così da poter esaminare il problema clinico.

Anatomia dell'udito

Come si può notare in figura 2, l'anatomia del sistema uditivo è costituita da tre parti principali: orecchio esterno, orecchio medio ed orecchio interno. Ora le andremo a descrivere in maggiore dettaglio nei paragrafi successivi.

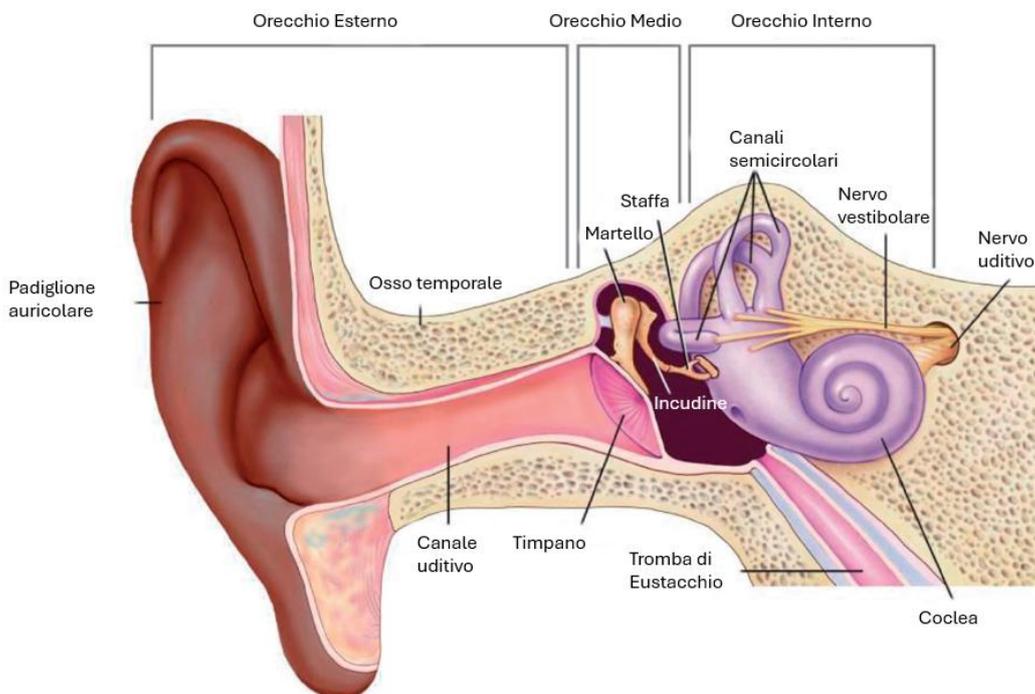


Figura 2: Anatomia dell'orecchio. Andando dall'esterno verso l'interno il sistema uditivo è composto da orecchio esterno, compreso di padiglione auricolare e canale uditivo, orecchio medio, composto da membrana timpanica, ossicini e tromba di Eustacchio ed infine orecchio interno composto da coclea, canali semicircolari e loro relativi nervi (nervo vestibolococleare). Figura modificata da [2].

ORECCHIO ESTERNO

L'orecchio esterno è costituito da due parti: padiglione auricolare e canale uditivo. Il padiglione auricolare è la parte più esterna dell'orecchio. Esso sporge ai lati della testa ed è costituito interamente da cartilagine ricoperta da pelle. Il canale uditivo è un canale lungo circa 4 centimetri composto da due parti: la parte esterna che svolge un ruolo di barriera protettiva e disinfettante, grazie alla presenza di peli e ghiandole sudoripare e sebacee oleose adibite alla produzione di cerume; la parte interna, invece, è costituita da pelle che diventa sempre più sottile in modo da condurre il suono verso la fine dell'orecchio esterno, cioè verso la membrana timpanica [3].

ORECCHIO MEDIO

L'orecchio medio è anch'esso costituito da tre parti: membrana timpanica, ossi uditivi e tromba di Eustacchio. Anatomicamente, l'orecchio medio inizia alla fine dell'orecchio esterno, con la membrana timpanica, e finisce all'inizio dell'orecchio interno, con la finestra ovale della coclea.

La membrana timpanica è costituita da due strati: uno strato esterno che altro non è che il continuo del condotto uditivo composto da epitelio squamoso stratificato con proprietà autopulenti e da uno strato interno composto da epitelio squamoso. Dal punto di vista morfologico, la membrana timpanica si può dividere in due regioni: la pars tensa, una struttura tesa adatta a vibrare, e la pars flaccida dorsale, una struttura elastica. La membrana timpanica è un'apertura rotonda a forma di cono di altoparlante avente spessore di 0,1 mm e diametro di circa 1 cm [2].

Muovendosi verso l'interno dell'orecchio, dopo la membrana timpanica, seguono gli ossi uditivi o ossicini. Sono composti da tre piccoli ossi (i più piccoli che si possono trovare all'interno del corpo umano) disposti in questo ordine: il martello, l'incudine e la staffa. Il martello da un lato è incastonato nella membrana timpanica mentre dall'altro lato forma una connessione rigida con l'incudine. L'incudine a sua volta forma una connessione flessibile con la staffa. Infine, vi è la staffa che è connessa a sua volta con la finestra ovale.

La tromba di Eustacchio, posta nella parte inferiore dell'orecchio medio, collega quest'ultimo con la rinofaringe così che possa svolgere il suo compito principale: ventilare l'orecchio medio e mantenere la pressione costante ai lati del timpano [4].

ORECCHIO INTERNO

L'orecchio interno è costituito dalla coclea e dal labirinto vestibolare, la prima è l'organo di senso necessario per l'udito, mentre il secondo è l'organo dell'equilibrio. Entrambi sono collegati ai rispettivi nervi: nervo vestibolare per i canali semicircolari ed il nervo uditivo per la coclea. L'anatomia completa si può osservare in figura 3:

- Canali semicircolari: sono composti da utricolo e sacco
- Coclea: comunica direttamente con l'orecchio medio tramite la finestra ovale e la finestra rotonda. La finestra ovale è collegata con la staffa degli ossicini in modo che possa ricevere il movimento di quest'ultima per trasmettere il suono.

La coclea è una struttura a spirale simile ad un guscio di lumaca lunga circa 35 mm, con un avvolgimento di circa 2,75 giri. Al suo interno è vi è del liquido il cui volume è pari a circa 0,2 ml.

La coclea è composta da tre compartimenti interni che corrono per tutta la sua lunghezza: scala vestibolare, scala timpanica e scala media o dotto cocleare. Lateralmente vi sono la scala vestibolare e la scala timpanica, entrambe comunicanti con l'orecchio medio tramite due finestre: la finestra ovale per la scala vestibolare e la finestra rotonda per la scala timpanica. Le due scale sono comunicanti e si incontrano all'estremità della coclea, nell'elicotrema [5]. L'elicotrema costituisce l'apice della coclea mentre la parte comunicante con l'orecchio medio ne costituisce la base.

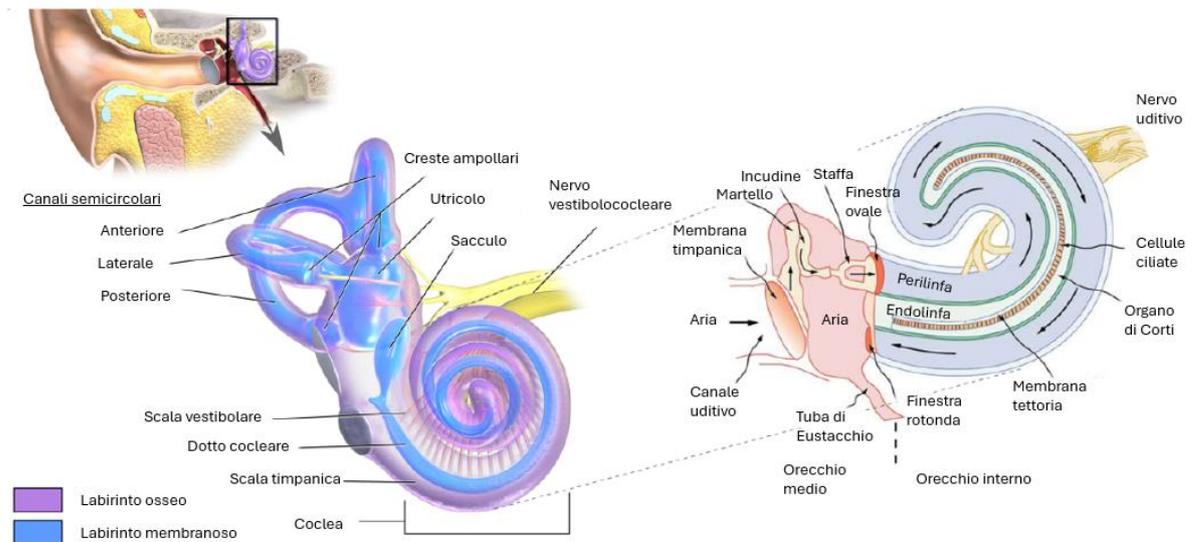


Figura 3: Anatomia dell'orecchio interno. L'orecchio interno è composto da due organi di senso. I canali semicircolari composti da utricolo e sacculo necessari per mantenere l'equilibrio e la coclea necessaria per l'udito. Figura modificata da [4]

La scala vestibolare e la scala timpanica sono separate tra loro dal dotto cocleare. La scala vestibolare è separata dalla scala media tramite la membrana di Reissner mentre la scala timpanica è separata dalla scala media tramite la membrana basilare (si veda figura 4). Ciò significa che il dotto cocleare è composto da due membrane differenti per tutta la sua lunghezza: membrana di Reissner e membrana basilare. Sulla parete laterale del dotto cocleare vi è un compartimento detto Stria vascularis, necessario per la produzione dell'endolinfa. Tuttavia, dal punto di vista della meccanica cocleare, solo la membrana basilare svolge un ruolo utile: la pressione generata dal movimento della staffa presente nell'orecchio medio arriva all'orecchio interno tramite la finestra ovale e viene rilasciata dalla finestra rotonda, quindi la differenza di pressione che si viene a creare guida il movimento della membrana basilare.

All'interno del dotto cocleare, più precisamente in cima alla membrana basilare, vi è l'organo di Corti. Grazie alla presenza di 30.000 cellule ciliate e le 19.000 fibre nervose l'organo di Corti è il vero e proprio organo sensoriale del sistema uditivo.

Le cellule ciliate sono coinvolte nella trasduzione mecano-elettrica del suono. Sono di due tipi: le cellule ciliate interne e le cellule ciliate esterne. Le prime sono poste lungo un'unica fila mentre le cellule ciliate esterne sono poste lungo tre file [6].

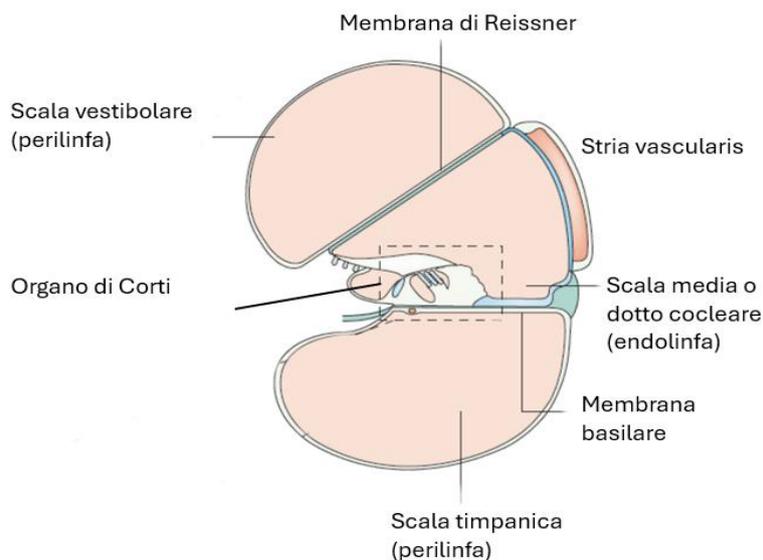


Figura 4: Sezione della coclea. Partendo dal basso verso l'alto si possono distinguere chiaramente le tre scale e le loro relative componenti. Il corridoio inferiore e superiore, rispettivamente la scala timpanica e la scala vestibolare sono riempiti di un fluido chiamato perilinfa; il corridoio intermedio è detto dotto cocleare ed è riempito di endolinfa. Figura modificata da [7].

Le cellule ciliate interne e le cellule ciliate esterne differiscono tra loro sia anatomicamente sia al livello di innervazione.

Le cellule ciliate interne svolgono un ruolo di trasmissione afferente mentre le esterne hanno un ruolo modulatorio. Le cellule ciliate interne forniscono le informazioni al sistema nervoso centrale mentre le cellule ciliate esterne realizzano l'amplificazione del segnale nervoso. In particolare, le cellule ciliate interne comunicano con il sistema nervoso centrale tramite il rilascio di un neurotrasmettitore eccitatorio, il glutammato [8].

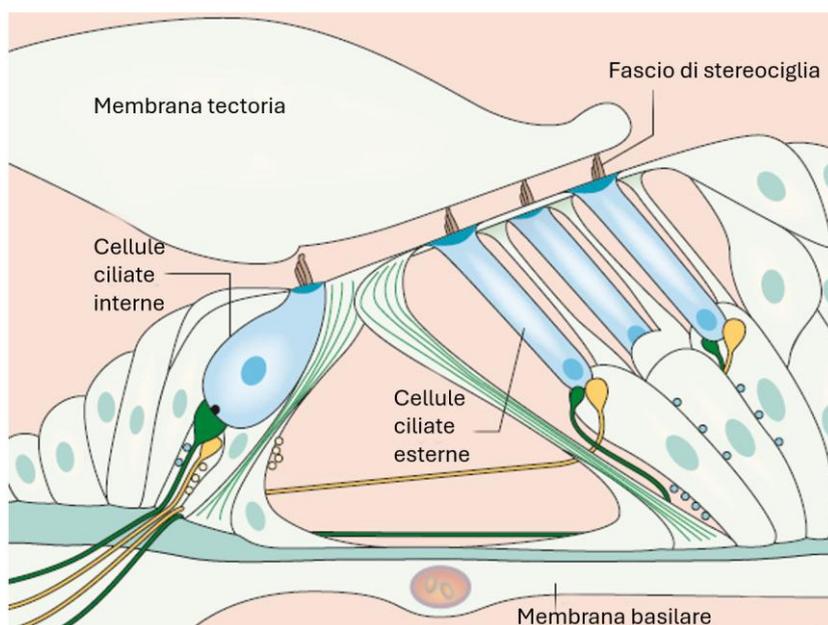


Figura 5: Anatomia dell'organo di Corti. L'organo di Corti è composto da cellule ciliate collegate al sistema nervoso tramite il ganglio spirale. Figura modificata da [7].

La coclea presenta tre fluidi extracellulari: l'endolinfa, la perilinfia ed il fluido intrastriale. L'endolinfa è contenuta all'interno della scala media, il liquido intrastriale all'interno della stria vascularis ed infine la perilinfia è presente nella scala vestibolare e timpanica. L'endolinfa ha una composizione simile al fluido intracellulare nonostante sia un fluido extracellulare, ciò è dovuto all'elevata concentrazione di ioni K^+ utili per la trasduzione sensoriale; infatti, questi ioni sono i principali trasportatori di carica. Oltre ad avere un'insolita concentrazione di K^+ , l'endolinfa presenta una bassa concentrazione di Na^+ , Ca^{2+} e proteine ed un'elevata concentrazione di HCO_3^- . Queste insolite concentrazioni sono necessarie per tamponare il pH (HCO_3^-) e per la generazione della corrente di traduzione meccanoelettrica (Ca^{2+}) [9]. L'endolinfa possiede un suo potenziale che è pari a circa 80 mV rispetto alla perilinfia presente nella scala timpanica e vestibolare. Tale potenziale dipende sia dalla composizione ionica presente nel fluido sia dall'attività della stria vascularis, una struttura presente nella parete esterna laterale del dotto cocleare [10]. L'endolinfa svolge il suo ruolo sia nel sistema uditivo che nei canali semicircolari, il suo ruolo nel sistema uditivo riguarda la stimolazione dei recettori grazie alle onde che si propagano all'interno di questo fluido.

Le cellule ciliate sono le cellule specializzate nella traduzione meccanoelettrica degli stimoli uditivi. Ciò che permette loro di effettuare la trasduzione è il fascio di ciglia che sporge sulla superficie apicale della cellula, composto da decine o centinaia di stereociglia, aste cilindriche di actina disposte in modo esagonale ed unite tutte insieme in un unico punto a formare una punta che, quando poste in vibrazione dall'onda sonora, vengono deviate l'una sull'altra tramite un movimento di taglio che permette l'apertura dei canali di trasduzione meccanoelettrica [11]. Le stereociglia possono assumere tre posizioni differenti come mostrato in figura 6.

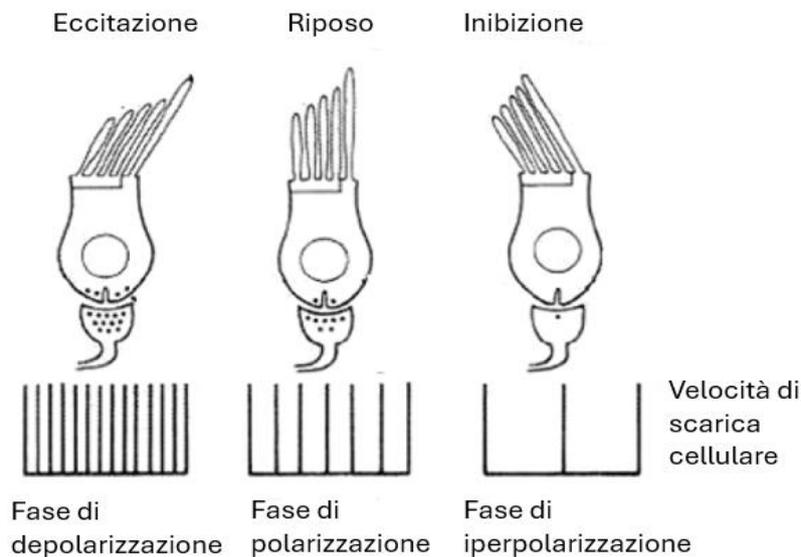


Figura 6: Anatomia delle stereociglia. Le stereociglia sporgono dalla superficie superiore delle cellule ciliate e in fase di riposo sono parallele alla cellula (figura centrale). Quando messe in movimento dall'onda viaggiante, le stereociglia possono muoversi verso destra o sinistra generando l'apertura dei canali legati al K^+ . Gli ioni K^+ depolarizzano la cellula (figura a sinistra) eccitandola. Quando poi il fascio si muove in direzione opposta (figura a destra) si è in fase di iperpolarizzazione e la cellula chiude tutti i canali. Figura modificata da [12]

Fisiologia dell'udito

Il compito del sistema uditivo è convertire l'onda sonora in un impulso nervoso così che possa essere elaborato dal cervello e quindi interpretato.

L'anatomia dell'orecchio e la sua suddivisione in tre parti risulta evidente osservandone la sua fisiologia. La conversione del suono in impulso nervoso avviene tramite due meccanismi: il meccanismo di conduzione e il meccanismo di trasduzione.

Il primo avviene nell'orecchio esterno e medio; il secondo nell'orecchio interno. I due meccanismi sono uno successivo all'altro.

Nel meccanismo di conduzione si ha un'amplificazione del suono e un cambio del mezzo acustico: l'onda sonora passa dall'aria al fluido intracocleare.

Nel meccanismo di trasduzione, l'onda generata all'interno del liquido cocleare viene trasportata in un impulso elettrico diretto all'ottavo nervo cranico (nervo stato-acustico).

Il percorso della vibrazione sonora è mostrato nella figura 7:

1. Le onde sonore entrano dall'orecchio esterno e tramite esso vengono incanalate nell'orecchio medio, dove tramite il sistema di ossicini, la vibrazione verrà trasportata all'orecchio interno.
2. La membrana basilare verrà messa in movimento, tramite il fenomeno dell'onda viaggiante che andrà ad attivare le cellule sensoriali, le cellule ciliate presenti nell'organo di Corti.
3. Queste ultime comunicano direttamente con il nervo uditivo tramite il ganglio spirale che manderà il segnale, cioè l'impulso nervoso, al tronco encefalico e di conseguenza alla corteccia uditiva dove il suono verrà interpretato.

Il processo non è a senso unico, infatti, il tronco encefalico comunica con l'orecchio interno e l'orecchio medio tramite il riflesso acustico e il sistema efferente.

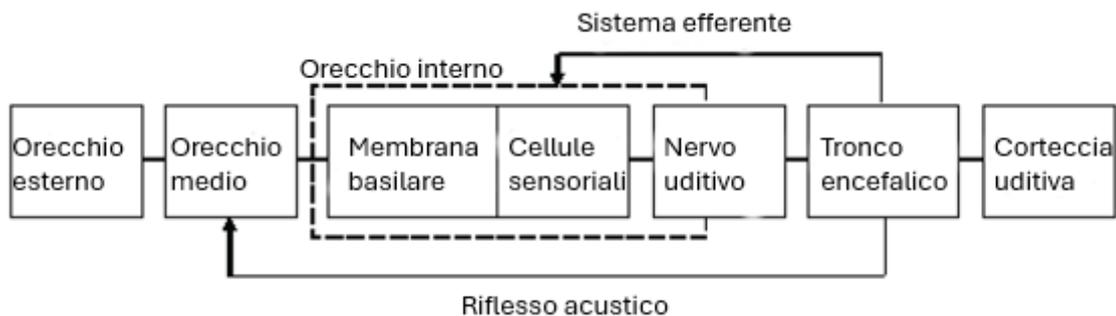


Figura 7: Percorso del suono. Figura modificata da [\[12\]](#)

MECCANISMO DI CONDUZIONE

Il suono raggiunge l'orecchio tramite il padiglione auricolare ed entra nel canale uditivo dove viene incanalato verso il timpano. Qui, l'onda sonora incidente mette in vibrazione la membrana timpanica che subirà una deflessione alla stessa frequenza dell'onda sonora. Il movimento della membrana timpanica muove la catena di ossicini dell'orecchio medio (figura 8).

Il martello, attaccato alla membrana timpanica, viene messo in movimento dalla deflessione di quest'ultima e così facendo sposta l'incudine che a sua volta muoverà l'ultimo ossicino, la staffa. La staffa si comporta come un pistone, cioè spinge avanti e indietro la finestra ovale della coclea deformandola trasportando così la vibrazione sonora. Si ha quindi che, il suono proveniente e viaggiante nell'aria, viene condotto nel fluido cocleare dove si generano delle onde dovute al movimento oscillatorio della staffa.

Queste onde dalla finestra ovale si propagano nella perilinfina innescando un meccanismo a catena: dalla scala vestibolare, l'onda viaggia verso l'elicotrema per poi ritornare verso la scala timpanica dove raggiungerà la finestra rotonda. Questo meccanismo è permesso grazie alla presenza del liquido all'interno della coclea che, essendo incomprimibile, fa sì che il volume spostato debba essere compensato. Di conseguenza il movimento che parte dalla deflessione della finestra ovale arriva a deflettere la finestra rotonda [13].

In questo meccanismo, l'orecchio esterno svolge un ruolo incanalizzatore del suono, invece l'orecchio medio svolge due importanti ruoli: amplifica il suono e svolge un ruolo di adattatore di impedenza tra aria e fluido intracocleare.

Il sistema di adattamento dell'impedenza svolto dall'orecchio medio è permesso da due meccanismi separati: la trasformazione del rapporto d'area e l'azione della leva della catena di ossicini.

Il primo meccanismo, il più efficace, è dovuto alla differenza di area tra la membrana timpanica e la finestra ovale. L'area della membrana timpanica è circa 17,2 volte più grande dell'area della finestra ovale, ciò si traduce in un aumento della pressione di circa 25 dB.

Nel secondo meccanismo si ha che la lunghezza del martello è circa 1,3 volte maggiore della lunghezza della staffa. Tale differenza porta ad un aumento della forza di circa 2 dB.

In totale, si ottiene quindi che, per compensare il disadattamento dell'impedenza aria-coclea, questi meccanismi portano ad un aumento della pressione sonora di circa 40-45 volte (32-33 dB) la pressione sonora presente sulla membrana timpanica [12].

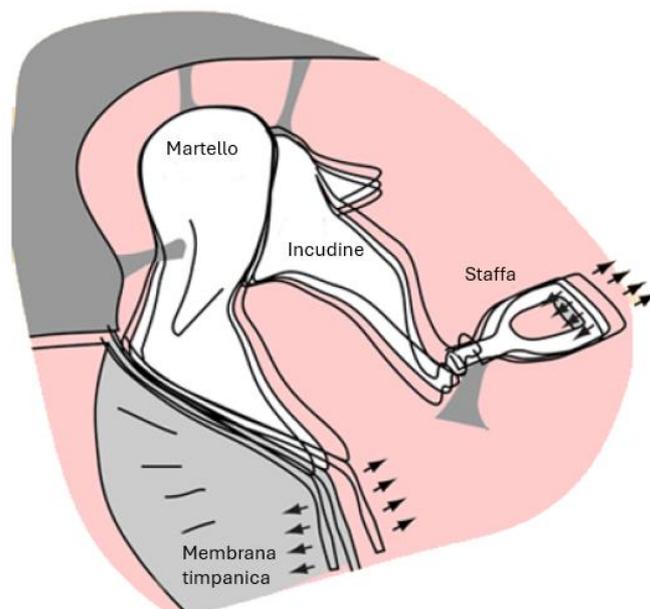


Figura 8: Meccanismo di conduzione nell'orecchio medio. Dalla figura si può notare bene il movimento vibratorio di tutto il sistema dell'orecchio medio e anche i due meccanismi che generano l'amplificazione e l'adattamento di impedenza. La membrana timpanica ha una superficie maggiore della base della staffa così come la lunghezza del martello in confronto alla lunghezza totale della staffa. L'angolo di 90° tra la staffa e l'incudine permette alla staffa di poggiare perfettamente sulla superficie della finestra ovale così che possa muoversi avanti e indietro deformandola. Figura modificata da [4].

MECCANISMO DI TRASDUZIONE

La staffa deforma la finestra ovale percuotendola e generando un'onda viaggiante all'interno della coclea, e generando di conseguenza un'onda che si propaga lungo la membrana basilare della coclea.

L'udito umano copre un range di frequenze compreso tra circa 20 Hz e 20 kHz. Questo intervallo di frequenze è determinato dalle proprietà meccaniche della membrana basilare presente nella coclea, che si comporta da analizzatore di frequenza in tempo reale.

Quando l'onda viaggiante generata dalla staffa si muove lungo la membrana basilare, a causa delle proprietà fisiche della membrana basilare, aumenta in ampiezza e diminuisce in lunghezza d'onda fino a svanire. Infatti, le proprietà fisiche della membrana basilare variano per tutta la lunghezza della membrana: spostandosi dalla base all'apice della coclea, si parte da una membrana stretta, leggera e tesa ad una membrana sempre più larga, pesante e molle. Ciò significa che ad ogni punto della membrana vi è associato un determinato punto di risonanza; quindi, un'onda viaggiante ad una determinata frequenza, quando raggiunge un determinato punto della membrana, corrispondente al punto di risonanza della sua frequenza, termina la sua propagazione (diviene risonante).

Per il range di frequenze tipiche dell'udito umano, le alte frequenze sono percepite dalla base della coclea (circa 20 kHz) mentre le basse frequenze dall'apice (circa 20 Hz) [7]. Range di frequenze e percorso dell'onda sono rappresentati in figura 9.

La deflessione della membrana basilare dovuta all'onda viaggiante genera una forza di taglio che muove le stereociglia delle cellule ciliate presenti nell'organo di Corti aprendone i canali ionici e producendo quindi una risposta elettrica.

Ciò avviene grazie a due caratteristiche della scala media che permettono la meccanotrasduzione delle cellule ciliate: la scala media presenta l'endolinfa, una soluzione carica di ioni K^+ , che favorisce la corrente attraverso i canali; inoltre, la scala media mantiene un potenziale endococleare pari a circa 80 mV che favorisce lo spostamento dei cationi [14].

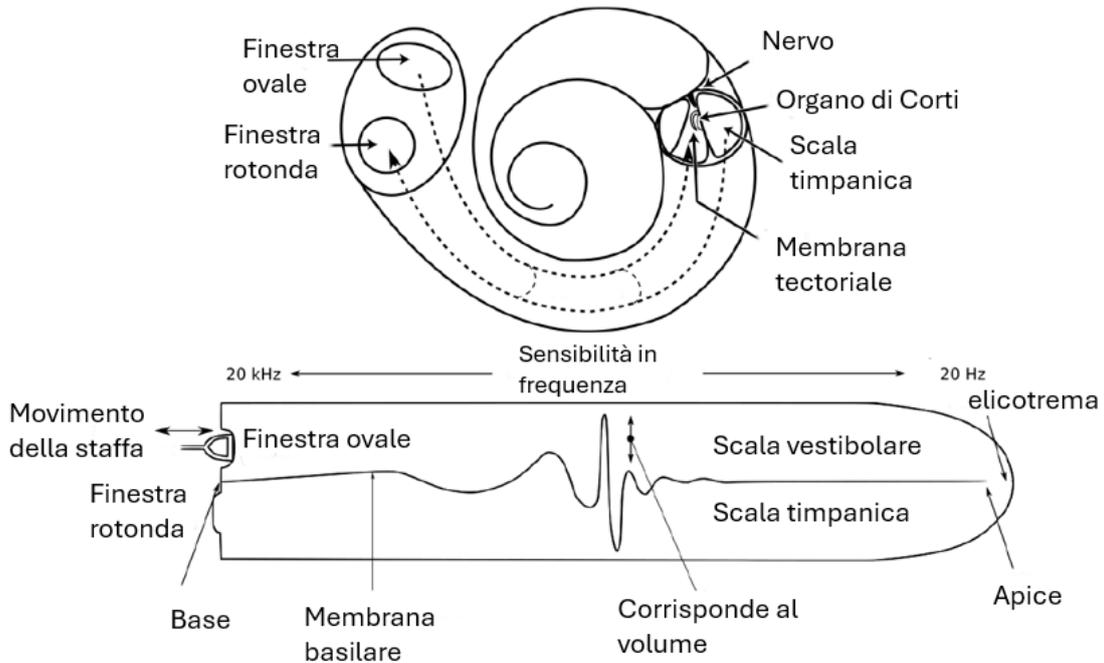


Figura 9: Tonotopia della coclea. Le linee tratteggiate all'interno della figura della coclea rappresentano il percorso dell'onda viaggiante nel fluido cocleare. La staffa preme sulla finestra ovale in modo oscillatorio generando un'onda viaggiante che si sposterà lungo la scala vestibolare e ritornerà verso la finestra rotonda tramite la scala timpanica. Come mostrato nella figura in basso, quest'onda si propaga lungo la membrana basilare. Figura modificata da [4].

Perdita uditiva

La perdita uditiva o ipoacusia consiste in un danneggiamento di una delle parti dell'orecchio. Tale danno non si limita alla sola compromissione della comunicazione, ma ha riscontri negativi anche da un punto di vista sociale, psicologico ed economico.

In base alle zone dell'orecchio colpite, l'ipoacusia viene classificata come conduttiva, neurosensoriale e mista:

- L'**ipoacusia conduttiva** riguarda l'orecchio esterno e l'orecchio medio, essa va ad interferire con il meccanismo di conduzione del suono;
- L'**ipoacusia neurosensoriale** riguarda l'orecchio interno, in particolare si presenta quando vi è un danneggiamento permanente delle cellule ciliate;
- La **perdita uditiva mista** si presenta quando si è di fronte ad una combinazione delle due. [15]

Come si può osservare dalla figura 10, ogni zona del sistema uditivo può essere colpita da una specifica malattia.

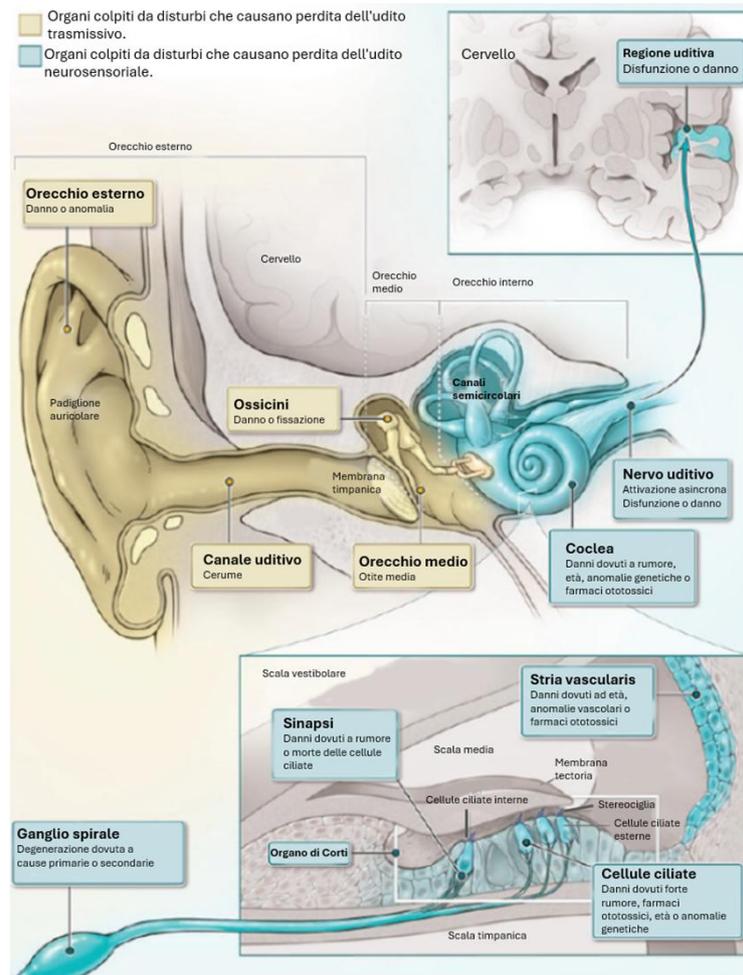


Figura 10: Malattie del sistema uditivo. La figura mostra in giallo le zone dell'orecchio che possono indurre ad ipoacusia conduttiva e in celeste le zone che possono indurre a ipoacusia neurosensoriale. Si nota come i fattori scatenanti più comuni sono l'esposizione al rumore, utilizzo di farmaci ototossici, età e malattie genetiche. Figura modificata da [15].

Per poter diagnosticare il tipo di perdita uditiva e di conseguenza scegliere il tipo di terapia da intraprendere è necessario svolgere il test audiometrico, cioè una serie di test svolti da un audiologo in un ambiente controllato. Tali prove consistono nell'udire toni puri a diverse frequenze per valutare la soglia uditiva del paziente, riconoscere parole specifiche per valutare la percentuale di parole monosillabiche che il paziente può ripetere, misurare la soglia di ricezione del parlato ed infine test di conduzione ossea, riflessi acustici e timpanometria per rilevare l'eventuale presenza di disturbi specifici [16].

Le cause che portano alla perdita uditiva sono diverse e variano in base alla posizione dell'orecchio: ostruzione del canale esterno dovuta al cerume o ad un corpo estraneo, otite, perforazione del timpano, timpanosclerosi; colesteatoma, otite media con versamento, otosclerosi, rottura della catena ossiculare; malattie autoimmuni, tumori, infezioni, esposizione al rumore, farmaci ototossici, presbiacusia o un trauma.

La gravità della perdita uditiva viene classificata in diversi modi. Secondo la WHO esistono 5 livelli di gravità, definiti in base al grado di perdita uditiva nel miglior orecchio del paziente. Tramite il test audiometrico è possibile stimare la capacità uditiva di ciascun orecchio. L'orecchio che presenta la miglior capacità residua è utilizzato come riferimento per la classificazione, così divisa:

- normale (< 25 dB), possono essere uditi suoni come il normale respiro;
- medio (da 26 a 40 dB), si inizia ad avere difficoltà nell'udito quando si è in presenza di discussioni a bassa voce, suoni deboli in ambienti poco rumorosi, discussioni a distanza o in presenza di rumore di fondo;
- moderato (da 41 a 60 dB), difficoltà nell'udire una discussione regolare anche a distanza ravvicinata o il suono del frigorifero;
- severo (da 61 a 80 dB), non può essere udita una normale conversazione a meno che non sia ad un elevato volume;
- profondo (>81 dB), possono essere percepiti dei suoni elevati come delle vibrazioni [17].

In base alla severità e alla classificazione della perdita uditiva viene scelto il metodo per trattare il problema clinico. Tali metodi possono essere farmacologici, interventi chirurgici e tecnologia acustica. Nella tecnologia acustica vi sono gli apparecchi acustici, gli impianti cocleari e le protesi impiantabili. Tale tecnologia da sola non basta, infatti deve far parte di un percorso riabilitativo. Generalmente, chi soffre di ipoacusia da lieve a moderata utilizza gli apparecchi acustici. Gli apparecchi acustici hanno il compito di amplificare e trasmettere il suono. Si dividono in analogici e digitali, entrambi catturano l'energia sonora, la trasformano in segnali elettrici e una volta amplificati li trasmettono verso la membrana timpanica. La differenza tra i due sta nella possibilità, da parte degli apparecchi digitali, di poter essere programmati in modo da adattarsi meglio alla persona. Quando gli apparecchi acustici non ottengono nessun risultato, entrano in gioco gli impianti cocleari. Essi trattano l'ipoacusia neurosensoriale bilaterale con un grado di severità compreso tra moderato e profondo e vanno a stimolare direttamente il nervo uditivo di una persona non udente [1].

2 Stato dell'arte

Dispositivo

L'impianto cocleare è una neuroprotesi in grado ripristinare l'udito in persone affette da ipoacusia neurosensoriale grave o profonda. È un metodo di trattamento approvato sin dagli anni 80 e spesso interviene quando gli apparecchi acustici non sono in grado di risolvere il problema uditivo. Tale dispositivo è in grado di bypassare l'orecchio esterno e medio così che possa stimolare direttamente il nervo uditivo dell'orecchio interno tramite una stimolazione elettrica. Per far ciò, l'impianto cocleare è costituito da una parte esterna detta processore del suono e da una parte interna impiantata chirurgicamente (figura 11). Il processore del suono è tipicamente indossato dietro l'orecchio ed è collegato con l'impianto interno tramite un magnete [18].

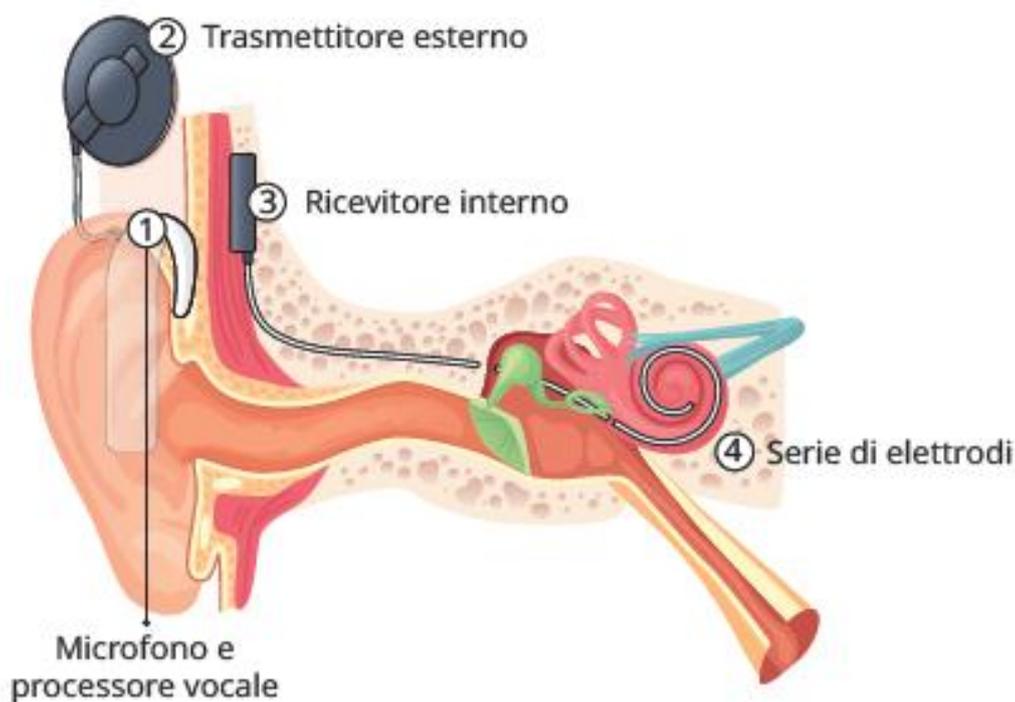


Figura 11: Parti fondamentali dell'impianto cocleare. I componenti base di un impianto cocleare sono fondamentalmente 4:1. Processore vocale e microfono, 2. Trasmettitore esterno, 3. Ricevitore interno e 4. Array di elettrodi. Le parti 1 e 2 compongono la parte esterna dell'impianto; invece, le parti 3 e 4 sono la parte impiantata chirurgicamente. Figura presa da [\[1\]](#).

Il ruolo del processore esterno è captare il suono proveniente dall'esterno, elaborarlo e trasformare questo segnale acustico in segnale elettrico. Il segnale elettrico viene così inviato nella parte impiantata della neuroprotesi tramite un segnale a radiofrequenza.

Per svolgere il suo ruolo, come mostrato in figura 12, il processore sonoro è così composto:

- uno o più microfoni
- un'unità di elaborazione del segnale digitale (DSP)
- un amplificatore di potenza
- un trasmettitore a radiofrequenza
- una batteria.

Il percorso di elaborazione inizia dai microfoni; essi captano le onde sonore e le inviano al DSP. Quest'ultimo estrae le caratteristiche del suono convertendole in un segnale digitale che viene così inviato all'impianto interno tramite l'antenna esterna a radiofrequenza.

Il DSP non si limita solo ad estrarre le caratteristiche specifiche del suono, esso permette inoltre di memorizzare delle informazioni specifiche per ogni paziente così che possa essere adattato ad ogni tipo di utente.

Come il DSP, anche il trasmettitore a radiofrequenza ha una duplice funzione: non si limita solo ad inviare i pattern di stimolazione creati dal DSP ma, essendo l'impianto interno sprovvisto di batteria, l'antenna trasmittitrice invia anche l'alimentazione [19].



Figura 12: Esempio di processore del suono (Processore sonoro Naida CI M90 di Advanced Bionics). Nella figura si possono notare i componenti fondamentali del processore esterno: il trasmettitore esterno a radiofrequenza, la batteria, il DSP e i microfoni. Figura presa e modificata da [20].

I segnali contenenti il tipo di stimolazione e l'alimentazione vengono inviati dall'antenna trasmittitrice esterna all'antenna ricevitrice interna tramite un collegamento a radiofrequenza. Le due antenne sono mantenute insieme e allineate grazie ad un magnete. I segnali, una volta arrivati nell'impianto interno, vengono usati per stimolare il nervo uditivo.

Per tale scopo, l'impianto cocleare interno è composto dalle seguenti parti (figura 13):

- antenna ricevitrice a radiofrequenza accoppiata induttivamente con il trasmettitore esterno
- stimolatore ermeticamente sigillato
- array di elettrodi.

Lo stimolatore è composto da circuiti elettrici attivi che decodificano i segnali provenienti dalla parte esterna e li convertono al fine di generare le correnti di stimolazione che viaggeranno lungo l'array di elettrodi.

L'array di elettrodi è solitamente composto da 12-22 elettrodi ed è lungo circa 2 cm; esso viene impiantato chirurgicamente all'interno della coclea per il primo giro e mezzo. L'array di elettrodi è così composto per rispettare la tonotopia biologica della coclea: ogni contatto dell'elettrodo è posto vicino alle fibre uditive in modo da rispettare la loro sensibilità alla frequenza; quindi, la stimolazione è specifica per ogni elettrodo [21].

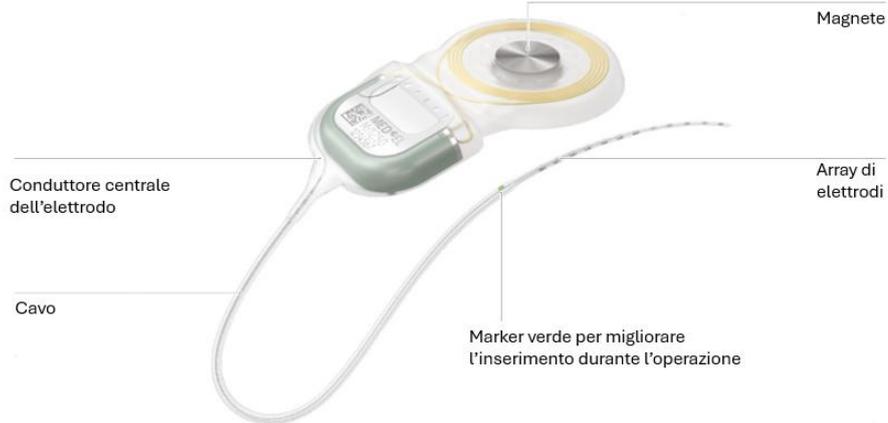


Figura 13: Esempio di impianto interno (SYNCHRONY 2 di Med-El). Il ricevitore avvolge il magnete ed è collegato con lo stimolatore ermeticamente sigillato. Da qui parte il cavo contenente l'array di elettrodi. È importante notare come il dispositivo sia pensato per l'operazione chirurgica; infatti, contiene un marker di colore verde per aiutare tale operazione. Figura presa e modificata da [22].

Tutta l'architettura di un impianto cocleare si può riassumere nella figura 14.

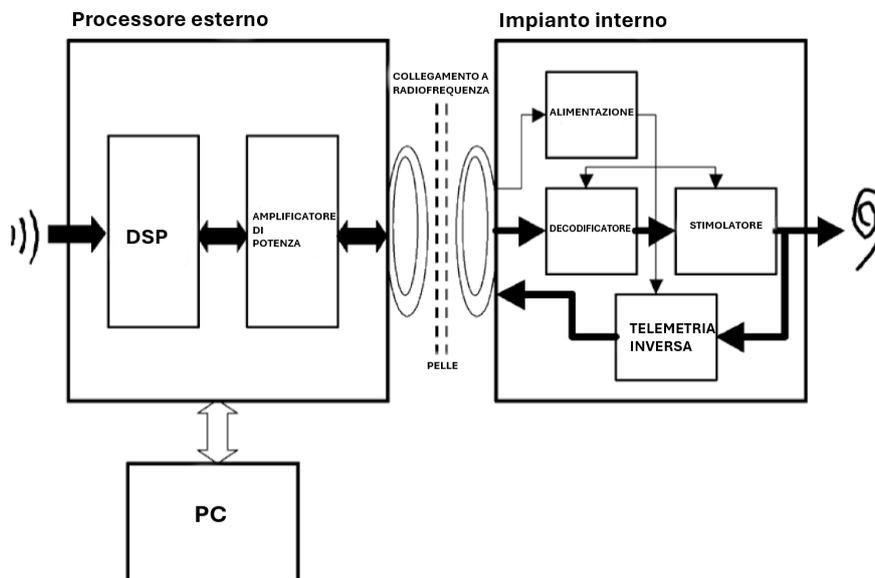


Figura 14: Architettura a blocchi di un tipico impianto cocleare. Tra le onde sonore, l'inizio del percorso del suono, e l'array di elettrodi, la fine del percorso, ci sono tre macro-blocchi: il processore esterno, composto dal DSP e dall'amplificatore di potenza; il PC, necessario per la programmazione del processore esterno e l'impianto interno. L'impianto interno è composto da più blocchi: l'alimentazione, il decodificatore, lo stimolatore e la telemetria inversa. L'alimentazione proviene dall'impianto esterno ed è necessaria per alimentare gli altri tre blocchi. Il decodificatore o decoder interpreta e converte i segnali provenienti dall'esterno. Lo stimolatore genera le correnti da mandare all'array di elettrodi. L'impianto interno non ha una comunicazione unidirezionale con l'impianto esterno, infatti, tramite un sistema di telemetria inversa esso comunica con la parte esterna. Figura presa e modificata da [19].

Aziende produttrici e costi

Le principali aziende produttrici di impianti cocleari sono 4:

- Cochlear Ltd. [\[23\]](#)
- MED-EL [\[22\]](#)
- Advanced Bionics [\[20\]](#)
- Oticon Medical [\[24\]](#)

Le prime tre sono le aziende predominanti sul mercato italiano. Di seguito sono riportati gli attuali impianti presentati da ognuna delle tre aziende per trattare il problema dell'ipoacusia neurosensoriale.

In questa piccola analisi verrà data importanza all'autonomia energetica di ognuno di essi così da ottenere un confronto con il sensore tonotopico passivo presentato in questo elaborato di tesi. È importante notare come l'autonomia energetica, indipendentemente dalla marca, sia solo un valore di riferimento in quanto varia molto dall'utilizzo che si fa dell'impianto stesso. Inoltre, è importante definire sin da subito delle differenze che vi possono essere tra i processori del suono appartenenti ad una stessa azienda. In generale il processore del suono può essere retroauricolare (detto anche BTE) oppure a bottone (o off-the-ear). La principale differenza sta nella forma e nella modalità di alimentazione.

Il BTE è il classico processore esterno indossabile intorno all'orecchio e avente antenna trasmittitrice al di fuori del processore impiantata sulla cute della testa e quindi collegata al processore tramite un filo; quindi, l'antenna è a vista. Tendenzialmente possiede due possibilità di scelta dal punto di vista dell'alimentazione: batterie usa e getta o ricaricabili.

Il processore del suono a bottone è un dispositivo compatto, cioè tutti i componenti sono uniti in un solo dispositivo, di conseguenza l'antenna non è collegata all'esterno del processore tramite un filo esterno ma è incorporata all'interno del processore sonoro. Tendenzialmente sono dispositivi con batterie ricaricabili, quindi è possibile staccare l'intero dispositivo per effettuare la ricarica.

COCHLEAR Ltd.

Per quanto riguarda la Cochlear, attualmente offre due tipi di processore del suono (figura 15): Nucleus 8 e Nucleus Kanso 2. Si possono notare due importanti differenze. La prima differenza è nella forma del processore: il primo è il classico retroauricolare o BTE; invece, il secondo è un processore del suono a bottone o off-the-ear. L'altra importante differenza risiede nella batteria:

- Il Nucleus 8 ha la possibilità di usare batterie usa e getta zinco-aria avente autonomia di 35 ore, invece se alimentato con batterie ricaricabili, queste hanno durata di 10 o 20 ore a seconda dell'impostazione scelta dall'utilizzatore;
- Il Kanso 2 ha solo una batteria integrata ricaricabile della durata massima di 18 ore, non dispone di batterie usa e getta.



Figura 15: A sinistra il processore del suono Nucleus 8 di tipo BTE, a destra il Kanso 2 di tipo off-the-ear. Immagini prese da [23].

Dal punto di vista dell'impianto interno, Cochlear dispone di due serie, la Nucleus Profile e la Profile Plus. Entrambe hanno lunghezze differenti in modo da adattarsi ai diversi pazienti. La caratteristica in comune tra tutti è la presenza 22 canali attivi.

Med-EI

Anche Med-EI come Cochlear propone due tipi di processore del suono (figura 16): Il SONNET 3 e il RONDO 3. Il primo è BTE, il secondo è off-the-ear. Anche qui la differenza principale è sempre a livello di autonomia energetica:

- Il SONNET 3 ha quattro possibilità di scelta: 2 batterie usa e getta zinco-aria 675 con autonomia pari a 60 ore oppure batterie ricaricabili di tre diverse misure (micro, standard e max). Passando dalle micro alle max aumenta la dimensione e il peso. Dal punto di vista dell'autonomia, le batterie micro presentano un'autonomia di sole 7 ore, le standard 10 ore, infine le max 17 ore. Dal punto di vista dei cicli carica e scarica le micro e le standard presentano un valore maggiore a 1100 cicli, le max invece 1800 cicli.
- Il RONDO 3 utilizza batterie ricaricabili, in particolare una batteria agli ioni di litio integrata con autonomia di 24 ore e cicli scarica e carica superiori a 2000.



Figura 16: A sinistra il processore del suono SONNET 3, a destra il RONDO 3. Immagine presa da: [22].

Per l'impianto interno vi è il SYNCHRONY 2. Anche qui ci sono diverse forme in base alle necessità del paziente. Med-El presenta la serie FLEX per utenti con coclee normali e la serie FORM per utenti aventi coclee malformate o per situazioni in cui si prevedano perdite di fluido cerebrospinale (CSF); entrambe le serie presentano un catalogo con lunghezze diverse per adattarsi alla persona. La serie FLEX presenta 19 elettrodi in platino attivi mentre la serie FORM possiede 24 elettrodi in platino attivi. Questo impianto presenta due possibilità di stimolazione: stimolazione non simultanea in sequenza a 12 canali o stimolazione simultanea da 2 a 12 canali. L'elettrodo di riferimento della stimolazione è posto sull'involucro in titanio. Dal punto di vista del tipo di impulso, esso è disponibile sia per tipi di impulso bifasi, trifasici simmetrici e trifasici.

Advanced Bionics

L'Advanced Bionics presenta 5 tipi di processori sonori: Marvel, Naída, Neptune, Chorus, Harmony.

Considerando il processore BTE Naída CI M90 progettato per persone adulte (figura 12), è possibile osservare diverse scelte per l'alimentazione prendendo in considerazione la loro durata con un certo tipo di strategia di processing (HiRes Optima):

- Batterie small: con una capacità di 90 mAh garantiscono più di 9 ore di autonomia
- Batterie medium: con una capacità di 150 mAh garantiscono più di 16 ore di autonomia
- Batterie large: con una capacità di 220 mAh garantiscono più di 16 ore di autonomia anche con altri tipi di strategie
- Batterie standard: con una capacità di 160 mAh garantiscono più di 16 ore di autonomia
- Batterie impermeabili: con una capacità di 225 mAh garantiscono più di 18 ore di autonomia
- Batterie Zinco-Aria: garantiscono più di 16 ore di autonomia

Ogni tipo di batteria avrà peso e dimensioni differenti.

Invece come impianto interno, l'ultimo modello è l'HiRes-Ultra 3D il quale presenta 16 elettrodi in platino.

CLASSE CE E COSTO IN ITALIA

Analizzando questi dispositivi tramite l'Elenco dei dispositivi medici [\[25\]](#) presenti in Italia è stato possibile risalire alla loro classe di rischio, normativa e CND (Classificazione Nazionale dei Dispositivi medici), riassunti nella seguente tabella. Si nota come il processore esterno ed interno abbiano CND differenti. La classe di rischio invece risulta essere uguale per entrambi i processori e risulta essere la più alta in tale scala. Uno stesso dispositivo può presentare due classi e due normative differenti, ciò è dovuto all'anno in cui tale dispositivo è stato prodotto per la prima volta.

TIPO DI IMPIANTO	AZIENDA	NOME DISPOSITIVO	CND	NORMATIVA	CLASSE CE
Processore del suono	Cochlear	Nucleus 8	J0380 - IMPIANTI AURICOLARI ATTIVI - ACCESSORI	Reg. UE 2017/745	R3 - CLASSE III
Processore del suono	Cochlear	Kanso 2	J0380 - IMPIANTI AURICOLARI ATTIVI - ACCESSORI	Reg. UE 2017/745	R3 - CLASSE III
			J0301 - IMPIANTI COCLEARI	D.L.vo 507/92 attuazione Dir. CE 90/385	IA - Impiantabili attivi
Processore del suono	Med-El	Sonnet 3	J0380 - IMPIANTI AURICOLARI ATTIVI - ACCESSORI	Reg. UE 2017/745	R3 - CLASSE III
Processore del suono	Med-El	Rondo 3	J0380 - IMPIANTI AURICOLARI ATTIVI - ACCESSORI	D.L.vo 507/92 attuazione Dir. CE 90/385	IA - Impiantabili attivi
Processore del suono	Advanced Bionics	Naida CI M90	J0380 - IMPIANTI AURICOLARI ATTIVI - ACCESSORI	D.L.vo 507/92 attuazione Dir. CE 90/385	IA - Impiantabili attivi
Impianto interno	Cochlear	Profile plus	J0301 - IMPIANTI COCLEARI	D.L.vo 507/92 attuazione Dir. CE 90/385	IA - Impiantabili attivi
Impianto interno	Med-El	Synchrony 2	J0301 - IMPIANTI COCLEARI	Reg. UE 2017/745	R3 - CLASSE III
				D.L.vo 507/92 attuazione Dir. CE 90/385	IA - Impiantabili attivi
Impianto interno	Advanced Bionics	HiRes Ultra 3D	J0301 - IMPIANTI COCLEARI	D.L.vo 507/92 attuazione Dir. CE 90/385	IA - Impiantabili attivi

Tabella 1: CND, normativa e classe CE degli attuali impianti in commercio.

In Italia la gestione di un impianto cocleare avviene tramite dei bandi di gara. Da una analisi di questi, il costo di un impianto totale si aggira intorno ai 15.000 euro, dove la maggior parte del costo è ricoperto dal processore esterno.

Processore del suono (microfoni e strategie di elaborazione)

Il processore del suono, come già detto, è quella parte della neuroprotesi che si occupa di captare il suono ed elaborarlo. Per quanto la parte interna sia la più impattante a causa dell'intervento chirurgico, il processore esterno è la parte fondamentale di tutto l'impianto. Infatti, esso è il vero e proprio "cervello" di tutta la neuroprotesi e da esso dipende la funzione principale dell'impianto cocleare, cioè ripristinare l'udito.

Il suono non deve essere solo captato, esso deve anche essere elaborato in modo da svolgere il suo ruolo funzionale nella vita di tutti i giorni dal punto di vista sociale: udire le conversazioni.

È necessario quindi che il parlato venga compreso in modo fluente. Proprio come avviene biologicamente, è necessario filtrare il rumore di fondo, cioè porre attenzione a ciò che è veramente importante fra i tanti suoni presenti nell'ambiente. È necessario quindi che il dispositivo si adatti alle diverse situazioni in modo da evitare il fenomeno del "cocktail party", cioè l'incapacità di distinguere la sorgente sonora di interesse a causa della presenza di altri suoni come può avvenire quando si è in una stanza con più interlocutori.

Per questo motivo, negli anni, la parte che è stata più modificata dell'impianto cocleare è stata la strategia di elaborazione del suono in modo da migliorare la comprensione del parlato.

MICROFONI E PRE-ELABORAZIONE

Indipendentemente dal tipo di strategia scelta, una caratteristica fondamentale è la pre-elaborazione del suono presente in tutti i dispositivi. Tale pre-elaborazione avviene tramite i microfoni e specifiche strategie finalizzate a ridurre il rumore. Negli impianti cocleari non viene mai usato un singolo microfono per captare il suono, ma bensì vengono usati più microfoni di tipo direzionale. Tali microfoni non si limitano a individuare la posizione della sorgente sonora ma l'obiettivo principale è separare le sorgenti sonore in base alla loro posizione spaziale in modo da poter separare ciò che è realmente utile, come il parlato, dal rumore di fondo.

Il suono, una volta ricevuto dai microfoni, viene ripulito dal rumore in modo da alzare il rapporto segnale-rumore (signal-noise ratio, SNR). Tale riduzione del rumore avviene tramite tecniche tradizionali o più recentemente tramite algoritmi basati su Machine Learning. Le tecniche tradizionali ipotizzano delle differenze tra parlato e rumore: viene stimato lo spettro di potenza del rumore, quindi, viene calcolato un SNR a priori per ogni unità di tempo, e tramite questo rapporto segnale-rumore si implementa un filtro per enfatizzare il parlato rispetto al rumore di fondo. Invece, le moderne tecniche di Machine Learning utilizzano algoritmi basati su modelli allenati su dei dati di tipo acustico generalizzate, ad esempio sono addestrati su dati composti da diverse voci coperte da uno specifico rumore di fondo [26].

Una volta che è stata effettuata questa pre-elaborazione è possibile passare alla strategia di elaborazione del suono vera e propria, necessaria per comandare la parte interna dell'impianto cocleare e quindi gestire il pattern di stimolazione.

STRATEGIE DI ELABORAZIONE DEL PARLATO

L'impianto cocleare non si limita alla sola percezione sonora. Affinché avvenga la completa reintegrazione del paziente nella vita sociale, è necessario che la percezione del parlato sia ottimizzata. Affinché una persona portatrice di impianto cocleare comprenda fluentemente un discorso, è fondamentale che la neuroprotesi sia implementata con un adeguata strategia di elaborazione del suono. Per discutere del funzionamento di tali strategie, è necessario definire come viene analizzato un segnale vocale.

Quando si parla di segnale vocale, esso può essere analizzato dal punto di vista della produzione del parlato o dal punto di vista della percezione del parlato.

Nel primo caso, ciò che si va ad osservare sono:

- Periodicità della forma d'onda: definisce se un suono è sonoro o non sonoro (vocali vs consonanti) in quanto analizza lo stato vibrazionale delle corde vocali;
- Durata temporale del suono e intervalli silenziosi: definisco l'articolazione, cioè come l'aria viene modificata passando lungo il tratto vocale (consonante nasale, fricativa, occlusiva, affricate, ecc)
- Le formanti definite in frequenza: definiscono il luogo di articolazione

Nella percezione del parlato, ciò che si osserva è la modulazione in ampiezza o in frequenza, in particolare:

- nella modulazione in ampiezza si osservano gli involuppi in determinate bande di frequenza
- nella modulazione in frequenza si osserva la struttura fine del segnale.

Tutti gli impianti cocleari hanno adottato la modulazione in ampiezza, Med-El si è concentrata molto sulla modulazione in frequenza, quindi sull'analisi della struttura fine temporale del segnale [\[27\]](#).

Indipendentemente dalla strategia scelta, ciò che il processore del suono deve mandare all'impianto interno è il pattern di stimolazione. Tale pattern deve contenere le informazioni riguardanti il suono, tali informazioni sono:

- Rappresentazione in frequenza e durata del suono: per fornire tali informazioni viene sfruttata la tonotopia della coclea. Viene creato ciò che viene detto cocleogramma: una mappa tempo-frequenza della stimolazione. Ogni elettrodo stimola un determinata zona cocleare in un certo momento. I canali di stimolazione non sono quindi attivati contemporaneamente per tutta la durata del suono. In questo modo viene rispettata la mappa posizione-frequenza.
- Altezza del suono: questa informazione è decisa dalla periodicità delle modulazioni temporali applicata ai treni di impulsi e descrive quanto un suono è acuto o grave. Più veloce sarà la modulazione, più l'altezza è elevata;
- Intensità del suono: viene stabilita dalla quantità di carica rilasciata dal singolo elettrodo, più il suono è forte, maggiore sarà la corrente di stimolazione. Questo perché aumentando la quantità di carica aumenta il reclutamento delle fibre nervose o aumenta l'efficacia delle fibre che comunque verrebbero stimulate [\[21\]](#).

L'invio di queste informazioni viene esemplificato nella figura 17. Ogni suono è caratterizzato da una durata e da un contenuto in frequenza. Estrahendo queste due informazioni è possibile pilotare correttamente gli elettrodi intracocleari così da inviare le informazioni riguardo a frequenza, altezza e intensità sonore.

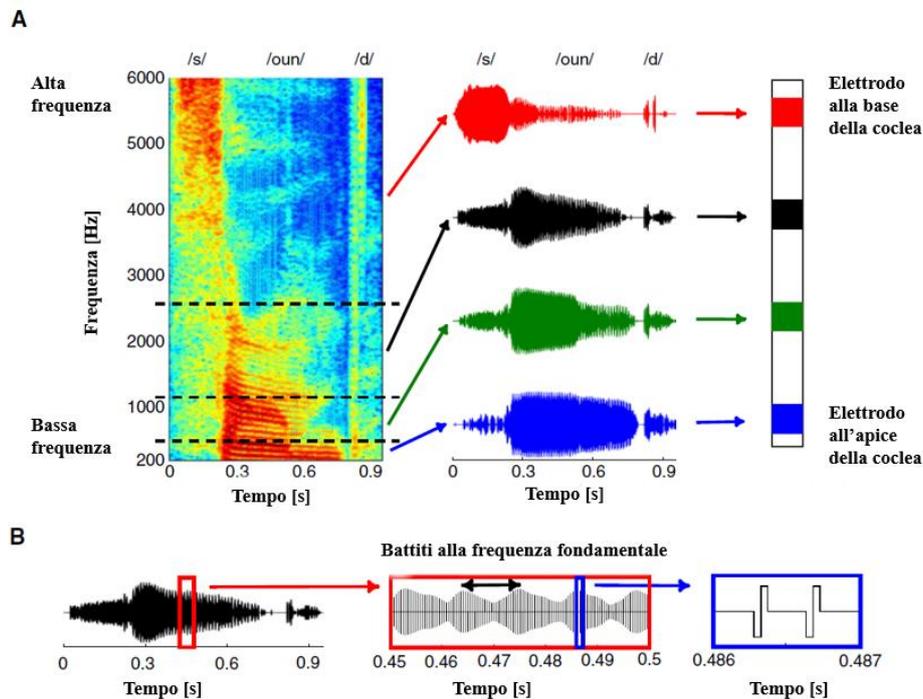


Figura 17: Trasformazione del suono in segnale elettrico per l'impianto interno. La parte superiore (A) rappresenta uno spettrogramma, cioè una rappresentazione tempo-frequenza, della parola "sound". Lo spettrogramma mostra che ogni parte costituenti della parola è caratterizzata da una specifica banda in frequenza. Tale informazione viene utilizzata per gestire il tipo di stimolazione di ogni elettrodo: ogni elettrodo avrà una stimolazione differente in base alla sua posizione cocleare. Tramite questo metodo viene inviata l'informazione sulla frequenza del suono. Nel pannello B viene mostrato un singolo segnale mandato ad un singolo elettrodo. Qui si può notare l'informazione dell'altezza e dell'intensità del suono. L'altezza del suono, e quindi della voce di chi parla, viene rappresentata dalla modulazione temporale degli impulsi, in questo caso pari a 100 Hz. Invece, l'intensità viene mandata tramite gli impulsi bifasici e simmetrici. La variazione dell'intensità è rappresentata dalla variazione in ampiezza degli impulsi. Figura presa e modificata da [21].

Dopo questo breve riepilogo sull'analisi del segnale vocale, verrà illustrato come le aziende abbiano sfruttato tali analisi negli impianti cocleari. Le strategie di elaborazione vocale si dividono principalmente in due categorie: i dispositivi che analizzano le caratteristiche grossolane del segnale e i dispositivi che vanno più a fondo analizzando le caratteristiche spettrali fini. L'analisi delle caratteristiche spettrali grossolane altro non è che un'analisi del segnale basata sulla modulazione in ampiezza, cioè trovare l'involuppo del segnale nel dominio del tempo o della frequenza. Invece, l'analisi delle caratteristiche spettrali fini corrisponde ad una modulazione in frequenza, cioè un'analisi della struttura fine del segnale sia essa nel tempo o nella frequenza. Alcune aziende hanno optato per dar maggior risalto ad un'analisi basata sulla modulazione in ampiezza mentre

altre sulla modulazione in frequenza. A prescindere dal tipo di analisi condotta, l'obiettivo di ogni azienda è mandare le informazioni descrittive del suono (sopra citate) alla coclea.

Ogni azienda ha presentato negli anni diversi tipi di strategie di elaborazione del suono. È possibile classificare le più importanti in base al tipo di tecnica utilizzata per l'analisi vocale:

- **INVILUPPO SPETTRALE:** commutazione della frequenza (F0, F1, F2), filtraggio in bande di frequenza (FIS);
- **INVILUPPO TEMPORALE:** CIS, strategia n-of-m (elaborata nelle strategie commerciali SPEAK e ACE)
- **ANALISI DELLA STRUTTURA FINE:** HiRes, FSP.

Ognuna di queste è stata sviluppata ed utilizzata da una delle tre aziende nominate nel paragrafo precedente: CIS, SPEAK e ACE sono sviluppate dalla Cochlear, la HiRes dall'Advanced Bionics e la FSP da Med-El.

Sebbene le strategie siano in continua evoluzione, nell'ambito degli impianti cocleari, una strategia maggiormente implementata non è detto che sostituisca una più datata. La scelta finale della strategia da adoperare spetta all'utilizzatore. Tale decisione sarà presa in base alle sue esigenze: un strategia apparentemente più semplice può dimostrarsi più affidabile rispetto ad una strategia più evoluta.

In figura 18 viene mostrata l'evoluzione delle strategie di elaborazione del suono.

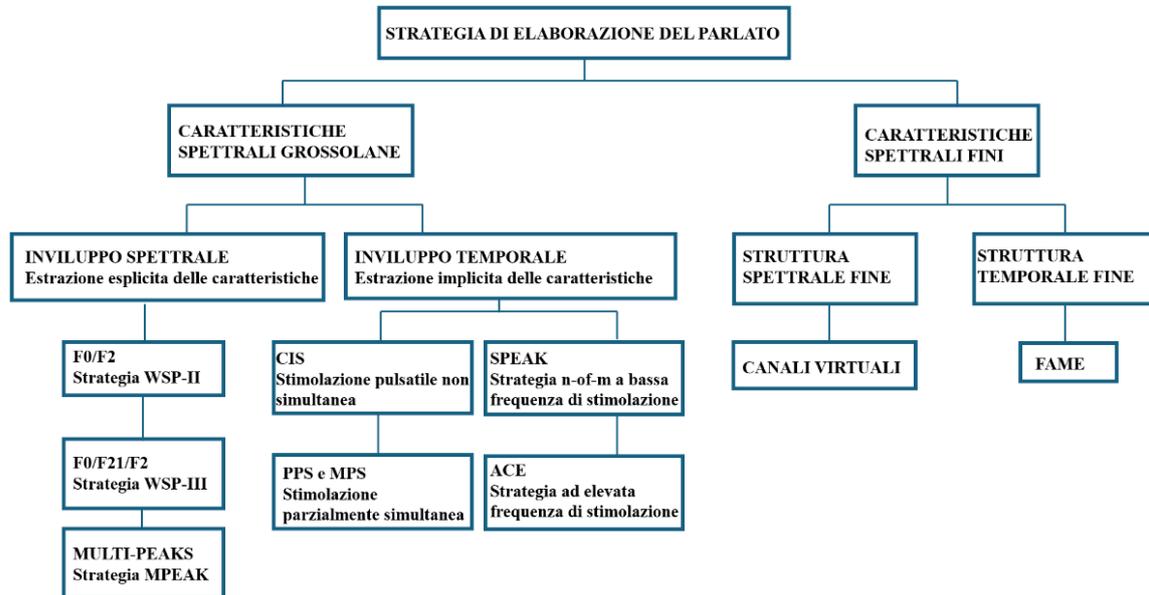


Figura 18: Evoluzione delle strategie di elaborazione. Figura modificata da [19].

Segue, un'analisi delle strategie più importanti. Le strategie sono state raggruppate in base alla tecnica utilizzata e in base all'evoluzione temporale.

INVILUPPO SPETTRALE

Il primo impianto cocleare, chiamato 3M/House era composto da un singolo elettrodo, solo successivamente si è passati a sistemi con più elettrodi per rispettare la tonotopia della coclea. Questo impianto comprimeva l'onda sonora analogica in ampiezza per poi effettuare una modulazione in ampiezza tramite una portante sinusoidale a 16 kHz che veniva mandata al singolo elettrodo per effettuare la stimolazione [28]. Successivamente si è passati a sistemi multielettrodo. Questi dispositivi, risalenti agli anni 80, basavano la loro analisi del segnale sull'inviluppo spettrale, cioè veniva estratta la frequenza fondamentale detta F0 e la seconda formante detta F2. Dalla F0 è possibile ricavare il tono della voce dello speaker; quindi, ad esempio, permette di riconoscere un uomo da una donna. Questa strategia si è evoluta nel tempo estraendo la prima formante F1 e successivamente estraendo i tre picchi compresi tra 2000 e 8000 Hz nella strategia MULTI-PEAKS [19].

INVILUPPO TEMPORALE

Con il passare del tempo, è cambiata la filosofia di pensiero decidendo di estrarre le caratteristiche tramite l'inviluppo temporale. In questo modo il segnale viene separato in tante bande specifiche per ogni singolo elettrodo. Quindi ad ogni elettrodo arriverà un'analisi dello stesso segnale ma diversa in base alla sua posizione lungo la coclea. In figura 17 viene rappresentato il concetto alla base di tale analisi, nell'immagine viene mostrato come per ogni elettrodo cambia il segnale di stimolazione nonostante il segnale sonoro di partenza sia lo stesso. Ciò viene implementato in modo da enfatizzare le caratteristiche spettrali riguardanti lo specifico elettrodo. Nella classificazione dell'inviluppo temporale, le strategie predominanti sono la CIS, la SPEAK e la ACE.

CIS: la Continuous Interleaved Sampling (CIS) è stata sviluppata dalla Cochlear. L'obiettivo di tale strategia è utilizzare l'inviluppo dei segnali uscenti da diversi filtri passa-banda per modulare in ampiezza gli impulsi bifasici uscenti dagli elettrodi. Per far ciò vengono seguiti i seguenti step:

1. Il suono in ingresso entra in un banco di filtri passa-banda (solitamente un numero compreso tra 4 e 22 filtri);
2. Dal segnale uscente da ogni filtro viene calcolato l'inviluppo temporale
3. Ogni inviluppo viene compresso tramite una funzione di mappatura non lineare. Questo passaggio è necessario per compensare l'elevato intervallo dinamico, cioè la differenza tra il suono più debole e il suono più forte che un determinato sistema può percepire. Nel caso degli impianti cocleari si ha che l'intervallo dinamico dei suoni provenienti dall'ambiente esterno è di circa 90 dB al contrario dell'intervallo dinamico dell'udito generato dall'impianto che è pari a 10-20 dB.
4. L'inviluppo compresso viene utilizzato per modulare in ampiezza gli impulsi bifasici mandati a tutti gli elettrodi.
5. Ogni segnale uscente da un determinato filtro passa-banda è specifico per un singolo elettrodo: il segnale uscente dal filtro passa-banda che tratta le basse frequenze sarà indirizzato all'elettrodo che tratta le basse frequenze (elettrodo più apicale) e ciò avviene per tutti gli elettrodi. La stimolazione degli elettrodi non è simultanea ma al contrario ogni stimolo verrà mandato in tempi diversi così da non creare un'interazione tra i campi elettrici.

Maggiori dettagli sulla tecnica sono forniti nella figura 19.

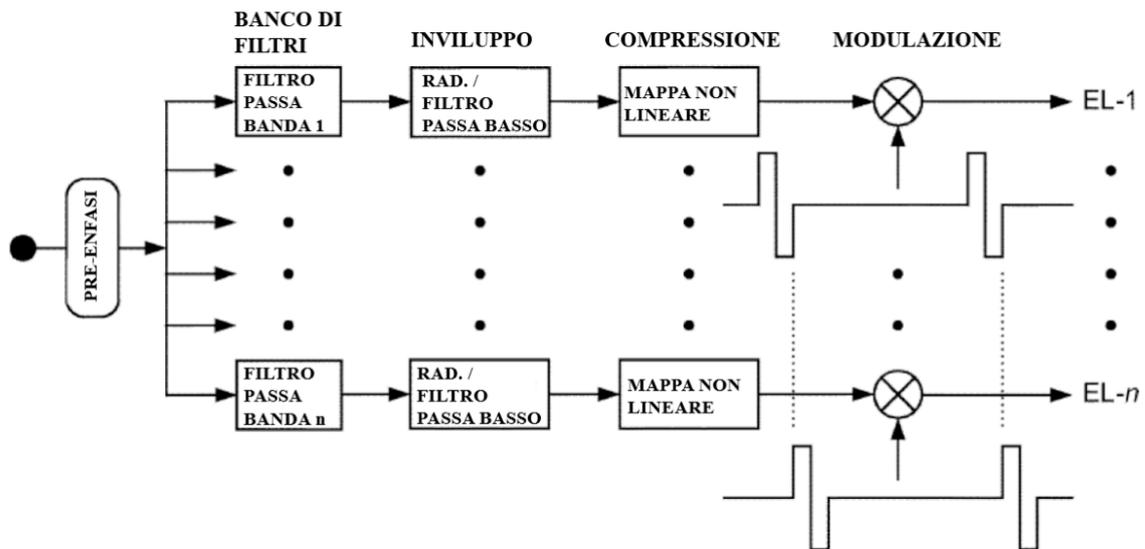


Figura 19: Schema di funzionamento della strategia CIS. Osservando la figura da sinistra verso destra si può osservare il percorso del suono nella strategia CIS. Il primo step prevede l'utilizzo di un filtro di pre-enfasi per attenuare le componenti del parlato sotto 1,2 kHz. Segue il banco di filtri passa-banda e subito dopo il calcolo dell'involuppo. L'involuppo può essere calcolato in diversi modi: generalmente viene utilizzato un raddrizzatore a onda intera o a mezza onda seguito da un filtro passa-basso a 200 Hz così da rappresentare le frequenze fondamentali dei suoni vocali. Esistono altri due metodi per calcolare l'involuppo: la trasformata di Hilbert o semplicemente un raddrizzatore a mezza onda. L'ultimo step è la modulazione in ampiezza del treno di impulsi bifasici. Figura modificata da [29].

Successivamente alla CIS è stata elaborata una nuova strategia, fondata sempre sull'involuppo temporale, denominata n-of-m.

n-of-m: nella strategia n-of-m ci sono più filtri passa-banda rispetto alla CIS per un totale di m filtri e di conseguenza m elettrodi. In questa strategia vengono scelti dei frame temporali della durata compresa tra 2 e 4 ms in cui vengono selezionati n involuppi uscenti dai filtri passa-banda. Tali involuppi corrispondono agli involuppi con maggior ampiezza. Questi n involuppi definiscono gli n elettrodi che verranno stimolati in quel determinato frame temporale. Nella strategia n-of-m dunque non vengono stimolati tutti gli elettrodi presenti come avviene nella CIS (infatti se in un determinato frame si avesse $n=m$ la strategia sarebbe esattamente uguale alla CIS) ma bensì solo gli elettrodi scelti volta per volta [19].

La strategia n-of-m nasce come strategia di ricerca. Commercialmente è stata implementata in due strategie commerciali, la SPEAK e la ACE, sebbene solo la prima sia la più rappresentativa in quanto la ACE utilizza solo dei principi simili ma ha una codifica del segnale differente.

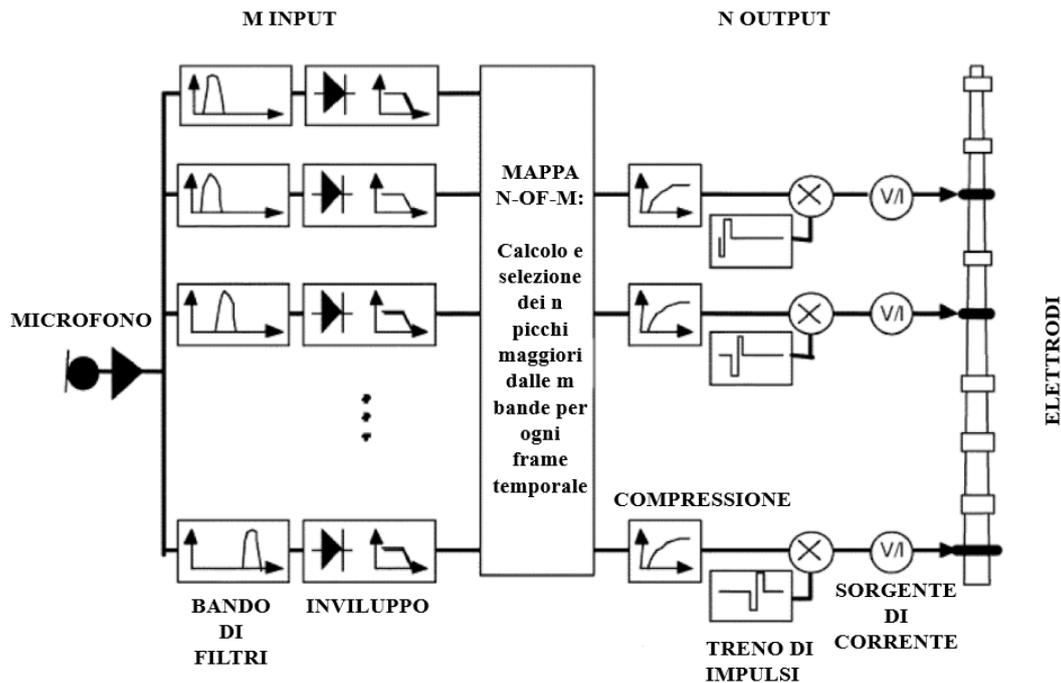


Figura 20: Strategia n-of-m. Nella figura è possibile osservare come partendo da m input corrispondenti alle m bande si ottengano n output che saranno gli n elettrodi che verranno stimolati nel frame in analisi. Questa strategia contiene gli stessi blocchi funzionali della strategia CIS, ad eccezione della MAPPA N-OF-M. Figura modificata da [19].

SPEAK: la Spectral Peak Detection (SPEAK), o peak picking, usa come nella CIS diverse bande per ottenere gli involucri temporali, ma in questo caso, utilizza solo le bande che presentano maggior energia per effettuare la stimolazione [27]. Di fatto, con la strategia SPEAK, il processore del suono analizza continuamente il segnale acustico in ingresso andando ad osservare i massimi spettrali, cioè le regioni nel dominio in frequenza che hanno maggior energia [30]. Più precisamente questa strategia seleziona i 6-8 picchi maggiori ed utilizza una frequenza fissa per ciascun canale pari a 250 Hz [19].

ACE: la Advanced Combined Encoder (ACE) unisce i punti di forza della CIS e della SPEAK. Della prima ne prende la rappresentazione spettrale mentre dalla SPEAK prende l'alta frequenza di stimolazione. Questa combinazione crea un approccio con due punti chiave:

1. Vengono stimolati solo gli elettrodi appartenenti alle bande con maggior energia spettrale
2. Gli elettrodi stimolanti sono più numerosi rispetto a quelli usati nella SPEAK: per ogni ciclo di stimolazione vengono usati da 2 a 20 elettrodi.

Un'evoluzione dell'ACE è la MP3000 in cui si utilizzano meno picchi così da ridurre la stimolazione e di conseguenza risparmiare sulla batteria [30].

In [31] gli autori hanno effettuato un confronto tra le strategie CIS e ACE ottenendo la figura 21 che rende bene l'idea di come la forma d'onda di un segnale vocale venga elaborato tramite le due strategie.

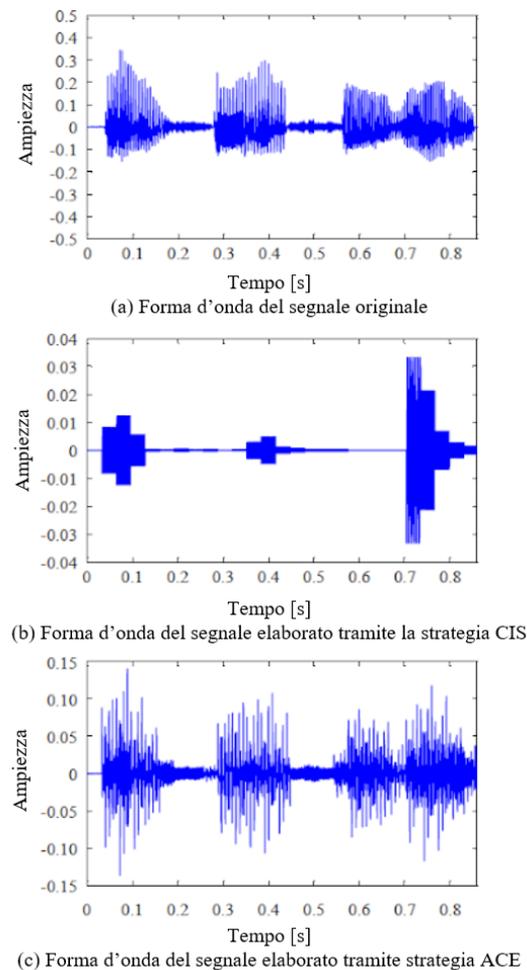


Figura 21: Confronto tra strategia ACE e CIS. Il segnale originale, raffigurato nel pannello (a), è la rappresentazione della parola “ICICM” pronunciata da una donna in inglese. Il pannello (b) rappresenta l'elaborazione di questo segnale tramite la strategia CIS, invece il pannello (c), rappresenta lo stesso ma con la strategia ACE. Si nota chiaramente come entrambe le strategie vanno a distorcere il segnale, ma la strategia ACE porta più informazioni rispetto alla CIS. Figura modificata da [31].

STRUTTURA FINE

HiRes: la strategia High Resolution (HiRes) è stata sviluppata da Advanced Bionics e utilizza ciò che viene detto uno schema di media dinamica, in cui il valore medio di un insieme di dati, in questo caso l'intensità del segnale, viene calcolato e aggiornato continuamente in base ai cambiamenti nel segnale stesso. In pratica, con questo metodo non si utilizza più l'involuppo del segnale che fornisce solo informazioni sulle variazioni lente dell'ampiezza del segnale, e che quindi utilizza una media fissa, ma in questo caso si utilizza una media che cambia continuamente con i cambiamenti del segnale, ottenendo informazioni più accurate sul segnale stesso.

L'evoluzione dell'HiRes è l'HiRes 120 basata sull'utilizzo del current steering e i canali virtuali. Il current steering è un metodo di stimolazione degli elettrodi che consiste nello stimolare gli elettrodi adiacenti invece che un elettrodo alla volta, con l'effetto di risultare un metodo di controllo della corrente molto preciso. Questo metodo permette di regolare il flusso di corrente tra due elettrodi generando ciò che viene detto canale virtuale. Utilizzando 15 coppie di elettrodi adiacenti si creano 7 canali virtuali, cioè 7 percezioni uditive per ogni coppia. Di conseguenza si può spiegare da dove deriva il nome HiRes 120: avendo 15 coppie di elettrodi che generano 7 canali virtuali si ottiene un totale di 105 canali virtuali a cui bisogna sommare i 15 canali fisici. Di conseguenza questa tecnica permette di generare ben 120 canali percettivi [30].

FSP: la Fine Structure Processing (FSP) è stata sviluppata da Med-El ed è focalizzata ad ottenere un maggior numero di informazioni della struttura fine del segnale vocale per quanto riguarda le informazioni ad alta frequenza. Per far ciò vengono utilizzati filtri a campana o bell-shaped filters. Questa strategia fornisce diversi compiti ai diversi elettrodi:

1. La stimolazione di due elettrodi adiacenti è sequenziale (ciò comporta un aumento della risoluzione spettrale);
2. Gli elettrodi apicali (1-4) specializzati nelle basse frequenze codificano sia l'involuppo che la struttura fine del segnale e inoltre la loro frequenza della stimolazione viene modulata in tempo reale per seguire la struttura fine del segnale;
3. La frequenza di taglio inferiore viene abbassata a 100 Hz invece che 250 Hz così da prendere più informazioni nelle basse frequenze (i suoni più gravi);
4. Tutti gli altri elettrodi, cioè i canali adibiti alle alte frequenze utilizzano una strategia detta HDCIS che altro non è che una strategia CIS modificata appositamente.

L'evoluzione della FSP è la FS4 in cui la principale modifica è nella velocità di stimolazione dei quattro canali apicali (6000 impulsi al secondo) [30].

Impianto interno (array di elettrodi)

Il cuore dell'impianto interno è l'array di elettrodi, la parte adibita alla stimolazione elettrica del nervo. L'array di elettrodi viene inserito all'interno della coclea (figura 22) tramite un'operazione in cui il chirurgo ha due possibilità di inserimento: creare un'apertura direttamente nella finestra rotonda oppure creare una nuova apertura detta "cocleostomia". In entrambi i casi l'array verrà inserito all'interno della scala timpanica senza però ricoprire la sua interna lunghezza. Infatti, esistono diverse possibilità di scelta relativa alla lunghezza dell'array che sarà specifico per ogni persona. Indipendentemente da ciò, non si arriva mai alla copertura totale della coclea, di conseguenza non ci sarà una copertura tonotopica completa e quindi si darà maggior risalto alle frequenze tipiche della base cocleare.

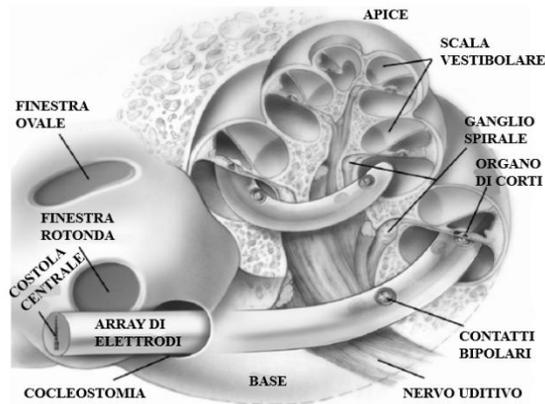


Figura 22: Inserimento dell'array di elettrodi. Si può osservare in basso a sinistra della figura le due possibilità di inserimento: finestra rotonda e cocleostomia. L'array viaggia per tutta la scala timpanica senza però raggiungere l'apice. Figura modificata da [29].

Gli impulsi elettrici erogati dall'array di elettrodi sono tipicamente di tipo bifasico a predominanza catodica sebbene siano state provate anche altre varianti come gli impulsi trifasici. Si definisce invece la modalità di stimolazione il percorso seguito dalla corrente. Le due modalità più utilizzate sono la monopolare e la bipolare: nella monopolare vi è un elettrodo di ritorno posto all'esterno della coclea, invece, nella bipolare l'elettrodo di ritorno è un elettrodo intracocleare vicino [26].

Oltre alla modalità di stimolazione, esistono altri 4 parametri che descrivono il treno di impulsi: la durata dell'impulso, l'intervallo interfase, l'intervallo interimpulso, e l'ampiezza della corrente erogata. Tali parametri sono riportati in figura 23. Modificando questi parametri è possibile cambiare la sensazione uditiva fornita dall'impianto e possono essere adattati al paziente. Ci sono alcuni aspetti da tener presente nel controllare questi parametri specialmente dal punto di vista fisiologico. Infatti, un aumento della durata dell'impulso o dell'intensità della corrente non garantisce una maggior sensazione in quanto bisogna tener conto di molti aspetti come la refrattarietà dei neuroni, la quantità di cellule ancora vive o il loro stato di salute. Per questi motivi il tipo di stimolazione è molto legata al singolo paziente.

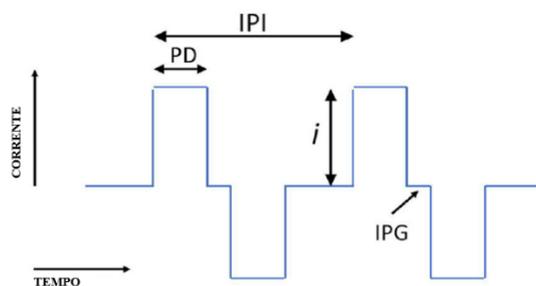


Figura 23: Caratteristiche descrittive del treno di impulsi bifasico a predominanza catodica. Oltre alla modalità di stimolazione, i parametri fondamentali di un impulso sono 4: la durata dell'impulso (PD), l'intervallo interfase (IPG), l'intervallo interimpulso (IPI), e l'ampiezza della corrente erogata (rappresentata da "i" in figura). Figura modificata da [32].

3 Materiali e metodi

Il secondo obiettivo di questo elaborato di tesi consiste nel caratterizzare un elemento di base (risonatore meccanico) che potrebbe essere utilizzato come elemento centrale di un nuovo dispositivo con la funzione di coclea artificiale. Questa caratterizzazione mira a generare un dataset contenente dati elaborati dal dispositivo stesso, per valutarne l' idoneità per il riconoscimento del parlato sia attraverso metodi tradizionali che tecniche di machine learning.

La caratterizzazione del dispositivo si è concentrata principalmente sul verificare la tonotopia del sensore così che possa garantire un pre-processing efficace per lo scopo preposto. Per ottenere tale obiettivo, sono state condotte misurazioni sperimentali in laboratorio su un prototipo (di qui in avanti denominato "coclea artificiale" per brevità), sviluppato in precedenza dal gruppo in cui ho lavorato. Queste misure mirano a stimolare la coclea artificiale tramite un input meccanico multifrequenza, osservandone la risposta elastica in diverse posizioni in seguito a tale eccitazione. Ogni risposta costituisce il dataset.

Nei successivi paragrafi verrà descritta la coclea artificiale, gli strumenti utilizzati durante il lavoro e il percorso logico seguito per lo svolgimento dello stesso.

Descrizione coclea artificiale

Qui nel seguito viene descritto il prototipo utilizzato.

CARATTERISTICHE ED UTILIZZO DEL DISPOSITIVO

Viene qui presentato il sensore tonotopico passivo bio-ispinato realizzato in polimero tramite stampa 3D che verrà utilizzato come elemento centrale delle misure:

- Sensore tonotopico: il dispositivo, proprio come tutti i sensori, è in grado di rilevare una grandezza fisica (un'onda meccanica) ed elaborarla, fornendo in output la stessa grandezza ma separata in frequenza. Infatti, il ruolo della coclea artificiale è quello di un analizzatore/separatore di frequenza. Questo è permesso dal suo comportamento naturalmente tonotopico che permette di differenziare il segnale in ingresso in base alla posizione lungo il dispositivo.
- Passivo: per svolgere il ruolo di analizzatore di frequenza il sensore non ha bisogno di un'alimentazione, non è quindi un sensore attivo. Si sfruttano a tale scopo solo le caratteristiche di vibrazione della struttura.
- Bio-ispinato: la coclea artificiale ricorda la geometria della coclea biologica, segue una spirale logaritmica da cui deriva in parte la proprietà di tonotopia. Inoltre, il suo ruolo da analizzatore di frequenza è la simulazione del comportamento cocleare biologico.

Da tutti questi aspetti, oltre che ad essere un analizzatore in frequenza, il dispositivo viene utilizzato come trasduttore acustico-meccanico e risonatore elettro-meccanico senza bisogno di alimentazione tramite l'utilizzo di sensori piezoelettrici applicati in superficie.

COSTRUZIONE E GEOMETRIA DEL DISPOSITIVO

La realizzazione del dispositivo è così ottenuta:

Partendo da una sezione rettangolare $b_0 \times h_0$ con b_0 base iniziale e h_0 altezza iniziale, viene seguita la curvatura della spirale dall'esterno verso l'interno. Lungo questo percorso la larghezza va sempre diminuendo, quindi b_0 è la larghezza massima della coclea ed è posta nel punto più esterno. La larghezza minima invece si ottiene nel punto più interno della spirale. L'altezza della sezione rimane invariata.

In figura 24 è mostrata la variazione della base del rettangolo della sezione lungo la curvatura della spirale. Il modello visto dall'alto è mostrato in figura 25.

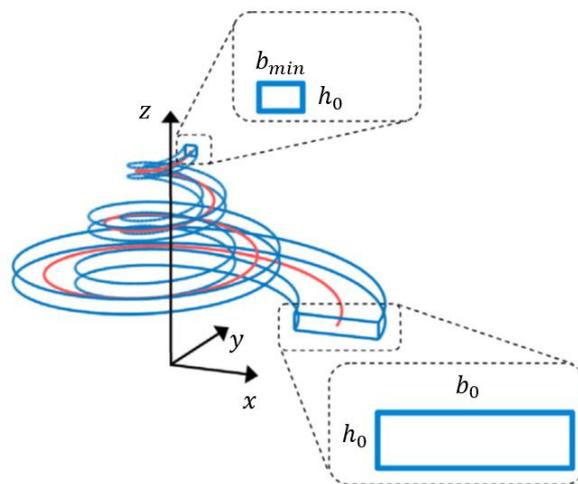


Figura 24: Costruzione della coclea artificiale. La sezione del dispositivo è rettangolare. Tale sezione parte con una certa larghezza b_0 che andrà man mano diminuendo seguendo la curvatura della spirale, rappresentata dalla linea rossa nell'immagine. L'area minore si ottiene quindi all'apice cocleare, nel punto più interno del dispositivo. Figura modificata da [33].

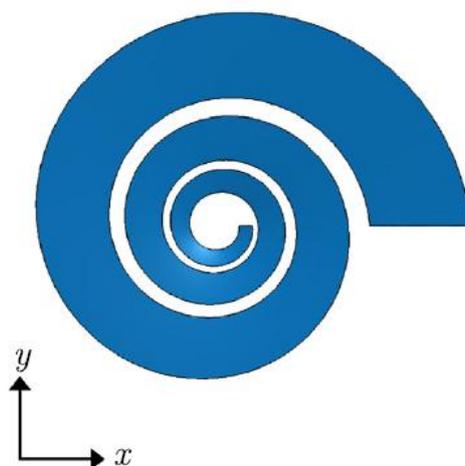


Figura 25: Geometria della coclea artificiale visualizzata dall'alto. Osservando la figura dall'alto (in un piano xy) è possibile notare la forma a spirale logaritmica. La spirale parte con una larghezza iniziale che diminuisce man mano muovendosi verso l'interno. La base della coclea corrisponde alla parte più esterna in cui la sezione è maggiore, l'apice coincide con la parte più interna dove la sezione è minima. Figura presa da [33].

La coclea artificiale utilizzata durante l'intero lavoro svolto presenta le seguenti caratteristiche:

- Lunghezza pari a circa 5 cm
- Larghezza pari a circa 4 cm
- Larghezza della base pari a circa 2 cm
- Larghezza dell'apice pari a circa 2 mm

TONOTOPIA

Le scelte geometriche fatte per il design derivano da un processo di ottimizzazione condotto allo scopo di discriminare il contenuto in frequenza di un'onda elastica che viaggia lungo la spirale in base alla posizione, cioè accentuare il più possibile la caratteristica di tonotopia della struttura. È infatti possibile vedere che iniettando segnali di input (burst gaussiani) con la frequenza centrale crescente, la risposta presenta un massimo di vibrazione in una posizione che varia progressivamente verso il centro della spirale, così come ci si aspetta per una struttura tonotopica (la risposta ai burst e il relativo comportamento tonotopico è mostrato in figura 26). Tale comportamento simula la tonotopia cocleare biologica, ottenendo così un ampio range di frequenze per poter analizzare un'onda elastica viaggiante.

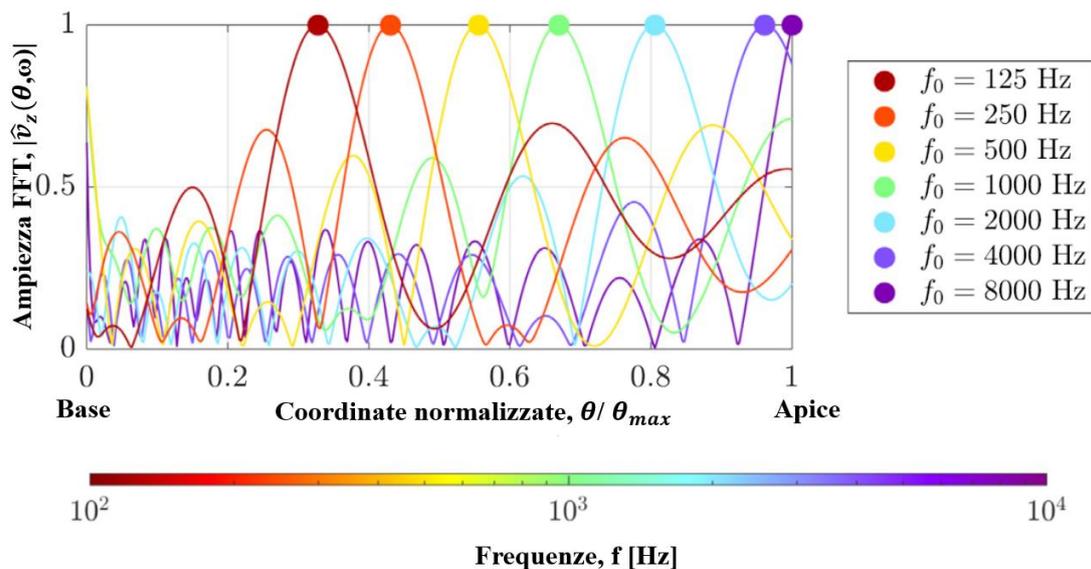


Figura 26: Posizione dei massimi di spostamento lungo la coclea artificiale ottenuti da una simulazione numerica. L'immagine mostra sull'asse x le coordinate normalizzate della coclea e sull'asse y l'ampiezza delle Fast Fourier Transform relative alle velocità fuori piano. Figura modificata da [33].

Un dettaglio non indifferente di questo sensore è l'orientazione della tonotopia: la coclea artificiale presenta una tonotopia inversa rispetto a quella biologica. Ciò è dovuto a diversi motivi riportati in [33]. La realizzazione di un'orientazione tonotopica uguale a quella biologica è possibile ma non è efficiente dal punto di vista dell'ottimizzazione degli spazi, dimensione del dispositivo e peggiore discriminazione dei massimi di spostamento, anche se soluzioni alternative sono allo studio. Il range di frequenze di cui dispone il sensore è compreso tra 100 Hz e 5 kHz, adatto all'obiettivo da raggiungere

nella tesi: effettuare riconoscimento del parlato. Un confronto tra la coclea biologica ed artificiale è mostrato in figura 27.

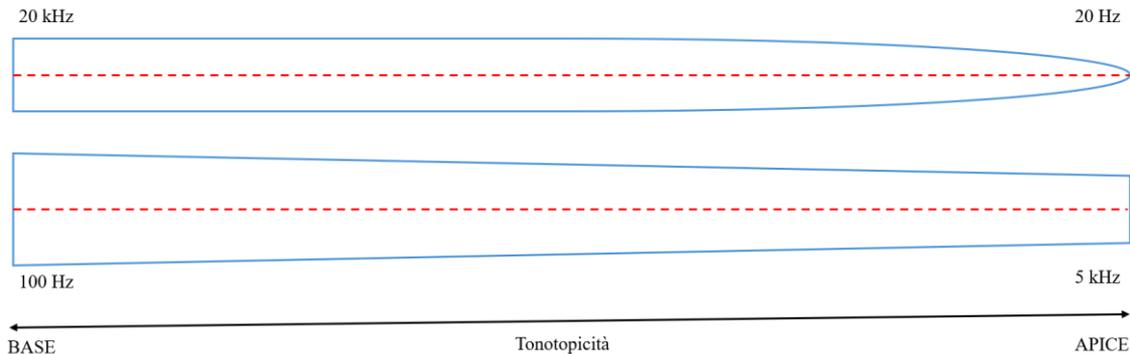


Figura 27: Confronto del comportamento tonotopico tra coclea biologica ed artificiale. Ipotizzando di poter prendere le due coclee per la loro lunghezza, si può osservare che l'apice (a destra) e la base (a sinistra) delle due coclee sono invertiti dal punto di vista tonotopico sia in posizione che in ampiezza. La coclea biologica (riportata sopra) presenta un range compreso tra 20 kHz e 20 Hz tra base e apice e la tonotopia è data dalla membrana basilare rappresentata come una linea rossa tratteggiata. La coclea artificiale, al contrario, presenta un range compreso tra 100 Hz e 5 kHz tra base e apice (comportamento inverso) e tale comportamento è fornito dai massimi di spostamento presenti sulla linea centrale della spirale logaritmica rappresentata come una linea rossa tratteggiata in figura.

Strumenti utilizzati

Per analizzare il sensore tonotopico passivo è stato necessario utilizzare un set di strumenti adatto a questo scopo: sensori piezoelettrici, PC, PXI, oscilloscopio, vibrometro laser e amplificatori di tensione.

L'obiettivo principale di tutti questi strumenti è stato creare un metodo compatto per poter effettuare le misure in modo riproducibile. Infatti, la prima parte del lavoro svolto in laboratorio è stata incentrata sul programmare uno strumento chiamato PXI tramite il software Matlab al fine di ottenere un sistema di misura completo: generare i segnali, acquisire le risposte cocleari a tali segnali ed elaborare i risultati ottenuti il tutto tramite l'utilizzo di un solo software. Il sistema di misura è stato programmato ed automatizzato per permettere la misura di un grande numero di registrazioni vocali, contenute nel dataset di riferimento. Inoltre, per ogni dataset, è stato necessario compiere due volte le misure a causa del limitato numero di canali di lettura della PXI (8 canali) rispetto al numero di canali necessari per coprire l'intera coclea (16 canali).

Il vibrometro laser è stato utilizzato per verificare il corretto funzionamento dei piezoelettrici prima che questi venissero montati sul dispositivo; infatti, tramite l'utilizzo di questo strumento è stato possibile scartare i sensori piezoelettrici che non funzionavano correttamente oppure capire quali di essi non funzionavano a causa di un'errata saldatura dei fili e di conseguenza rieffettuare la saldatura degli stessi.

L'oscilloscopio (limitato a 4 canali) è stato fondamentale nella parte di programmazione della PXI; infatti, tramite esso è stato possibile controllare che la PXI sia stata programmata correttamente sia in fase di generazione che di acquisizione. Senza l'oscilloscopio non si sarebbe potuto programmare la PXI perché non si sarebbe avuto un confronto.

L'amplificatore di tensione è stato lo strumento che ha accompagnato tutte le misure. Tramite esso è stato possibile visualizzare correttamente i dati raccolti, sfruttando il suo potere di amplificazione pari a 20 per amplificare i segnali generati dalla PXI e quindi da inviare alla coclea artificiale.

Il PC, come l'amplificatore, è stato parte di tutto il lavoro svolto in laboratorio. Tramite esso ed in particolare tramite l'utilizzo del software Matlab è stato possibile realizzare tutto: programmare la PXI, raccogliere le misure ed effettuare l'analisi e l'elaborazione dei risultati.

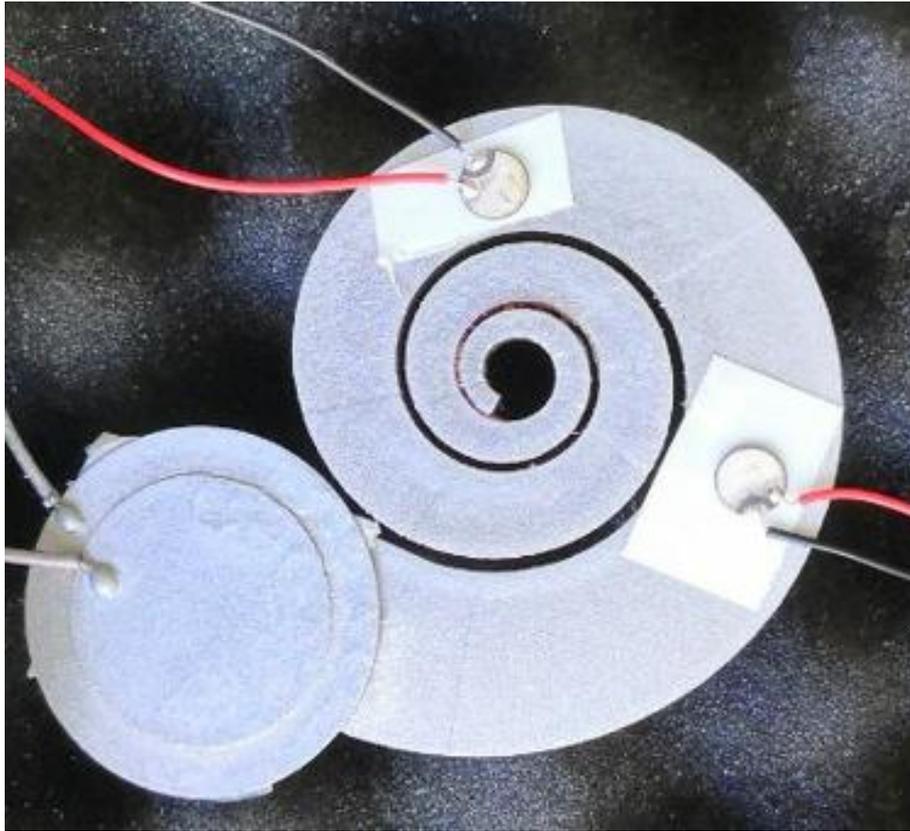
Di seguito vi è la descrizione dei due componenti fondamentali.

SENSORI PIEZOELETRICI

Come si può osservare dalla figura 28, i sensori piezoelettrici utilizzati sono di due tipi definiti dalle loro dimensioni che caratterizza la frequenza caratteristica di lavoro. Il sensore di maggiori dimensioni (25 mm di diametro), da qui in avanti definito stimolatore, permette di fornire l'input acustico alla coclea artificiale; invece, il sensore di dimensioni minori (5 mm di diametro), da qui in avanti definito rilevatore, è necessario per la lettura della misura. Mentre per lo stimolatore è servito un solo sensore, per effettuare la lettura della risposta cocleare sono stati necessari più rilevatori per un numero totale di 16 che costituiscono i 16 canali di lettura del segnale uscente dalla coclea. Entrambi sono stati incollati sulla coclea. Lo stimolatore essendo sempre posto alla base cocleare è stato incollato tramite una colla epossidica; invece, i rilevatori sono stati incollati tramite un biadesivo durante le varie prove; solo nelle misure finali sono stati incollati tramite la colla.

Lo stimolatore viene utilizzato per la trasduzione elettro-meccanica, infatti l'input generato dalla PXI e amplificato dall'amplificatore arriva come segnale elettrico al sensore, di conseguenza il suo compito è convertire questo segnale elettrico in un'onda meccanica che possa viaggiare lungo la coclea. Questa trasduzione avviene tramite il fenomeno fisico detto "effetto piezoelettrico inverso" in cui il segnale elettrico e quindi una differenza di potenziale deforma il sensore che risponderà generando un'onda meccanica.

I sensori di rilevamento, invece, devono catturare la vibrazione generata dalla coclea al passaggio dell'onda meccanica incidente e convertirla in un segnale elettrico così che possa tornare alla PXI per effettuare la lettura della misura. Questo compito di trasduzione elettro-meccanica avviene tramite il fenomeno detto "effetto piezoelettrico" in cui l'onda meccanica genera una deformazione del sensore e di conseguenza una variazione del suo potenziale elettrico. Il corretto funzionamento di questi sensori piezoelettrici è stato sempre testato tramite il vibrometro laser prima che questi fossero incollati lungo la linea centrale della spirale cocleare.



***Figura 28:** Coclea utilizzata per le prime misurazioni. Nell'immagine è rappresentata la coclea insieme ai sensori piezoelettrici utilizzati per le prime misurazioni riguardanti l'impostazione della PXI. Sulla base cocleare vi è incollato il sensore stimolatore; invece, lungo la linea centrale sono incollati due sensori di rilevamento.*

PXI

La PXI è una piattaforma di strumentazione elettronica modulare progettata dalla National Instruments. Tramite questo strumento, programmabile tramite PC, è possibile generare un segnale e allo stesso tempo acquisire la sua lettura; quindi, svolge il ruolo sia di generatore di funzioni che di oscilloscopio. Nel nostro caso la PXI è stata programmata tramite utilizzo del software Matlab. Precisamente la PXI utilizzata è la PXIe-1082 (figura 29) contenente all'interno 3 schede:

- PXIe-5413 necessaria per la generazione dei segnali.
- PXIe-5172 necessaria per l'acquisizione dei segnali.
- PXIe-8301 necessaria per l'alimentazione ed il collegamento al PC.



Figura 29: PXIe-1082 utilizzata. La scheda di generazione dei segnali PXIe-5413 dispone di due canali di generazione di segnali, sebbene nel nostro caso è stato utilizzato un solo canale. Una parte fondamentale di questa scheda sono i due led. I led sono necessari per il controllo della misura da parte dell'operatore in quanto fornisce informazioni sulla generazione del segnale. La scheda di acquisizione dei segnali PXIe-5172 svolge il ruolo di oscilloscopio. È dotata di 8 canali di lettura da poter usare contemporaneamente.

Descrizione delle misure (primo e secondo dataset)

Le misurazioni svolte sono atte a fornire uno stimolo acustico (una parola pronunciata da vari speaker) alla coclea artificiale. Tali misure sono necessarie per valutare il pre-processing passivo svolto dal dispositivo e quindi ottenere un dataset sufficientemente ampio per il successivo sistema di riconoscimento del parlato.

Per ottenere gli obiettivi prefissati sono state svolte due misure complete su due dataset differenti. Il primo dataset [34] è composto da 3000 file audio di tipo wav campionati a 8 kHz. Il dataset è composto da 3000 parole in inglese pronunciate da 6 diversi speaker: ogni speaker pronuncia 10 digit (da 0 a 9) per 50 volte. Il secondo dataset [35] è composto da 30000 file audio di tipo wav pronunciate in inglese, campionate a 48 kHz. Anche in questo caso ogni persona, per un totale di 60 speaker, pronuncia le prime dieci cifre numeriche (da 0 a 9) ripetendole per 50 volte.

Il primo dataset è stato considerato nella sua interezza, invece per il secondo sono state considerate solo le prime 6 persone in modo da avere un numero di campioni pari al primo dataset e quindi confrontabile con esso. Nel primo dataset i sei speaker sono tutti uomini con accenti differenti (USA, francese, tedesco e greco), invece nel secondo dataset i primi sei speaker sono anche qui tutti uomini ma con lo stesso accento (tedesco). Le differenze principali tra i due dataset sono gli speaker e la frequenza di campionamento. Essendo la PXI dotata di solo 8 canali di acquisizione per ogni dataset la misura completa è stata ripetuta due volte: la prima volta per i canali più esterni, la

seconda per i canali più interni. Tutto il sistema è stato gestito dal PC tramite l'utilizzo di un apposito algoritmo implementato su Matlab.

Tale algoritmo è stato progettato nel seguente modo:

1. Viene presa la singola parola dal dataset e viene inviata sotto forma di segnale elettrico tramite la scheda di generazione di segnali della PXI
2. La scheda di acquisizione della PXI svolge il ruolo di oscilloscopio e preleva il segnale dagli 8 canali collegati alla coclea artificiale
3. Viene salvato il risultato ottenuto in un'apposita cartella
4. Si ripete il processo con la parola successiva

Avendo ogni parola lunghezza diversa è stato necessario porre tutte le parole di ugual lunghezza; per far ciò è stata presa la parola più lunga di tutto il dataset e tutte le altre sono state adattate alla sua lunghezza andando ad aggiungere degli zeri per compensare la lunghezza. Inoltre, l'algoritmo si occupa della programmazione della PXI: frequenza di campionamento, configurazione del trigger, apertura e chiusura del canale di generazione.

Quindi, considerando il singolo segnale vocale, il percorso svolto da quest'ultimo (figura 30) è il seguente:

1. Il segnale vocale viene generato tramite la PXI in fase di generazione sotto forma di segnale elettrico;
2. Il segnale giunge al sensore piezoelettrico stimolatore incollato alla base del sensore tonotopico dove avverrà la trasduzione mecano-elettrica: tramite l'effetto piezoelettrico inverso il segnale elettrico viene trasdotto in un'onda meccanica che viaggerà lungo la coclea;
3. La struttura tonotopica inversa svolge il suo ruolo: ogni frequenza presente nell'onda meccanica si ferma nel punto di risonanza del sensore;
4. Viene prelevato il segnale da 16 punti diversi della coclea lungo la sua linea centrale. Tale lettura avviene tramite i 16 sensori di rilevamento sfruttando l'effetto piezoelettrico e arriverà alla PXI in fase di acquisizione. In realtà ci sono solo 8 canali nella PXI quindi per ogni parola bisogna ripetere la procedura due volte (canali esterni ed interni).

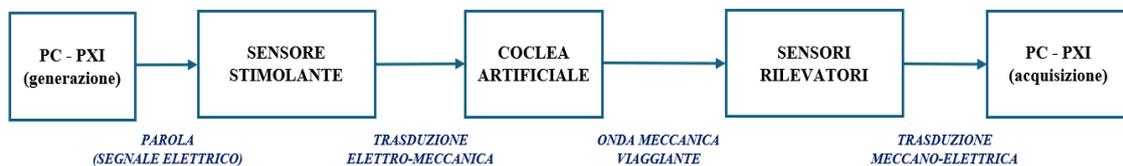


Figura 30: Percorso condotto da ogni singolo segnale vocale presente nei dataset. Tramite il PC viene gestito l'algoritmo implementato su Matlab adibito al controllo della PXI.

Prima di poter ottenere le misure complete su entrambi i dataset è stato necessario svolgere delle prove per verificare il corretto funzionamento degli strumenti. In particolare, abbiamo utilizzato un oscilloscopio per confrontare i risultati. Per quanto riguarda la scheda di generazione dei segnali della PXI, la sua uscita è stata collegata al canale di lettura dell'oscilloscopio a cui sono stati mandati in input delle forme d'onda standard e una parola.

Verificato il corretto funzionamento della scheda, si è passati alla scheda di acquisizione dei segnali. Utilizzando la configurazione della coclea in figura 28 sono stati mandati in input sempre delle forme d'onda e delle parole presenti nel dataset e sono state visualizzate le risposte da parte dei due sensori piezoelettrici di lettura sia sul PC (collegato alla PXI tramite Matlab) che sull'oscilloscopio.

I risultati sono stati sovrapponibili dimostrando il corretto funzionamento della PXI. È stato quindi possibile passare alla prima prova di misura sulla coclea basata su più canali e su un dataset intero.

Il setup completo dell'esperimento è mostrato in figura 31.

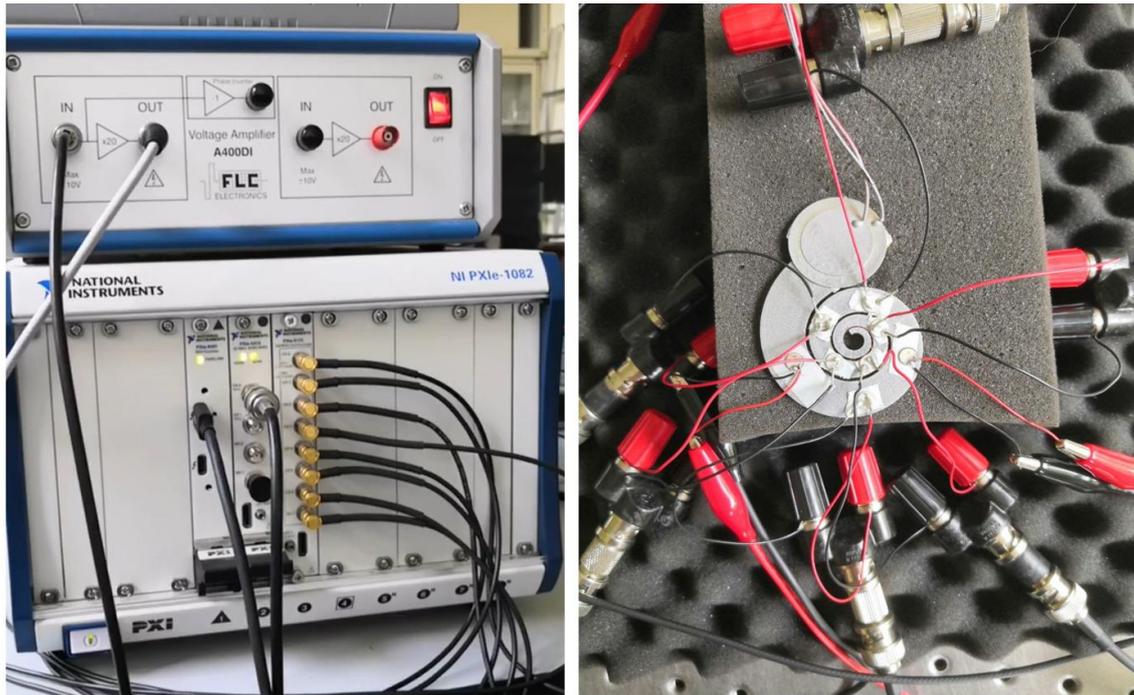


Figura 31: Setup completo del primo esperimento. Sulla sinistra è raffigurata la PXI e l'amplificatore di tensione, sulla destra la configurazione della coclea artificiale per questo esperimento. Il canale di uscita del generatore di funzioni della PXI è collegato all'amplificatore di tensione la cui uscita è collegata al sensore stimolatore. Lungo la coclea sono attaccati tramite un biadesivo 8 sensori di lettura collegati alla scheda di acquisizione della PXI.

Per tutte le successive misurazioni il setup della PXI e dell'amplificatore non verrà cambiato. Ora il sistema è pronto per la misura completa vera e propria. Per tutta la lunghezza della coclea sono stati incollati 16 sensori piezoelettrici di lettura e sono state svolte due misure complete sull'intero dataset. La nuova configurazione della coclea è mostrata in figura 32.

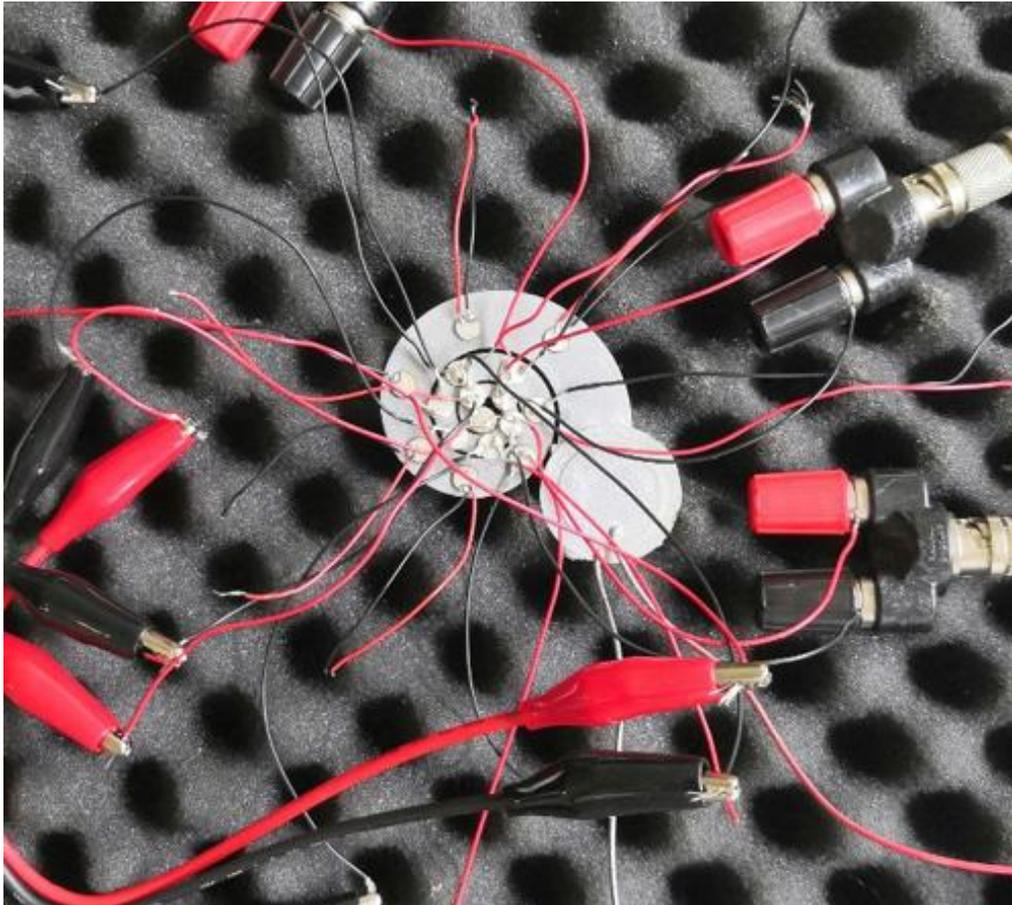


Figura 32: Configurazione finale della coclea. La configurazione non cambierà più anche per il secondo dataset. L'unico componente che cambierà sarà il sensore piezoelettrico stimolatore in quanto verrà sostituito a causa di un suo malfunzionamento durante le prove svolte con il secondo dataset.

Completata la prima misura, è stata implementata la seconda misura basata su un altro dataset. Lo schema di processo è sempre lo stesso. La misura è stata ripetuta due volte in quanto durante una prova il sensore stimolatore si è rotto e quindi è stato necessario sostituirlo. Effettuata la sostituzione è stato possibile effettuare la misura completa sul nuovo dataset (canali interni e canali esterni).

METODI USATI PER L'ANALISI DELLE MISURE

Tutti i metodi matematici utilizzati per l'analisi dei risultati sono stati implementati sul software Matlab. L'idea di base di questi metodi è l'analisi in frequenza tramite Trasformata di Fourier:

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-i2\pi ft} dt \quad (1)$$

Nel nostro caso, questo operatore è stato implementato tramite l'algoritmo presente su Matlab chiamato Fast Fourier Transform (FFT) che ne calcola la trasformata in forma discreta su un numero limitato di campioni. Per ognuno dei 16 canali di lettura viene calcolato il modulo della FFT al fine di osservare i picchi presenti nel dominio della frequenza e quindi confrontarli tra loro e con il segnale in input.

Per verificare in modo più marcato la tonotopia della coclea artificiale viene effettuata una prima operazione di “denoising”: lo spettro in frequenza viene suddiviso in diverse bande e per ognuna di esse viene calcolato l’integrale [dB]. Lo step finale di ogni banda è di 5 Hz, risultato di diverse prove. Per ogni banda di frequenze viene calcolato il seguente integrale:

$$I_i = \int_{f_{i,min}}^{f_{i,max}} |X(f)|_i df \quad (2)$$

Dove i è la banda, $f_{i,max}$ e $f_{i,min}$ sono gli estremi della banda e $|X(f)|_i$ è il modulo della Trasformata di Fourier. Lavorando nel discreto, l’integrale è una sommatoria discreta e la Trasformata di Fourier è il modulo della FFT del segnale. L’intero canale I sarà formato quindi dai vari I_i e verrà convertito in dB come $20 \log_{10} I$.

Ogni canale presenta una baseline diversa per vari motivi, come la vicinanza al sensore piezoelettrico stimolante o un cattivo incollaggio del sensore di lettura. Per questo motivo, è necessario allineare i canali tra loro. Viene eseguita così una seconda operazione di denoising: poiché il segnale vocale generalmente non supera i 4000 Hz, è stato scelto un valore di frequenza in cui il segnale è assente (10000 Hz) per allineare i canali a un valore di 0 dB.

Per ottenere una descrizione più significativa dal punto di vista statistico delle misure è stata applicata la media su ogni canale sulle 50 ripetizioni. Ciò è avvenuto per ogni parola e per ogni speaker. Tramite tutte queste analisi è stato possibile ottenere dei grafici che confrontano i singoli canali tra loro e delle mappe che confrontano tutti e 16 i canali.

4 Risultati ed analisi dati

Per verificare se gli obiettivi preposti siano stati raggiunti o meno sono state condotte le seguenti analisi dei risultati:

1. per verificare la tonotopia dei risultati sono stati messi a confronto i singoli canali a coppie (canali più esterni vs canali più interni);
2. per poter riconoscere il parlato (discriminare uno speaker dall’altro ed un digit dall’altro) sono state costruite delle mappe contenenti tutti i canali;

La prima analisi condotta consiste nel verificare la risposta attesa dei canali sia nel dominio temporale che nel dominio della frequenza.

DOMINIO TEMPORALE

Nel tempo, nei canali di lettura ci si aspetta un segnale simile all’originale nella forma e nella durata temporale ma anche delle lievi distorsioni. Dalle figure 34, 35 e 36 è possibile osservare un confronto visivo tra i canali e il segnale originale. È importante notare che il segnale originale è stato amplificato di un fattore 20 tramite l’algoritmo implementato in Matlab in modo da simulare l’amplificatore presente nella strumentazione reale (il segnale all’uscita del sensore stimolatore non è stato prelevato). Da queste figure si nota come la riproduzione della forma del segnale sia presente seppur distorta; inoltre è possibile notare come ogni canale ha un’ampiezza diversa, in particolare muovendosi dalla base cocleare verso l’apice è possibile osservare una

diminuzione dell'ampiezza. Ciò è dovuto alla lontananza rispetto al sensore stimolante. Inoltre, muovendo dai canali più esterni verso i più interni, essi diventano leggermente più frastagliati, cioè le variazioni del segnale diventano sempre più veloci come ci si dovrebbe aspettare.

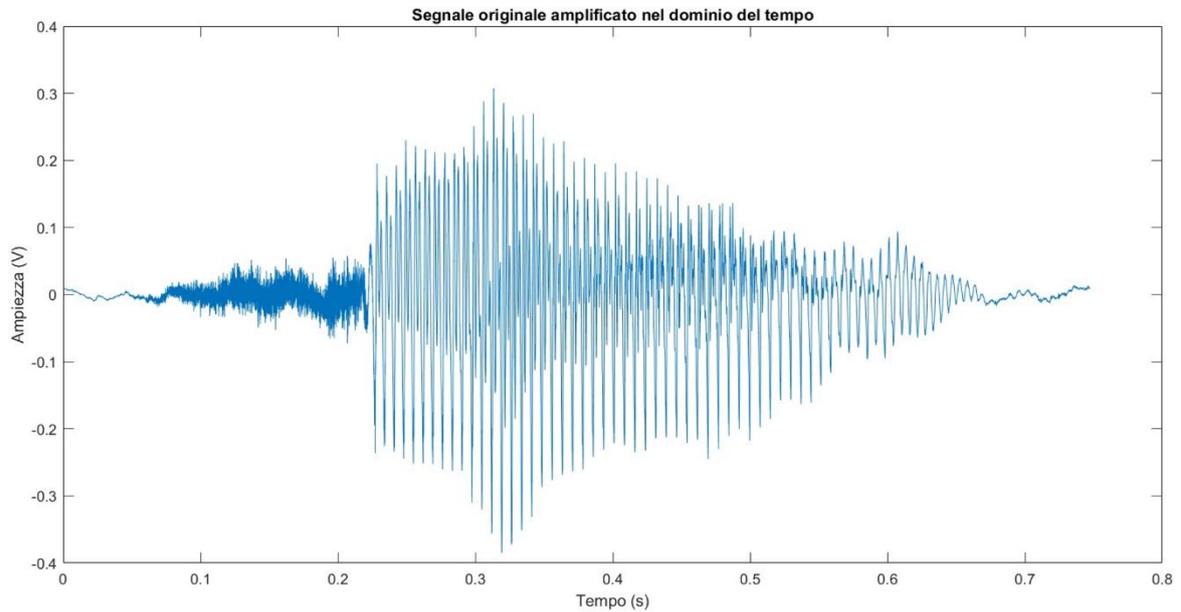


Figura 34: Segnale vocale in input. Tale segnale rappresenta la riproduzione della parola "zero" pronunciata da un uomo del secondo dataset (lo speaker 1).

CANALI ESTERNI

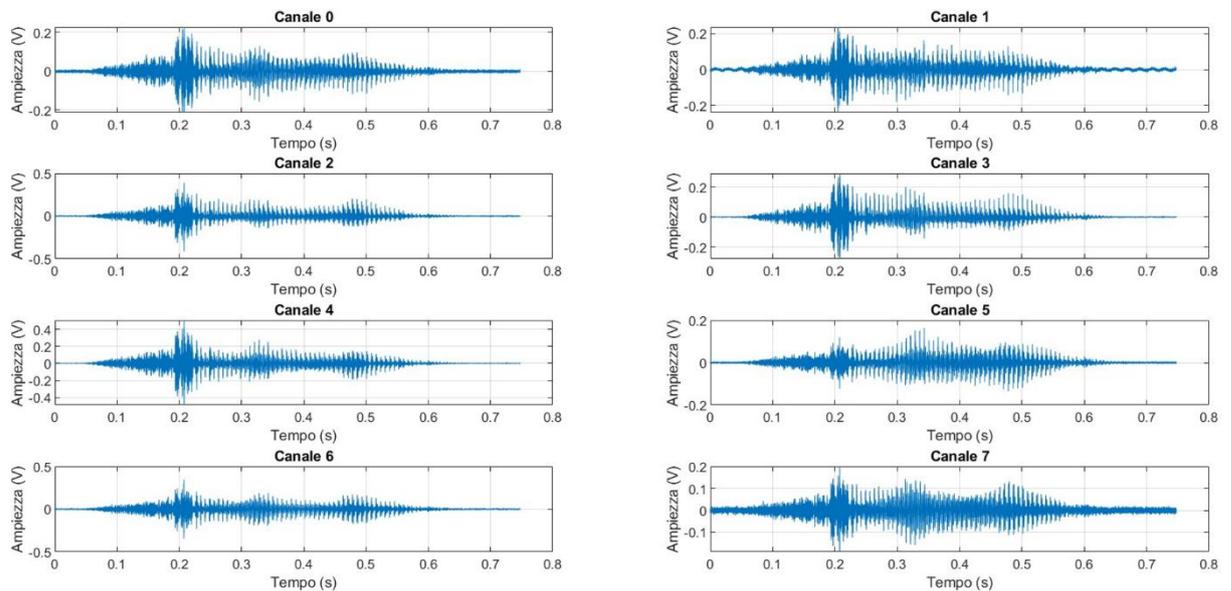


Figura 35: Primi otto canali più esterni di lettura del segnale.

CANALI INTERNI

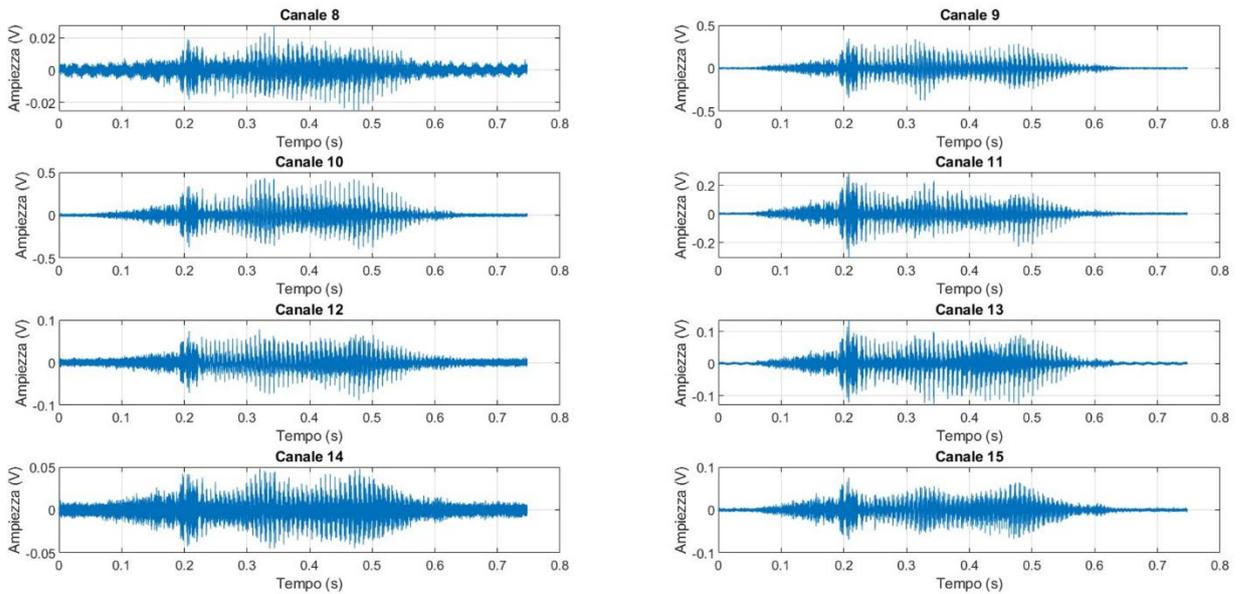


Figura 36: Ultimi otto canali più interni di lettura.

Nelle seguenti figure 37, 38 e 39 viene mostrato un confronto visivo tra il segnale originale e i canali più esterni e più interni oltre che a un confronto tra gli ultimi due. In tutte e tre le figure i segnali sono sovrapponibili, la differenza principale è nell'ampiezza. Tale differenza è dovuta, come già detto precedentemente, alla simulazione dell'amplificazione effettuata successivamente tramite l'algoritmo e soprattutto al sistema di prelievo dei canali di lettura.

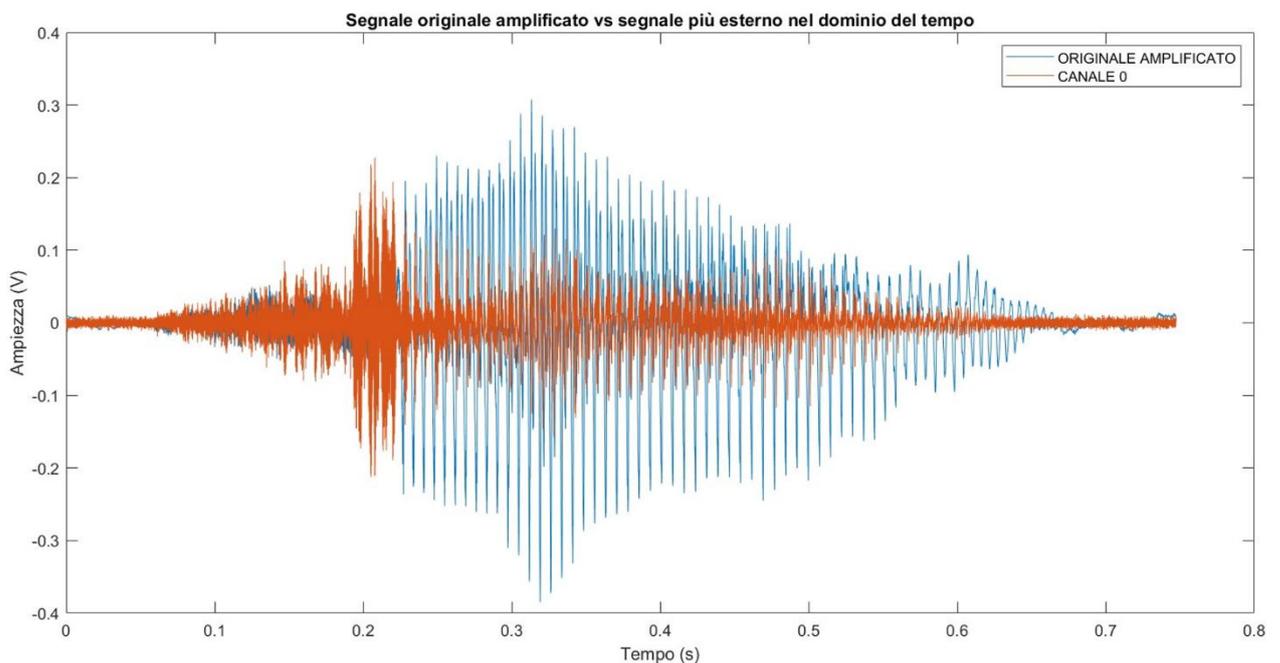


Figura 37: Confronto tra canale più esterno e segnale originale.

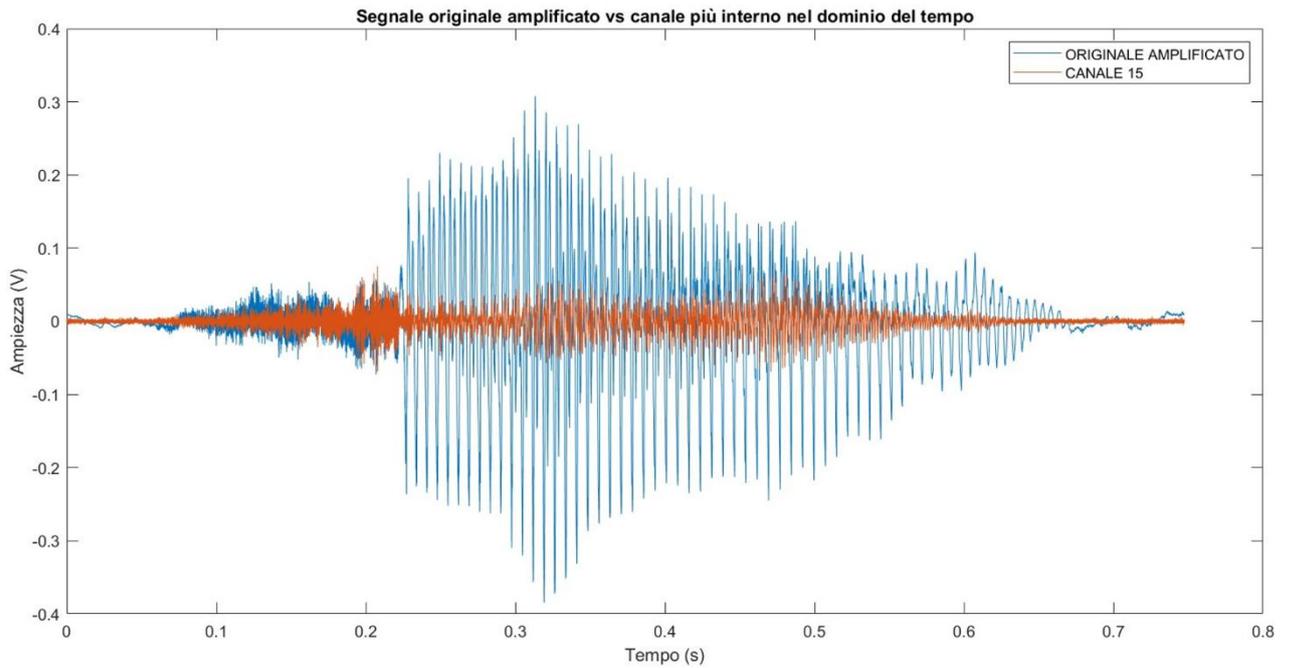


Figura 38: Confronto tra canale più interno e segnale originale.

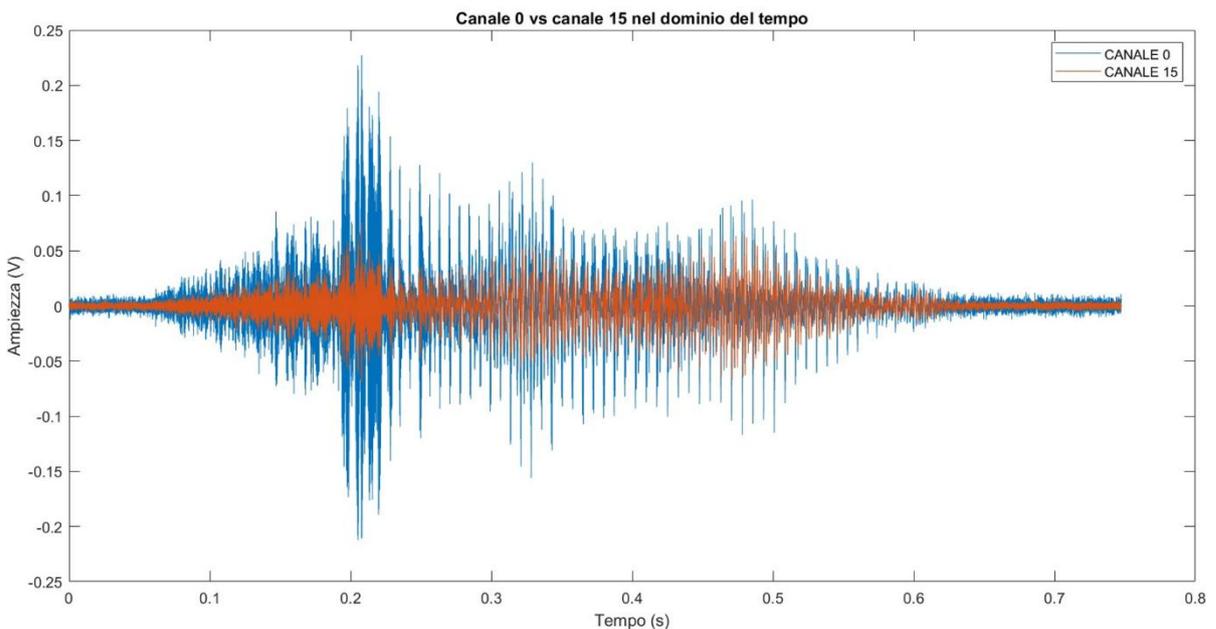


Figura 39: Confronto tra canale più interno e canale più esterno.

RISPOSTA IN FREQUENZA

In figura 40, 41 e 42 sono riportati i risultati ottenuti da una prima analisi in frequenza. La figura 40 mostra il modulo della FFT del segnale vocale. Con questa immagine è possibile osservare i picchi più importanti del segnale nonché le bande di frequenza che lo caratterizzano. Il segnale vocale è caratterizzato principalmente fino a 1000 Hz. Le due successive immagini rappresentano lo stesso lavoro svolto sui canali di lettura. Tutti i canali riportano gli stessi picchi fondamentali del segnale originale con un'eccezione sempre presente: un picco intorno ai 1800 Hz. Tale picco, molto ampio, è dovuto alla

sensibilità del sensore piezoelettrico stimolatore intorno a quel range di frequenza. Tale picco sarà sempre presente in tutti i canali anche se con ampiezze e larghezze di banda differenti. In particolare, tale picco sarà sempre meno prevalente spostandosi verso i canali più interni in quanto, come nel dominio del tempo, questi sono più lontani dal sensore.

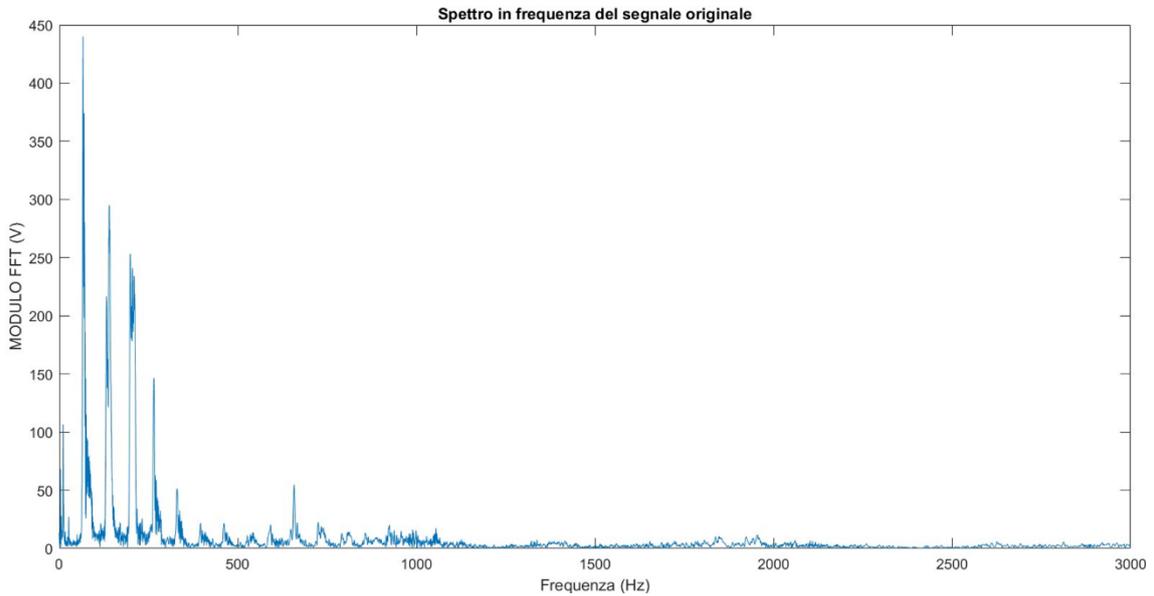


Figura 40: Spettro in frequenza del segnale vocale.

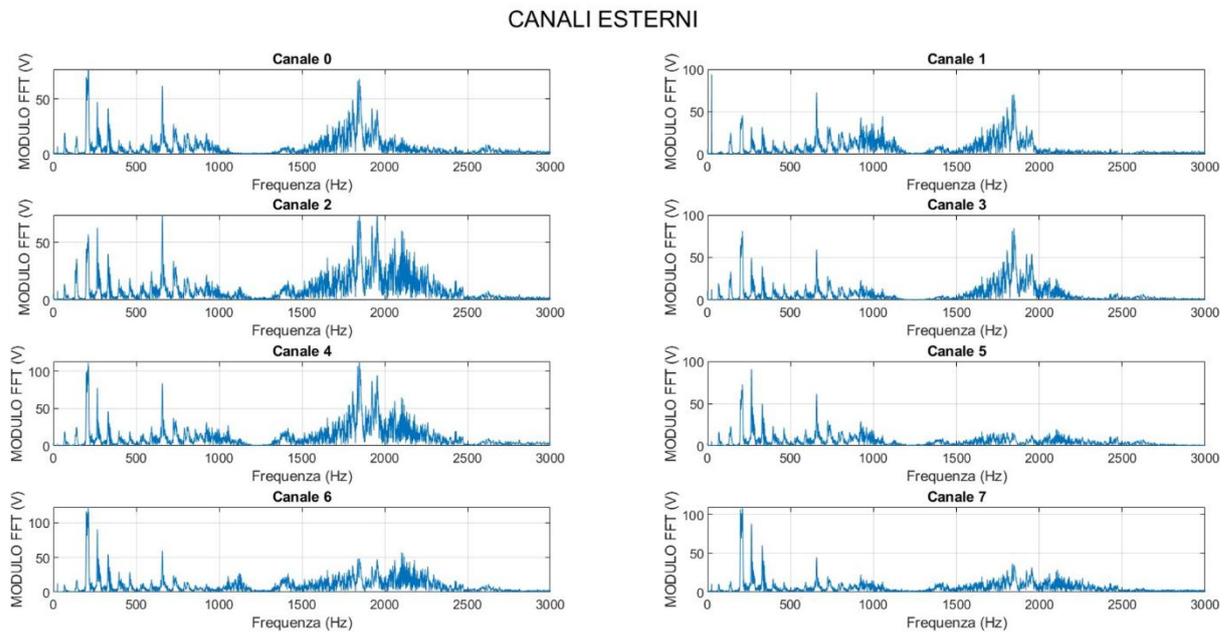


Figura 41: Spettri in frequenza dei primi otto canali più esterni.

CANALI INTERNI

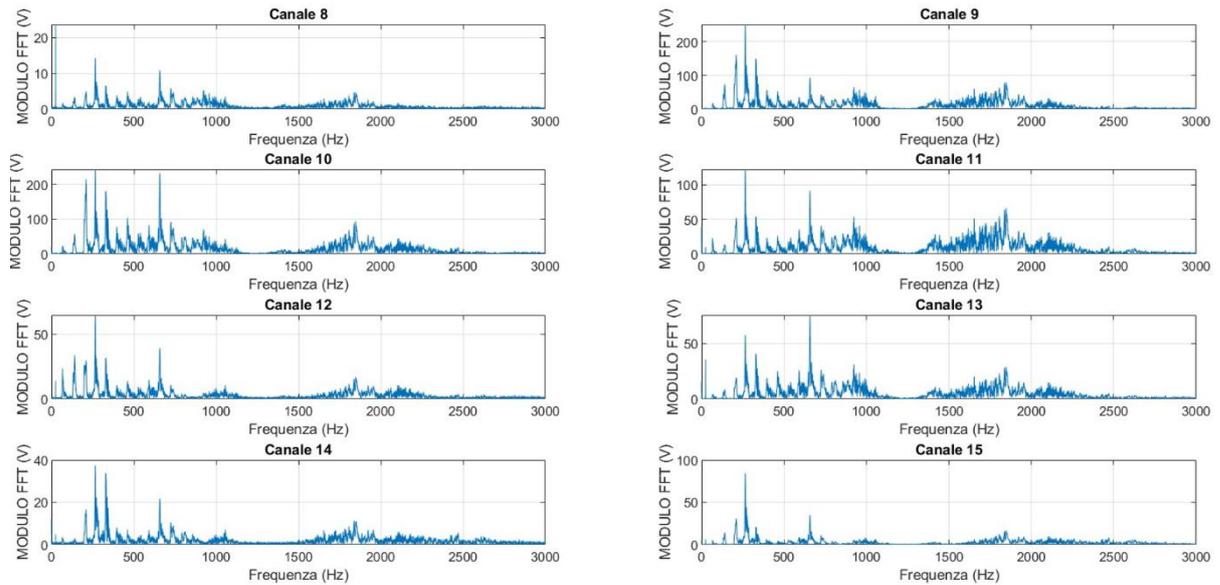


Figura 42: Spettri in frequenza dei canali più interni.

Come per l'analisi nel tempo, è stato effettuato un confronto tra i canali più esterni e il segnale originale. Tale confronto è riportato nelle figure 43, 44 e 45. Osservando il segnale vocale originale e il canale più esterno si osserva come quest'ultimo presenti gli stessi picchi del segnale originale (a patto del picco dovuto al sensore, come già detto). Lo stesso avviene per il canale più interno.

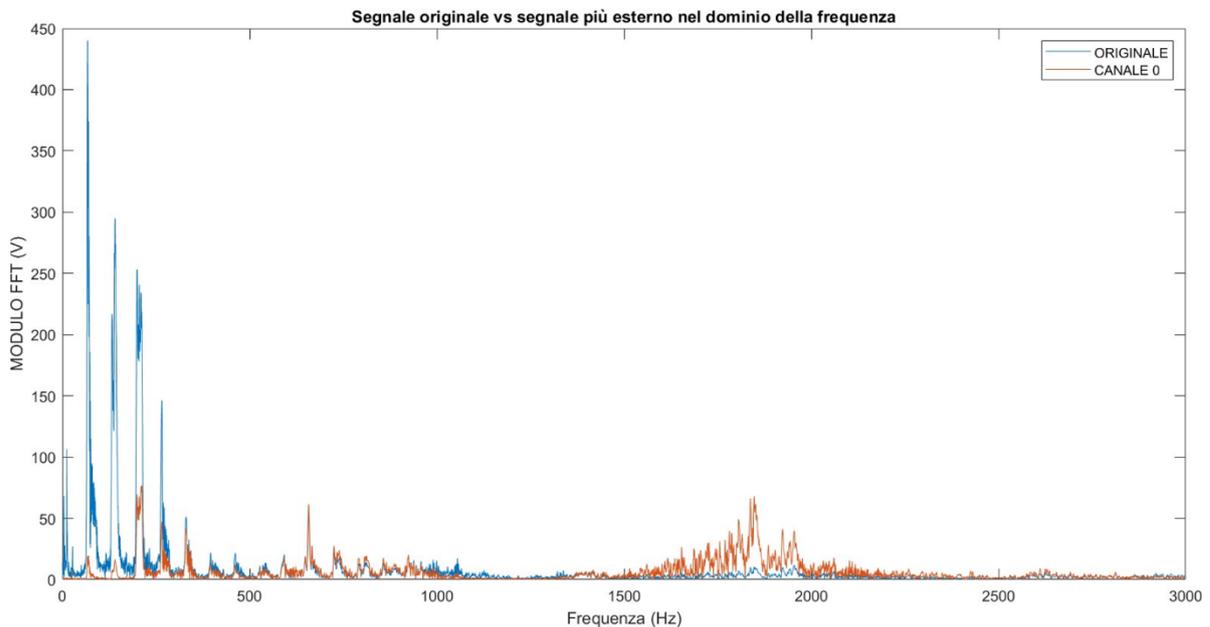


Figura 43: Spettro in frequenza del segnale vocale e del canale più esterno sovrapposti.

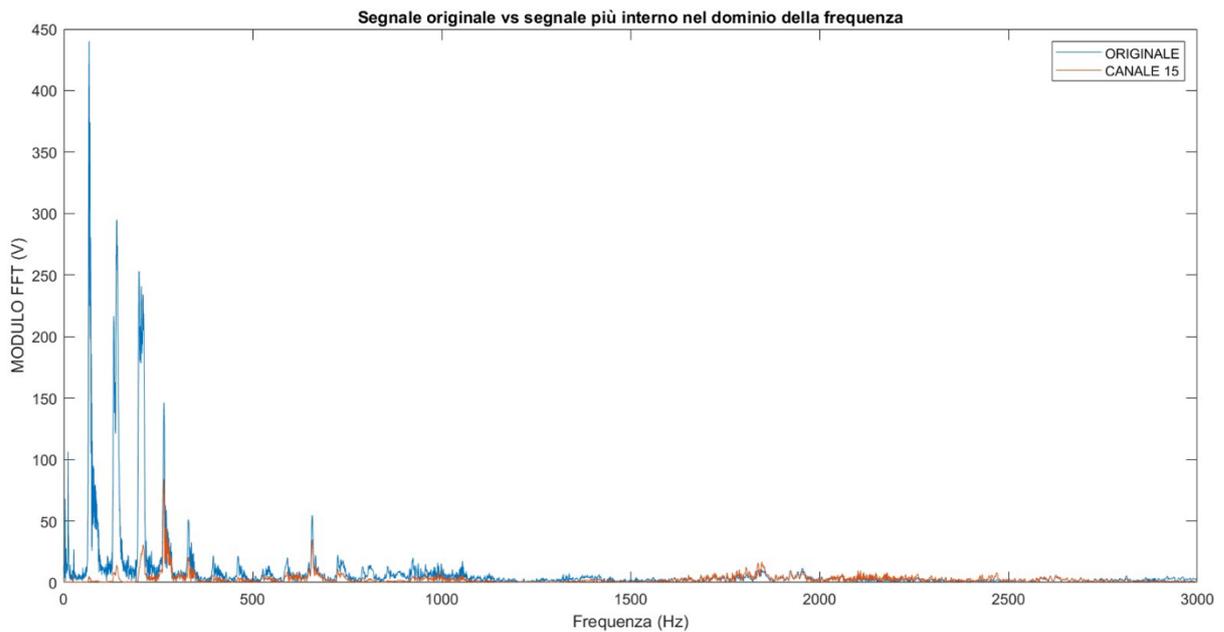


Figura 44: Spettro in frequenza del segnale vocale e del canale più interno sovrapposti.

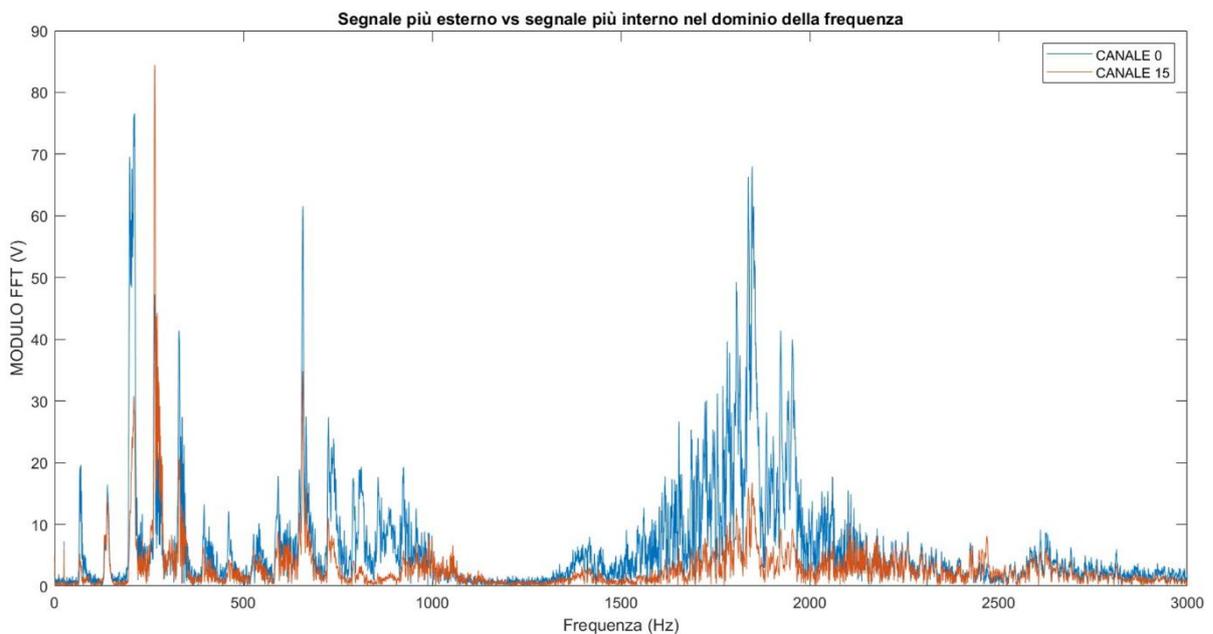


Figura 45: Spettro in frequenza del canale più esterno e del canale più interno sovrapposti.

OPERAZIONI DI DENOISING

Verificata la risposta del dispositivo nel dominio del tempo ed il dominio della frequenza è possibile passare ad un'analisi più dettagliata della tonotopia. Per ottenere un confronto tra i canali, come descritto nella sezione di questa tesi denominata "Materiali e metodi", si è optato per implementare un'integrale del segnale in bande da 5 Hz ciascuna ottenendo così la figura 46. Ogni integrale è stato convertito in dB, ma nell'immagine non viene riportata tale scala così da poter offrire una visione più chiara di questa prima operazione di denoising (l'integrale mostrato nella figura 46 è il risultato dell'operazione di denoising ottenuto dal canale 0 mostrato in figura 45).

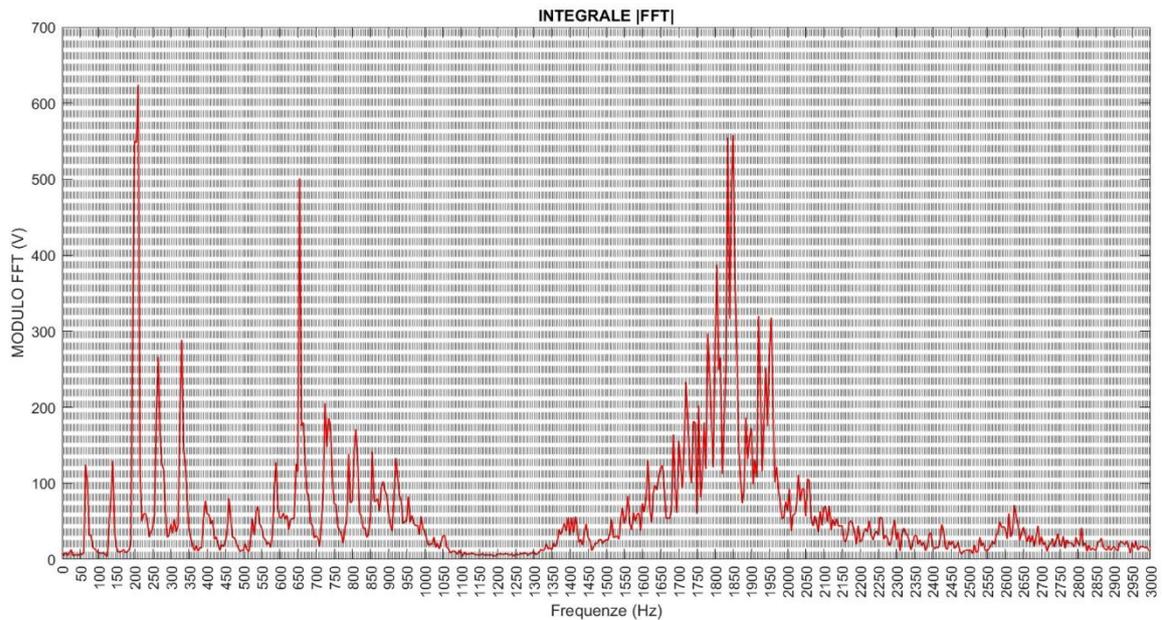


Figura 46: rappresentazione delle bande più rilevanti a step di 5 Hz. In questa figura è rappresentato l'integrale del modulo della trasformata di Fourier per bande di 5 Hz svolto nel canale più esterno. Le linee tratteggiate separano le varie bande.

In figura 47 è riportata la seconda operazione di denoising effettuata per tutti i canali: l'allineamento dei canali tramite baseline comune. Come si può notare dall'immagine, questa operazione è necessaria anche per i canali vicini a causa dei molteplici fattori scatenanti questo tipo di problema. Tramite questa operazione è possibile confrontare senza errori i canali tra loro e quindi poter trovare i range di frequenza che caratterizzano un canale rispetto all'altro.

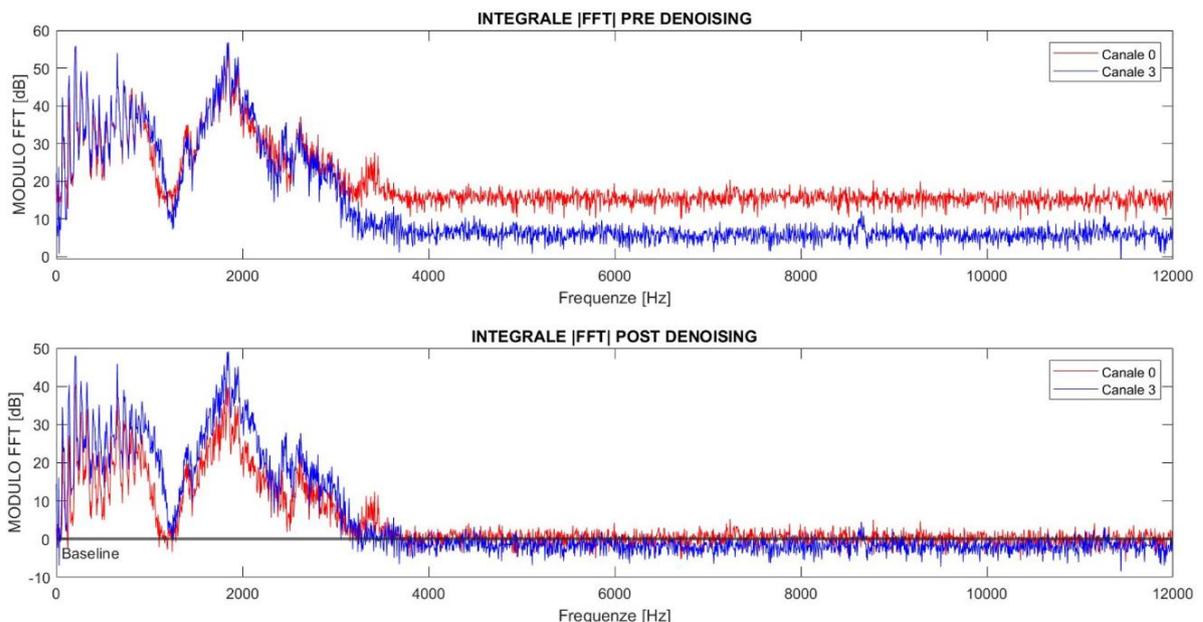


Figura 47: allineamento dei canali.

CONFRONTO TRA I SINGOLI CANALI

Al fine di avere una analisi statisticamente più rilevante, abbiamo condotto l'analisi mediando i segnali su 50 ripetizioni. In questa analisi vengono messi a confronto canali con posizioni sulla coclea molto diverse fra loro, così facendo ci si aspetta di trovare gli effetti della tonotopia più marcati. Per ottenere questa analisi sono stati confrontati prima i canali più distanti tra loro a coppie e successivamente l'insieme di tutti i canali tramite la creazione di mappe.

I risultati del confronto tra i singoli canali sono riportati nelle figure 48, 49 e 50. Le immagini rappresentano tre speakers diversi (03, 04 e 06) che pronunciano tre diverse cifre (5, 6 e 9). Da queste tre immagini è possibile confrontare il canale più esterno (il canale 2) e il canale più interno (il canale 16). Confrontando le immagini si può notare come la forma dell'integrale è diverso sia per lo speaker che per il digit (come ci si aspetta) oltre che la sensibilità dei canali alle frequenze. Sono stati scelti questi canali in quanto meno rumorosi e quindi più rappresentativi. Data la sensibilità dei sensori di misura a determinate bande di frequenze, l'effetto tonotopico del dispositivo è distinguibile nel range di frequenze compreso tra 700 e 1300 Hz e tra 2000 e 2800 Hz. Come ci si aspetta, nel primo range di frequenze il canale esterno è leggermente più marcato rispetto a quello più interno; invece, nel secondo range di frequenze, i canali invertono il loro comportamento. In particolare, osservando i singoli canali, per il digit 5, il canale esterno enfatizza le basse frequenze rispetto alle alte mentre il canale interno non presenta questo squilibrio (le alte frequenze sono ben enfatizzate rispetto al canale esterno ma non rispetto alle basse frequenze del canale stesso). Lo stesso accade per il digit 9. Per il digit 6, il comportamento è inverso: i due canali sono sempre distinguibili tra loro ma presi singolarmente si può osservare come il canale esterno non enfatizzi molto le alte frequenze rispetto alle basse al contrario del canale più interno che sopra i 2000 Hz è in grado di enfatizzare la risposta rispetto alle frequenze sotto i 1300 Hz. Confrontando le cifre tra loro si ottiene che per il digit 6, tra 2000 e 2900 Hz, vi è un picco importante e ben ampio intorno a 2400 Hz al contrario del digit 5 che possiede una risposta abbastanza piatta in quel range di frequenze (la risposta non supera i 2600 Hz). Invece, per le basse frequenze, entrambi i digit hanno un comportamento marcato tra 700 e 1200 Hz che è simile per il canale più esterno ma leggermente diverso per il canale più interno. Tutte queste caratteristiche finora osservate permettono di distinguere i digit e gli speaker. Tali differenze sono maggiormente distinguibili tramite la creazione delle mappe.

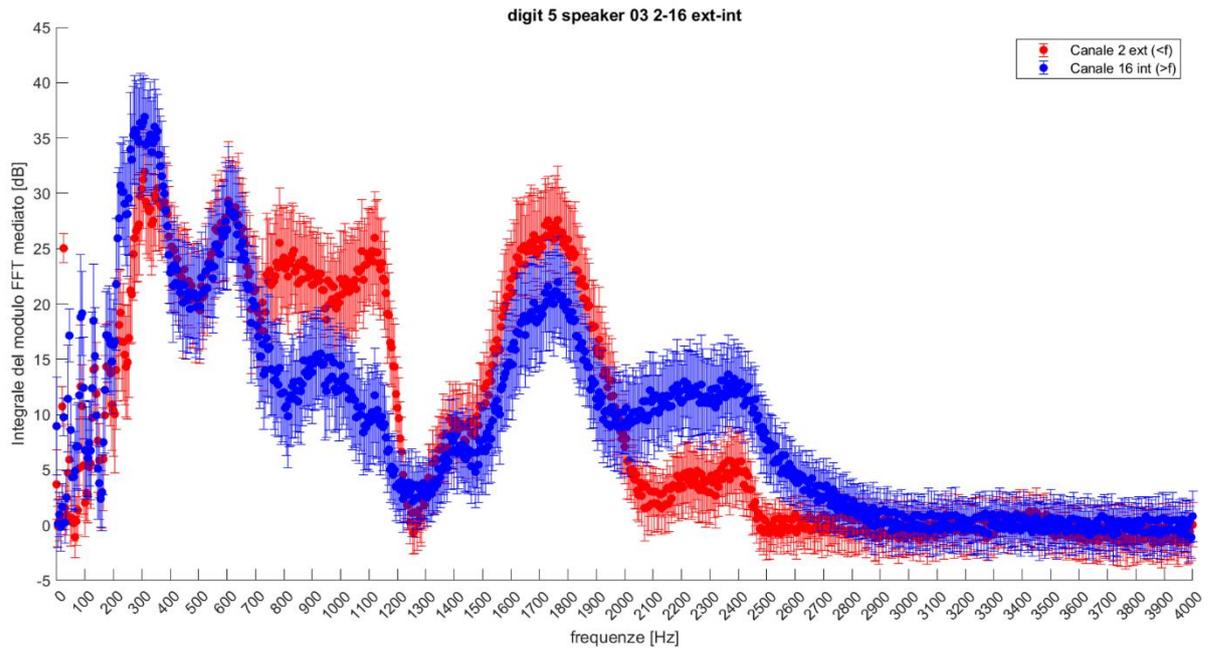


Figura 48: Confronto tra canale 2 e 16 della cifra 5 pronunciata in inglese dallo speaker 3.

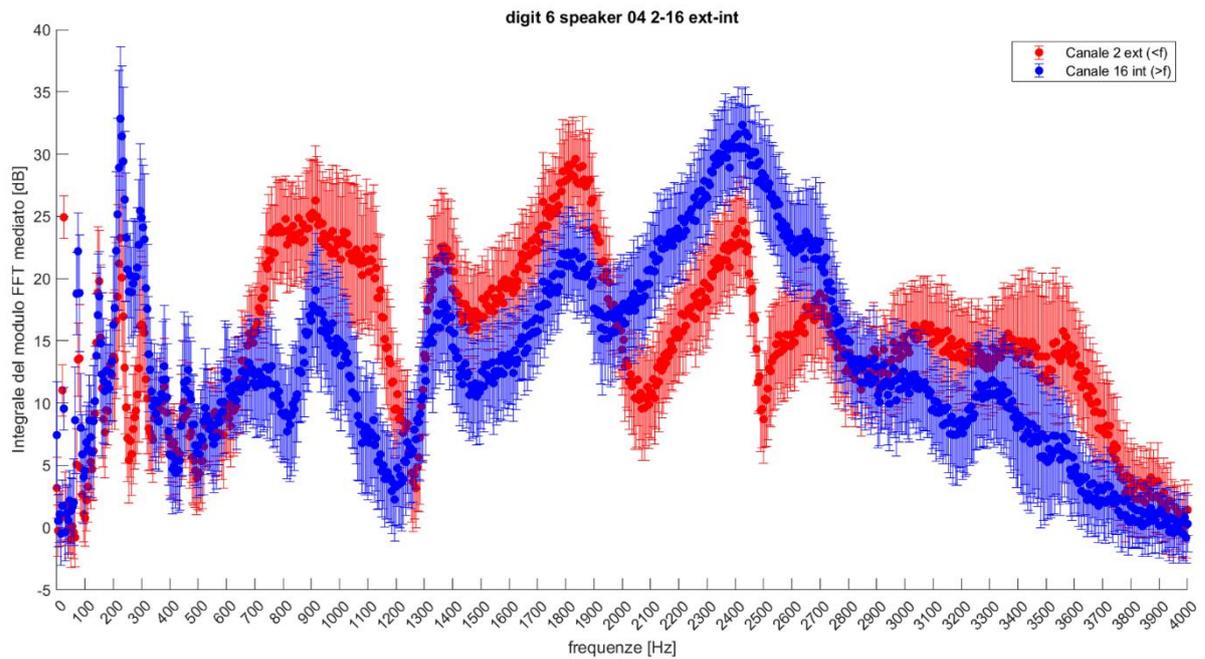


Figura 49: Confronto tra canale 2 e 16 della cifra 6 pronunciata in inglese dallo speaker 4.

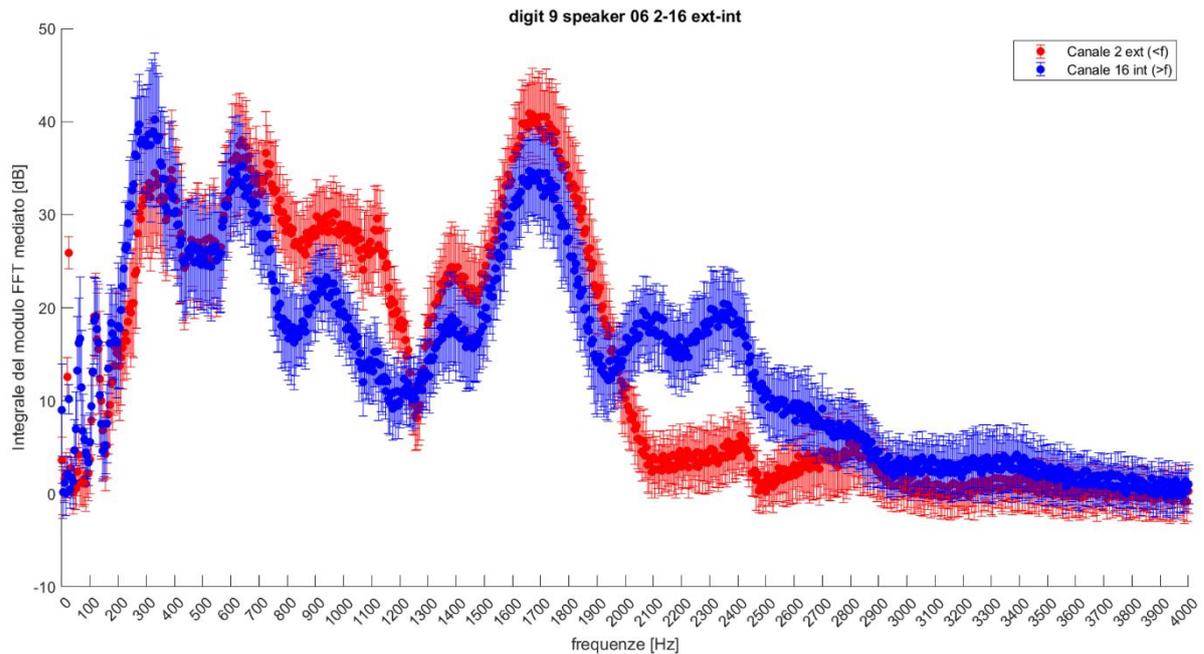


Figura 50: Confronto tra canale 2 e 16 della cifra 9 pronunciata in inglese dallo speaker 6.

MAPPE

Per verificare che il dispositivo sia in grado di fornire informazioni sufficienti a distinguere uno speaker dall'altro e un digit dall'altro, è stato necessario confrontare tutti i canali tramite la creazione di mappe basate su una codifica a colori. Da queste mappe è stata ottenuta la differenza per ottenere un maggior confronto tra gli speaker e i digit. I confronti delle differenze tra gli speaker sono riportati nelle figure 51, 52, 53, 54 e 55. Ogni figura rappresenta due mappe, ciascuna riferita a due speaker diversi che pronunciano lo stesso digit, e la loro differenza. Le persone non sono necessariamente le stesse in ogni figura, ma per ciascun digit sono stati selezionati i due speaker con caratteristiche più differenti tra loro. Invece, i confronti delle differenze che vi sono tra due digit pronunciati dalla stessa persona sono riportati nelle figure 56 e 57.

Anche nel caso delle mappe vi sarà sempre la presenza di un picco in tutti i canali intorno ai 1800 Hz, quindi da qui in poi non verrà più nominato per favorire la lettura ma è importante ricordare la sua presenza durante la lettura dei risultati.

Per il digit "0" sono stati confrontati gli speaker 1 e 5 (vedere figura 51). Lo speaker 1 è ben distinguibile dal 5 nel range di frequenze compreso tra 200 e 300 Hz e nelle alte frequenze (sopra i 2500 Hz). Nelle alte frequenze la differenza tra i due speaker è omogenea nei canali più esterni mentre nei canali più interni la differenza è meno marcata. Nelle frequenze comprese tra 300 e 2500 Hz il comportamento delle due persone è abbastanza simile con delle eccezioni: tra 450 e 600 Hz e intorno a 1600 Hz lo speaker 5 è più caratterizzato; mentre tra 600 e 800 Hz è leggermente più enfatizzato lo speaker 1.

Per il digit "1" sono stati confrontati gli speaker 3 e 6 (vedere figura 52). Le frequenze caratteristiche del primo speaker rispetto al secondo sono intorno ai 200 Hz e tra 1100 e 1200 Hz; invece per il secondo speaker il range è compreso tra 400 e 750 Hz, tra 1300 e

1800 Hz. Questi range sono molto variabili spostandosi lungo i canali: nelle frequenze superiori a 2000 Hz, lo speaker 3 è leggermente più caratterizzato nei canali più esterni ma spostandosi verso i canali più interni, lo speaker 6 è più enfatizzato rispetto al 3. osservando quindi globalmente i canali si può notare come lo speaker 3 è caratterizzato per i canali più esterni, quindi per le frequenze minori, invece lo speaker 6, sebbene caratterizzato da quasi tutti i canali, presenta delle bande in frequenza più alte spostandosi nei canali più interni.

Per il digit “2”, riportato in figura 53, sono stati confrontati gli stessi speaker del digit precedente. In questo caso, le alte frequenze sono caratterizzanti lo speaker 6 per tutti i canali che rende così lo speaker ben distinguibile rispetto allo speaker 3.

In figura 54 viene mostrato il digit “8” pronunciato dagli speaker 1 e 6. Il primo speaker è ben distinguibile fino a 900 Hz, tra 1100 e 1200 Hz e tra 1800 e 2200 Hz. Al contrario, il secondo speaker ha frequenze caratterizzanti intorno a 1000 Hz, tra 1300 e 1700 Hz e nelle alte frequenze comprese tra 2500 e 3400 Hz. In quest’ultimo range la distribuzione dei canali non è omogenea: nei canali più interni sono presenti le frequenze più vicine a 2500 Hz invece nei canali più esterni sono più rappresentati le frequenze vicine a 3500 Hz.

Per il digit “9”, riportato in figura 55, sono stati confrontati gli speaker 2 e 6. Lo speaker 2 è ben distinguibile intorno a 500 Hz e tra 2100 e 2500 Hz (con comportamento variabile tra i canali). Il secondo speaker è distinguibile intorno a 300 Hz, tra 900 e 1200 Hz ed infine tra 1500 e 1800 Hz.

Osservando le figure 56 e 57 si può osservare come cambia la mappa tra una parola e l’altra nonostante gli speaker siano gli stessi. Ad esempio considerando lo speaker 3 si può notare come nel pronunciare il digit “1” le frequenze caratteristiche sono comprese tra 200 e 500 Hz, invece per il digit “2” sono prevalenti le frequenze intorno a 900 Hz e tra 1400 e 2800 Hz. Sempre per il digit “2”, nei canali più esterni sono comprese anche le frequenze tra 3100 e 3500 Hz. Invece, per lo speaker 6 si ha che il range di frequenze compreso per il digit “1” è più ampio rispetto allo speaker 3 (da 300 a 700 Hz) così come per il digit “2” (il range delle basse frequenze è compreso tra 700 e 1300 Hz con delle lievi variazioni tra i canali mentre il range per le alte frequenze è compreso tra 1800 e 3700 Hz con una distribuzione più omogenea tra i canali rispetto allo speaker 3).

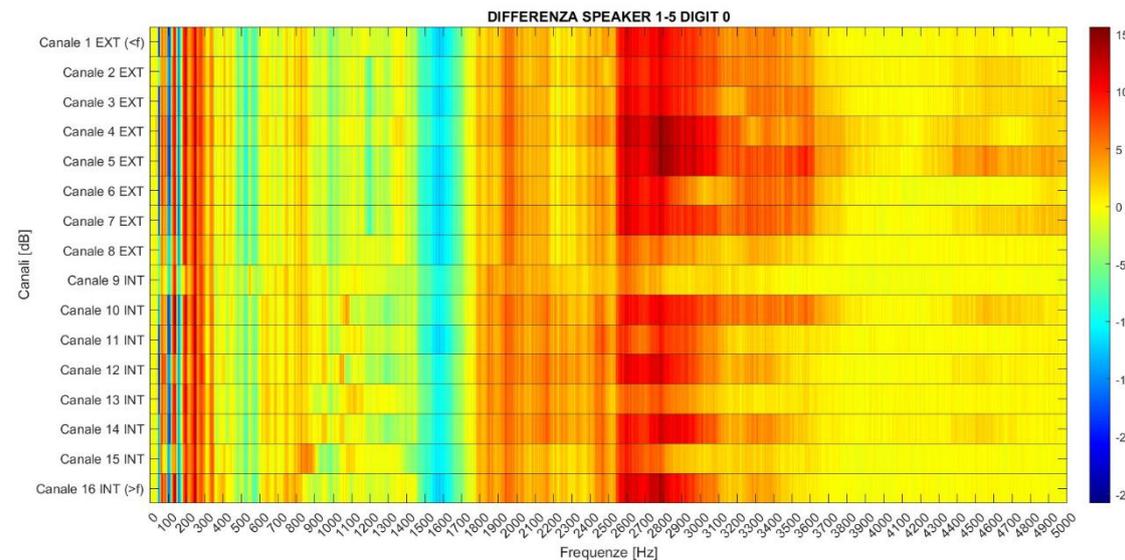
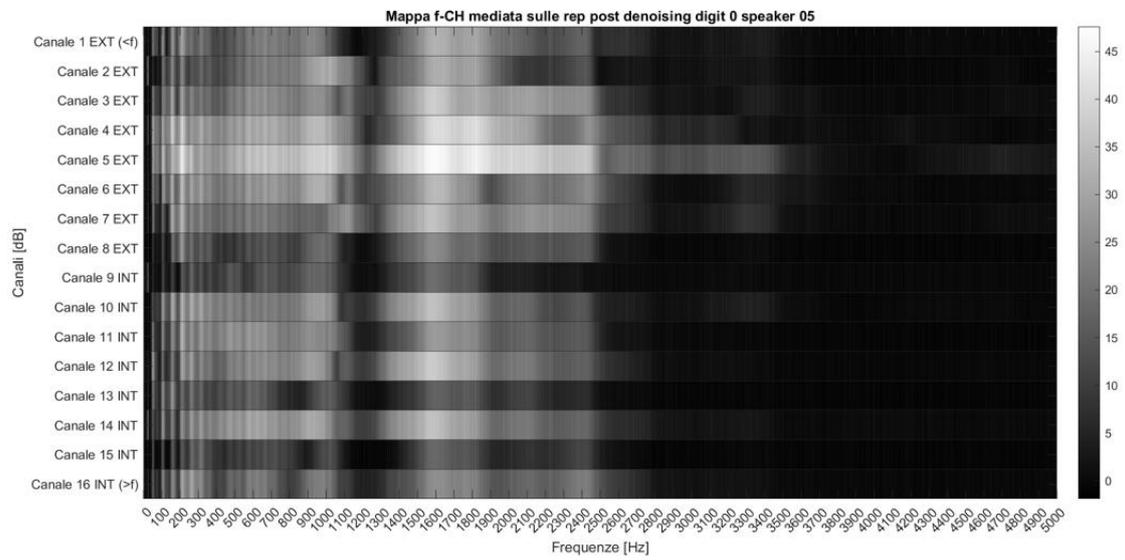
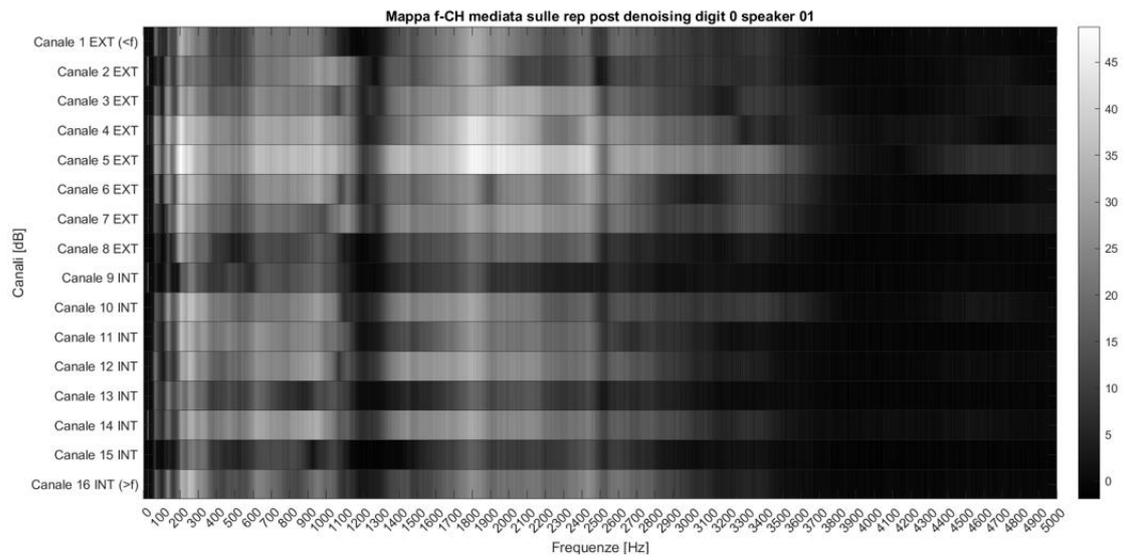


Figura 51: Mappe e differenza del digit "0" pronunciate dagli speaker 01 e 05.

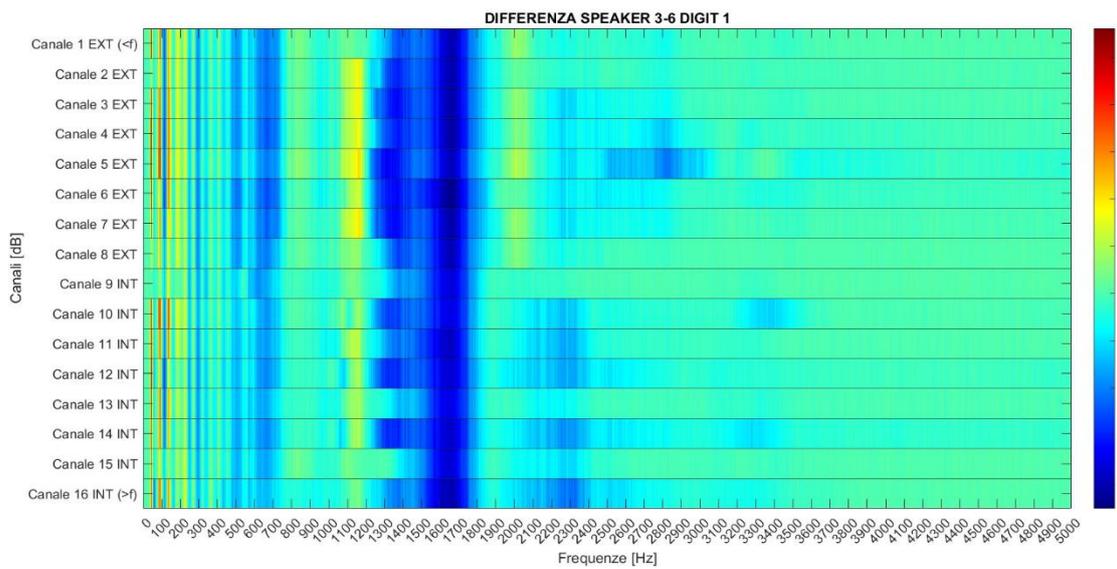
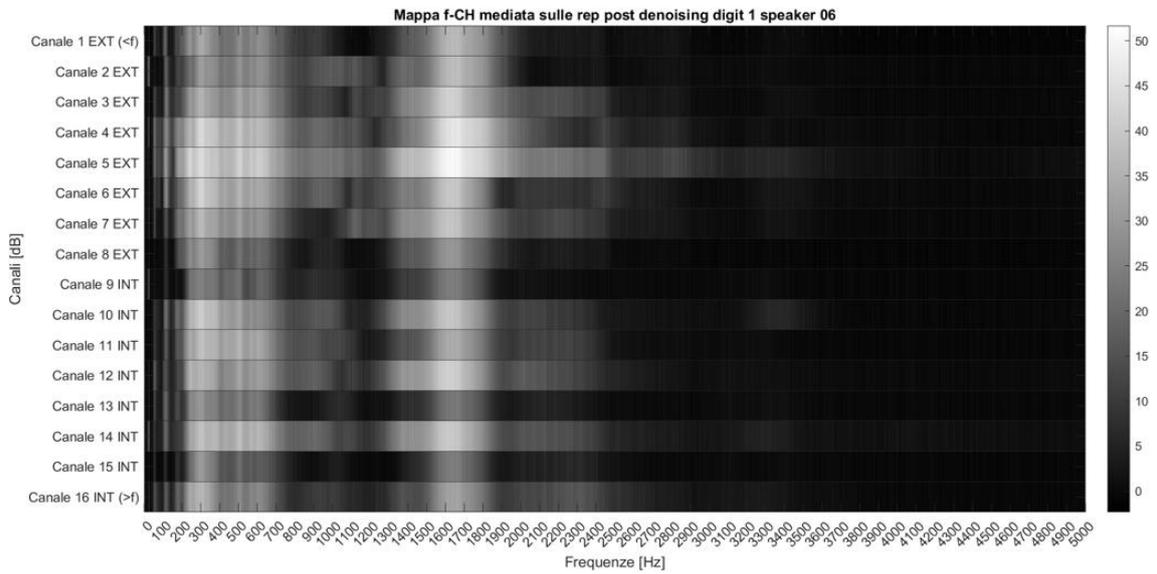
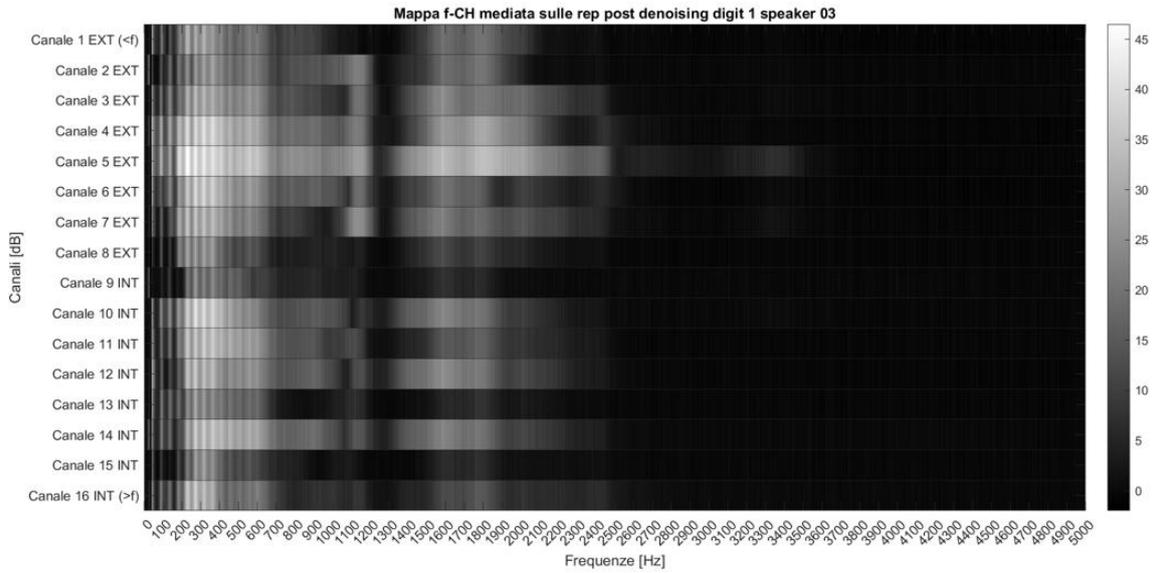


Figura 52: Mappe e differenza del digit "1" pronunciate dagli speaker 03 e 06.

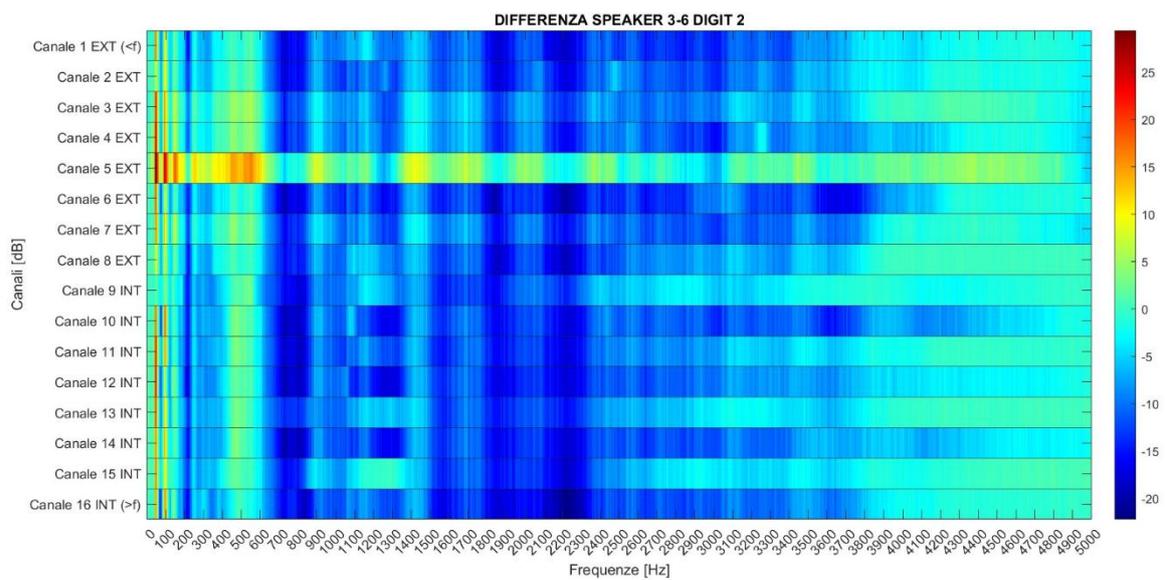
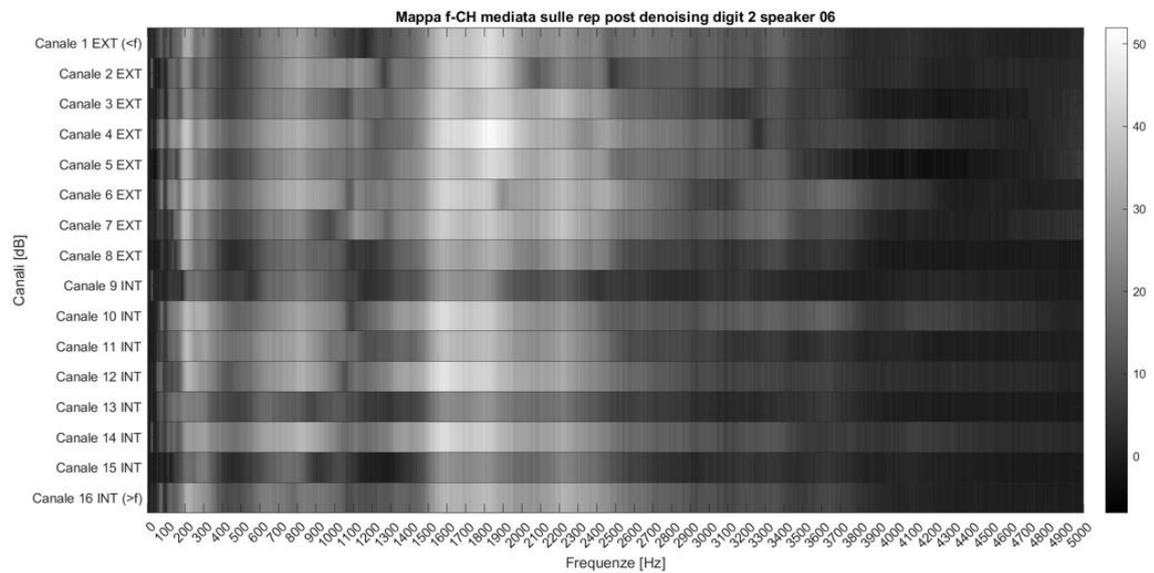
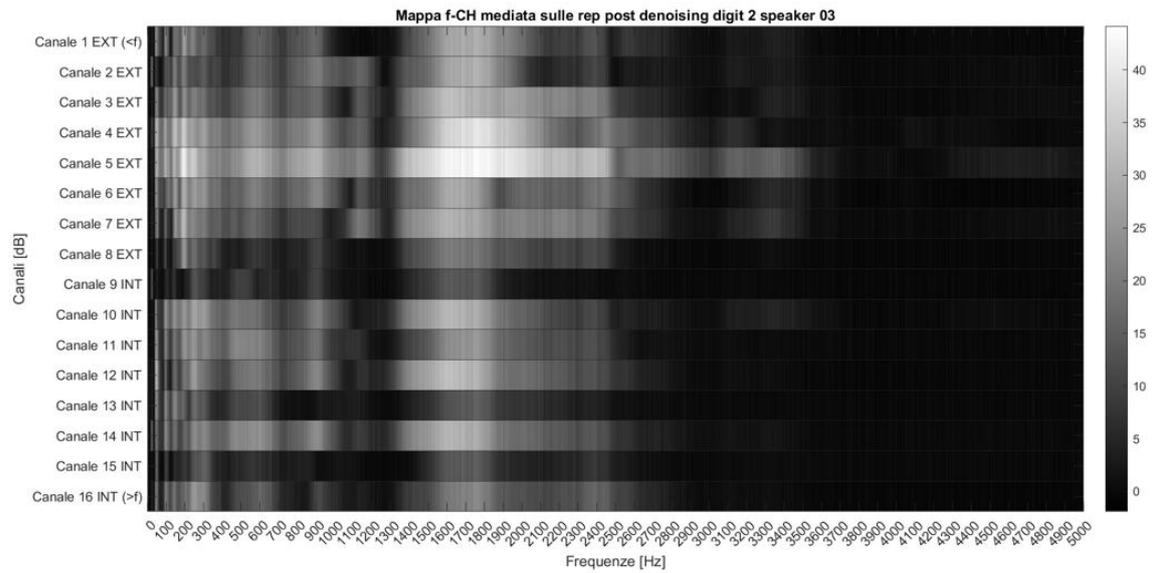


Figura 53: Mappe e differenza del digit "2" pronunciate dagli speaker 03 e 06.

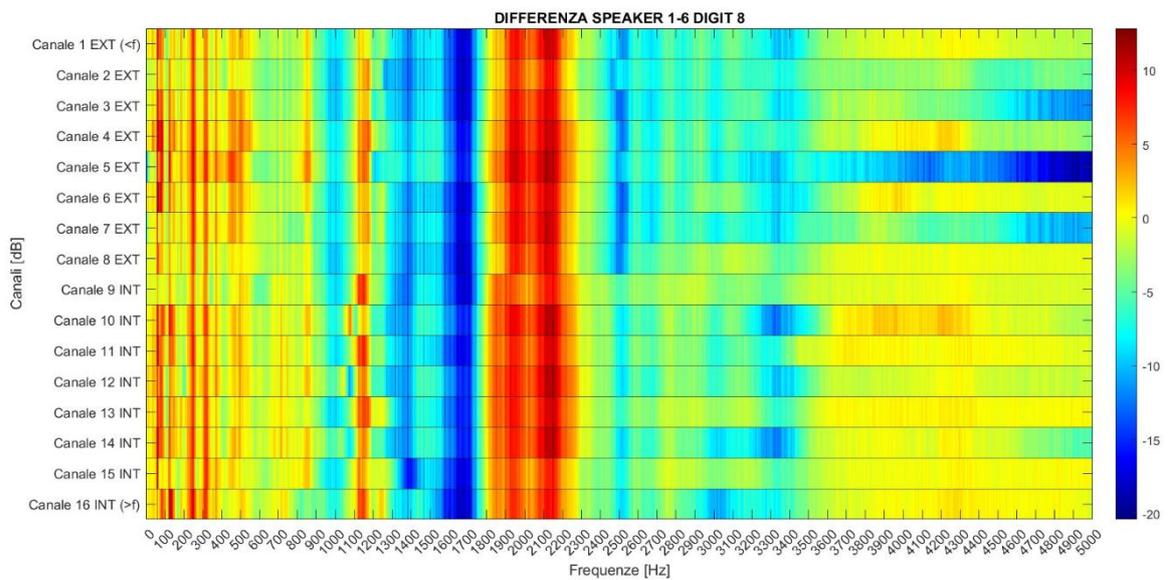
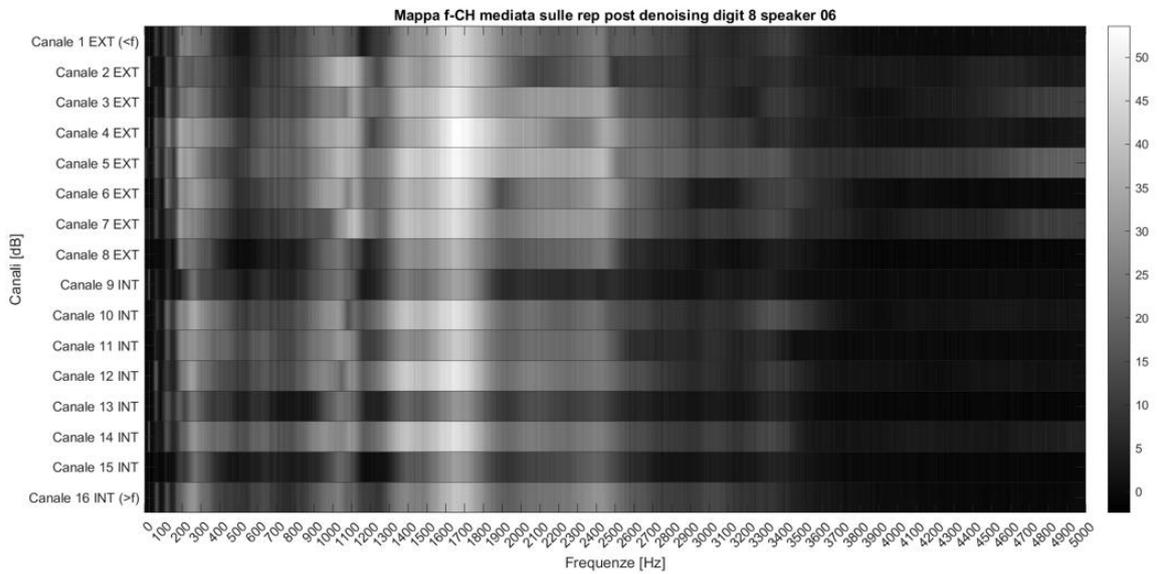
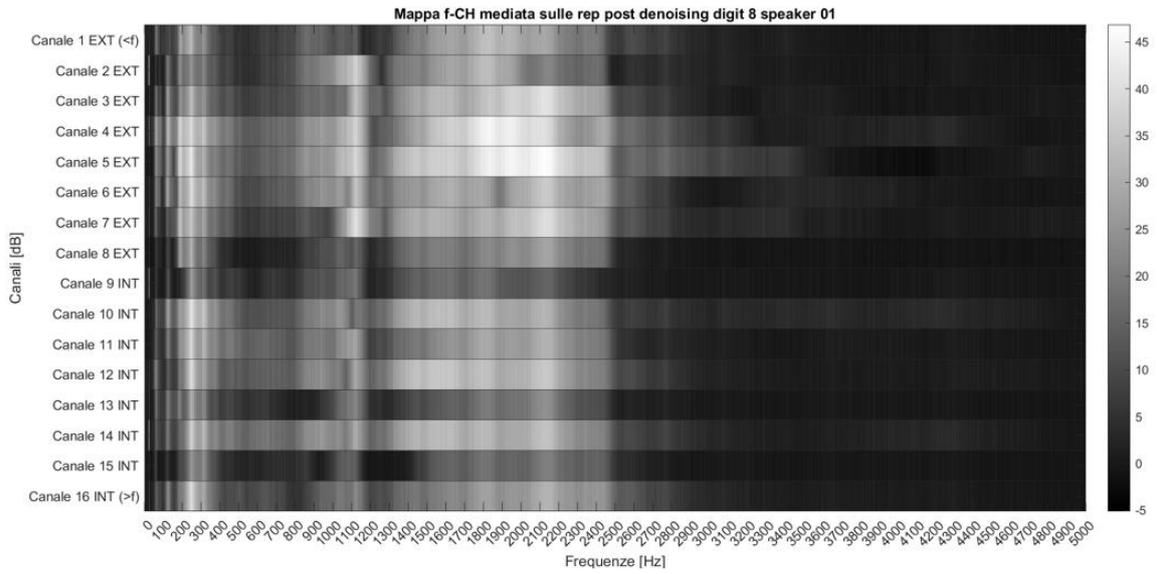


Figura 54: Mappe e differenze del digit “8” pronunciate dagli speaker 01 e 06.

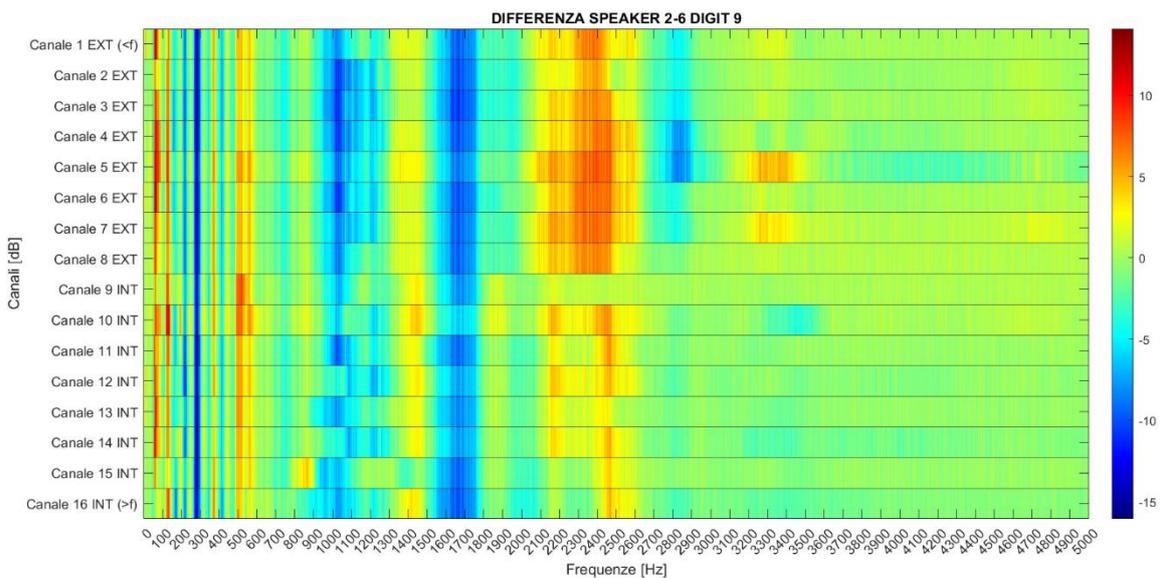
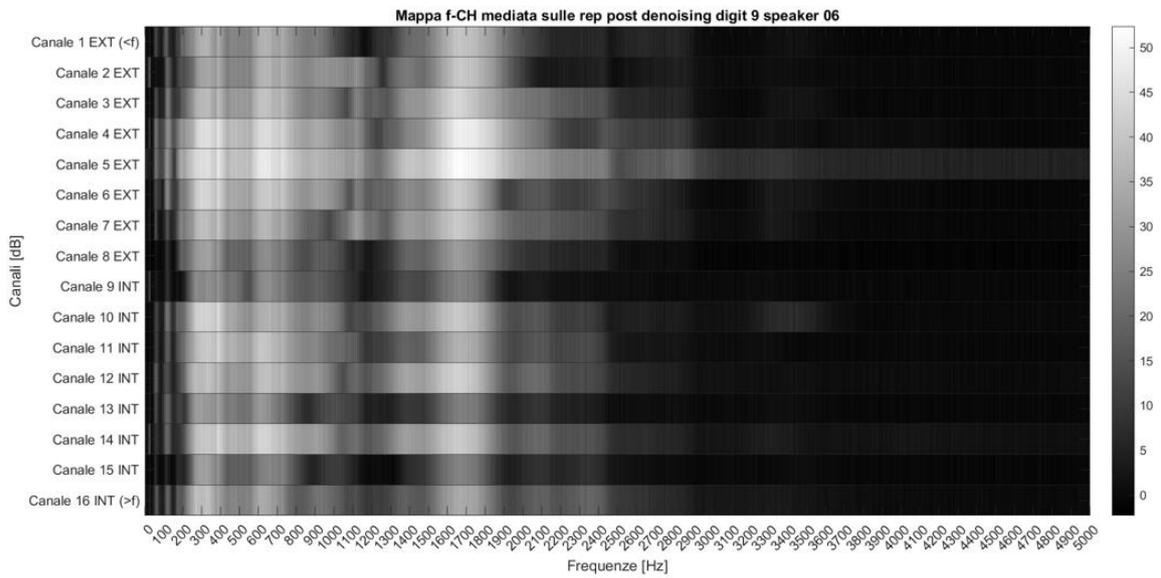
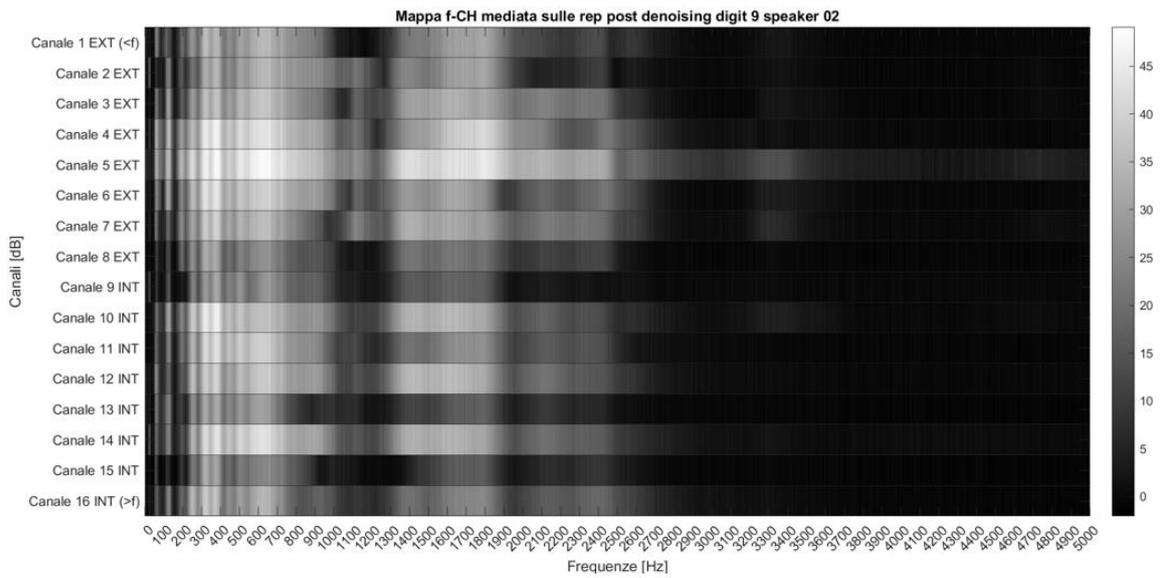


Figura 55: Mappe e differenza del digit “9” pronunciate dagli speaker 02 e 06.

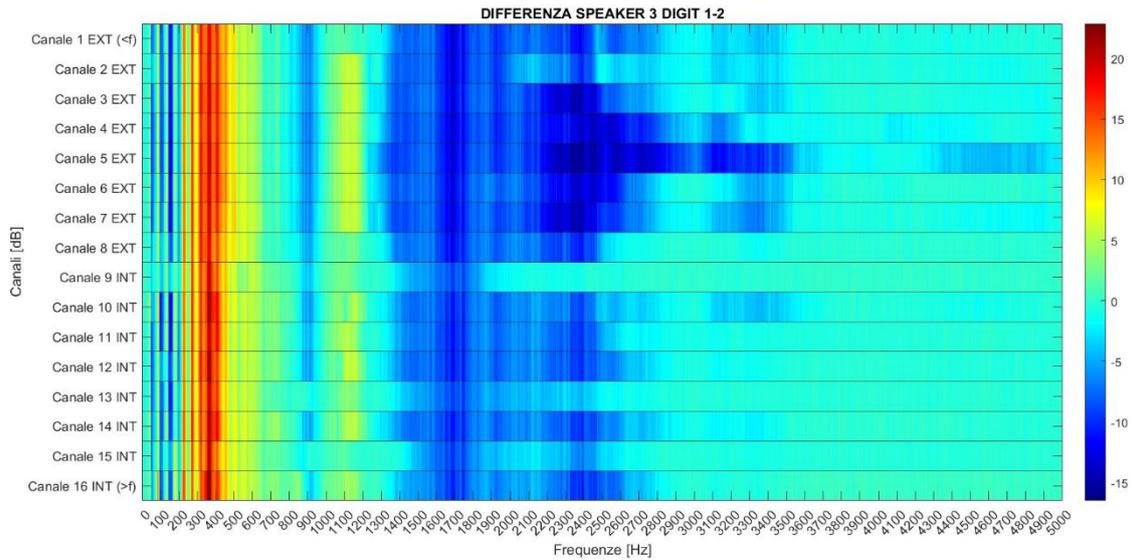


Figura 56: Differenza tra le mappe dello speaker 3 che pronuncia il digit 1 e 2.

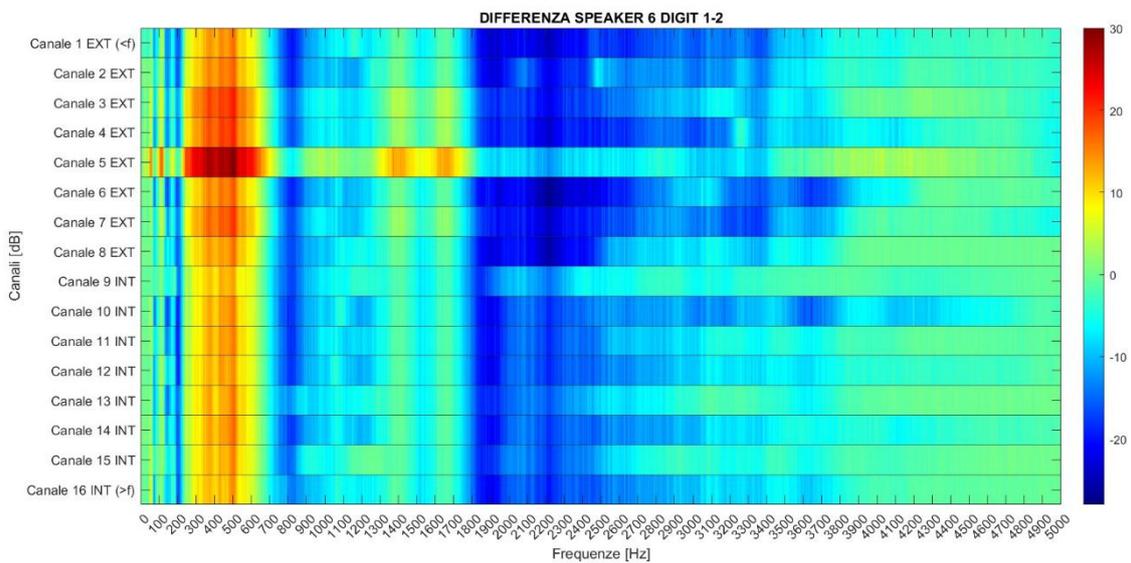


Figura 57: Differenza tra le mappe dello speaker 6 che pronuncia il digit 1 e 2.

5 Conclusioni

I macro-obiettivi prefissati per questo lavoro di tesi sono stati due: una rassegna della tecnologia attualmente disponibile sul mercato per gli impianti cocleari e la caratterizzazione di un sensore bio-ispirato da poter utilizzare in un possibile futuro come alternativa ai classici impianti citati nel primo obiettivo.

Il primo obiettivo, seppur di carattere bibliografico, è stato fondamentale per tutto il lavoro svolto durante la tesi. Infatti, lo scopo finale di questo progetto è poter valutare eventuali vantaggi competitivi del possibile nuovo impianto cocleare proposto rispetto a quelli esistenti, evidenziando le caratteristiche più significative. Ciò significa che la parte di analisi del funzionamento di queste neuroprotesi è fondamentale per poter offrire un confronto ragionato e competitivo con altre tecnologie già esistenti. Tramite questa ricerca è stato possibile ottenere molte informazioni utili a questo scopo.

L'analisi svolta ha permesso di evidenziare dove l'impianto cocleare interviene e la fascia di popolazione interessata. L'impianto cocleare è considerato una delle neuroprotesi più efficaci in commercio perché riesce a ricostruire un organo di senso fondamentale come l'udito. Considerare l'udito come la sola percezione di suoni è fortemente riduttivo. Infatti, chi è affetto da perdita uditiva è fortemente svantaggiato nella vita di tutti i giorni. Gli impianti cocleari non si limitano a risolvere solo l'aspetto medico, tramite il ripristino dell'udito, ma permettono alle persone di superare le limitazioni imposte dalla sordità in diversi ambiti: sociale, economico, psicologico, ecc. Questo emerge dall'analisi della documentazione fornita dalle case produttrici degli impianti cocleari, in cui la persona è messa sempre al primo posto, e fornendo una serie di accorgimenti che rendano la persona portatrice di impianto cocleare inserita al meglio nella società. Nell'ambito di un futuro sviluppo di impianto cocleare, tutti questi aspetti della vita quotidiana sono fondamentali per la corretta progettazione di una neuroprotesi perché possono determinare la scelta o meno del dispositivo da parte del paziente. Un altro importante aspetto da considerare è il punto di vista sanitario. Essendo una protesi impiantabile, necessita di un intervento chirurgico. Perciò, è importante realizzare un dispositivo che faciliti il lavoro del chirurgo. Questo importante aspetto è una delle priorità per le aziende produttrici. Un modello cocleare che renda l'intervento più semplice o che si adatti meglio al singolo paziente è anch'esso un fattore che può indirizzare la persona ad una scelta più mirata.

L'analisi condotta inoltre ha permesso di studiare più nel dettaglio e descrivere la tecnologia utilizzata da questi dispositivi. Se una parte fondamentale è l'impianto interno, altrettanto importante è il processore esterno e soprattutto la durata della batteria. Questi due aspetti sono direttamente collegati all'utilizzatore: il processore esterno può fornire più scelte per modificare il tipo di ascolto in base alle situazioni e di conseguenza cambia l'autonomia dell'alimentazione del dispositivo. L'autonomia energetica del dispositivo è sicuramente impattante sulla vita della persona perché garantisce la durata del dispositivo e quindi effettivamente la durata del sistema uditivo. La dipendenza da un sistema di ricarica può condizionare in maniera pesante la vita del paziente. D'altro canto, la condizione ideale sarebbe quella di ottenere un sistema sempre attivo che possa simulare il sistema uditivo in persone sane, quindi dare la possibilità alla persona portatrice di impianto cocleare di non sentirsi "diverso". Nel caso delle batterie ricaricabili attualmente utilizzate, difficilmente esse riescono a coprire un giorno intero di utilizzo. Avere quindi un dispositivo che offra buone prestazioni di ascolto e una lunga durata della batteria è fondamentale.

Durante gli anni sono state sviluppate molte tecniche per analizzare il suono e quest'ultime sono tutt'ora in continuo aggiornamento al fine di massimizzare l'esperienza dell'utente. Questi due aspetti devono essere considerati per un possibile progetto futuro.

Si può affermare che, dal primo obiettivo è stato possibile ottenere tutte le informazioni necessarie per capire il funzionamento dei dispositivi attualmente in commercio, capire le loro limitazioni, e anche ottenere informazioni nell'ambito dei dispositivi medici in Italia. Tutto ciò è stato utile per il secondo obiettivo: la caratterizzazione di una nuova coclea artificiale, ideata e realizzata dal gruppo in cui ho lavorato per la tesi.

La parte fondamentale del possibile impianto cocleare è il sensore tonotopico bio-ispirato analizzato in questa tesi. Per poter ipoteticamente sostituire l'impianto interno di una neuroprotesi commerciale, esso deve rispecchiare la tonotopia della coclea biologica. Tale dispositivo, infatti, è stato realizzato in modo da presentare una caratteristica tonotopica naturale, grazie a cui poter effettuare un'analisi in frequenza senza bisogno di un post-processing aggiuntivo e quindi consentendo un risparmio energetico rilevante, che, come detto, è uno degli obiettivi per la progettazione di una nuova generazione di coclee artificiali.

La tonotopia del sensore è stato il primo aspetto analizzato durante la tesi. Infatti, una volta ottenuti i risultati da entrambi i dataset è stato possibile condurre un'analisi sui canali come ripotato nel capitolo conclusivo della tesi. Dalle analisi condotte è stato possibile osservare l'effetto tonotopico su tutti i canali, ovviamente considerando tutto ciò che ha causato rumore come il non perfetto incollaggio dei sensori piezoelettrici di rilevamento, l'eccessiva durata della prova, il segnale prelevato è più pulito vicino al sensore piezoelettrico stimolatore rispetto che al sensore di rilevamento, ecc.

Avuta la conferma sul corretto funzionamento dei canali e quindi dell'effetto tonotopico del sensore è stato possibile analizzare se la coclea fosse in grado di riconoscere il parlato (quindi riconoscere delle caratteristiche precipue per ognuno dei "digit" proposti) ed anche poter riconoscere una persona dall'altra. Per verificare questa possibilità sono state effettuate analisi specifiche sui canali in funzione della frequenza rilevando come, a seconda della posizione di rilevamento del segnale (definita da un diverso canale) fosse osservabile un diverso contenuto in frequenza. L'analisi svolta determina un elevato grado di affidabilità nel riconoscimento di segnali acustici diversi tramite la coclea artificiale proposta, fornendo un'indicazione sull'utilizzabilità della stessa come biosensore.

In conclusione, unendo i due obiettivi della tesi, si può affermare che ipoteticamente il dispositivo può essere potenzialmente un ottimo concorrente agli attuali impianti cocleari; infatti, la sua struttura potrebbe allungare la durata della batteria confermando allo stesso tempo delle ottime prestazioni nel riconoscimento del parlato.

Tutto ciò che è stato svolto in questa tesi è solo il primo passo per un più ampio progetto sull'argomento. Infatti, tramite questo lavoro è stato anche possibile individuare i punti di maggiore criticità che devono essere presi in considerazione in uno sviluppo futuro. Dal punto di vista dell'analisi del dispositivo ci sono due aspetti da poter migliorare: l'utilizzo dei sensori piezoelettrici può essere migliorato o addirittura sostituito con strumenti più affidabili e precisi per il nostro caso specifico; così come può essere migliorata il tipo di analisi dei risultati, ad esempio tramite diverse tecniche di analisi del segnale. Dal punto di vista dell'integrazione del sensore con la neuroprotesi, bisogna considerare diversi aspetti. Nella sua forma attuale, il sensore non può essere impiantato direttamente nel paziente, quindi è indispensabile un'interfaccia che ne garantisca la biocompatibilità e la stabilità a lungo termine. Questa interfaccia deve rendere il dispositivo sicuro per l'organismo umano, prevenire reazioni infiammatorie e garantire la resistenza ai fluidi dell'orecchio, evitando il riconoscimento come corpo estraneo da parte del sistema immunitario (rigetto). Un possibile approccio consiste nell'utilizzo di una membrana bio-compatibile che ricopra completamente il sensore. Inoltre, l'interfaccia

deve fungere da ponte tra l'accoppiamento acustico-meccanico, che si genera nel processo di captazione del suono, e la successiva stimolazione diretta del nervo cocleare. Per questo, per il primo accoppiamento è necessario integrare un sistema di trasduzione che converta il segnale elettrico derivante dal microfono esterno in una vibrazione meccanica per il sensore tonotopico; attraverso ad esempio sensori piezoelettrici, capacitori, microfoni MEMS o altri tipi di tecnologia. Queste soluzioni tecnologiche possono essere utilizzate anche per il secondo accoppiamento che necessita di un comportamento inverso in cui la vibrazione meccanica della coclea viene convertita in segnali elettrici. Questi segnali, essendo già processati dalla coclea artificiale devono essere trasmessi al nervo cocleare mediante una matrice di elettrodi come avviene negli impianti tradizionali.

6 Bibliografia

1. World Health Organization. (2021). *World report on hearing*. World Health Organization. <https://iris.who.int/handle/10665/339913>
2. Sheikh, A., Bint-e-Zainab, Shabbir, K., & Imtiaz, A. (2022). *Structure and physiology of human ear involved in hearing*. IntechOpen. <https://doi.org/10.5772/intechopen.105466>
3. Alberti, P. W. (2001). *The anatomy and physiology of the ear and hearing*. In *Occupational exposure to noise: Evaluation, prevention, and control* (pp. 53–62).
4. Tris Atmaja, B. (2019). *The physiology, mechanism, and nonlinearities of hearing*. <https://doi.org/10.13140/RG.2.2.24895.97444>
5. Kunchur, M. N. (2023). *The human auditory system and audio*. *Applied Acoustics*, 211, 109507. <https://doi.org/10.1016/j.apacoust.2023.109507>
6. Ni, G., Elliott, S. J., Ayat, M., & Teal, P. D. (2014). *Modelling cochlear mechanics*. *BioMed Research International*, 2014, 150637. <https://doi.org/10.1155/2014/150637>
7. Hudspeth, A. (2014). *Integrating the active process of hair cells with cochlear function*. *Nature Reviews Neuroscience*, 15, 600–614. <https://doi.org/10.1038/nrn3786>
8. Zhang, Q., & Duan, M. L. (2009). *Anatomy and physiology of peripheral auditory system and common causes of hearing loss*. *Journal of Otology*, 4(1), 7–14. [https://doi.org/10.1016/S1672-2930\(09\)50002-5](https://doi.org/10.1016/S1672-2930(09)50002-5)
9. Wangemann, P. (2006). *Supporting sensory transduction: Cochlear fluid homeostasis and the endocochlear potential*. *Journal of Physiology*, 576(1), 11–21. <https://doi.org/10.1113/jphysiol.2006.112888>
10. Hall, J. W. 3rd. (2000). *Development of the ear and hearing*. *Journal of Perinatology*, 20(8 Pt 2), S12–S20. <https://doi.org/10.1038/sj.jp.7200439>
11. Hudspeth, A. J. (1997). *How hearing happens*. *Neuron*, 19(5), 947–950. [https://doi.org/10.1016/S0896-6273\(00\)80385-2](https://doi.org/10.1016/S0896-6273(00)80385-2)
12. Emanuel, D., Maroonroge, S., & Letowski, T. (2009). *Auditory function: Physiology and function of the hearing system*.
13. Reichenbach, T., & Hudspeth, A. J. (2014). *Rep. Prog. Phys.*, 77, 076601.
14. Reichenbach, T., Stefanovic, A., Nin, F., & Hudspeth, A. J. (2012). *Waves on Reissner's membrane: A mechanism for the propagation of otoacoustic emissions from the cochlea*. *Cell Reports*, 1(4), 374–384. <https://doi.org/10.1016/j.celrep.2012.02.013>
15. Cunningham, L. L., & Tucci, D. L. (2017). *Hearing loss in adults*. *New England Journal of Medicine*, 377(25), 2465–2473. <https://doi.org/10.1056/NEJMra1616601>
16. Yueh, B., Shapiro, N., MacLean, C. H., & Shekelle, P. G. (2003). *Screening and management of adult hearing loss in primary care: Scientific review*. *JAMA*, 289(15), 1976–1985. <https://doi.org/10.1001/jama.289.15.1976>

17. Michels, T. C., Duffy, M. T., & Rogers, D. J. (2019). *Hearing loss in adults: Differential diagnosis and treatment*. *American Family Physician*, 100(2), 98–108.
18. Hainarosie, M., Zainea, V., & Hainarosie, R. (2014). *The evolution of cochlear implant technology and its clinical relevance*. *Journal of Medicine and Life*, 7(Spec No. 2), 1–4.
19. Zeng, F. G., Rebscher, S., Harrison, W., Sun, X., & Feng, H. (2008). *Cochlear implants: System design, integration, and evaluation*. *IEEE Reviews in Biomedical Engineering*, 1, 115–142. <https://doi.org/10.1109/RBME.2008.2008250>
20. Advanced Bionics. (n.d.). *Home*. <https://www.advancedbionics.com/it/it/home.html>
21. Macherey, O., & Carlyon, R. P. (2014). *Cochlear implants*. *Current Biology*, 24(18), R878–R884. <https://doi.org/10.1016/j.cub.2014.06.053>
22. Medel. (n.d.). *Home*. <https://www.medel.com>
23. Cochlear. (n.d.). *Home*. <https://www.cochlear.com>
24. Oticon Medical. (n.d.). *Home*. <https://www.oticonmedical.com>
25. Ministero della Salute. (n.d.). *Interrogazione dispositivi medici*. <https://www.salute.gov.it/interrogazioneDispositivi>
26. Carlyon, R. P., & Goehring, T. (2021). *Cochlear implant research and development in the twenty-first century: A critical update*. *Journal of the Association for Research in Otolaryngology*, 22(5), 481–508. <https://doi.org/10.1007/s10162-021-00811-5>
27. Nie, K., Stickney, G., & Zeng, F. G. (2005). *Encoding frequency modulation to improve cochlear implant performance in noise*. *IEEE Transactions on Biomedical Engineering*, 52(1), 64–73. <https://doi.org/10.1109/TBME.2004.839799>
28. Zeng, F. G. (2004). *Trends in cochlear implants*. *Trends in Amplification*, 8(1), 1–34. <https://doi.org/10.1177/108471380400800102>
29. Wilson, B., & Dorman, M. (2008). *Interfacing sensors with the nervous system: Lessons from the development and success of the cochlear implant*. *IEEE Sensors Journal*, 8, 131–147. <https://doi.org/10.1109/JSEN.2007.912917>
30. Melnikov, N. S., & Kozlov, A. G. (2023). *Analysis of coding strategies in cochlear implant systems*. *IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT)*. <https://doi.org/10.1109/USBREIT58508.2023.10158841>
31. Chen, Y., Chen, Y., & Xue, G. (2018). *ACE speech processing strategy and feature analysis in CI device*. *2018 IEEE 3rd International Conference on Integrated Circuits and Microsystems (ICICM)*, 318–321. <https://doi.org/10.1109/ICAM.2018.8596414>
32. McKay, C. M. (2021). *Applications of phenomenological loudness models to cochlear implants*. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.611517>

33. Dal Poggetto, V. F., Bosia, F., Urban, D., Beoletto, P. H., Torgersen, J., Pugno, N. M., & Gliozzi, A. S. (2023). *Cochlea-inspired tonotopic resonators*. *Materials & Design*, 227, 111712. <https://doi.org/10.1016/j.matdes.2023.111712>
34. Jackson, Z., Souza, C., Flaks, J., Pan, Y., Nicolas, H., & Thite, A. (2018). *Jakobovski/free-spoken-digit-dataset: v1.0.8 (v1.0.8)* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.1342401>
35. Becker, S., Vielhaben, J., Ackermann, M., Müller, K.-R., Lapuschkin, S., & Samek, W. (2024). *AudioMNIST: Exploring explainable artificial intelligence for audio analysis on a simple benchmark*. *Journal of the Franklin Institute*, 361(1), 418–428. <https://doi.org/10.1016/j.jfranklin.2023.11.038>