

POLITECNICO DI MILANO

Facoltà di Ingegneria

Consiglio di Corso di Studio in Ingegneria Biomedica  
Corso di Laurea (CL) in Ingegneria Biomedica



Trattamento del segnale vocale ed effetti sul clear speech:  
sviluppo di un simulatore di protesi acustica digitale.

Relatori: prof. Gabriella TOGNOLA

prof. Ferdinando GRANDORI

Correlatore: ing. Alessia PAGLIALONGA

Laureando: Giulia BIGOGNO

matricola 739775

Anno Accademico 2010 - 2011

## STRUTTURA TESI

---

	Pag.
SOMMARIO	8
ABSTRACT	9
INTRODUZIONE	10
CAPITOLO 1: GLI APPARECCHI ACUSTICI	12
1. Classificazione degli apparecchi acustici	17
2. Apparecchi acustici digitali	20
2.1. Principio di funzionamento e componenti interni	21
2.1.1. Processore digitale	25
2.1.2. Banco di filtri	26
2.1.3. <i>Output – limiting</i> e compressione	29
2.1.4. Controlli	32
3. Fenomeni che degradano la performance di un apparecchio acustico	33
3.1. Problema del range dinamico	33
3.2. Problema della scarsa discriminazione soprasoglia	35
3.3. Problema della distorsione	35
4. Distorsioni elettroacustiche	36
5. Campi della ricerca attuale	38
CAPITOLO 2: IL CLEAR SPEECH	41
1. Introduzione	41
2. Cosa rende più intelligibile il <i>clear speech</i> ?	44
2.1. Espansione del <i>vowels space</i>	45
2.2. Diminuzione della velocità e aumento delle pause nel parlato	47
2.3. Inviluppo temporale	48
3. Considerazioni generali e possibili sviluppi futuri	50
CAPITOLO 3: TEORIA DEI SIMULATORI DI PROTESI ACUSTICHE	52
1. Considerazioni generali	52
2. Esempio di simulatore	53
3. Esempio di simulazione	55
CAPITOLO 4: MATERIALI E METODI	59
1. Considerazioni generali	59

2. Descrizione del set-up sperimentale	60
2.1. Microfono	61
2.2. Processore digitale	62
2.3. Banco di filtri	63
2.4. <i>Automatic Gain Control</i> (AGC)	65
2.5. Feedback acustico	70
3. Materiale vocale	72
4. Descrizione delle <i>features</i> estratte	75
4.1. <i>Modulation Index</i>	75
4.2. Formanti e <i>vowels space</i>	75
5. Protocollo di <i>speech processing</i>	76
5.1. Estrazione del <i>Modulation Index</i>	78
5.2. Estrazione delle formanti e costruzione del <i>vowels space</i>	80
CAPITOLO 5: RISULTATI	81
1. Parole “salto” e “cinque”: <i>Modulation Index</i>	81
1.1. AGC da 65dB e 75dB sempre in funzione	81
1.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente	85
2. Vocali singole: <i>vowels space</i>	88
2.1. AGC da 65dB e 75dB sempre in funzione	88
2.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente	89
CAPITOLO 6: DISCUSSIONE	90
1. Parole “salto” e “cinque”: <i>Modulation Index</i>	90
1.1. AGC da 65dB e 75dB sempre in funzione	90
1.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente	92
2. Vocali singole: <i>vowels space</i>	93
2.1. AGC da 65dB e 75dB sempre in funzione	93
2.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente	94
CONCLUSIONI	95
RINGRAZIAMENTI	97
BIBLIOGRAFIA	98

## INDICE DELLE FIGURE

---

	Pag.
1. Schema a blocchi di un semplice apparecchio acustico analogico. I componenti principali sono il microfono, il ricevitore, un preamplificatore, un filtro analogico e un amplificatore di potenza: questi ultimi tre sono monitorati attraverso controlli analogici, variabili manualmente tramite interruttori o rotelle meccaniche.	13
2. Schema a blocchi di un apparecchio digitale con controllo analogico. La differenza con le protesi analogiche tradizionale sta nella possibilità di essere controllati digitalmente dall'audioprotesista attraverso una memoria programmabile.	14
3. Schema a blocchi di un apparecchio acustico digitale, caratterizzato da un microfono, un ADC, un processore digitale, un DAC e un ricevitore, componenti sotto il diretto controllo di un minicalcolatore.	15
4. Tipologie di apparecchi acustici in base al posizionamento sull'orecchio: <i>completely-in-the-canal</i> (CIC), <i>in-the-canal</i> (ITC), <i>in-the-ear</i> (ITE) e <i>behind-the-ear</i> (BTE).	17
5. A sinistra: apparecchio acustico di tipo <i>body-worn</i> , tascabile; a destra apparecchi acustici integrati con occhiali da vista.	19
6. Sezione trasversale di un apparecchio acustico ITE.	20
7. Schema di un microfono <i>electret</i> .	21
8. Diagramma polare di un microfono omnidirezionale.	22
9. Diagramma polare di un microfono direzionale di tipo cardioide.	22
10. Rappresentazione schematica delle fasi di speech processing che avvengono all'interno di un apparecchio acustico digitale. Il segnale d'ingresso subisce l'azione di un primo filtro <i>anti-aliasing</i> , viene convertito in forma digitale, processato, riconvertito in forma analogica e, dopo l'azione di un filtro <i>anti-imaging</i> , è pronto per essere inviato ad un sistema di trasduzione di pressione sonora.	24
11. Intensità delle principali frequenze del parlato, comprese tipicamente tra 125 e 8000 Hz.	25
12. Esempio di audiogramma tonale binaurale.	27
13. Risposta del guadagno di un processore a tre bande.	28
14. Tecnica del <i>peak-clipping</i> . La parte di segnale al di sopra della soglia prestabilita viene letteralmente rimossa.	29

15.	Curva ingresso – uscita di un AGC, in scala logaritmica. Si osserva una fase lineare, una di compressione e una di limitazione dell’uscita ad un valore di limite di intensità sonora.	30
16.	Controllo del volume.	32
17.	<i>Vowels space</i> costruito su quattro vocali della lingua inglese.	46
18.	Inviluppo temporale di un segnale.	48
19.	(a) segnale originale; (b) segnale filtrato nella banda 283-336Hz; (c) inviluppo del segnale in (b).	49
20.	Esempio di un modello di simulazione di apparecchio acustico digitale presentato in letteratura.	53
21.	Schema a blocchi delle possibili fasi di elaborazione del segnale di speech all’interno di un simulatore di apparecchio acustico.	56
22.	Schema a blocchi di un processo di riduzione del rumore di sottofondo in un Simulatore di apparecchio acustico digitale.	57
23.	Schema del simulatore implementato. Sono visibili il microfono, il banco di filtri, un sistema di peak-clipping e il blocco che simula il feedback acustico.	60
24.	Risposta in frequenza del microfono.	61
25.	Schema dei componenti all’interno di ciascuno dei sedici canali.	62
26.	Risposta in frequenza dei filtri passa-banda dei sedici canali.	64
27.	Circuito rilevatore di picchi ( <i>peak detector</i> ). Il segnale in uscita dal filtro viene rettificato e permette di caricare il condensatore se la durata del picco è maggiore del tempo di attacco ( $R1 \cdot C$ ). Al diminuire dell’ampiezza del segnale, il condensatore si scaricherà sulla resistenza $R2$ , con un tempo di rilascio pari a $R2 \cdot C$ . L’uscita del peak detector è la differenza di potenziale ai capi del condensatore.	66
28.	Circuito compressore. Se il segnale in uscita dal peak detector è superiore alla soglia stabilita, viene compresso tramite l’applicazione di un fattore moltiplicativo minore di 1, in funzione dell’intensità del segnale in uscita dal filtro ( $u$ ) e del segnale rettificato in uscita dal peak detector ( $ u $ ). In caso contrario il segnale si propaga inalterato.	66
29.	Nucleo della compressione. Il segnale $u$ viene moltiplicato per un fattore dipendente dal segnale rettificato in uscita dal PD.	67

## INDICE DELLE TABELLE

---

	Pag.
1. Frequenze centrali e rispettive larghezze di banda dei 16 filtri passa-banda del banco. * Il filtro n°1 è un passa-basso del 109°ordine con frequenza di taglio 250 Hz.	64
2. Elenco delle sedici configurazioni di AGC studiate.	67

## INDICE DEI GRAFICI

---

	Pag.
1. Diversi tipi di compressione al variare della soglia e del rapporto di compressione (RC).	68
2. Guadagno della compressione in funzione dell'ingresso.	69
3. Diagramma di Bode di ampiezza e fase del feedback acustico.	71
4. Forma d'onda della vocale A.	72
5. Forma d'onda della vocale E.	72
6. Forma d'onda della vocale I.	73
7. Forma d'onda della vocale O.	73
8. Forma d'onda della vocale U	73
9. Forma d'onda delle parole "cinque" (in alto) e "salto" (in basso).	74
10. Entità della compressione (su scala normalizzata) subita dalla parola "salto" in funzione delle sedici diverse configurazioni del simulatore di apparecchio acustico digitale, con tutti gli AGC in funzione. Sull'asse delle ascisse sono riportate le sedici configurazioni studiate: prendendo a titolo esemplificativo la prima a destra, i primi due valori (75 e 2:1) rappresentano la soglia in dB e il rapporto di compressione degli AGC dei canali a bassa frequenza (canali 1-7), mentre gli ultimi due la soglia e il rapporto degli AGC dei canali ad alta frequenza (canali 8-16).	82
11. Entità della compressione (su scala normalizzata) subita dalla parola "cinque" in funzione delle diverse configurazioni del simulatore di apparecchio acustico	82

- digitale, con tutti gli AGC in funzione.
12. Spettro in frequenza dell'inviluppo temporale della parola "salto" con tutti gli AGC in funzione: confronto tra il segnale non elaborato (in nero) e quelli ottenuti nelle sedici diverse configurazioni del simulatore. 83
  13. Spettro in frequenza dell'inviluppo temporale della parola "cinque" con tutti gli AGC in funzione: confronto tra il segnale non elaborato (in nero) e quelli ottenuti nelle sedici diverse configurazioni del simulatore. 84
  14. Entità della compressione (su scala normalizzata) subita dalla parola "salto" in funzione delle sedici diverse configurazioni del simulatore di apparecchio acustico digitale, con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente. 86
  15. Entità della compressione (su scala normalizzata) subita dalla parola "cinque" in funzione delle sedici diverse configurazioni del simulatore di apparecchio acustico digitale, con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente. 86
  16. Spettro in frequenza dell'inviluppo temporale della parola "salto" con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente: confronto tra il segnale originale (in nero) e tre tipi di configurazioni, rivelatesi equivalenti. 87
  17. Spettro in frequenza dell'inviluppo temporale della parola "cinque" con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente: confronto tra il segnale originale (in nero) e quattro tipi di configurazioni, rivelatesi equivalenti. 88
  18. *Vowels spaces* ottenuti dalle sedici configurazioni del simulatore in condizioni di attivazione continua di tutti gli AGC, sovrapposti al vowels space di "controllo" (linea e indicatori neri). 89
  19. *Vowels spaces* ottenuti dalle sedici configurazioni del simulatore in condizioni di attivazione continua degli AGC da 65dB e funzione intermittente degli AGC con soglia, sovrapposti al vowels space di "controllo". 89

## SOMMARIO

Gli apparecchi acustici sono dispositivi elettronici in grado di compensare perdite uditive trasmissive e neurosensoriali, fornendo un guadagno di amplificazione selettivo in frequenza, sulla base dell'audiogramma di un individuo ipoacusico.

Tali dispositivi nascono per migliorare la qualità della comunicazione verbale di soggetti affetti da difficoltà percettive, ma sono tuttavia accompagnati da fattori che degradano la performance dell'apparecchio acustico stesso, il cui risultato consiste in una perdita globale di intelligibilità.

Il presente studio si propone di analizzare l'effetto della compressione fornita da un apparecchio acustico simulato su due caratteristiche spettrali del segnale vocale, responsabili dell'intelligibilità dello speech stesso.

Il primo capitolo fornisce una visione globale dello stato dell'arte delle protesi acustiche moderne, analizzandone il principio di funzionamento, i principali componenti interni e le caratteristiche di elaborazione del segnale vocale.

Il secondo capitolo riguarda l'ambito della ricerca sul clear speech, illustrando le caratteristiche acustiche e spettrali responsabili della produzione di un clear speech di alta qualità.

Il capitolo 3 mostra invece le potenzialità delle simulazioni computerizzate e delle tecniche di elaborazione digitale nella progettazione di nuovi apparecchi acustici e nel perfezionamento di quelli già esistenti.

I capitoli successivi sono invece relativi alla parte sperimentale: il quarto capitolo descrive il modello di apparecchio acustico implementato, il materiale vocale usato, le *features* estratte e il protocollo di *speech processing*, il quinto contiene i grafici dei risultati ottenuti discussi nel sesto e ultimo capitolo.



## **ABSTRACT**

A hearing aid is an electronic device able to make up for transmissive and sensorineural hearing impairments, providing a frequency – selective gain based on the audiogram of a hypoacusic patient.

The need of these devices arises in order to improve verbal communication quality of people affected by perceptual difficulties; nevertheless they are accompanied by drawbacks that deteriorate hearing aid performance itself, whose outcome consists of an overall loss of intelligibility.

The present study intends to examine the compression effect of a simulated hearing aid on a couple of vocal signal spectral features, thought responsible of speech intelligibility.

Chapter one provides with a comprehensive view of the current hearing aids state-of-art, analyzing its working principles, its main internal components and speech processing features.

Chapter two concerns clear speech research, describing the acoustic and spectral characteristics responsible of a high quality clear speech production.

Chapter three, instead, shows the potential of computerized simulations and digital signal processing in the design of new hearing aids and in the improvement of those already existing.

Following chapters concern the experimental part of this work: chapter four illustrates the hearing aid model built for this purpose, the vocal material used, the features explored and the speech processing protocol, the fifth one contains the obtained charts, discussed in the sixth and last chapter.

## INTRODUZIONE

Uno dei principali campi applicativi della bioingegneria si occupa dello sviluppo e perfezionamento di dispositivi in grado di far fronte alle diverse tipologie di perdite uditive e ha dato così alla luce i primi impianti cocleari negli anni '70 e, ancor prima, i primi apparecchi acustici analogici.

Basicamente, una protesi acustica è un piccolo apparecchio elettromedicale costituito da un microfono, che capta le onde sonore incidenti sulla sua membrana e le converte in segnali elettrici, da una serie di filtri e amplificatori che permettono un'amplificazione selettiva in frequenza di tali segnali, e da un ricevitore, la cui funzione è quella di trasdurre il l'onda amplificata in segnale pressorio, in modo che sul timpano del paziente incida un suono di intensità facilmente percepibile.

Grazie all'instancabile innovazione tecnologica, l'evoluzione della microelettronica ha portato, nel corso delle ultime decadi, a passi da gigante nel campo audioprotesico, e alla realizzazione di apparecchi acustici digitali miniaturizzati, con vantaggi incomparabili rispetto ai dispositivi analogici convenzionali.

La tecnologia si è portata tuttavia con sé nuove difficoltà, tra cui le principali riguardano l'eliminazione del rumore di sottofondo, la generazione del fastidioso segnale di feedback acustico e la scelta di un opportuno algoritmo di compressione. Tutti questi fattori giocano un ruolo di primaria importanza nella produzione di un *clear speech* di buona qualità, caratterizzato da un'alta intelligibilità dei segnali di speech. È in questo scenario che si delinea un ulteriore campo di ricerca, focalizzato sull'indagine delle caratteristiche acustico-fonetiche del parlato responsabili dell'alta intelligibilità del linguaggio: è questo l'ambito della ricerca sul *clear speech*, un particolare stile di linguaggio adottato quando sorge la necessità di farsi comprendere nella comunicazione in situazioni di difficile percezione, dovute alla presenza di rumore, ipoacusia o semplicemente di persone di lingua madre differente.

Risulta ora evidente la forte interconnessione tra l'ambito di ricerca riguardante le audioprotesi e quello del *clear speech*, dal momento che il segnale processato da codesti dispositivi, affinché sia intelligibile (e non solo banalmente amplificato), deve mantenere il più possibile tutte le caratteristiche, spettrali e temporali, ritenute essenziali nella produzione di un *clear speech* di alta qualità.

A questo scopo, nel presente studio è stato ricostruito un apparecchio acustico virtuale, per molti tratti simile alle protesi digitali attualmente in commercio. Tale modello è stato poi utilizzato per valutare il processo di elaborazione del segnale vocale che avviene al suo interno e quantificare come tale elaborazione del parlato va ad inficiare le caratteristiche salienti del *clear speech*.

Data la ridondanza delle variabili in gioco di un apparecchio acustico, la molteplicità dei gradi di libertà nella sua programmazione e l'elevata numerosità delle *features* acustiche deputate alla produzione di un buon *clear speech*, si è deciso di studiare come il processo di compressione del segnale, rappresentato dalla soglia degli AGC (*Automatic Gain Control*) e dal loro rapporto di compressione, influisce su due caratteristiche basilari nell'ambito del *clear speech*.

Tali *features* sono il *modulation index* e il *vowels space*: il primo parametro riguarda le fluttuazioni dell'involuppo temporale del segnale di speech, mentre il secondo si riferisce alle prime due formanti delle cinque vocali della lingua italiana.

I risultati sperimentali del presente studio potranno dunque fornire elementi utili in fase di prescrizione e programmazione di apparecchi acustici digitali, tenendo in considerazione le conclusioni che verranno tratte riguardo al rapporto compressione – intelligibilità.

## CAPITOLO 1

### Gli apparecchi acustici

Il problema dell'ipoacusia è nato insieme all'uomo ed essendo un problema più sociale che fisico, sin dai tempi più lontani si è cercato di porvi rimedio. L'uomo, infatti, ha sempre cercato di sperimentare tutte le soluzioni possibili per far fronte al disagio di non poter comunicare. Uno dei primi espedienti utilizzati dall'uomo per risolvere tale problema fu quello di mettere la mano a conchiglia dietro il padiglione auricolare, aumentando così la sua superficie ricettiva. In questo modo, con il fenomeno della risonanza della cavità del condotto uditivo, si viene a creare una specie di "amplificazione naturale" che conferisce all'ipoacusico la possibilità di comunicare.

Con il passare degli anni l'uomo ha sperimentato nuove tecniche per dare a tutti l'opportunità di sentire e di conseguenza non isolarsi dalle relazioni sociali. Primo fra tutti fu realizzato nel 1878 un dispositivo consistente in un microfono a carbone, una batteria ed un ricevitore telefonico. Successivamente, fu inventato il "tubo vuoto", che si trattava di un contenitore sigillato in vetro o in acciaio contenente due elettrodi, in mezzo ai quali ne era posizionato un terzo, detto "griglia", la cui funzione era quella di controllare l'intensità di corrente che passava nel tubo. In funzione della differenza di potenziale tra i due elettrodi, si produceva un flusso elettronico lungo la griglia, la cui variazione determinava una variazione di tensione della griglia e di conseguenza un'amplificazione del segnale iniziale. Nel 1924 fu sviluppato un grande apparecchio acustico che utilizzava un microfono magnetico, costituito da una membrana vibrante collegata ad una bobina in movimento. Dal suo movimento solidale con la membrana del microfono, si variava l'intensità di corrente che attraversava il solenoide e di conseguenza determinava una differenza di potenziale ai capi del microfono stesso e quindi del segnale da amplificare e inviare, attraverso il ricevitore, al condotto uditivo del soggetto. Nella seconda metà degli anni '40 grazie alla miniaturizzazione delle valvole, gli apparecchi acustici diventano "tascabili". Il 1952 segna l'inizio dell'era delle protesi acustiche contemporanee, caratterizzate dall'invenzione del transistor, che determinò notevoli progressi tecnologici. Il passaggio dai tubi vuoti ai transistor non ha determinato cambiamenti sostanziali nella prestazione dell'apparecchio acustico, ma ha ridotto di molto il costo computazionale e la dimensione dell'alimentazione della batteria. Nascono così le prime *protesi acustiche analogiche*, di

cui è stata riportata una semplice schematizzazione in figura 1. Queste sono costituite da un microfono, un blocco preamplificatore, un filtro, un amplificatore di potenza e un ricevitore. Tutti i componenti interni sono monitorati attraverso controlli analogici, la cui impostazione è esclusivamente di tipo meccanico manuale.

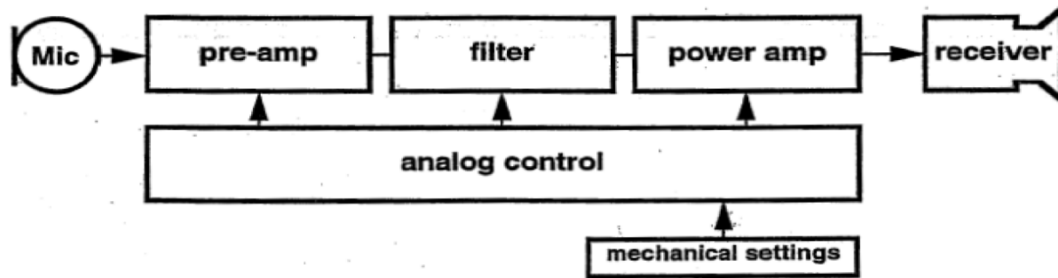


Figura 1: schema a blocchi di un semplice apparecchio acustico analogico. I componenti principali sono il microfono, il ricevitore, un preamplificatore, un filtro analogico e un amplificatore di potenza: questi ultimi tre sono monitorati attraverso controlli analogici, variabili manualmente tramite interruttori o rotelle meccaniche.

Negli anni Novanta la tecnologia sviluppa apparecchi acustici con contenuti innovativi di interesse sempre maggiore: compaiono i primi apparecchi programmabili a due/tre canali di amplificazione ed un programma di funzionamento, seguiti a breve distanza da apparecchi analoghi ma con un numero maggiore di programmi. Un vantaggio offerto dagli *apparecchi analogici con controllo digitale* (figura 2), pur mantenendo una circuiteria di tipo analogico, è la possibilità di essere “controllati” da una sorgente esterna digitale (computer), che agisce direttamente sui dati registrati all’interno di una memoria programmabile, inserita nella protesi. Sono apparecchi acustici che permettono all’audioprotesista di intervenire su più controlli per cui risultano molto più flessibili e di conseguenza meglio adattabili alle esigenze del paziente, rispetto agli apparecchi analogici tradizionali.

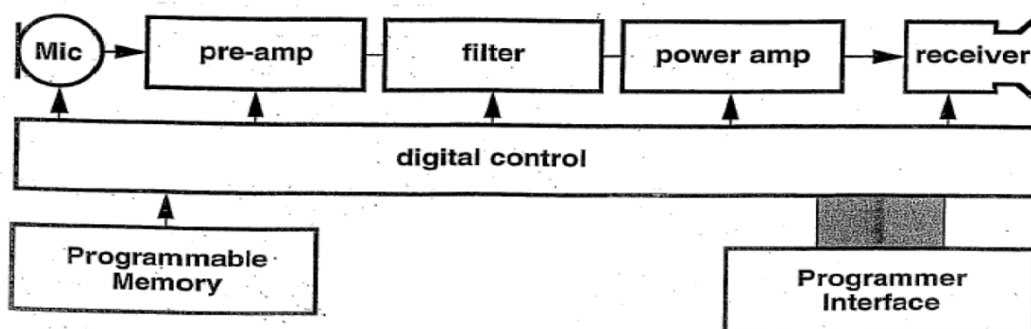


Figura 2: schema a blocchi di un apparecchio digitale con controllo analogico. La differenza con le protesi analogiche tradizionale sta nella possibilità di essere controllati digitalmente dall'audioprotesista attraverso una memoria programmabile.

Dal 1995 la tecnologia non fa che offrire miglioramenti sempre più accentuati: dal primo prodotto digitale a tre canali di amplificazione si passa ad apparecchi progressivamente più complessi, e si aggiunge la possibilità di gestire uno o due microfoni in maniera manuale, semi automatica o totalmente automatica. Questa era tecnologica è caratterizzata dal fatto che il guadagno, l'uscita massima e la risposta in frequenza dell'apparecchio acustico sono gestiti da microprocessori, col fine di elaborare il segnale sonoro in modo tale da soddisfare i requisiti indicati dall'audiogramma del paziente. Le impostazioni volute delle caratteristiche elettroacustiche dell'apparecchio acustico si ottengono collegando il computer all'apparecchio stesso sfruttando dei programmi specifici che utilizzano vari sistemi di adattamento.

Negli *apparecchi acustici digitali* (figura 3), caratterizzati da dimensioni miniaturizzate, sia il processing dei segnali audio sia il controllo del processing stesso è realizzato attraverso tecniche digitali. Il segnale sonoro viene infatti campionato e poi convertito in forma binaria. In questo modo è possibile trattare il segnale comodamente tramite specifici algoritmi di elaborazione digitale. All'uscita del complesso di elaborazione, il segnale è ancora una sequenza numerica, diversa da quella originaria, che successivamente dovrà essere riconvertita dal ricevitore in un segnale analogico di pressione sonora. Così facendo si è reso più semplice analizzare i vari tipi di suoni entranti nel microfono dell'apparecchio, separando in maniera specifica le componenti del rumore da quelle del parlato, riuscendo di conseguenza a dare un comfort superiore all'utente ipoacusico.

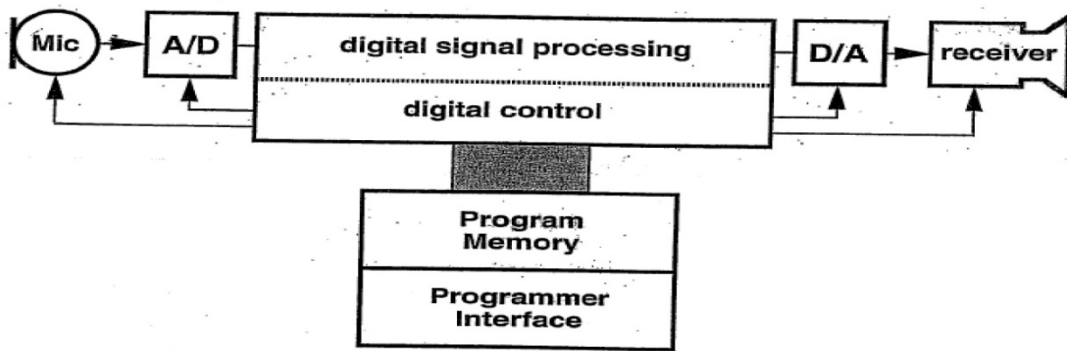


Figura 3: schema a blocchi di un apparecchio acustico digitale, caratterizzato da un microfono, un ADC, un processore digitale, un DAC e un ricevitore, componenti sotto il diretto controllo di un minicalcolatore.

L'innovazione tecnologica rappresentata dall'approccio digitale presenta innumerevoli vantaggi rispetto ad un approccio analogico tradizionale, i quali possono essere classificati secondo tre ampie categorie:

- 1) capacità affini di elaborazione del segnale, ma superiori, a quelle offerte dagli apparecchi acustici convenzionali;
- 2) capacità di elaborazione del segnale uniche dei sistemi digitali, che non possono essere implementate in apparecchi acustici analogici;
- 3) metodi di elaborazione e controllo del segnale che modificano il nostro modo di pensare a come gli apparecchi acustici dovrebbero essere progettati, prescritti e adattati.

Il terzo tipo di vantaggi è il più delicato e dunque il più importante. Un apparecchio acustico digitale può essere non solo programmato per amplificare, ma anche per generare segnali audio. In quanto tale, lo strumento può venire programmato per funzionare come audiometro, in modo da facilitare il rilevamento delle caratteristiche audiologiche fondamentali per una successiva determinazione dei parametri prescrittivi della protesi. Usando questo approccio, è possibile eludere il difficile problema della correzione dell'intensità del suono, modifica necessaria a causa delle differenze frequenza-dipendenti tra le cuffie usate nell'audiometria tradizionale e il ricevitore dell'apparecchio acustico del paziente.

Un'altra peculiarità dei sistemi digitali riguarda la possibilità di utilizzare simulazioni al computer: lo sviluppo di tecniche di simulazioni computerizzate per facilitare il progetto e lo sviluppo di vocoder e di altri sistemi di *speech-processing* ha generato l'idea geniale di

utilizzare i calcolatori anche per simulare gli apparecchi acustici. Dal momento che questo è risultato possibile, si è pensato di utilizzare i sistemi digitali di *signal-processing* direttamente all'interno dei dispositivi stessi.

Per essere più chiari, i vantaggi che offrono gli apparecchi acustici digitali rispetto a quelli convenzionali possono essere così elencati:

- maggiore flessibilità di funzionamento;
- migliore gestione del guadagno e della compressione dei segnali d'ingresso su un numero maggiore di canali per una migliore definizione della curva di risposta;
- cancellazione automatica e/o adattiva del feedback;
- riduzione del rumore ed intensificazione del parlato con notevoli incrementi del SNR (*signal-to-noise-ratio*);
- riduzione del rumore interno generato dal microfono per mezzo dell'espansione del segnale d'ingresso;
- possibilità di gestire vari tipi di direzionalità con differenti configurazioni di funzionamento del microfono (o più microfoni);
- possibilità di gestire differenti programmi indipendenti tra loro;
- possibilità di eseguire prove *in situ* per un più preciso adattamento alle reali necessità dell'utente.

Tutto questo per garantire alte prestazioni ed un ascolto estremamente confortevole, andando incontro in maniera quasi ottimale alle diverse esigenze dell'ipoacusico. I progressi della tecnologia conducono alla creazione di circuiti elettronici miniaturizzati all'estremo, capaci di svolgere ormai le milioni di funzioni di un computer.



## 1. Classificazione degli apparecchi acustici

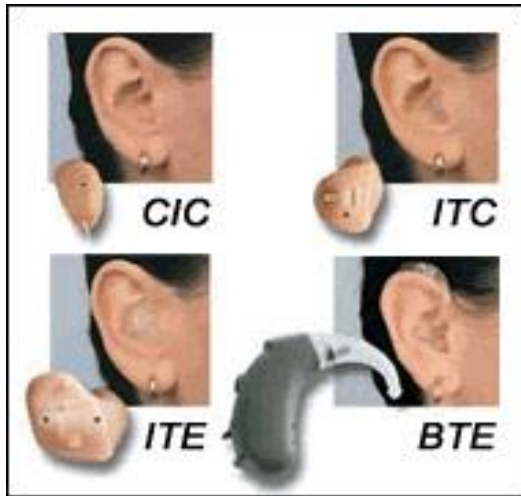


Figura 4: tipologie di apparecchi acustici in base al posizionamento sull'orecchio: *completely-in-the-canal* (CIC), *in-the-canal* (ITC), *in-the-ear* (ITE) e *behind-the-ear* (BTE)

Oltre ad un'ampia classificazione basata sul tipo di componenti interni e di elaborazione del segnale, è possibile raggruppare gli apparecchi acustici anche secondo il tipo di posizionamento nell'orecchio. Le principali modalità di costruzione di un apparecchio acustico prevedono l'inserimento completo della protesi nel condotto uditivo (CIC) o parziale (ITC), il posizionamento dell'apparecchio nella conca (ITE) o sulla pinna (BTE).

Il modello *completely-in-the-canal* (CIC) (fig. 4), posizionato totalmente all'interno del condotto uditivo, è quello di dimensioni minori (quasi invisibile) e che riproduce il suono nel modo più naturale possibile. Grazie alla profondità di inserzione, questo tipo di design offre numerosi benefici al portatore: è richiesto infatti un minor rapporto guadagno/uscita, offre un guadagno maggiore alle alte frequenze, consente la riduzione o addirittura l'eliminazione dell'effetto di occlusione, una buona stabilità di posizionamento all'interno del condotto, prestazioni migliori in ambienti rumorosi, riduzione del feedback, possibilità di utilizzare normalmente il telefono e, infine ma non meno importante, un'estetica migliore.

Un apparecchio di questo tipo può tuttavia essere prescritto solamente a soggetti affetti da ipoacusia lieve o moderata, in quanto non esisterebbe lo spazio materiale per contenere i componenti elettronici necessari a fornire un'amplificazione maggiore.

I modelli *in-the-canal* (ITC) presentano due versioni di costruzione: perimeatali e peritimpanici.

Usando componenti miniaturizzati, gli apparecchi acustici perimeatali hanno la maggior parte dei componenti posizionati all'interno della conca e alcuni nella porzione cartilaginea del condotto uditivo, chiamata appunto meato acustico. L'apertura del microfono è

posizionata nella porzione esterna della conca. A causa della loro dimensione e della posizione, questo tipo di dispositivi non ne permette un ampio uso ma forniscono tuttavia alcuni vantaggi nel guadagno alle alte frequenze dovuti alla risonanza acustica nella porzione vuota di conca e alla profondità di inserzione nel condotto uditivo.

Gli apparecchi peritimpanici racchiudono tutti i loro componenti all'interno del condotto uditivo esterno e terminano in prossimità della membrana timpanica: possono essere quindi assimilati alla categoria *completely-in-the-canal* precedentemente descritta.

I modelli *in-the-ear* (ITE) sono costituiti da un guscio realizzato sulla base dell'impronta auricolare del paziente, all'interno del quale vengono inseriti i componenti. Questa tipologia di protesi può essere di tipo *full concha*, *low profile* o *half concha*, in base al posizionamento fisico e alle dimensioni all'interno della conca.

Il tipo *full concha* è il tipo più comunemente impiegato. L'utilizzo dell'intero volume della conca fornisce il massimo spazio e permette quindi la massima flessibilità nella costruzione dell'apparecchio. Le sue dimensioni e la sua forma consentono di ridurre i problemi di feedback e di mantenerlo al sicuro all'interno dell'orecchio. La tipologia *low profile* presenta la stessa configurazione del *full concha* ma sporge in maniera minore dalla conca. Di conseguenza è disponibile meno spazio al suo interno, riducendo quindi la sua abilità a contenere una complessa circuiteria. Infine, la versione *half concha* occupa solo una piccola parte della conca, riducendo così ulteriormente le sue capacità elettroniche.

I modelli di tipo *behind-the-ear* (BTE) sono progettati per essere posizionati dietro la pinna e sono generalmente indicati per sopperire ad una perdita uditiva severa o profonda in quanto, massimizzando la distanza tra microfono e ricevitore, viene minimizzato il segnale di feedback acustico.

Il feedback acustico, uno dei principali problemi riscontrati dalla quasi totalità dei portatori, è il risultato di una parte di segnale in uscita dal ricevitore (quindi elaborato ed amplificato) che viene ricaptato dal microfono e dunque rielaborato. Come conseguenza il soggetto sentirà un fastidioso fischio ad alta frequenza, di intensità crescente con l'ipoacusia del portatore in quanto sarà maggiore anche il guadagno necessario.

È stato dimostrato che l'intensità del feedback acustico è inversamente proporzionale alla distanza tra microfono e ricevitore, per cui una protesi del tipo BTE, essendo di dimensioni

maggiori, riduce l'effetto del feedback acustico e dispone dello spazio necessario per un'amplificazione maggiore.

Per concludere, esistono anche altre soluzioni tecnologiche, seppur di scarso utilizzo, come apparecchi acustici integrati a occhiali da vista e protesi *body-worn* (fig.5) costituite da un auricolare collegato tramite un tubicino a vuoto alla "scatola" contenente i componenti elettronici, da tenere in tasca.



Figura 5: a sinistra: apparecchio acustico di tipo *body-worn*, tascabile; a destra apparecchi acustici integrati con occhiali da vista.

Gli apparecchi acustici possono inoltre essere classificati in base al tipo di trasduzione dell'uscita: **conduzione aerea** o **conduzione ossea**. Entrambe le tipologie sono simili, ad esclusione dell'ultimo stadio di trasduzione del segnale amplificato, il quale deve essere convertito in stimoli utilizzabili dal portatore.

Le *protesi a conduzione aerea* sono progettate in modo da riconvertire l'energia elettrica in energia acustica e indirizzarla verso il condotto uditivo; la maggioranza degli apparecchi acustici sono di questo tipo.

Gli *apparecchi a conduzione ossea* sono state invece ideate per trasformare l'energia elettrica in vibrazioni meccaniche che, quando applicate direttamente alla testa, stimolano l'intero cranio. L'uso dei dispositivi a conduzione ossea è limitato alle situazioni in cui esiste un'ampia differenza (30-50 dB) tra perdita uditiva aerea e ossea o quando è presente una scarica cronica da un orecchio che impedisce l'utilizzo di un ricevitore a trasmissione aerea. Tuttavia, anche in situazioni del primo tipo, è preferibile l'uso di una protesi a conduzione aerea molto potente rispetto ad un dispositivo a conduzione ossea. Questi

ultimi sono infatti disponibili in un numero limitato di modelli: BTE, *body-worn* e occhiali da vista integrati a protesi acustica.

In definitiva, la scelta del modello di apparecchio acustico dovrà trovare un compromesso tra le preferenze del soggetto, le sue caratteristiche anatomiche e audiometriche.

## 2. Apparecchi acustici digitali

Un apparecchio acustico digitale moderno è basicamente costituito da un microfono direzionale, un banco di filtri distribuito su un numero variabile di canali (da 2 per le protesi meno recenti a 20 per i sistemi di nuova generazione), un sistema di controllo automatico del guadagno (AGC), un sistema di amplificazione e un ricevitore. È inoltre necessaria la presenza di una batteria che fornisca la corrente di alimentazione, di un foro che permetta la ventilazione ed eviti fenomeni di occlusione e di alcuni sistemi di controllo come il controllo automatico del guadagno in ingresso, il controllo di volume e il controllo del guadagno in uscita.

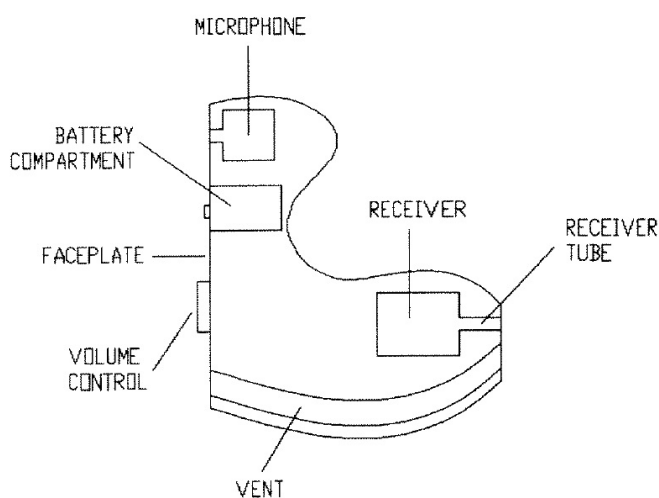


Figura 6: sezione trasversale di un apparecchio acustico ITE.

In figura 6 è rappresentata la sezione trasversale di un apparecchio acustico di tipo ITE. Il microfono è posizionato all'apice del dispositivo, al di sopra del controllo del volume e del compartimento contenente la batteria. La presa d'aria per la ventilazione (*vent*) si trova alla base dell'apparecchio, scorre dalla piastra frontale (*faceplate*) fino al condotto uditivo e può assumere diverse

forme, anche se un tubicino avente sezione trasversale circolare sia la soluzione che meglio si adatta alla geometria auricolare. Il ricevitore è posizionato nella porzione del dispositivo che rimane nel condotto uditivo e la sua uscita è connessa al canale per mezzo di un corto tubicino. Infine, il ricevitore viene posizionato in modo tale da massimizzare la distanza dal microfono, per minimizzare l'inconveniente del feedback acustico.

## 2.1. Principio di funzionamento e componenti interni

Di primaria importanza per il funzionamento di un apparecchio acustico è la presenza di trasduttori, ossia dispositivi che convertono il segnale in ingresso in un segnale di forma differente: questi sono il microfono e il ricevitore.

Il trasduttore d'ingresso, costituito dal **microfono**, converte il segnale acustico, sotto forma di pressione sonora, incidente sul suo diaframma in segnali elettrici analogici.

Negli apparecchi acustici moderni, il microfono utilizzato su scala praticamente universale

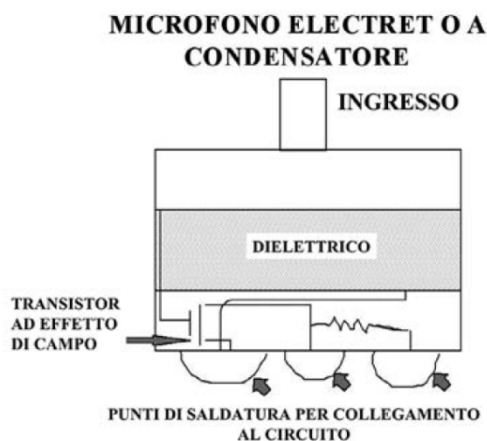


Figura 7: schema di un microfono *electret*.

è quello di tipo *electret* o a condensatore, grazie alla sua alta sensibilità, alla sua eccellente risposta in frequenza a banda larga, con conseguente ottima qualità del suono riprodotto, alla sua affidabilità e insensibilità alle vibrazioni meccaniche.

Il microfono *electret* è rappresentato da due membrane, delle quali una è fissa e l'altra è libera di vibrare in sincronia col suono incidente, come si osserva in figura 7. Le due

membrane sono affacciate a costituire le superfici di un condensatore il quale presenta una capacità  $C$  proporzionale all'area  $A$  delle facce del condensatore, alla distanza  $d$  fra di esse e al tipo di dielettrico fra esse interposto ( $C = \epsilon_0 \epsilon_r A / d$ ). Partendo dal presupposto che il dielettrico non può essere modificato e che la superficie delle membrane è fissa, risulta chiaro che la variazione di potenziale viene ottenuta in seguito alle variazioni della distanza tra le membrane, dovute alle vibrazioni della membrana mobile. Questa piccola tensione elettrica viene poi amplificata da un transistor ad effetto di campo (FET, *field effect transistor*) posizionato all'interno del microfono e poi inviata ai terminali d'ingresso dell'amplificatore principale.

I microfoni *electret* permettono di plasmare la risposta in frequenza fornendo un'ampia varietà di frequenze di taglio, con l'obiettivo un'attenuazione selettiva degli ingressi a bassa frequenza e sono lo strumento principale utilizzato per alterare la risposta in frequenza degli apparecchi acustici.

I microfoni moderni possono essere di tipo omnidirezionale o direzionale.

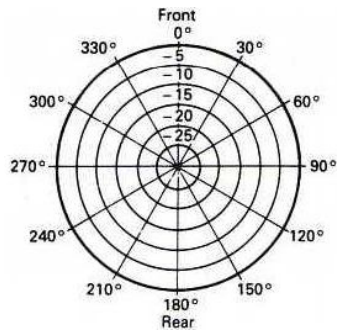


Figura 8: diagramma polare di un microfono omnidirezionale

I primi sono i più semplici: presentano un'unica apertura che indirizza il suono frontalmente verso il diaframma del microfono. Se testato a campo libero, il microfono omnidirezionale rileva il suono proveniente da tutte le direzioni nello stesso modo ed è perciò caratterizzato da un pattern polare di selettività circolare (figura 8).

Il microfono direzionale ha invece la particolarità di avere due ingressi per il segnale sonoro: una anteriore, come di norma, e una posteriore, che costringe il segnale a percorrere una sorta di labirinto all'interno del microfono. Il segnale che entra da questo ingresso è costretto a giungere sulla membrana vibrante con un certo ritardo rispetto al segnale che entra nella parte anteriore. Ciò fa sì che i due segnali sonori, incidendo in tempi diversi sulla membrana, causino uno sfasamento del segnale, utile a sopprimere i rumori più gravi, determinando

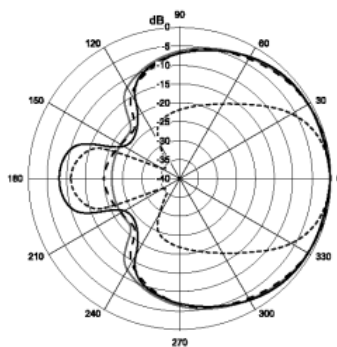


Figura 9: diagramma polare di un microfono direzionale di tipo cardioide.

una miglior intelligibilità del segnale e consentendo di individuare la provenienza del suono, di estrema importanza nel caso di applicazione binaurale. Ciascun microfono direzionale avrà un pattern polare di selettività differente, tra cui il più usato è quello a cardioide (figura 9): tale tipologia di microfono ha sensibilità massima ai suoni provenienti dalla direzione frontale, mentre è solo lievemente sensibile ai suoni provenienti da direzioni posteriori.

Da alcuni anni sono presenti sul mercato anche apparecchi con due microfoni, con l'obiettivo di migliorare significativamente la capacità di discriminazione dello *speech* anche in presenza di rumore di fondo. In questo caso la direzionalità è ottenuta sottraendo il segnale captato dal microfono posteriore da quello frontale, orientando la direzione della risposta nulla per mezzo dell'aggiunta di un ritardo al microfono posteriore. Con questo procedimento, la risposta in frequenza frontale presenta una caratteristica passa-alto, con pendenza di 6dB/ottava che rende il suono leggermente metallico e inaudibili i segnali a bassa frequenza. Inoltre, a causa del posizionamento dei due microfoni, la direzionalità viene esaltata anche grazie al cosiddetto effetto *headshadow*.

Il miglioramento nell'intelligibilità è di circa 5dB per un segnale di *speech* frontale all'ascoltatore e un rumore di sottofondo diffuso, il che sta ad indicare che il *signal-to-noise ratio* (SNR) è di 5dB inferiore a quello necessario per ottenere lo stesso livello di intelligibilità con un microfono omnidirezionale.

La direzionalità diventa tuttavia meno efficace quando sono presenti differenze nella fase e nell'ampiezza dei due microfoni e l'effetto di tale disaccoppiamento cresce al diminuire della distanza tra i due microfoni. Da qui la necessità di un perfetto accoppiamento tra i due dispositivi. La sostituzione del microfono omnidirezionale posteriore con uno di tipo direzionale consente inoltre un'efficace riduzione del rumore di sottofondo.

Passiamo ora al trasduttore d'uscita, rappresentato dal **ricevitore**. Tale dispositivo svolge basicamente la funzione di un altoparlante che converte i segnali elettrici in uscita dall'amplificatore finale in un segnale acustico o vibratorio. I ricevitori, come già detto, possono infatti essere a conduzione aerea o, più raramente, a conduzione ossea. I primi sono generalmente di tipo magnetico e ben si adattano a funzionare con amplificatori a bassa tensione di alimentazione, in quanto vengono azionati direttamente dallo stadio d'uscita. Gli avvolgimenti dei ricevitori magnetici possono essere realizzati in modo da avere l'impedenza elettrica desiderata per diversi tipi di amplificatori.

I ricevitori per apparecchi acustici di tipo ITE, ITC, BTE ed *eyeglass* sono solitamente di tipo magnetico ad armature bilanciate grazie alle elevate prestazioni che possono essere raggiunte in uno spazio incredibilmente piccolo.

Un'armatura di metallo magneticamente permeabile viene resa un elettromagnete per mezzo di un segnale di corrente alternata che scorre all'interno della bobina avvolta intorno ad essa. L'estremità libera dell'armatura viene alternativamente attratta o allontanata dai piccoli magneti permanenti; la vibrazione risultante verrà poi trasmessa, attraverso una barretta metallica, al diaframma e il suono verrà poi sviluppato nella cavità adiacente, da cui è trasmesso all'orecchio attraverso un apposito sistema di accoppiamento. L'output acustico della membrana del ricevitore corrisponde all'input originale nella forma d'onda, ma risulta intensificato grazie allo stadio amplificatore dell'apparecchio acustico. I ricevitori più piccoli sono caratterizzati da un ampio range e possono amplificare fino a oltre 6000Hz. Nonostante ciò sia considerato da molti un fattore positivo per una migliore riproduzione del suono e discriminazione dello *speech*, questi ricevitori ad ampio range sono meno sensibili e maggiormente soggetti al fenomeno del feedback acustico. Tuttavia,

grazie alle loro dimensioni miniaturizzate, ottenute principalmente attraverso la realizzazione di circuiti integrati contenenti anche lo stadio amplificatore d'uscita, sono quelli più largamente usati negli apparecchi di tipo *in-the-ear* e *in-the-canal*.

Entrando ora nel cuore di un apparecchio acustico digitale, è possibile descrivere il procedimento di elaborazione del suono in fasi sequenziali, rappresentate schematicamente in figura 10.

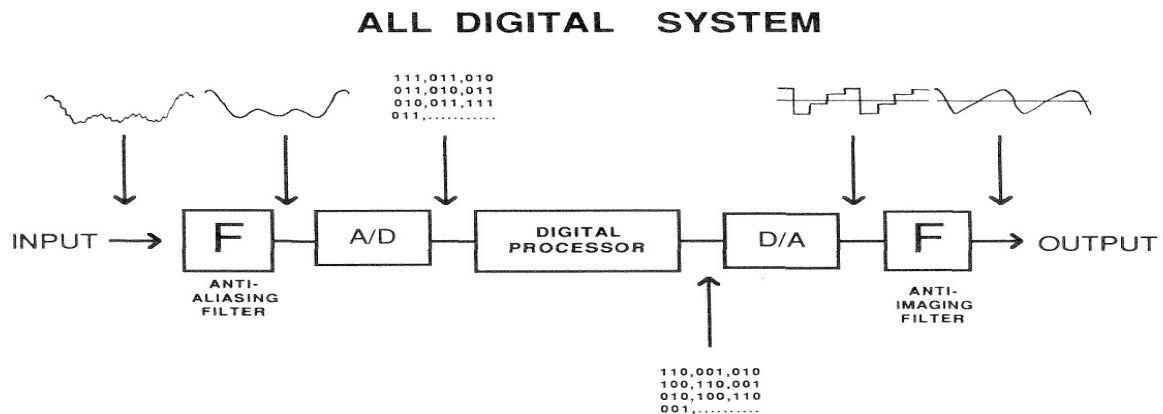


Figura 10: rappresentazione schematica delle fasi di speech processing che avvengono all'interno di un apparecchio acustico digitale. Il segnale d'ingresso subisce l'azione di un primo filtro *anti-aliasing*, viene convertito in forma digitale, processato, riconvertito in forma analogica e, dopo l'azione di un filtro *anti-imaging*, è pronto per essere inviato ad un sistema di trasduzione di pressione sonora.

Il segnale sonoro incidente sul diaframma del microfono, sotto forma di onda di pressione, viene convertito in un segnale di tipo elettrico, al quale viene applicato un filtro *anti-aliasing* immediatamente prima dell'operazione di campionamento. Un filtro anti-aliasing è tipicamente un filtro passa-basso che permette la totale attenuazione delle frequenze superiori alla frequenza di taglio  $f_c$ , la quale, secondo il teorema di Shannon, deve essere minore della metà della frequenza di campionamento  $f_s$ , ossia

$$f_c < \frac{1}{2} f_s$$

Come mostra la figura 11, le principali frequenze del parlato sono tipicamente comprese tra 125 e 8000Hz, mentre una classica frequenza di campionamento in problemi di *speech-processing* è di 20kHz. Utilizzando a titolo di esempio una frequenza di taglio pari a



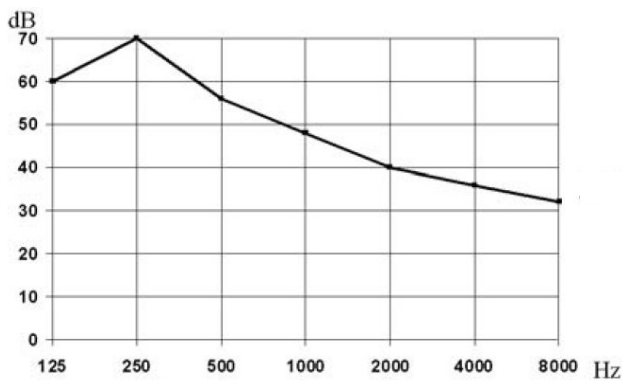


Figura 11: intensità delle principali frequenze del parlato, comprese tipicamente tra 125 e 8000 Hz.

10kHz, è possibile eliminare segnali a frequenza superiore, non appartenenti al segnale di speech ma ad esso sovrapposti in quanto captati dal microfono.

Il segnale così filtrato viene dunque digitalizzato, ossia convertito in forma

binaria, e costituirà l'ingresso del processore digitale, all'interno del quale

viene filtrato, eventualmente compresso e opportunamente amplificato.

Analogamente a quanto avviene per l'*anti-aliasing*, per evitare la generazione di forme d'onda spurie nel momento in cui la sequenza di dati campionata viene riconvertita in forma analogica, una pratica comune è quella di usare un filtro *anti-imaging* prima della conversione da digitale ad analogico. Generalmente, un filtro di questo tipo presenta le stesse caratteristiche del filtro passa-basso *anti-aliasing* corrispondente.

A questo punto, il segnale elettrico risultante verrà sottoposto ad una fase di *clipping* o di compressione per evitare la generazione di un'uscita sonora troppo intensa la quale, dopo essere stata riconvertita dal ricevitore in forma d'onda pressoria, andrà ad incidere sul timpano del portatore.

### 2.1.1. Processore digitale

Il processore digitale rappresenta il nucleo centrale di una protesi acustica moderna, in cui avvengono tutte le fasi di elaborazione del segnale in uscita dal filtro anti-aliasing. Un processore moderno per apparecchi acustici digitali è tipicamente costituito da un numero variabile di canali indipendenti (generalmente da 12 a 24) in cui il segnale elettrico viene frammentato e diversamente elaborato.

Ad eccezione del primo canale, contenente un filtro passa-alto, e dell'ultimo, contenente un passa-basso, ogni canale è costituito da un filtro passa-banda, da un sistema di compressione e da un amplificatore: questo approccio permette di scomporre il segnale in un numero prestabilito di bande di frequenza, in modo tale da poterlo elaborare in modo differenziato e da fornire un'amplificazione specifica per ciascun range di frequenza.

L'importanza di questo procedimento sta nella diversa sensibilità dell'orecchio umano, sia normoudente sia ipoacusico, alle frequenze sonore.

Il fine ultimo di un apparecchio acustico è quello di amplificare i suoni fino ad un'intensità che risulti ben comprensibile dal portatore, ma tale da non superare la soglia di fastidio del portatore e da degradare il meno possibile la sua intelligibilità. Questo può avvenire solamente in presenza di un sistema di elaborazione del suono *multi-channel*: scomponendo infatti lo speech in varie bande di frequenza, è possibile trattare ognuna di queste in modo diverso in base alla perdita uditiva del paziente risultante dall'audiogramma tonale, al suo range dinamico, alla soglia di fastidio e a tutti gli altri parametri che emergono dai test psicoacustici eseguiti sui soggetti ipoacusici.

L'output di ogni canale è rappresentato da un segnale caratterizzato da una ristretta banda di frequenza, che andrà sommato alle uscite degli altri canali, in modo tale da ricostruire il segnale di *speech* che verrà udito dal portatore.

#### 2.1.2. Banco di filtri

I filtri digitali impiegati nel processore agiscono su forme d'onda campionate all'incirca nello stesso modo (ma non *esattamente* nello stesso modo) in cui un filtro elettronico opera su un segnale elettrico e questa operazione di filtraggio è analoga all'esecuzione della media aritmetica di una sequenza di dati.

Un importante vantaggio del filtraggio digitale rispetto a quello di tipo elettronico convenzionale è la possibilità di ottenere una maggiore precisione, che può essere vari ordini di grandezza più grande rispetto a quella ottenibile nella pratica mediante l'uso di componenti analogiche. Per esempio, non è difficile realizzare un filtro digitale con una pendenza molto elevata, come 100dB/ottava, il che sarebbe quasi impossibile utilizzando componenti elettronici.

Un secondo importante vantaggio è dato dal fatto che i filtri digitali possono essere riprogrammati, senza la necessità di modifiche nell'hardware, al fine di avere caratteristiche ampiamente differenti; la riprogrammazione, in aggiunta, può aver luogo in poche frazioni di secondo.

Un terzo vantaggio fondamentale dei filtri digitali è rappresentato dal fatto che questi possono essere programmati per includere operazioni logiche, in modo da modificare le proprie caratteristiche in risposta a eventi del segnale preselezionati.

Le frequenze centrali dei filtri passa-banda di ciascun canale del processore sono selezionate in modo tale da essere equispaziate nella scala di *bark*. Il bark costituisce l'unità di misura della scala uditiva a bande critiche, attualmente considerata la più efficace per la classificazione delle vocali. La natura di tale scala uditiva è essenzialmente fisiologica in quanto la coclea, il principale organo deputato alla trasduzione meccano-elettrica del suono, presenta una distribuzione tonotopica, ossia dipendente dalla frequenza. Ciò vuol dire che alla base della coclea vengono codificati suoni a frequenza più elevata mentre all'apice si ha una maggiore sensibilità ai suoni a bassa frequenza. Più precisamente, il nostro sistema uditivo analizza i suoni come successioni in parallelo di più filtri passa-banda. Tali filtri non presentano una larghezza di banda costante, bensì variabile a seconda della frequenza centrale occupata dalla banda medesima: essa sarà più stretta nelle fasce di frequenza più gravi e più ampia per le frequenze più acute. La larghezza di banda di ogni filtro corrisponde per l'appunto ad un bark, ovvero ad una banda critica.

Anche la perdita uditiva neurosensoriale è caratterizzata da una diminuita sensibilità che varia in funzione della frequenza: è per questo motivo che le prove audiometriche vengono eseguite per diverse frequenze e diversi valori di intensità dello stimolo sonoro.

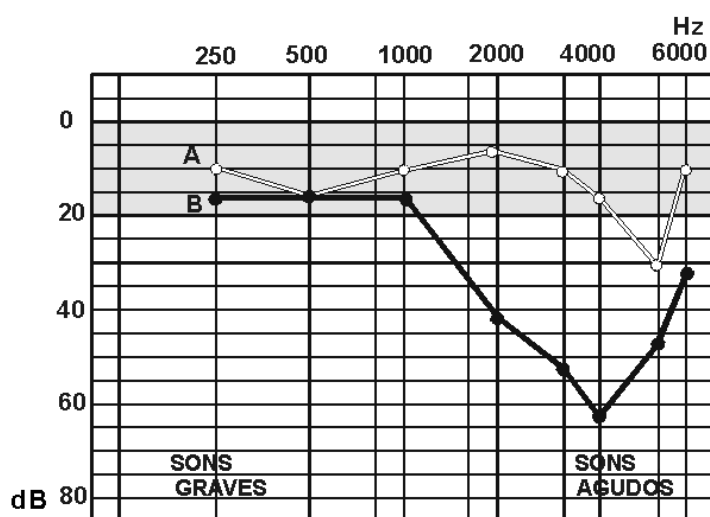


Figura 12: esempio di audiogramma tonale binaurale.

A titolo esemplificativo, è stata riportata l'audiometria tonale (fig. 12) di un soggetto che presenta, per un solo orecchio (linea nera), un'ipoacusia moderata per frequenze elevate, mentre, per suoni gravi, è presente solo una lievissima perdita. Dato che l'obiettivo di un apparecchio

acustico è quello di ristabilire la *loudness* a livelli normali, il guadagno fornito dovrà dipendere dall'intensità e dalla frequenza del segnale. Da qui nasce l'esigenza di filtrare il segnale in specifiche bande di frequenza e di applicare un diverso guadagno al segnale in ogni banda.

A differenza dei primi apparecchi acustici digitali che scomponavano semplicemente il segnale in un paio di bande, i dispositivi moderni mirano invece a filtrare il segnale sonoro similmente a quanto avviene in una coclea sana: le frequenze centrali dei filtri sono quindi tipicamente equispaziate al di sotto dei 2000Hz mentre per frequenze superiori la *bandwidth* cresce con andamento logaritmico. Quest'ultima non ha un valore fisso per tutti gli apparecchi acustici usati, ma varia in funzione del numero di canali del dispositivo (generalmente 12,16, 20 o 24).

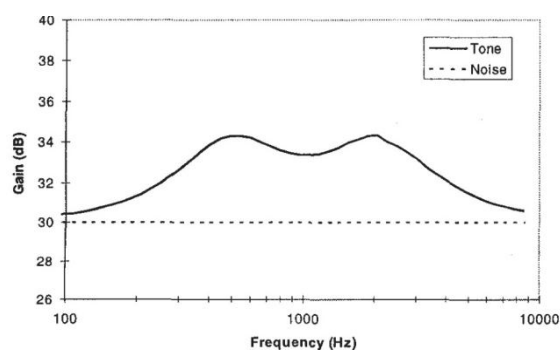


Figura 13: risposta del guadagno di un processore a tre bande.

Un inconveniente di questo sistema è illustrato in figura 13, la quale mostra la risposta del guadagno di un processore a tre bande progettato per fornire uguale guadagno a tutte le frequenze. Vengono usate funzioni I/O identiche per ogni banda, in modo che debba essere applicato lo stesso guadagno per tutte le frequenze del segnale. A causa delle

pendenze dei filtri usati, il processamento non lineare vicino alle frequenze di crossover tra due bande fornisce maggiore guadagno per segnali a banda stretta rispetto a segnali a banda larga e viola i requisiti di progetto di una funzione di guadagno piatta. Questo effetto peggiora all'aumentare del numero di bande, dal momento che aumenta anche il numero di regioni di crossover. Potrebbero inoltre sorgere ulteriori problemi all'avvicinarsi della *bandwidth* alla spaziatura in frequenza del segnale dato che le armoniche potrebbero cadere vicino alle regioni di transizione tra bande e il guadagno applicato ad ogni armonica dipenderà da quante armoniche cadono all'interno di una data banda. Questo problema viene risolto da una sufficiente sovrapposizione delle bande, in modo che il guadagno ad ogni frequenza data sia determinato da una molteplicità di bande. Il filtraggio passa-banda può essere trattato come il campionamento dello spettro in potenza e quindi è richiesta una frequenza di campionamento minima per eliminare il fenomeno dell'aliasing nel dominio di autocorrelazione. Tramite la sovrapposizione delle bande, l'effetto di crossover non viene solamente ridotto ma anche gli inconvenienti risultanti, per esempio, da una variazione del guadagno in funzione del numero di armoniche all'interno di una banda viene drasticamente ridotto.

### 2.1.3. *Output – limiting* e compressione

Dopo l'operazione di filtraggio, quindi di selezione della banda di frequenza, è spesso necessario limitare il segnale all'interno di ogni canale e, successivamente, applicare un determinato guadagno al segnale così elaborato.

In linea generale, è consigliabile fornire un guadagno maggiore per segnali a bassa intensità e quindi difficilmente percepibili, e un guadagno minore a segnali ad alta intensità, che necessitano dunque di poca amplificazione affinché vengano uditi. Può tuttavia accadere che, talvolta, i segnali amplificati possano presentare un'intensità superiore alla soglia di fastidio del portatore dell'apparecchio acustico: nell'elaborazione del segnale sonoro è dunque necessaria una fase di limitazione dell'ampiezza.

Esistono basicamente due tecniche di *output-limiting*: il *peak-clipping* e la compressione.

Il *peak-clipping* è la tecnica più semplice di *output-limiting*, che veniva impiegata



Figura 14: tecnica del *peak-clipping*. La parte di segnale al di sopra della soglia prestabilita viene letteralmente rimossa.

prevalentemente negli apparecchi acustici tradizionali. Può essere definita come la rimozione, tramite mezzi elettronici, di uno (*peak-clipping* asimmetrico) o entrambi (*peak-clipping* simmetrico) gli estremi del segnale che eccedono una soglia limite di intensità prefissata. Per chiarire meglio il concetto si

veda la figura 14: dopo aver stabilito una soglia massima di intensità, al segnale originale vengono letteralmente tagliate le parti che eccedono la soglia di clipping, le quali avranno, dopo il clipping, un'intensità pari alla soglia.

Naturalmente questo tipo di procedimento genera una forte distorsione armonica e di intermodulazione, che risultano in una diminuita intelligibilità dello speech udito e quindi in una bassa performance dell'apparecchio acustico.

Un metodo alternativo e più efficace del *peak-clipping* è costituito dalla cosiddetta compressione, che sta alla base del principio di funzionamento di un AGC (*Automatic Gain Control*), sistema di fondamentale importanza nei moderni apparecchi acustici digitali.

### Automatic Gain Control (AGC)

Questi sistemi possiedono un circuito di monitoraggio integrato che riduce automaticamente il guadagno in funzione dell'ampiezza del segnale da amplificare. I principali propositi di questi sistemi sono:

- 1) aggiustare il range dinamico del segnale in ingresso all'apparecchio acustico;
- 2) ridurre il guadagno del sistema all'aumentare dell'ingresso (misurato in dB SPL), in modo tale che la capacità d'uscita del sistema non ecceda il limite prestabilito e che la distorsione venga mantenuta bassa;
- 3) ridurre il range dinamico del segnale d'uscita.

Per soddisfare il primo requisito viene spesso inserito un AGC ad azione lenta, con tempi

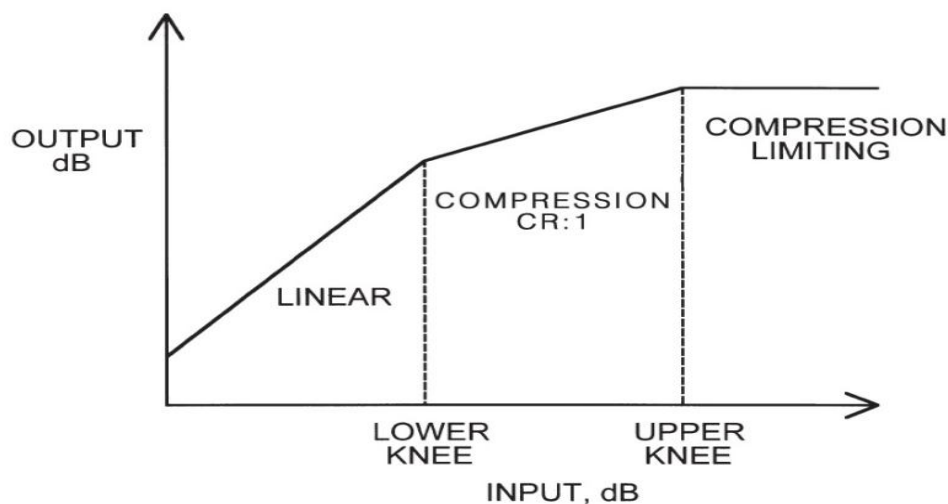


Figura 15: curva ingresso – uscita di un AGC, in scala logaritmica. Si osserva una fase lineare, una di compressione e una di limitazione dell'uscita ad un valore di limite di intensità sonora.

di attacco e recupero dell'ordine di 5 e 300 ms rispettivamente, tra il microfono e il banco di filtri, in modo tale da aumentare il range dinamico del soggetto ipoacusico, il quale è generalmente ridotto a causa di una minor ampiezza dell'intervallo soglia uditiva – soglia di fastidio rispetto ad un individuo normoudente.

Per rispondere alle esigenze definite al punto 2 e 3 vengono invece impiegati AGC ad azione rapida all'interno di ogni canale dell'apparecchio acustico, a valle del rispettivo filtro passabanda.

La curva input/output generalizzata di un AGC (figura 15) presenta tre componenti principali: una sezione lineare dove bassi valori di input (in dB SPL) causano incrementi uguali (lineari) dell'output (sempre in dB SPL); una sezione di compressione, dove incrementi dell'input causano minori incrementi dell'output, e una sezione in cui un aumento dell'ingresso non produce aumenti significativi dell'output.

In generale, un sistema AGC viene identificato dai seguenti parametri:

- **Limiting level:** livello di saturazione dell'output.
- **Ginocchio dell'AGC** (knee point): conosciuto anche come breakaway point, soglia di controllo, soglia di compressione o soglia dell'AGC. È il valore in dB dell'ingresso oltre il quale viene attivato il meccanismo di compressione.
- **Rapporto di compressione:** è dato dal rapporto tra la variazione dell'ingresso e la relativa variazione dell'uscita, misurate in dB SPL. È un indice della pendenza della curva compresa nella sezione di compressione.
- **Slew rate:** velocità di cambiamento del guadagno, in ms/dB, che il circuito può manipolare.
- **Costanti temporali:** sono i ritardi temporali, causati dal circuito di feedback, nella stabilizzazione ad un nuovo valore di guadagno. Il *tempo di attacco* si riferisce al tempo richiesto al circuito di feedback per impostare un nuovo guadagno a seguito di un forte segnale di ingresso; il *tempo di rilascio* è invece relativo al tempo necessario per tornare alla normale amplificazione dopo che il forte segnale d'ingresso non è più presente. Il tempo di rilascio deve essere più lungo del tempo di attacco per evitare un fenomeno chiamato "AGC flutter": se troppo veloce, l'azione di compressione seguirebbe l'ampiezza istantanea dei singoli cicli, introducendo così una forte distorsione della forma d'onda.

Esistono molte variazioni della curva generalizzata: per esempio, potrebbero essere presenti solo la sezione lineare e quella di saturazione, con un alto rapporto di compressione. Il punto di ginocchio è generalmente compreso tra 60 e 80 dB SPL, ma in alcune situazioni potrebbe anche trovarsi intorno ai 40 dB SPL, con un rapporto di compressione di 2:1 o 3:1 o, viceversa, esistono condizioni in cui è sufficiente un punto di ginocchio molto elevato, come 90 dB SPL.

Le caratteristiche input/output possono inoltre variare in base al range di frequenza: di fatto, ogni canale di un apparecchio acustico *multi-channel* presenta generalmente sistemi AGC con soglie e rapporti di compressione differenti.

#### 2.1.4. Controlli

In aggiunta ai componenti essenziali per il funzionamento delle protesi acustiche, tali dispositivi dispongono sempre di elementi aggiuntivi che ne perfezionano le prestazioni, adattandole alle esigenze e alle condizioni ambientali del portatore: i cosiddetti “controlli”.

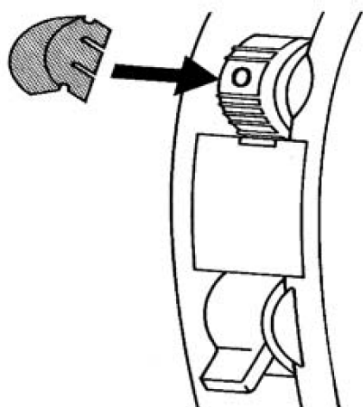


Figura 16: controllo del volume.

Di primaria importanza è il **controllo del volume**, che consente di variare il livello d’uscita dell’apparecchio acustico facendo giungere sull’amplificatore finale una percentuale maggiore o minore di segnale, determinando così la potenza erogata. Tale potenza, oltre ad essere direttamente proporzionale all’intensità del segnale amplificato, dipende anche dal livello del segnale in ingresso al microfono. In figura 16 è mostrato il “cursore” del potenziometro, ossia la rotella

meccanica che, nei potenziometri convenzionali, ha la funzione di prelevare il segnale originale ad una certa percentuale di rotazione, determinando anche, di conseguenza, la percentuale di segnale da inviare allo stadio finale.

Con l’avvento dei circuiti ibridi, anche il controllo del volume può subire profondi mutamenti ed essere sostituito da funzioni elettroniche digitali che lo sostituiscono completamente.

Lo stesso avviene anche per i “trimmer”, ossia dei potenziometri semifissi il cui accesso è consentito solo al tecnico audioprotesista: tali componenti possono essere sostituiti da circuiti digitali, con il vantaggio di occupare meno spazio di un trimmer convenzionale e di essere programmabili dall’esterno. La loro funzione è quella di aggiustare ulteriormente il guadagno della protesi. Quando impostato a valori bassi, questo controllo può essere utilizzato per mantenere il guadagno dello strumento al di sotto del *feedback*, anche quando l’utente attiva al massimo il volume.



Esistono poi i **controlli elettronici di tono**, che alterano la morfologia della risposta in frequenza dell'apparecchio acustico, per portarla a meglio rispondere alle esigenze di ascolto del paziente. Le variazioni della risposta in frequenza avvengono in impostazioni discrete ottenute tramite interruttori o l'utilizzo di strumenti meccanici come cacciaviti. La rete di filtri utilizzata a questo scopo spazia da semplici filtri passivi del prim'ordine agenti sulle frequenze acute a filtri attivi di ordine maggiore che permettono una maggiore soppressione delle basse frequenze o la soppressione di qualsiasi banda di un apparecchio *multi-channel*.

Altri controlli delle protesi acustiche sono costituiti dalla preregolazione del guadagno massimo, dai controlli sul ginocchio degli AGC e da circuiti volti a ridurre l'effetto del *feedback* acustico.

### **3. Fenomeni che degradano la performance di un apparecchio acustico**

Nonostante i numerosi vantaggi offerti dalle protesi acustiche a soggetti audiolesi, è comunque doveroso ricordare che tali dispositivi, agendo direttamente sul contenuto spettrale dei segnali vocali, possono introdurre distorsioni armoniche e produrre effetti negativi, che si traducono in un deterioramento della performance dell'apparecchio acustico stesso. In questo paragrafo verranno brevemente illustrati i principali svantaggi derivanti dall'utilizzo delle protesi acustiche e i metodi più usati per la loro riduzione.

#### 3.1. Problema del range dinamico

I soggetti audio-lesionati con *loudness recruitment* presentano un range dinamico ridotto tra la soglia uditiva e quella di fastidio. Inoltre, il segnale di *speech* che viene presentato appena al di sopra della soglia non è generalmente intelligibile. Come conseguenza, il range dinamico effettivo per il parlato (dalla soglia di intelligibilità a quella di fastidio) è minore di quello dei toni.

Il range dinamico potrebbe inoltre variare notevolmente in funzione della frequenza. Gli apparecchi acustici dovrebbero processare il parlato in modo tale che tutti gli elementi importanti del segnale vocale siano sopra-soglia, ma si trovino comunque al di sotto del livello che provoca fastidio. Nonostante questo requisito appaia semplice, non è facilmente

realizzabile a causa della complessità con cui varia nel tempo lo spettro a breve termine dello *speech*.

Esistono due ragioni per cui un range dinamico ridotto crea difficoltà al soggetto ipoacusico.

Prima di tutto, l'intensità globale dello *speech* potrebbe variare in un range di 30 dB da una situazione ad un'altra: per far fronte a questo problema, può essere utilizzato, prima che il segnale venga filtrato nei vari canali del processore, un AGC ad azione lenta, caratterizzato per esempio da una soglia di compressione di 75 dB applicata bruscamente sul segnale (*compression limiting*) e da tempi di attacco e recupero di 5 e 300 ms, rispettivamente. Un amplificatore di questo tipo ha dunque la funzione di compensare variazioni nel livello globale del parlato da una situazione ad un'altra, mandando il segnale di *speech* ad un livello confortevole, senza riguardo del livello di ingresso.

In secondo luogo, anche per uno *speech* ad un'intensità media costante, i livelli degli elementi acustici individuali del parlato potrebbero variare in un range di 30 dB. In generale, le correlazioni acustiche tra consonanti sono meno intense di quelle delle vocali, così che un soggetto ipoacusico potrebbe essere in grado di udire le vocali ma non le consonanti. È dunque necessaria in questo caso la cosiddetta "compressione sillabica" realizzata dagli AGC ad azione rapida, agenti preferibilmente in modo indipendente nelle diverse bande di frequenza, in modo da permettere la variazione del range dinamico del paziente in funzione della frequenza.

Un'ulteriore tecnica per evitare i fastidi derivanti da un ridotto range dinamico è quindi quella di scomporre il segnale in varie bande di frequenza, pratica ormai di uso comune nelle protesi acustiche moderne. Questo offre la possibilità di applicare una quantità di compressione diversa in ogni banda di frequenza, compensando variazioni nel range dinamico dovute alle frequenze ed evitando che intense componenti a bassa frequenza (come le vocali) influenzino il guadagno ad alte frequenze (riguardante prevalentemente l'udibilità delle consonanti). Sfortunatamente, la compressione multibanda introduce una serie di problemi a sé stanti. Prima di tutto è intrinsecamente non lineare e introduce quindi distorsione armonica e di intermodulazione. La distorsione può diventare grave quando le costanti temporali del sistema compressore sono minori di vari periodi della minima frequenza da trattare. In secondo luogo, una compressione multibanda produce uno *smoothing* dello spettro del suono. A causa di ciò, l'ascoltatore potrebbe trovare più

difficile l'estrazione delle caratteristiche salienti dello spettro, come i picchi delle formanti, difficoltà aumentata dalla ridotta selettività nel dominio delle frequenze che accompagna una perdita uditiva cocleare. Per ultimo, la compressione multibanda potrebbe introdurre cambiamenti spuri nello spettro e nell'involuppo temporale del suono. È stato sperimentalmente dimostrato che la percezione del parlato dipende più dai cambiamenti nello spettro del suono che negli spettri stazionari. Da qui, anche leggeri cambiamenti introdotti dalla compressione potrebbero portare ad effetti deleteri nella percezione dello *speech*. Questi problemi potrebbero tenere conto del fatto che non è stato in generale rilevato che la compressione multibanda migliori l'intelligibilità in condizioni di rumore in confronto ad un'amplificazione lineare accompagnata dal cosiddetto "*frequency shaping*". Tutto ciò suggerisce che potrebbero esistere svariati svantaggi nell'avere troppi canali in un apparecchio acustico. I problemi associati allo smoothing dello spettro e all'introduzione di cambiamenti spettrali spuri potrebbero essere minimizzati solo usando un basso numero di canali, riducendo però di gran lunga la selettività in frequenza.

### 3.2. Problema della scarsa discriminazione soprasoglia

I soggetti affetti da perdita uditiva neurosensoriale presentano una ridotta abilità nel discriminare suoni di abbondantemente soprasoglia. In particolare, la discriminazione e la selettività in frequenza, la risoluzione temporale e il processamento binaurale sono tutti danneggiati. L'unica dimensione a rimanere intatta è quella relativa alla discriminazione dell'intensità. Tali deficit psicoacustici costituiscono la causa principale delle difficoltà del soggetto audioleso nella comprensione del parlato e non vengono generalmente corretti dagli apparecchi acustici. Tuttavia, un alleviamento di questi disturbi può essere ottenuto tramite l'uso di microfoni direzionali, con direzione preferenziale frontale e, ancora, l'uso di AGC ad azione rapida, caratterizzati da bassi tempi di attacco e recupero.

### 3.3. Problema della distorsione

Gli apparecchi acustici introducono nel segnale vari tipi di distorsione (distorsione armonica e di intermodulazione, range delle frequenze limitato e risposta in frequenza irregolare), che hanno un effetto deleterio sull'intelligibilità del parlato. Anche se ogni tipo

di distorsione, da sola, potrebbe avere solo un piccolo effetto sull'intelligibilità dello *speech*, presi insieme potrebbero avere un effetto globale decisamente negativo.

I principali accorgimenti costruttivi per far fronte al problema della distorsione consistono in un controllo del volume, accessibile dall'utente, usato solamente per ridurre la distorsione generata da livelli di input sonoro molto elevati, e l'utilizzo di un amplificatore di potenza di alta qualità che applichi un forte smorzamento sul ricevitore, riducendo così la distorsione e le risonanze introdotte dal sistema tubicini/stampo auricolare/condotto uditivo.

#### **4. Distorsioni elettroacustiche**

La distorsione consiste nell'insuccesso di un sistema nel riprodurre o trasmettere con esattezza una forma d'onda ricevuta. Si riferisce a qualsiasi componente del segnale o caratteristica, presente all'uscita di un amplificatore, che era assente nel segnale in ingresso. Alcune tipologie di distorsioni, come la distorsione in frequenza, vengono intenzionalmente introdotte per risolvere problemi di perdita uditiva specifici. Altre distorsioni non vengono invece appositamente introdotte, ma potrebbero bensì influenzare negativamente le prestazioni dell'apparecchio acustico.

Vari tipi di distorsione del segnale acustico sono capaci di essere percepiti dall'orecchio umano: i più comuni sono la distorsione armonica e la distorsione di intermodulazione, entrambe non lineari. Oltre alle non lineari, esistono varie forme di distorsione, che includono il transiente, la frequenza, la fase e il rumore. Fortunatamente, l'influenza di questi tipi di distorsione sull'uso degli apparecchi acustici moderni è probabilmente minima, per lo meno in base alle misure correnti. Nel seguito verranno quindi discusse solamente le forme principali di distorsione.

I prodotti della **distorsione armonica** (segnali spuri) sono generati dal passaggio dei segnali attraverso un amplificatore non lineare, il quale corrompe il segnale catturando una porzione dell'energia dell'input e ridistribuendola successivamente in forma di nuovi segnali, detti appunto prodotti di distorsione, caratterizzati da frequenze multiple della frequenza fondamentale del segnale originale. Per esempio, un segnale d'ingresso con frequenza fondamentale di 500 Hz, soggetto all'azione di un amplificatore non lineare, può

generare prodotti di distorsione a frequenze di 1000, 1500, 2000 Hz, etc. Separando le armoniche dalla frequenza fondamentale nel segnale d'uscita e calcolando il rapporto tra il valore totale delle armoniche e della fondamentale, la distorsione può essere espressa per mezzo di una percentuale. All'aumentare della non linearità dell'amplificatore, cresce anche l'ampiezza dei prodotti di distorsione e diminuisce quindi la qualità del suono percepito. In aggiunta, nonostante la comparsa di distorsione negli apparecchi acustici possa indicare difetti funzionali (come trasduttori difettosi o amplificatori asimmetrici), alcuni accorgimenti usati per limitare suoni troppo alti possono provocare anch'essi distorsione.

La **distorsione di intermodulazione (IM)** è il rapporto tra la potenza del segnale d'uscita a frequenze diverse da quelle originali e la potenza dei segnali all'ingresso dell'apparecchio acustico. Essa include anche i toni risultanti dalla somma e differenza aritmetiche tra i segnali di ingresso e loro armoniche. La distorsione di intermodulazione può essere spiegata considerando due frequenze d'ingresso (per esempio 500 e 700 Hz) di uguale ampiezza ma non armonicamente correlate. Come risultato del passaggio attraverso un sistema non lineare, si può ottenere un'uscita complessa, comprendente le frequenze fondamentali (500 e 700 Hz), le rispettive frequenze multiple (1000, 1500, 2000 Hz e 1400, 2100, 2800 Hz, etc.) e quelle date dalla somma e differenza tra toni ( $700 - 500 = 200$  Hz,  $700 + 500 = 1200$  Hz, etc.).

Questo tipo di distorsione è presente quindi quando un segnale, come quello del parlato, contiene più di una frequenza; questo problema viene ancora più accentuato in presenza di rumore ambientale, in quanto un numero ancora maggiore di frequenze viene aggiunto al segnale di speech.

Le componenti della distorsione armonica giacciono a frequenze multiple della fondamentale, la cui udibilità può risultare abbastanza accettabile in quanto le componenti della distorsione armonica coincidono spesso con le componenti armoniche già presenti nello speech. In aggiunta, se la risposta ad alte frequenze dell'apparecchio acustico è ridotta, allora un aumento di distorsione può quindi essere tollerato. Tuttavia, la distorsione di intermodulazione dà luogo a componenti, derivanti da somma e differenza di toni, che non presentano relazioni armoniche musicali e che risultano così alquanto fastidiose.

## 5. Campi della ricerca attuale

Il *Digital Signal Processing* (DSP) gioca un ruolo fondamentale nella progettazione dei sistemi da applicare agli apparecchi acustici a causa dell'ampia innovazione tecnologica degli hardware avvenuta negli ultimi decenni. Al giorno d'oggi, i sistemi basati su DSP rappresentano le architetture più popolari nel 100% degli apparecchi acustici digitali: applicando tecniche di *signal processing*, l'intero sistema tenta di far fronte alla perdita uditiva del paziente. Tali tecniche devono tenere conto delle reali condizioni degli apparecchi acustici: in particolare il rumore aggiunto al segnale, la risposta reale dell'apparecchio acustico, la risposta particolare dell'orecchio interno e il segnale di feedback che può corroborare le prestazioni generali dell'intero apparecchio acustico.

I principali campi di ricerca nell'ambito delle tecniche di *signal processing* applicate agli apparecchi acustici sono basicamente la riduzione del rumore aggiunto al segnale, la riduzione del *feedback* acustico e lo sviluppo e miglioramento di specifici algoritmi di compressione indicati per ogni singolo paziente, tipo di sistema e di perdita uditiva.

La riduzione del rumore è uno stadio importante nell'elaborazione del segnale dell'apparecchio acustico, dal momento che i soggetti ipoacusici devono comprendere lo *speech* prevalentemente in condizioni con rumore di fondo. Il portatore di apparecchio acustico deve infatti essere in grado di focalizzare l'attenzione su una singola conversazione o una singola sorgente di segnali acustici anche nella cosiddetta condizione di *cocktail party*, che si verifica quando nello stesso ambiente sono presenti suoni, rumori e conversazioni provenienti da diverse sorgenti e caratterizzate da diverse intensità e frequenze.

L'algoritmo di compressione è caratterizzato principalmente da due parametri, la soglia e il rapporto di compressione, i quali possono variare in funzione della frequenza, e quindi essere differenti in ogni canale dell'apparecchio acustico. L'obiettivo del presente studio è proprio quello di trovare la combinazione migliore, per ogni canale di un apparecchio acustico simulato, del punto di ginocchio dell'AGC e del rapporto di compressione. La configurazione migliore è quella che permette la massima intelligibilità, ossia quella in grado di riprodurre un *clear speech* di elevata qualità, a prescindere dal tipo e grado di perdita uditiva del singolo individuo. Sulla base di studi precedenti, è stato rilevato che il grado di *clear speech* può essere dedotto tramite l'estrazione e l'analisi di numerosi indici

e parametri, principalmente di tipo spettrale. Tali *features*, supposte fondamentali per la produzione di un *clear speech* di alto livello, verranno ampiamente illustrate nel seguente capitolo.

Infine, la riduzione del *feedback* acustico è il terzo tipo di problema da risolvere nell'adattamento dell'apparecchio acustico al paziente. Tale fenomeno, come già detto in precedenza, ha origine quando il suono in uscita dall'altoparlante del dispositivo viene captato dal microfono, facendo letteralmente fischiare la protesi. Questo effetto viene ulteriormente amplificato dalla presenza di un oggetto portato vicino all'orecchio, come il ricevitore del telefono, un cappello o semplicemente una mano. Dato che i fori di ventilazione vengono spesso impiegati per prevenire un effetto di occlusione del segnale di *speech* del soggetto stesso, che lo farebbe "rimbombare" all'interno dell'orecchio, la quantità di guadagno da impiegare dovrebbe essere limitata al di sotto di quella necessaria per colmare la perdita uditiva del paziente. Una soluzione che molti produttori hanno da sempre adottato è semplicemente quella di ridurre il guadagno della protesi nell'ampia banda di frequenza che comprende il segnale di *feedback*. Un altro metodo impiegato è stato quello di applicare filtri *notch* centrati sulla frequenza di tale segnale. Entrambi questi accorgimenti presentano però lo svantaggio di limitare il guadagno massimo che può essere applicato al segnale acustico nella regione di frequenza più problematica, rendendo qualsiasi informazione dello *speech* contenuta in suddetta banda praticamente inaudibile al portatore.

Per far fronte a questo inconveniente sono stati presentati in letteratura vari metodi basati su algoritmi e filtri adattivi per la cancellazione del *feedback*, in procinto di essere applicati agli apparecchi acustici moderni.

Tutti i miglioramenti nelle tecniche di processing non solo conducono all'aumento di intelligibilità da parte del soggetto ipoacusico portatore di protesi, ma riducono anche il costo computazionale di ciascuna fase di elaborazione del suono. Uno dei problemi di questo genere di ricerca è dato dal fatto che deve essere condotta senza la condizione di ambiente reale. A questo punto, una simulazione supportata da un calcolatore diventa un passo utile, affidabile ed efficiente nel progetto e sviluppo di un dispositivo, in quanto risulta l'unica tecnica possibile per produrre segnali reali e modelli su cui lavorare. La quasi totalità degli studi in letteratura utilizza strumenti di simulazione computerizzata, dal momento che non sono disponibili dati reali su cui lavorare. I software di simulazione a

questo scopo più usati sono MATLAB e Pspice, che, lavorando a livello circuitale, forniscono modelli elettrici e virtuali del sistema e costituiscono uno stadio di primaria importanza nella progettazione di un apparecchio acustico.



## CAPITOLO 2

### Il *clear speech*

#### 1. Introduzione

Nella comunicazione di tutti i giorni, l'obiettivo degli interlocutori è quello di divulgare i propri messaggi in una forma che risulti *intelligibile* agli ascoltatori. Quando si è consapevoli di una difficoltà dell'ascoltatore nel percepire il parlato, dovuta al rumore di sottofondo, ad una lesione acustica o all'essere di lingua madre differente, il proprio modo di parlare viene naturalmente e spontaneamente modificato, in modo da favorire l'interlocutore nell'ascolto e nella comprensione del discorso. Viene infatti adottato uno stile distinto di linguaggio, chiamato "*clear speech*". Sforzandoci di renderci il più comprensibili possibile, parliamo di solito più lentamente, più forte e articoliamo i suoni in maniera "esagerata". Le modificazioni che hanno luogo nel passaggio da linguaggio discorsivo a *clear speech* mirano a fornire all'ascoltatore maggiori caratteristiche acustiche salienti del segnale di speech, che potrebbero migliorare la sua abilità di comprensione del messaggio.

I cambiamenti articolatori nel passaggio dello speech da "conversational" (discorsivo) a "clear" presentano una certa variabilità intersoggettiva, ma trovano tuttavia un compromesso tra chiarezza ed economia dello sforzo: risulta evidente la necessità di uno sforzo maggiore nella produzione di clear speech rispetto al parlato discorsivo naturale.

Le diverse modalità di clear speech sono orientate in primo luogo al tipo di interlocutore: il linguaggio discorsivo viene diversamente modificato se si sta parlando con un bambino, uno straniero, un soggetto ipoacusico o se ci si trova in un ambiente rumoroso. Il clear speech condivide inoltre alcune caratteristiche con la variazione dovuta al rafforzamento prosodico e alla diminuzione della velocità del parlato.

Per essere chiari, la prosodia è la parte della linguistica che studia l'intonazione, il ritmo, la durata e l'accento nel linguaggio parlato. Le caratteristiche prosodiche di un'unità di linguaggio parlato (si tratti di una sillaba, di una parola o di una frase) sono dette soprasegmentali, perché simultanee ai segmenti in cui può essere divisa quell'unità: le si può infatti idealmente rappresentare come "sovrapposte" ad essi. Alcuni di queste caratteristiche sono, ad esempio, la lunghezza della sillaba, il tono e l'accento.

Gli adattamenti dovuti allo stile del parlato, alla velocità e alla struttura prosodica e discorsiva, possono essere visti come fenomeni che coinvolgono cambiamenti dell'iperarticolazione, come la riduzione del *target undershoot* e il miglioramento dei contrasti fonemici. Il clear speech è unico tra questi stili di parlato *goal-oriented*, in quanto ha il fine ultimo di migliorare l'intelligibilità per interlocutori adulti con difficoltà percettive.

L'obiettivo della ricerca nell'ambito del clear speech è quello di identificare, da una parte, le caratteristiche articolatorie-acustiche salienti che caratterizzano la produzione di clear speech e, dall'altra, comprendere quali di queste modificazioni da parlato discorsivo a chiaro influiscono maggiormente sull'intelligibilità.

Attraverso una migliore comprensione dei fattori legati a *talker*, *listener* e segnale vocale, la ricerca nell'ambito del clear speech mira a far luce sulle interazioni fra fattori cognitivi ad alto livello, fattori sensoriali a basso livello e fattori percettivi che influenzano l'elaborazione del linguaggio.

Nel contesto del presente elaborato, il termine *talker* sarà riferito ai soggetti che, nel corso degli studi sperimentali eseguiti in laboratorio, hanno il compito di pronunciare sillabe, parole o frasi intere in diverse modalità, le quali costituiranno a loro volta il materiale acustico usato per le diverse prove sperimentali. Tali parole verranno registrate e fatte ascoltare ai *listeners*, che costituiscono la seconda tipologia di soggetti partecipanti alla ricerca sul clear speech. Talkers e listeners possono essere di età, sesso e lingua differente, presentare diversi tipi di ipoacusia o non avere alcun problema di udito.

Un ulteriore scopo della ricerca in questo ambito è quello di scoprire i meccanismi che stanno alla base della plasticità nella produzione di speech che permette ai talkers di adattare immediatamente il proprio linguaggio alle impostazioni comunicative stabilite e i meccanismi da cui dipende l'elaborazione dello speech che consente ai listeners di sfruttare queste modificazioni articolatorie-acustiche per segmentare il segnale di speech e arrivare così alla corretta struttura prosodica e sintattica della singola espressione pronunciata.

In generale la ricerca nell'ambito del clear speech mira principalmente a sviluppare algoritmi di *speech-enhancement*, per migliorare la qualità della comunicazione verbale di categorie specifiche di persone, come i soggetti ipoacusici e i non madre-lingua, e per applicazioni tecnologiche.

Sono già stati compiuti significativi progressi nell'identificazione di un'ampia gamma di accorgimenti acustico-articolatori che accompagnano la trasformazione del parlato da discorsivo a chiaro e nella dimostrazione dei benefici sull'intelligibilità dati dal clear speech.

Prima di iniziare a parlare delle scoperte fatte nell'ambito della ricerca, si intende chiarificare il concetto di linguaggio discorsivo (*conversational speech*) e di clear speech. Entrambi i termini si riferiscono alla produzione del parlato in laboratorio, secondo specifiche istruzioni date ai *talkers*. Tali istruzioni richiedono tipicamente che il soggetto legga lo stesso set di materiale due volte: la prima normalmente (*conversational speech*) e la seconda "come se stesse parlando ad un soggetto ipoacusico o ad uno straniero" (*clear speech*).

Il beneficio sull'intelligibilità indotto dal clear speech è stato indagato in numerosi studi tramite l'uso di materiale di diverso tipo: sillabe, parole sensate, parole senza senso e frasi di senso compiuto.

Dal momento che sarebbe impossibile includere una discussione che tratti tutte le misure ricavate e analizzate in questo ambito di ricerca, si possono inizialmente identificare due ampie categorie di misure: le **misure globali** e le **misure segmentali**.

Le prime comprendono la velocità del parlato, la frequenza e la durata delle pause, il range della frequenza fondamentale media, gli spettri a lungo termine (distribuzione dell'energia spettrale lungo l'intera pronuncia della parola o frase) e le modulazioni dell'involuppo temporale (spettro dell'involuppo temporale).

La seconda categoria di misure comprende invece i cambiamenti delle formanti (stato stazionario di una vocale e/o transizioni della frequenza della prima (F1) e seconda (F2) formante), lo spazio delle vocali, o *vowels space*, (area compresa tra le vocali, nello spazio determinato dalle coordinate F1 ed F2), la durata di ogni segmento, il rapporto consonante-vocale, il tempo d'attacco della voce (*voice onset time*, VOT), lo spettro a breve termine (spettro del segnale attorno ad un particolare istante temporale) e molti altri parametri di minor rilevanza.

## 2. Cosa rende più intelligibile il clear speech?

Nel corso delle indagini sugli effetti dello stile di linguaggio sull'intelligibilità, la ricerca nell'ambito del clear speech si è sforzata di identificare le caratteristiche acustico-fonetiche che caratterizzano la produzione di clear speech, contribuendo alla sua massima intelligibilità. I risultati derivanti da analisi acustiche mostrano che il clear speech prodotto naturalmente include un'ampia gamma di accorgimenti, tra cui i più importanti sono:

- espansione del cosiddetto *vowels space* (o spazio delle vocali), che si riflette anche in un movimento dinamico delle formanti;
- diminuzione della velocità del parlato, con pause più lunghe e ripetute e vocali di durata maggiore;
- maggiori valori di pressione sonora, principalmente delle vocali (il che significa semplicemente parlare ad un volume maggiore);
- maggiore energia nell'intervallo di frequenze 1000 – 3000 Hz dello spettro a lungo termine;
- aumento qualitativo della *modulation transfer function* (MTF), nel range di frequenze compreso tra 4 e 16 Hz;
- range dinamico della frequenza fondamentale F0 più ampio;
- maggiore intensità del rms (e quindi della potenza) nei tratti “non silenziosi” di alcune consonanti e aumento del rapporto consonante – vocale.

Nonostante siano state descritte numerose differenze acustiche tra clear e conversational speech, non è comunque ancora chiaro quale caratteristica contribuisca maggiormente all'intelligibilità del clear speech.

L'entità dei benefici derivanti dal clear speech pare invece essere indipendente dall'età del talker e dalla sua esperienza nell'interazione con soggetti audiolesi, mentre sembra che talkers di sesso femminile tendano comunque a produrre un clear speech più intelligibile rispetto a quello prodotto da individui di sesso maschile. Una simile variabilità tra talkers nell'intelligibilità del clear speech è stata riscontrata sia per le singole parole sia per intere frasi, suggerendo che la variabilità dell'intelligibilità è più una peculiarità del talker piuttosto che della singola vocale o parola.

Andiamo ora ad esaminare più da vicino gli studi sperimentali finora condotti e i rispettivi risultati dei principali accorgimenti che vengono solitamente adottati nella produzione del clear speech.

### 2.1. Espansione del *vowels space*

Una delle più importanti caratteristiche che si è visto accompagnare il clear speech naturalmente prodotto è l'espansione dello spazio delle vocali rispetto a quanto avviene per il parlato discorsivo. L'“iperarticolazione vocalica” associata al clear speech aumenta la distanza  $F1 \times F2$  tra le varie categorie di vocali e causa una migliore approssimazione del target atteso delle formanti, rendendo così le classi di vocali più distinte e meno confondibili a livello percettivo. L'espansione del *vowels space* nel clear speech può essere inoltre associato ad un vantaggio nell'intelligibilità sulla base delle differenze intersoggettive dei talkers nell'intelligibilità globale discorsiva: chi è di natura più intelligibile tende a produrre di conseguenza spazi vocalici più estesi.

L'espansione vocalica è inoltre solitamente accompagnata da un allungamento temporale delle vocali stesse, strategia che contribuisce ulteriormente all'intelligibilità del clear speech.

A livello pratico, per dimostrare i suddetti fenomeni, in un recente studio sono stati selezionati sette parametri per valutare le tre caratteristiche acustiche primarie considerate fondamentali per l'identificazione di una vocale nell'inglese americano: la durata temporale, identificata da un unico valore, i valori delle formanti in condizioni stazionarie, da cui si ricavano quattro parametri, e il movimento dinamico delle formanti (identificato da due parametri).

I quattro valori delle formanti sono stati scelti sulla base dell'assunzione, trovata in studi precedenti, che lo spazio delle vocali  $F1 \times F2$  è più ampio in condizione di clear speech. Da questi sono stati estratti i quattro parametri di interesse:

- 1) perimetro del poligono risultante nel piano  $F1-F2$ : l'intelligibilità migliora all'aumentare di questa misura;
- 2) range di  $F1$ : maggiore nel caso del clear speech;
- 3)  $F2_{front}$  (formante  $F2$  per le vocali frontali): maggiore in condizione di clear speech;

- 4) *F2 back* (formante F2 per le vocali posteriori): solo leggero aumento nel caso di clear speech.

Le dieci vocali della lingua inglese considerate sono state poi divise in due gruppi e ne sono state assegnate le coordinate nel piano F1-F2, sia nel caso del clear speech che del conversational speech. Sovrapponendo i grafici relativi per quattro vocali dello stesso tipo per i sei soggetti che hanno fornito i risultati migliori a livello di intelligibilità, si ottengono i quadrilateri mostrati in figura 17. Già a prima vista è possibile osservare che l'area dei quadrilateri relativi al clear speech è quasi sempre maggiore di quella relativa al conversational speech, il che risulta anche in un aumento del perimetro.

I risultati descritti dimostrano quindi che il clear speech è generalmente accompagnato dall'espansione, più o meno accentuata, del *vowels space*, traducibile in un aumento delle prime due formanti.

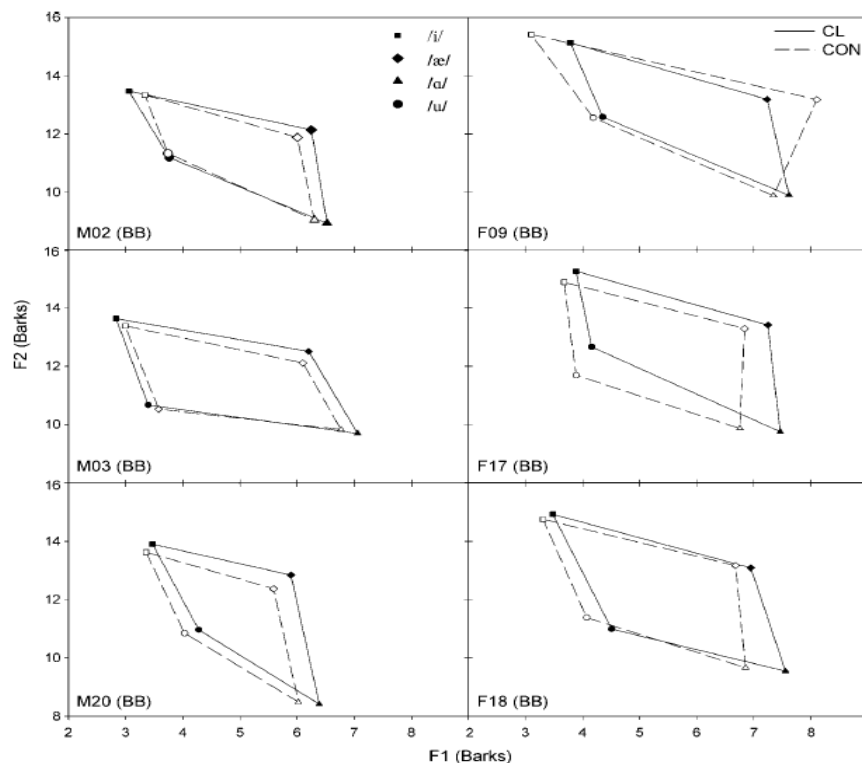


Figura 17: *vowels space* costruito su quattro vocali della lingua inglese

Globalmente, tali analisi indicano che, quando istruiti a parlare come se avessero di fronte un soggetto ipoacusico, la maggioranza dei talkers espande lo spazio vocalico, per lo meno di qualche grado. Sebbene tale affermazione sia coerente con la consapevolezza che il clear

speech è una forma di “*hyperspeech*”, suddetta espansione non è né uniforme né semplice e, per di più, la sua entità presenta una forte variabilità intersoggettiva.

Tutto ciò complica significativamente lo studio di come tutte le caratteristiche e i parametri legati all’espansione del vowels space influenzano globalmente l’intelligibilità.

## 2.2. Diminuzione della velocità e aumento delle pause nel parlato

A differenza dell’espansione dello spazio delle vocali, ormai definita come una delle caratteristiche che rendono lo speech effettivamente clear, è invece difficile stabilire una relazione diretta tra la variabilità di intelligibilità e le variazioni acustico-fonetiche dovute allo stile del parlato. Nonostante siano già state identificate numerose caratteristiche acustico-fonetiche delle trasformazioni che avvengono dal parlato discorsivo a quello chiaro, non è stato tuttavia ancora ben compreso come queste modificazioni influiscano sull’intelligibilità.

Come alternativa allo studio del clear speech naturalmente prodotto, comprendente numerose modificazioni a livello delle dimensioni acustiche e articolatorie che influenzano l’intelligibilità, molti studi hanno usato tecniche di *signal processing* per individuare il ruolo che le singole caratteristiche acustiche giocano sull’intelligibilità.

Una delle caratteristiche che accompagna tipicamente le trasformazioni “conversational-to-clear speech”, ossia la diminuzione della velocità del parlato, è stato ampiamente indagato in molti studi attraverso tecniche di elaborazione di segnali.

Il fenomeno della diminuzione della velocità per la produzione di clear speech è rappresentato da segmenti più lunghi e da pause più frequenti, ma in studi riguardanti la variabilità intersoggettiva nel parlato discorsivo non ha tuttavia mostrato una correlazione tra velocità globale del parlato e intelligibilità. Le modificazioni uniformi e non uniformi dello speech su scala temporale, ottenute accelerando il clear speech e rallentando il parlato discorsivo, hanno avuto come risultato una minore intelligibilità, rispetto al segnale di speech non processato, per soggetti ipoacusici, anziani, bambini e adulti normoudenti. Questi risultati confermano che il clear speech include cambiamenti acustici, oltre alla semplice velocità, che contribuiscono alla sua alta intelligibilità. Per di più, tali studi dimostrano che il rallentamento dovuto a modifiche nel modo di parlare non può essere ottenuto semplicemente per mezzo di un algoritmo di allungamento segmentale. Anche

l'eventuale aggiunta di pause artificiali nelle frasi di tipo discorsivo e l'eliminazione di pause esistenti dalle frasi in clear speech riducono in qualche modo l'intelligibilità.

Al contrario, studi più recenti hanno invece riscontrato un miglioramento dell'intelligibilità attraverso l'inserimento di intervalli temporali in frasi di tipo discorsivo, anche se non per tutti i valori di SNR, valorizzando l'ipotesi di un contributo significativo sull'intelligibilità da parte del rallentamento nello speech.

Ulteriori ricerche hanno individuato l'effetto della velocità del parlato sull'intelligibilità producendo un clear speech a velocità naturale, evitando in tal modo alcune "trappole" delle tecniche di *signal processing*. È stato infatti dimostrato che si può produrre, dopo un'opportuna fase di allenamento, un buon clear speech anche a velocità di parlato di tipo discorsivo, aumentando di conseguenza l'intelligibilità per soggetti normoudenti con perdita uditiva simulata (ossia in condizioni di rumore). Questi risultati suggeriscono che il fatto di parlare più lentamente non è l'unico fattore responsabile dei vantaggi sull'intelligibilità del clear speech, bensì che quest'ultimo possiede delle proprietà acustiche intrinseche indipendenti dalla velocità. Tuttavia, un leggero decremento nel vantaggio sull'intelligibilità del clear speech prodotto a velocità discorsive, in parallelo alla scoperta che un aumento del numero di pause produce un effetto benefico sull'intelligibilità, indica che il rallentamento del parlato, anche se non cruciale, è comunque un fattore che contribuisce ad una maggiore intelligibilità del clear speech.

### 2.3. Inviluppo temporale

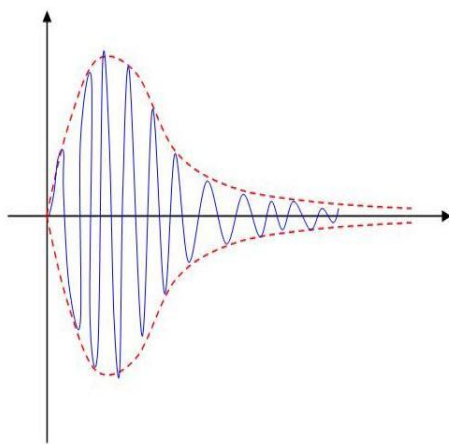


Figura 18: inviluppo temporale di un segnale.

L'inviluppo temporale di un segnale (mostrato in figura 18 dalla linea rossa tratteggiata) contiene una varietà di frequenze di modulazione, l'ampiezza delle quali può essere illustrata per mezzo di un'analisi spettrale. Tali modulazioni giocano un importante ruolo nella trasmissione dell'informazione contenuta nello speech: è stato infatti riscontrato un peggioramento in termini di intelligibilità se le fluttuazioni dell'inviluppo temporale non vengono trasferite adeguatamente.



Studi sulla percezione mostrano che l'involuppo temporale offre maggiori vantaggi sul clear speech per alti valori di SNR (ossia è critico per il riconoscimento dello speech in ambienti silenziosi), mentre la struttura temporale contribuisce maggiormente a bassi valori di SNR (è quindi critica per il riconoscimento dello speech in ambienti rumorosi).

L'importanza dell'involuppo temporale nella valutazione della qualità dello speech è stata elaborata nel concetto di *modulation transfer function* (MTF). La MTF è definita come il rapporto tra l'intensità degli spettri dei segnali dopo e prima il processing, ossia:

$$MTF = |H(f)| = \alpha \frac{|Y(f)|}{|X(f)|}$$

dove  $X(f)$  e  $Y(f)$  sono rispettivamente gli spettri degli involucri temporali del segnale d'ingresso e d'uscita.

Una misura diretta della MTF può essere quindi ottenuta confrontando gli spettri di modulazione dell'involuppo dell'input e dell'output,

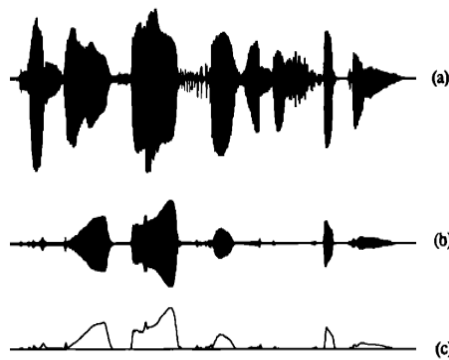


Figura 19: (a) segnale originale; (b) segnale filtrato nella banda 283-336Hz; (c) involuppo del segnale in (b).

per frequenze comprese tra 4 e 16 Hz. La ragione per cui viene analizzato questo ristretto intervallo di frequenze è mostrato in figura 19.

Nel tracciato (a) è rappresentata la forma d'onda del segnale originale; in (b) il segnale in uscita da un filtro passabanda con *bandwidth* pari a 53Hz e frequenza centrale 310 Hz e, in (c), l'involuppo del segnale in (b). Risulta dunque chiaro che le fluttuazioni dello spettro sono visibili per valori molto

bassi di frequenza.

In accordo con la nozione di MTF, sono stati adeguatamente misurati gli effetti dannosi del rumore e del riverbero sullo speech, in termini di riduzione della profondità di modulazione che producono.

In vari studi, i risultati deludenti sul beneficio della compressione negli apparecchi acustici sono stati discussi in termini di importanza delle modulazioni temporali. Il concetto di MTF è stato così usato per dimostrare che la compressione multicanale con brevi costanti temporali riduce i contrasti di intensità e porta quindi ad una diminuzione di intelligibilità.

Tuttavia, i meccanismi precisi con cui i cambiamenti della struttura e dell'involuppo temporale del segnale di speech influenzano l'elaborazione del linguaggio e l'intelligibilità rimangono tuttora poco chiari.

Combinare, tutte queste ricerche indicano chiaramente la necessità di una conoscenza più dettagliata dei principi che guidano la produzione di clear speech prima che si ottenga appieno la potenzialità dei sistemi di *speech processing* di migliorare la percezione dello speech.

### 3. Considerazioni generali e possibili sviluppi futuri

Tutti gli studi finora trattati hanno fornito numerosi indizi relativi alle caratteristiche e ai benefici del clear speech nella lingua inglese. Potrebbe comunque essere corretto affermare che la produzione e percezione di clear speech in qualsiasi altra lingua presentino tratti comuni all'inglese, nonostante tali studi abbiano ricevuto un'attenzione considerevolmente minore. Per esempio, il parlare più lentamente, più forte e in maniera "esagerata" caratterizza il linguaggio della maggioranza dei talkers quando mirano a rendersi più intelligibili, senza riguardo della loro lingua madre.

Un interessante quesito che un esame cross-linguistico ci permette di porre è se le modificazioni del clear speech vengano guidate da pattern linguistici specifici di iperarticolazione. Se così fosse, le regolazioni articolatorie del clear speech potrebbero aumentare la "distanza" tra categorie contrastive linguistiche specifiche, riflettendo proprietà fonologiche e prosodiche che variano nelle lingue. Una ricerca sistematica di pattern di clear speech cross-linguistici potrebbe fornire indizi inestimabili sul ruolo che le varie caratteristiche acustiche giocano nell'espressione e intensificazione dei contrasti fonologici e sull'interazione tra fattori generali uditivi-percettivi e strutturali che giocano un ruolo nell'incremento dell'intelligibilità del clear speech.

Studi focalizzati sulla lingua inglese mettono in evidenza la possibile influenza che la struttura fonologica manifesta sulla produzione di clear speech. Per esempio, è stato trovato che, nel clear speech inglese, le vocali "tese" vengono allungate più di quelle "mollie" e che lo stile di parlato influenza le variazioni della formante F2 in modo diverso a seconda che la vocale sia frontale o posteriore: come già accennato, le vocali frontali

manifestano una F2 maggiore nel clear speech rispetto al parlato discorsivo, mentre le vocali posteriori subiscono una minore variazione di F2.

Pare quindi che il clear speech sia in parte guidato da un principio di intensificazione del contrasto che massimizza la distanza tra categorie linguistiche sonore e che renda più saliente la struttura prosodica. Le caratteristiche acustico – fonetiche che accompagnano gli aggiustamenti occorrenti nel processo di “conversational-to-clear speech” potrebbero pertanto variare tra le diverse lingue, riflettendo la loro struttura fonologica sottostante.

Per concludere, i risultati finora analizzati sottolineano l’accentuata natura multidimensionale dello speech e suggeriscono l’esistenza di un’ampia varietà di fattori acustici che possono influenzare l’intelligibilità delle singole vocali e delle parole ma, con elevata probabilità, la vasta gamma di caratteristiche acustiche non è ancora stata esplorata a fondo. Studi sperimentali in questo ambito sono quindi tuttora in corso e il campo di ricerca relativo alle tecniche di *speech-processing* e *speech-enhancement* rimane aperto a nuovi sviluppi tecnologici e a nuove scoperte.

## CAPITOLO 3

### Teoria dei simulatori di apparecchi acustici

#### 1. Considerazioni generali

Le simulazioni computerizzate possono essere sfruttate come primo strumento nella progettazione di un nuovo apparecchio acustico, ai fini della ricerca e sviluppo e per applicazioni cliniche, come nelle procedure di impostazione del guadagno di inserzione per ciascun canale da fornire al paziente in base alla sua perdita uditiva in ogni banda di frequenza, rilevabile dall'audiogramma tonale.

Esistono basicamente tre approcci per simulare un apparecchio acustico digitale tramite calcolatore. Il primo è quello di progettare strumenti che riproducano essenzialmente le operazioni eseguite dagli apparecchi acustici convenzionali; il secondo è quello di sviluppare apparecchi acustici digitali che comprendano gli stessi principi fondamentali dei dispositivi analogici, facendo però uso di tecniche digitali che incorporano caratteristiche difficili e/o non pratiche da ottenere con i circuiti analogici tradizionali; il terzo approccio è infine quello di sviluppare strumenti digitali concettualmente differenti dalle protesi acustiche analogiche e contenenti caratteristiche che, ai fini pratici, possono essere ottenute solo per mezzo di tecniche di *digital signal processing* (DSP) avanzate.

Un esempio del primo approccio è un apparecchio acustico digitale costituito da un microfono e un preamplificatore tradizionali, un convertitore analogico-digitale (ADC), un filtro digitale, un convertitore digitale-analogico (DAC) e un amplificatore di potenza che piloti un ricevitore analogico.

Un modello del secondo approccio è un apparecchio acustico digitale che possa essere programmato per meglio incontrare le necessità dei singoli utenti. Un filtro digitale discretamente sofisticato potrebbe essere richiesto al fine di ottenere la caratteristica frequenza – guadagno ottimale per ogni individuo. Nonostante gli apparecchi acustici programmabili siano già stati sviluppati, esistono tuttavia alcune limitazioni relative al range e alla flessibilità dei filtri analogici che possono essere impiegati nelle applicazioni pratiche. I filtri digitali sono invece soggetti ad un numero minore di limitazioni. Tra i principali vantaggi degli strumenti digitali, degna di nota è la possibilità di progettare e dimensionare gli apparecchi acustici allo stesso tempo, tramite un'unica sessione di lavoro.

Il terzo e ultimo approccio offre il maggior potenziale per miglioramenti significativi nel supporto di soggetti audiolesi. I metodi innovativi di *signal processing* sono attualmente in esplorazione, con l'obiettivo principale di perfezionare le caratteristiche del parlato importanti dal punto di vista dell'intelligibilità, la riduzione automatica del rumore e del *feedback*.

## 2. Esempio di simulatore

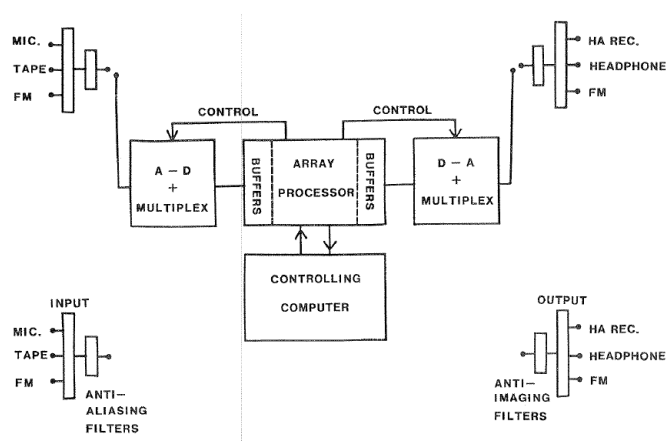


Figura 20: esempio di un modello di simulazione di apparecchio acustico digitale presentato in letteratura.

I principali componenti di un semplice sistema di simulazione, presentato in letteratura, è quello mostrato in figura 20.

I segnali d'ingresso possono provenire da un microfono, da un registratore o un ricevitore FM.

Il filtro anti-aliasing limita la larghezza di banda dei segnali processati dal sistema ad un valore

compatibile con la frequenza di campionamento del sistema digitale.

Nonostante possa in teoria essere utilizzata una frequenza di campionamento di 125kHz, il tempo richiesto per processare i segnali in tempo reale impone limiti pratici sulla selezione di tale frequenza. In generale, al diminuire della frequenza di campionamento, aumenta la quantità di *signal processing* che può essere ottenuta. Un utile e pratico *trade-off* per questa applicazione potrebbe essere una frequenza di campionamento pari a 12kHz, con una larghezza di banda risultante di 6kHz teoricamente utilizzabile, ritenuta adeguata per la simulazione di apparecchi acustici. In speciali applicazioni, come la simulazione di protesi acustiche a banda larga, può essere utilizzata una frequenza di campionamento di 25kHz, con una conseguente limitazione della quantità di *signal processing* ottenibile.

Il segnale proveniente dal filtro anti-aliasing viene poi digitalizzato tramite un ADC a 12 bit e salvato in uno *storage buffer*; in caso di simulazioni che coinvolgono input a più canali, come nel caso di apparecchi acustici binaurali, viene impiegato un *multiplexer*,

dispositivo capace di selezionare un singolo segnale fra diversi segnali in ingresso in base al valore degli ingressi di selezione.

Il nucleo centrale del sistema è un processore digitale che permette operazioni vettoriali ad alta velocità e ben si adatta al calcolo rapido di trasformate di Fourier e convoluzioni dei segnali in ingresso. A sua volta, il processore è controllato da un minicomputer che alimenta il processore con programmi che ne permettono contemporaneamente il suo funzionamento e controllo. Il computer di controllo può essere inoltre utilizzato per valutare le risposte del soggetto in esame, analizzare dati, verificare l'uscita del processore, guidare le simulazioni e implementare una strategia adattiva di valutazione, per mezzo della quale, in funzione delle risposte del soggetto, l'apparecchio acustico simulato converge verso i parametri ottimi per il singolo utente.

Le limitazioni di tale sistema sono determinate in primo luogo dal tempo computazionale del processore e dalla sua memoria. Nella modalità operativa in tempo reale è presente un ritardo temporale intrinseco tra ingresso ed uscita pari alla durata di due finestre temporali utilizzate nel calcolo della FFT (Fast Fourier Transform). Se tuttavia viene impiegato un filtraggio di tipo ricorsivo senza l'utilizzo di un buffer, il ritardo temporale si riduce al periodo di campionamento, pari a circa  $83\mu\text{s}$  per un *sampling rate* di 12kHz.

L'uscita del processore viene poi temporaneamente immagazzinata all'interno di buffer, riconvertita in forma analogica attraverso un DAC a 12 bit e, nel caso di segnali *multi-channel*, inviata ad un *demultiplexer*, che consente la ricostruzione del segnale.

L'output del convertitore digitale-analogico viene successivamente trasmesso ad un apposito sistema di trasduzione, il quale può essere rappresentato da cuffie, o un ricevitore per apparecchi acustici oppure un trasmettitore FM.

Il principale problema riguardante l'uso clinico di un apparecchio acustico simulato o fisicamente montato su appositi sostegni per simulazioni è legato alle intrinseche differenze acustiche tra dispositivi indossabili e non indossabili, in quanto questi ultimi non sono in grado di tenere conto dell'effetto di schermo acustico prodotto dalla testa, degli effetti dovuti geometria dell'orecchio e della complessa interazione tra i movimenti del capo e le percezioni uditive associate. Per rimuovere le differenze tra apparecchi indossabili e non indossabili, può essere utilizzato un segnale di trasmissione FM, in quanto è in grado di rendere la protesi acustica simulata acusticamente identica ad una effettivamente indossata da un individuo reale.

In aggiunta, per applicazioni che richiedono la simulazione di apparecchi acustici indossabili, vengono montati al dispositivo un microfono ed un ricevitore convenzionali. Lo stesso approccio può inoltre essere adottato per la simulazione di apparecchi ITE, che richiedono lo stampo auricolare personale del paziente. Per la simulazione postauricolare, l'uscita del ricevitore dell'apparecchio acustico viene condotta ad un'ipotetica chiocciola utilizzando un sistema standard di tubicini acustici.

In una tipica simulazione, il segnale elettrico generato dal microfono viene pre-amplificato ed inviato ad un trasmettitore FM tascabile, il cui segnale trasmesso viene rilevato da un ricevitore FM posizionato all'ingresso del computer. Dopo la fase di elaborazione, l'uscita di tale calcolatore viene ritrasmessa al soggetto in modo da essere captata da un ricevitore FM tascabile e sottoposta all'azione di un amplificatore, l'uscita del quale viene mandata direttamente al ricevitore dell'apparecchio acustico.

Ai fini pratici, gli stadi di input e output acustici dell'apparecchio acustico descritto sono identici a quelli di un dispositivo convenzionale reale, nonostante il trattamento elettronico dei segnali che avviene tra microfono e ricevitore venga trasferito ad un computer attraverso un sistema di trasmissione FM.

### **3. Esempio di simulazione**

In questo paragrafo, attraverso l'uso di un esempio di simulazione, viene dimostrato il potenziale di un sistema di simulazione di apparecchio acustico controllato da un computer.

In primo luogo, tale sistema può essere utilizzato per simulare un apparecchio acustico tradizionale sotto il diretto controllo di un calcolatore.

Oltre a fornire un'amplificazione selettiva per le frequenze, il sistema può anche essere programmato per funzionare come audiometro, per mezzo della generazione di impulsi a diversa frequenza e intensità, utilizzati nella pratica clinica per determinare la perdita uditiva dei soggetti ipoacusici.

Un possibile diagramma a blocchi di un apparecchio acustico simulato (figura 21) mostra che l'ingresso proveniente da ogni canale di questo sistema a due canali viene scomposto da un multiplexer e mandato a quattro buffer, due per ogni canale.

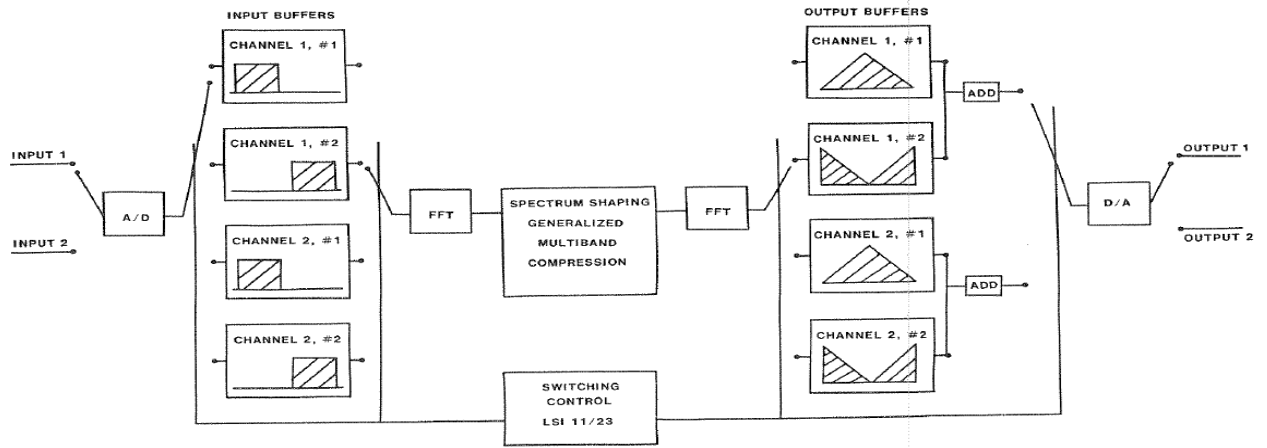


Figura 21: schema a blocchi delle possibili fasi di elaborazione del segnale di speech all'interno di un simulatore di apparecchio acustico.

Un approccio generale all'adattamento prescrittivo della caratteristica frequenza-guadagno è quello di incominciare con una prima stima della curva frequenza-guadagno ottima basata su considerazioni psicoacustiche, seguita da correzioni sistematiche, per arrivare ad una stima ottimale. La procedura di aggiustamento viene iterata fino a che, tramite il criterio prescritto, non possono essere ottenuti ulteriori miglioramenti.

In accordo con questa metodologia, la soglia uditiva, la soglia di fastidio e vari livelli intermedi vengono ottenuti tramite l'uso di bande di rumore da 1/3 d'ottava. Gli stimoli di test vengono generati e controllati attraverso il computer usando una strategia adattiva *up-down* per convergere efficacemente ai livelli indicati. La prima stima della caratteristica frequenza-guadagno ottima viene ottenuta trovando la curva frequenza-guadagno per cui l'rms del segnale di *speech* corrisponde al *comfort level* medio nel range di frequenze di interesse (200 – 5000 Hz). Le stime successive vengono poi ottenute con aggiustamenti sistematici della caratteristica frequenza-guadagno e verificando, di volta in volta, se l'intelligibilità è stata migliorata.

Lo stesso sistema può essere utilizzato per simulare un processo di riduzione del rumore di sottofondo. L'obiettivo è quello di progettare un apparecchio acustico che riduca automaticamente il cosiddetto *background noise*, particolarmente deleterio per l'intelligibilità in un soggetto audileso. Le fasi di elaborazione del segnale acustico in codesto procedimento di soppressione del rumore sono schematizzate in figura 22.



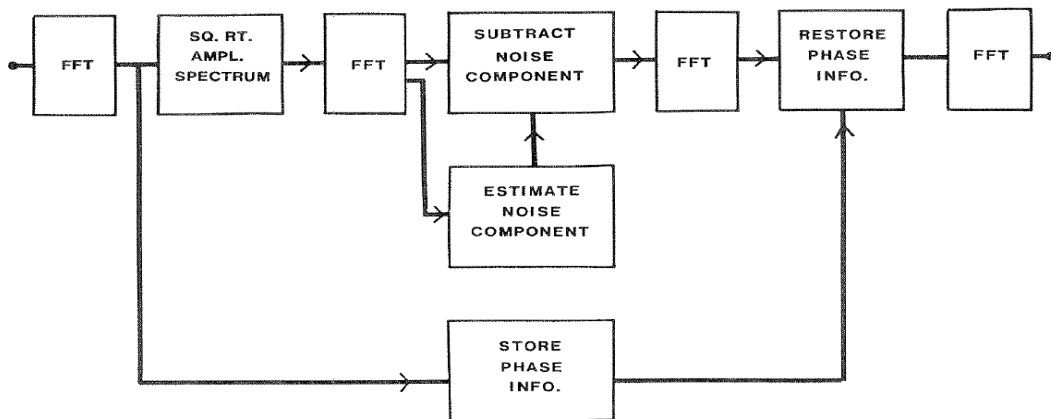


Figura 22: schema a blocchi di un processo di riduzione del rumore di sottofondo in un simulatore di apparecchio acustico digitale.

Gli stadi di ingresso e uscita del sistema sono gli stessi di quello descritto in precedenza, ad eccezione del fatto che il sistema è monoaurale, per cui il segnale trattato sarà a singolo canale. Dopo la prima trasformata di Fourier (FFT), l'informazione di fase viene conservata in una memoria per un suo successivo utilizzo, mentre la radice quadrata dello spettro in ampiezza viene riconvertito nel dominio del tempo per mezzo di una FFT inversa. Viene quindi eseguita una stima della componente di rumore di questo segnale elaborato attraverso una media pesata nel tempo su finestre successive. Lo spettro di rumore così ottenuto è poi sottratto dallo spettro relativo al segnale comprendente lo *speech* e il rumore, con il risultato di una rappresentazione migliorata del segnale, depurato da gran parte della componente rumorosa. Dopo un'ulteriore applicazione della trasformata di Fourier, allo spettro così ottenuto si aggiunge nuovamente la componente di fase, precedentemente salvata in una memoria.

Viene infine eseguita un'ultima FFT inversa per riportare il segnale nel dominio del tempo, il quale verrà poi inviato ai buffer d'uscita.

In conclusione, i due esempi appena descritti mostrano un paio tra gli innumerevoli vantaggi dell'utilizzo di simulazioni computerizzate nella fase di progettazione di un apparecchio acustico digitale.

Il primo esempio è di primaria importanza in ambito clinico e mostra come una simulazione al computer può fornire un semplice metodo che facilita il lungo processo dell'adattamento prescrittivo di un apparecchio acustico digitale. Il vantaggio offerto in questo campo applicativo è molteplice: in primo luogo, dal momento che l'apparecchio

acustico viene simulato, non è necessaria la realizzazione di un prototipo sperimentale, con conseguente risparmio di tempo e costi di produzione. In aggiunta, la vasta gamma e versatilità dei sistemi computerizzati supera di gran lunga quella ottenibile con i prototipi convenzionali di apparecchi acustici. Infine, essendo il sistema sotto il controllo diretto di un computer, la funzione di regolazione dell'apparecchio acustico risulta enormemente semplificata.

Il secondo esempio riportato mostra invece la raffinatezza che può essere raggiunta in simulazioni computerizzate real-time di protesi acustiche sperimentali. In questo caso, la progettazione e realizzazione del prototipo di un'unità deputata alla soppressione del rumore costituisce un passo tecnologico formidabile nello sviluppo di apparecchi acustici a prestazioni sempre più elevate.

La programmazione di un sistema di *speech-processing* agente in tempo reale, anche se considerevolmente meno complicato della costruzione di prototipi hardware, non è tuttavia opera di facile realizzazione. La ricerca scientifica in questo campo applicativo risulta quindi tutt'ora in via di sviluppo, dal momento che l'effetto deleterio del rumore sullo *speech* è maggiore per i soggetti ipoacusici che per i normoudenti, per cui qualsiasi lieve progresso nella soppressione del rumore risulta in un significativo beneficio per i soggetti audiolesi e portatori di protesi acustiche.

## CAPITOLO 4

### Materiali e metodi

#### 1. Considerazioni generali

Come illustrato nel capitolo dedicato alla teoria dei simulatori, in letteratura esistono basicamente tre modalità di implementazione di un apparecchio acustico digitale attraverso l'uso di un calcolatore.

Il primo è quello di progettare dispositivi caratterizzati da un microfono analogico, un convertitore analogico-digitale, filtri digitali, un convertitore digitale-analogico e un ricevitore convenzionale; il secondo è quello di sviluppare apparecchi acustici digitali programmabili, incorporando caratteristiche difficili da ottenere con i circuiti analogici tradizionali; il terzo approccio è infine quello di sviluppare strumenti digitali concettualmente differenti dalle protesi acustiche analogiche e contenenti caratteristiche che, ai fini pratici, possono essere ottenute solo per mezzo di tecniche di *digital signal processing* (DSP) avanzate.

Nel presente studio, si è deciso di riprodurre tramite il software Simulink il microfono e il *feedback* acustico tramite funzioni di trasferimento nel dominio di Laplace, tipiche della tecnologia analogica, mentre gli altri elementi del simulatore sono stati realizzati interamente tramite tecniche digitali. Di conseguenza, tenendo presente che i convertitori analogico-digitale e digitale-analogico sono stati evitati e sostituiti dall'uso di particolari accorgimenti (blocco denominato *rate transition* e specifici parametri della simulazione), si può dire che è stato utilizzato un approccio del primo tipo. I valori del guadagno di inserzione possono inoltre essere modificati con la finalità di adattare il dispositivo ad un generico paziente ipoacusico: tale simulatore è quindi programmabile, il che fa riferimento ad un approccio del secondo tipo. È infine presente anche una componente di *digital processing*, per cui è lecito affermare che l'approccio utilizzato in questo studio sperimentale è un "ibrido", dal momento che comprende caratteristiche appartenenti ai diversi tipi di metodologie presenti in letteratura.

Nei paragrafi successivi verrà dettagliatamente descritto il simulatore di apparecchio acustico digitale ideato, la cui realizzazione è avvenuta interamente tramite la piattaforma di Simulink, strettamente integrata al software di MATLAB, per mezzo del quale è stata effettuata l'analisi di segnali.

## 2. Descrizione del set-up sperimentale

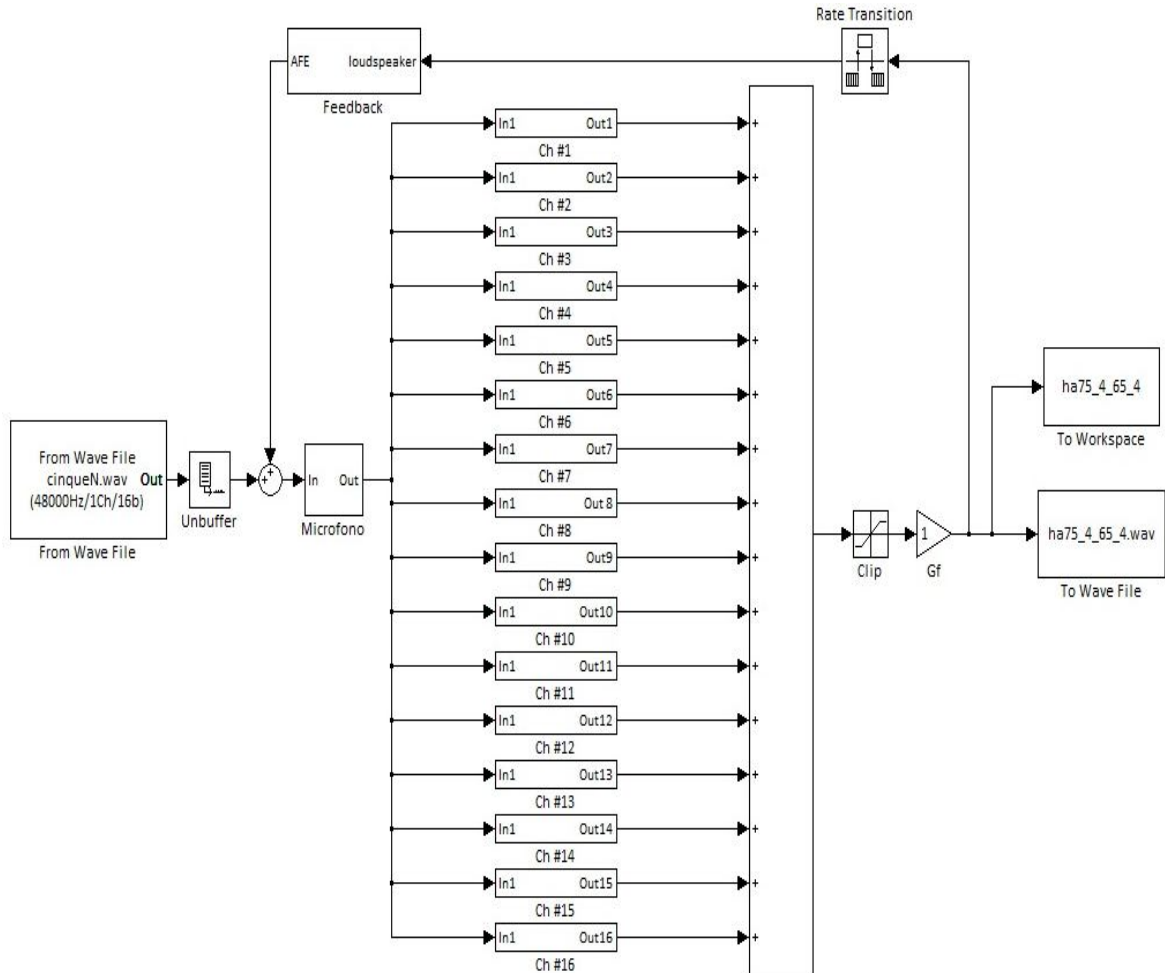


Figura 23: schema del simulatore implementato. Sono visibili il microfono, il banco di filtri, un sistema di peak-clipping e il blocco che simula il feedback acustico.

Il simulatore implementato, il cui schema è visibile in figura 23, rappresenta un modello *in-the-ear* (ITE) di apparecchio acustico digitale a sedici canali: un segnale di ingresso in formato *.wav* viene convertito, tramite un *unbuffer*, in un segnale utilizzabile da Simulink. Tale segnale, dopo aver subito una prima elaborazione dal microfono, è scomposto in 16 bande di frequenza, differenzialmente processato nei vari canali (indicati dal simbolo Ch# seguito dal rispettivo numero del canale) e, successivamente, ricostruito come somma dei segnali in uscita dai 16 canali. Il segnale così ricostruito può eventualmente essere soggetto ad una fase di *hard clipping*, necessaria per limitare il range dinamico dell'uscita e subire

un'ulteriore amplificazione. Come risultato si ottiene un segnale che viene indirizzato ad un ipotetico altoparlante, l'uscita del quale verrà registrata in due formati (.wav e .mat) e, contemporaneamente, costituirà l'ingresso del blocco denominato *Feedback*. Come si può intuire, tale blocco simula l'effetto del feedback acustico, sempre presente nelle protesi reali e dovuto al fatto che il segnale in uscita dall'altoparlante della protesi non solo incide sulla membrana del timpano del portatore, ma anche su quella del microfono, sommandosi al segnale acustico in ingresso.

## 2.1. Microfono

Il microfono implementato simula la funzione di trasferimento del microfono analogico Knowles EA-1842, rappresentato da un filtro passa-alto a un polo con frequenza di taglio pari a 300 Hz, in cascata a un filtro passa-basso a due poli, con frequenza di taglio pari a 5100 Hz e fattore di qualità Q di 1.55.

La funzione di trasferimento complessiva è rappresentata dalla seguente funzione nel dominio di Laplace:

$$M(s) = \frac{s}{s + 2\pi f_p} \frac{(\omega_c)^2}{s^2 + \frac{\omega_c}{Q}s + \omega_c^2} = \frac{s}{s + 600\pi} \frac{(10200\pi)^2}{s^2 + \frac{10200\pi}{1.55}s + (10200\pi)^2}$$

e la rispettiva risposta in frequenza è illustrata in figura 24

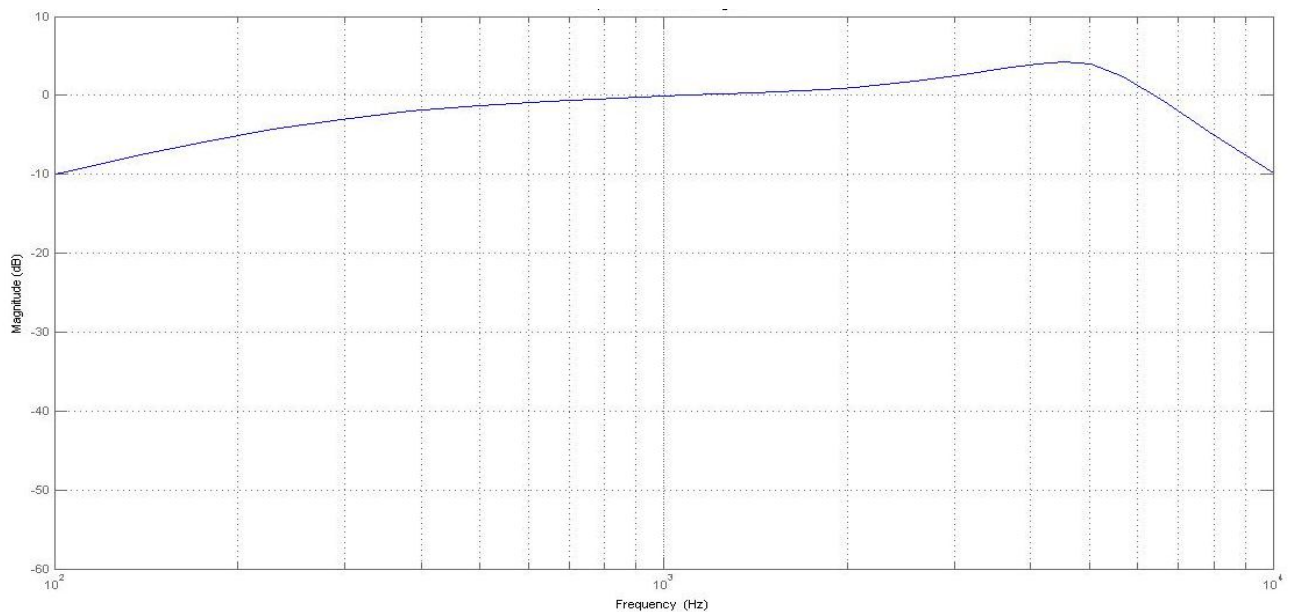


Figura 24: risposta in frequenza del microfono.

Un microfono di questo tipo, che presenta un'attenuazione di 0 dB a 1000 Hz, è in grado di fornire una risposta sufficientemente piatta su tutto il range delle frequenze del parlato.

## 2.2. Processore digitale

Il processore digitale è rappresentato da 16 canali, ognuno dei quali, come mostra la figura 25, contiene un filtro passa-banda ( $f(i)$  dove  $i$  rappresenta il numero del canale), un circuito rilevatore di picco ( $PD$ , uguale per tutti e 16 i canali), un sistema compressore e un amplificatore che fornisce il cosiddetto *guadagno d'inserzione*,  $G(i)$ . Il segnale in ciascun canale è quindi compresso e amplificato, in modo da adattarsi al range dinamico residuo del paziente nella banda di frequenza selezionata dal filtro passa-banda del canale. In linea generale, la selettività dei filtri non è uniforme su tutti i canali: se l'audiogramma mostra una significativa variazione della perdita uditiva nel dominio delle frequenze, ogni canale richiederà una selettività diversa rispetto a quanto avviene per lievi cambiamenti della perdita uditiva. È necessario quindi che i filtri siano programmabili. Nel presente studio, tuttavia, non si è fatto riferimento ad un soggetto ipoacusico specifico, per cui i 16 filtri passabanda, per quanto presentino larghezze di banda differenti, non vengono modificati nel corso degli esperimenti.

L'azione integrata dei circuiti rilevatore di picco e compressore rappresenta complessivamente un *Automatic Gain Control* (AGC), che permette di ottenere la compressione richiesta.

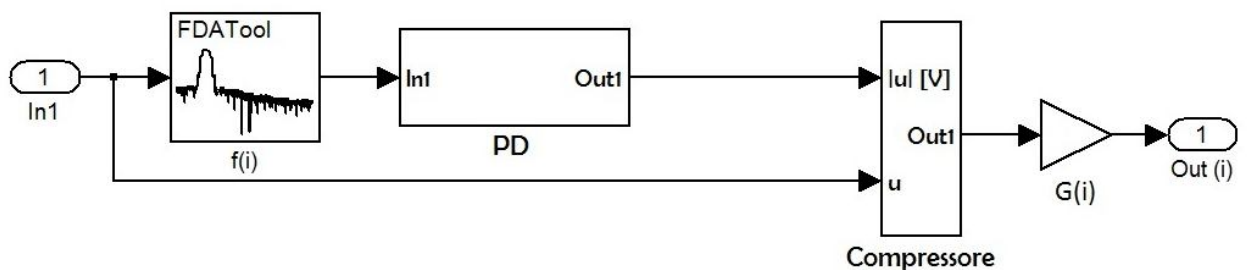


Figura 25: schema dei componenti all'interno di ciascuno dei sedici canali.

### 2.3. Banco di filtri

Il banco di filtri decompone il segnale d'ingresso in sedici "sotto-segnali" caratterizzati da un diverso contenuto in frequenza: la sua funzione primaria è quella di ottenere una risposta in frequenza programmabile dell'apparecchio acustico, in modo da compensare le elevate soglie uditive del portatore, ridurre il più possibile il rumore e realizzare una compressione multi-canale. Le specifiche del banco di filtri includono tipicamente l'attenuazione della stop-band (40-50 dB), il numero di canali e la spaziatura delle bande passanti, generalmente in scala di Bark, in modo da rispettare la tonotopicità della coclea. Come già accennato, nel presente studio sono stati utilizzati sedici canali, i cui filtri sono del 109° ordine, realizzati con la tecnica FIR (Finite Impulse Response) e una finestatura di tipo Hamming e campionati a 20kHz. È stata adottata, nello specifico, la tecnica LPFIR (Linear-Phase-Structured FIR) e la sua proprietà di simmetria principalmente per ridurre il costo computazionale. L'espressione di un filtro LPSFIR del 109° ordine è la seguente:

$$y^i[n] = \sum_{j=0}^{54} h^i[j] \cdot \{x[n-j] + x[n-109+j]\}$$

dove  $h^i[j]$  è il coefficiente per  $j$  (da 0 a 54) nel canale  $i$ ,  $y^i[n]$  è l'uscita del canale  $i$  al tempo  $n$  e  $x[n-j]$  è il campione in ingresso al tempo  $n-j$ .

Come è visibile anche dallo schema a blocchi, gli ingressi, costituiti dal segnale in uscita dal microfono, sono identici per tutti i canali.

Le frequenze centrali dei filtri passa-banda, mostrate in tabella 1, sono state scelte in modo tale da essere equispaziate nella scala di Bark. Le larghezze di banda dei primi quattro filtri sono costanti e pari a 250 Hz, mentre quelle dei filtri successivi crescono logicamente, proprio come avviene per la codifica delle frequenze all'interno della coclea.

N° canale	Frequenza centrale [Hz]	Bandwidth [Hz]
1	250 (frequenza di taglio) *	250
2	375	250
3	625	250
4	875	250
5	1137.5	275
6	1425	300
7	1737.5	325
8	2087.5	375
9	2475	400
10	2900	450
11	3375	500
12	3887.5	525
13	4500	700
14	5250	800
15	6125	950
16	7225	1250

Tabella 1: frequenze centrali e rispettive larghezze di banda dei 16 filtri passa-banda del banco.

\* Il filtro n°1 è un passa-basso del 109° ordine con frequenza di taglio 250 Hz.

Tali filtri sono inoltre progettati affinché si sovrappongano in corrispondenza di  $-6\text{dB}$  in quanto, durante la ricostruzione del segnale, le uscite dei singoli canali vengono sommate nel dominio del tempo. Pertanto, il vantaggio offerto da questi filtri è la possibilità di ottenere una risposta sufficientemente piatta per ogni valore di guadagno, proprietà non condivisa da filtri di tipo differente. La risposta in frequenza reale totale dei 16 filtri è illustrata in figura 26.

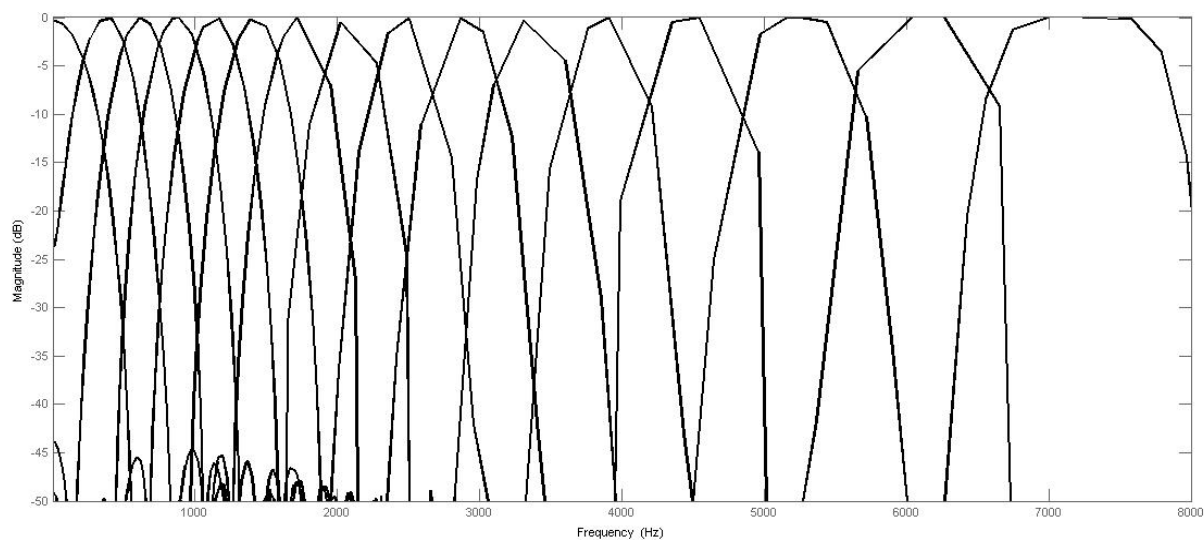


Figura 26: risposta in frequenza dei filtri passa-banda dei sedici canali.



#### 2.4. Automatic Gain Control (AGC)

Il controllo automatico del guadagno è fornito dall'azione complessiva del circuito rilevatore di picco e del sistema compressore.

Il segnale in uscita dal filtro del canale, dopo essere stato rettificato viene sottoposto ad un sistema di rilevatore di picchi, costituito da due resistenze, un diodo e un condensatore connessi come in figura 27. Affinché venga effettivamente rilevato un picco di durata superiore ad un valore prestabilito (la cui amplificazione sarebbe fastidiosa o dolorosa per il portatore), è necessario che il condensatore si carichi, fenomeno che avviene con un tempo pari al prodotto di  $R1$  per la capacità  $C$ . Al diminuire dell'ampiezza del segnale, il condensatore si scaricherà, grazie alla presenza del diodo, sulla resistenza  $R2$ , con un tempo di scarica dato dal prodotto di  $C$  ed  $R2$ . I tempi di carica e scarica del condensatore corrispondono ai tempi di attacco e rilascio dell'AGC.

Tale circuito è identico per tutti i canali ed è stato dimensionato con  $R1 = 1k\Omega$ ,  $R2 = 50k\Omega$  e  $C = 1pF$ , in modo da avere un tempo di attacco di 1ms e un tempo di rilascio di 50ms.

Per il corretto funzionamento dell'AGC, è di fondamentale importanza che il tempo di rilascio sia molto maggiore di quello d'attacco: è dunque essenziale la presenza del diodo, in assenza del quale il condensatore si scaricherebbe sulla resistenza  $R1$ , con il risultato di tempi di attacco e rilascio identici.

I blocchi identificati da *Voltage Source* e *Voltage Sensor* sono necessari per convertire il segnale "virtuale" usato nell'ambiente di Simulink in segnale elettrico e per riconvertirlo nella forma precedente, nuovamente disponibile per ulteriori elaborazioni. Il blocco identificato da *Solver Configuration* è invece essenziale per poter simulare, all'interno dell'ambiente di Simulink, una rete fisica reale, rappresentata da componenti elettronici e fili elettrici reali.

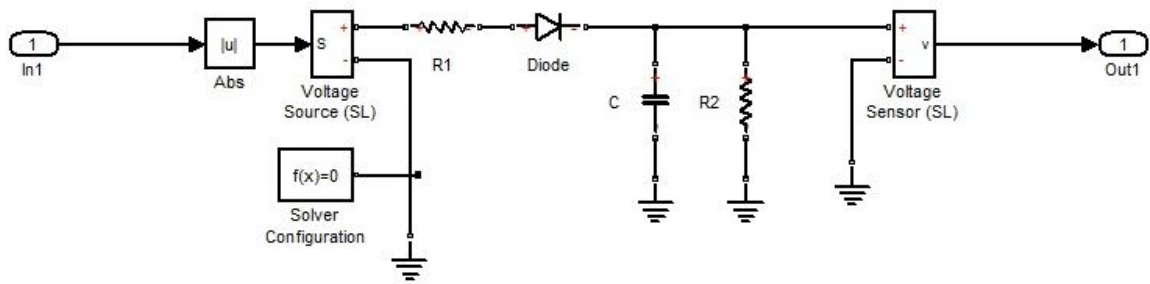


Figura 27: circuito rilevatore di picchi (*peak detector*). Il segnale in uscita dal filtro viene rettificato e permette di caricare il condensatore se la durata del picco è maggiore del tempo di attacco ( $R1 \cdot C$ ). Al diminuire dell'ampiezza del segnale, il condensatore si scaricherà sulla resistenza  $R2$ , con un tempo di rilascio pari a  $R2 \cdot C$ . L'uscita del *peak detector* è la differenza di potenziale ai capi del condensatore.

Il segnale in uscita dal *peak detector* (differenza di potenziale ai capi del condensatore) costituirà dunque l'ingresso di un circuito compressore, realizzato tramite la traduzione della struttura *if-then-else*, tipica del linguaggio C, nell'ambiente di Simulink. Il sistema implementato è illustrato in figura 28.

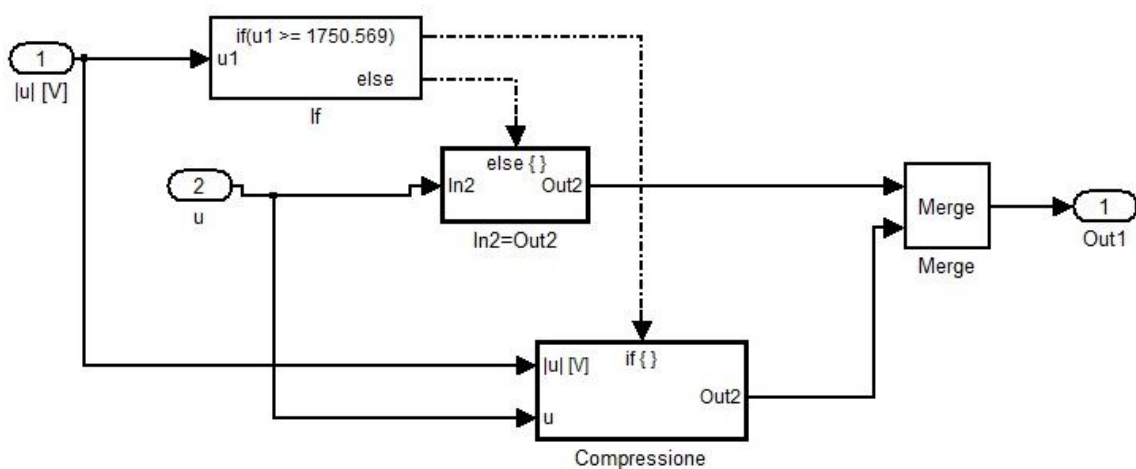


Figura 28: circuito compressore. Se il segnale in uscita dal *peak detector* è superiore alla soglia stabilita, viene compresso tramite l'applicazione di un fattore moltiplicativo minore di 1, in funzione dell'intensità del segnale in uscita dal filtro ( $u$ ) e del segnale rettificato in uscita dal *peak detector* ( $|u|$ ). In caso contrario il segnale si propaga inalterato.

Se l'ampiezza del segnale in uscita dal *peak detector* (indicata con  $|u|$ ) è inferiore alla soglia di compressione stabilita, il segnale in uscita dal filtro ( $u$ ) non viene compresso e si propaga inalterato, mantenendo il sistema lineare.

Se invece  $|u|$  supera la soglia dell'AGC, il segnale  $u$  entra nel blocco compressore e viene ad esso applicato un fattore moltiplicativo che è funzione del segnale  $|u|$  e del rapporto di compressione prescelto.

Le soglie di compressione utilizzate nel presente studio sono pari a 65 e 75 dB SPL e sono state specificate in termini di tensione equivalente in uscita dal microfono per un ingresso sinusoidale di frequenza 1kHz: i valori misurati sono 1750.569 e 5535.7853 V, che corrispondono rispettivamente a 65 e 75 dB SPL.

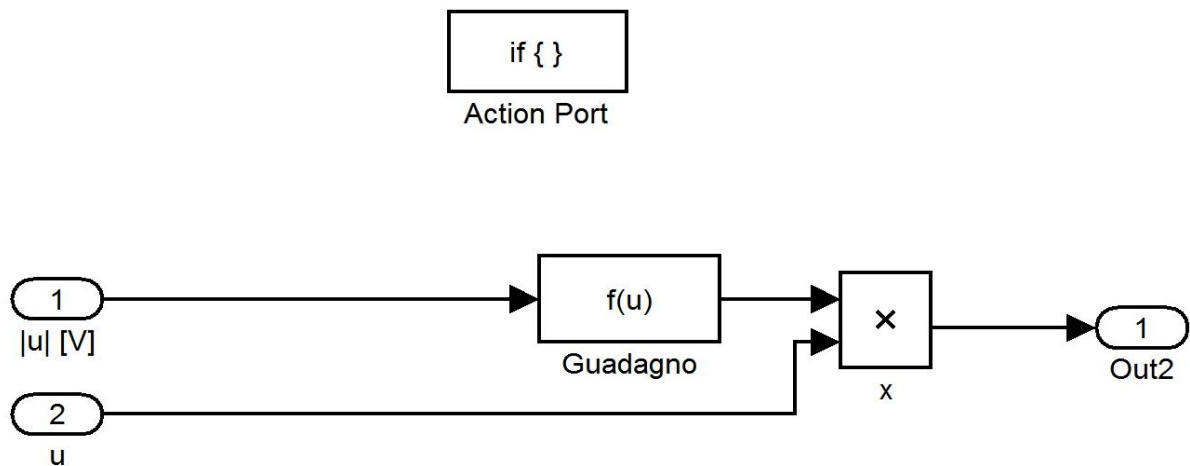


Figura 29: nucleo della compressione. Il segnale  $u$  viene moltiplicato per un fattore dipendente dal segnale rettificato in uscita dal PD.

In base alla soglia e al rapporto di compressione stabiliti, il guadagno (o fattore moltiplicativo) applicato al segnale filtrato può essere espresso per mezzo delle seguenti formule:

$$G = \frac{10^{1.625}}{\sqrt{|u|}} \quad (\text{Soglia} = 65 \text{ dB SPL; RC} = 2:1)$$

$$G = \frac{10^{2.3875}}{\sqrt[4]{|u|}} \quad (\text{Soglia} = 65 \text{ dB SPL; RC} = 4:1)$$

$$G = \frac{10^{1.8875}}{\sqrt{|u|}} \quad (\text{Soglia} = 75 \text{ dB SPL; RC} = 2:1)$$

$$G = \frac{10^{2.8125}}{\sqrt[4]{|u|}} \quad (\text{Soglia} = 75 \text{ dB SPL; RC} = 4:1)$$

In questo modo è stato dunque possibile ottenere diversi tipi di compressione: al di sotto della soglia dell'AGC, l'uscita corrisponde all'ingresso, il che equivale alla propagazione inalterata del segnale. Quando invece si supera la soglia di compressione, viene fornito un guadagno linearmente dipendente, in scala logaritmica, dall'ingresso (grafico 1). In generale, una maggiore compressione del segnale può essere ottenuta tramite un basso punto di ginocchio e un alto rapporto di compressione, come è possibile dedurre dalla curva verde, corrispondente ad una soglia di 65dB SPL e ad un rapporto di 4:1. Al contrario, una lieve compressione è invece fornita da un AGC che presenti una soglia piuttosto elevata e un basso rapporto di compressione, situazione rappresentata per esempio dalla curva viola, corrispondente ad una soglia di 75dB SPL e un rapporto pari a 2:1. In aggiunta, all'aumentare del rapporto di compressione, si può ottenere una situazione sempre più simile a quella dell'*hard clipping*, la cui curva caratteristica IN/OUT sopra soglia è rappresentata da una semiretta orizzontale.

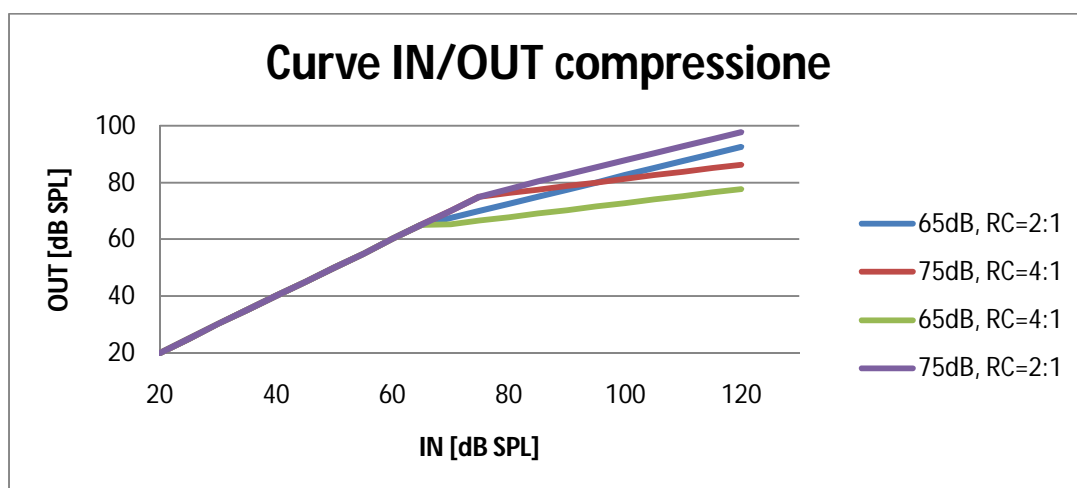


Grafico 1: diversi tipi di compressione al variare della soglia e del rapporto di compressione (RC).

Trasformando le unità logaritmiche in unità lineari, è possibile ricavare l'andamento del guadagno in funzione dell'ingresso: si vede infatti nel grafico 2 che, per bassi valori d'ingresso, il guadagno rimane costante 1, dopodiché, all'aumentare del livello di ingresso, decresce con andamenti di tipo logaritmico negativo.

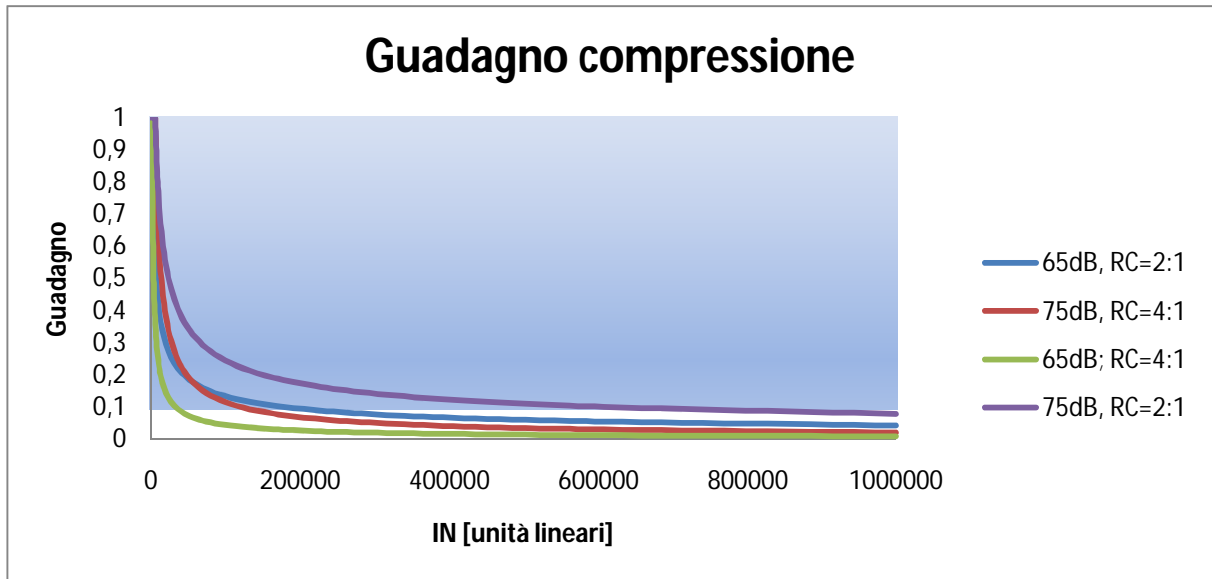


Grafico 2: guadagno della compressione in funzione dell'ingresso.

Tale guadagno non è tuttavia da confondere con il guadagno dell'amplificatore in serie al compressore, il cui valore è definito separatamente per ogni canale di frequenza, in base alla tecnica prescrittiva prescelta e all'audiogramma del paziente. In linea generale, verrà fornita un'amplificazione maggiore per le frequenze in cui è stata riscontrata una perdita uditiva rilevante.

Dopo la fase di elaborazione all'interno del processore digitale, i segnali provenienti da ciascun canale vengono sommati tra loro per ricostruire il segnale finale che sarà pronto per essere trasformato dal ricevitore in un'onda di pressione sonora, udibile da un ipotetico portatore della protesi acustica.

L'eventuale distorsione causata dall'amplificatore è stata simulata per mezzo di una fase di clipping simmetrico (blocco denominato *Clip*).

Nel presente studio non è stato necessario implementare un ricevitore, dal momento che l'obiettivo non era quello di progettare un chip per DSP, bensì un'analisi di segnali, con il fine ultimo di produrre un buon clear speech.

Il blocco nominato *Rate Transition*, presente nel percorso di feedback, è semplicemente un accorgimento che permette di connettere un sistema digitale, caratterizzato da una frequenza di campionamento definita, ad un blocco in cui è presente una funzione di trasferimento espressa nel dominio di Laplace, tipico dei sistemi analogici.

Discorso analogo vale per il blocco *unbuffer*, che ha la funzione di convertire il segnale d'ingresso, di tipo “*frame-based*”, in una forma compatibile con il microfono.

## 2.5. Feedback acustico

Il feedback acustico negli apparecchi acustici è provocato da svariati fattori, come la posizione relativa tra il microfono e l'altoparlante, il foro di ventilazione, i difetti dello stampo auricolare, le emissioni delle pareti dei tubicini, la trasmissione strutturale, le vibrazioni meccaniche del guscio del dispositivo, la trasmissione acustica e il feedback elettrico. Il parametro più importante è la distanza tra l'altoparlante e il microfono, per cui l'effetto di feedback è più accentuato per distanze minori. Questa è la ragione per cui gli apparecchi acustici *in-the-ear* (ITE) e *in-the-canal* (ITC) sono maggiormente soggetti a questo effetto rispetto ai modelli di tipo *behind-the-ear* (BTE).

È quindi possibile considerare un unico segnale di feedback acustico, ottenuto come combinazione di tutti i fattori agenti negli apparecchi acustici digitali. Questo si basa su considerazioni empiriche di studi precedenti, per le quali le risposte in frequenza del segnale di feedback di diverse categorie di apparecchi acustici condividono diverse proprietà. Il diagramma di Bode dell'intensità è approssimativamente piatta fino a 2kHz, presenta un picco di bassa qualità a 4kHz e un avvallamento a 7kHz. Il diagramma della fase è invece praticamente lineare su tutte le frequenze, ad eccezione di quelle più alte, dove la risposta subisce un leggero slittamento di fase.

È stato per cui possibile rappresentare il feedback acustico per gli apparecchi di categoria *in-the-ear* tramite la seguente funzione di trasferimento nel dominio di Laplace:

$$AFE(s) = Ke^{-sT} \frac{(s - z_1)(s - z_2)(s - z_2^*)}{(s - p_1)(s - p_1^*)(s - p_2)(s - p_2^*)}$$

dove  $K$  è un fattore di guadagno usato per controllare l'intensità (in questo caso impostato a 1),  $T$  è il ritardo di propagazione nel condotto uditivo dall'altoparlante al microfono,  $z_1$  e  $z_2$  gli zeri e  $p_1$  e  $p_2$  i poli del filtro. Lo zero singolo e una coppia di poli complessi coniugati sono necessari per controllare la pendenza fino a 4kHz, mentre la seconda coppia di poli complessi coniugati servono ad ottenere un picco a frequenze maggiori. Per dare un fattore di qualità migliore a questo picco, i poli devono essere posizionati vicino all'asse

immaginario. La coppia di zeri complessi coniugati, infine, consente di ottenere l'avvallamento alle frequenze più alte.

Considerando una distanza media  $d$  tra altoparlante e microfono di 17,6mm e la velocità del suono  $v$  pari a 340m/s, il ritardo di propagazione  $T$  sarà dato da  $T = d/v = 51.76 \mu\text{s}$ .

A zeri e poli sono invece stati attribuiti i seguenti valori:

$$z_1 = -1500000$$

$$z_2 = -40000 + 31400j$$

$$p_1 = -4500 + 25000j$$

$$p_2 = -2500 + 40800j$$

La risposta in frequenza del segnale di feedback simulato è riportata nel grafico 3, nel quale è chiaramente visibile l'andamento predetto dell'ampiezza e della fase.

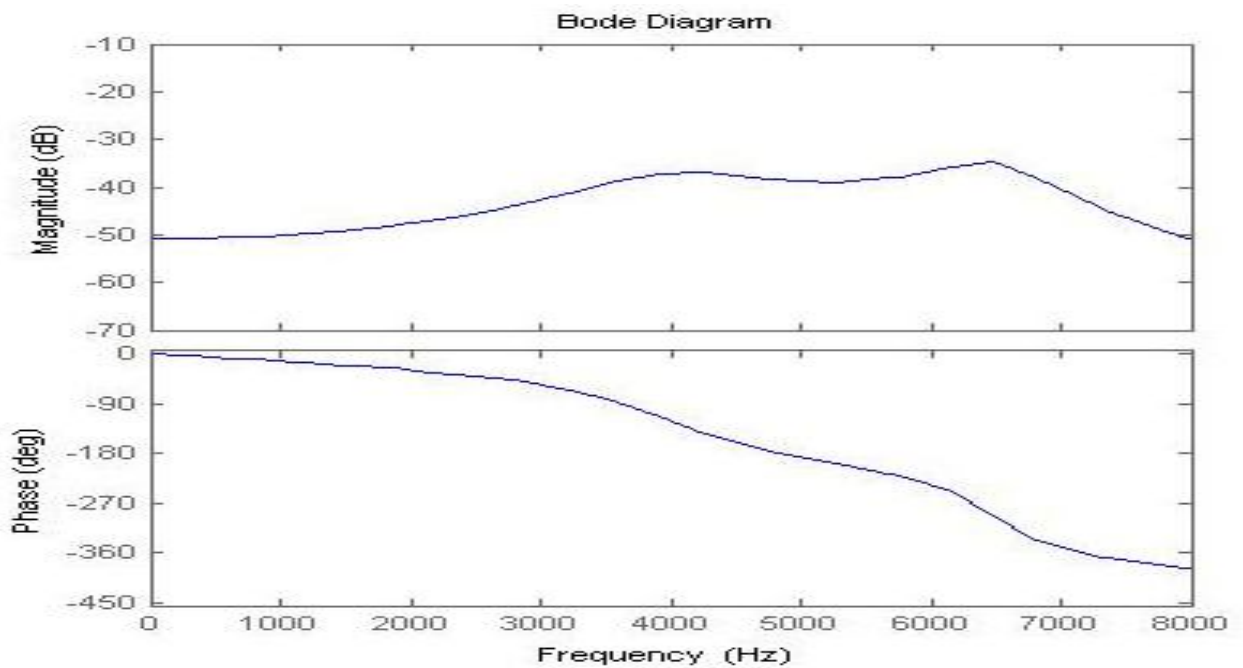


Grafico 3: diagramma di Bode di ampiezza e fase del feedback acustico.

### 3. Materiale vocale

Il materiale vocale utilizzato per l'estrazione delle *features* prestabilite (*modulation index* e formanti delle vocali) è costituito dalle cinque vocali della lingua italiana (a, e, i, o, u) e da due parole bisillabiche (salto e cinque).

Ciascuna vocale rappresenta la vocale accentata delle parole salto, felce, cinque, chiodi e uscio ed è stata estratta per mezzo del programma *Audacity 1.3 Beta*.

Si è scelto di utilizzare le vocali accentate in quanto sono quelle caratterizzate dalla massima durata e intensità sonora all'interno di una parola, quindi quelle maggiormente responsabili dell'intelligibilità.

Nei grafici 4, 5, 6, 7 e 8 è riportata la forma d'onda di ciascuna vocale: sull'asse delle ascisse è rappresentato il tempo mentre sull'asse delle ordinate l'intensità sonora.

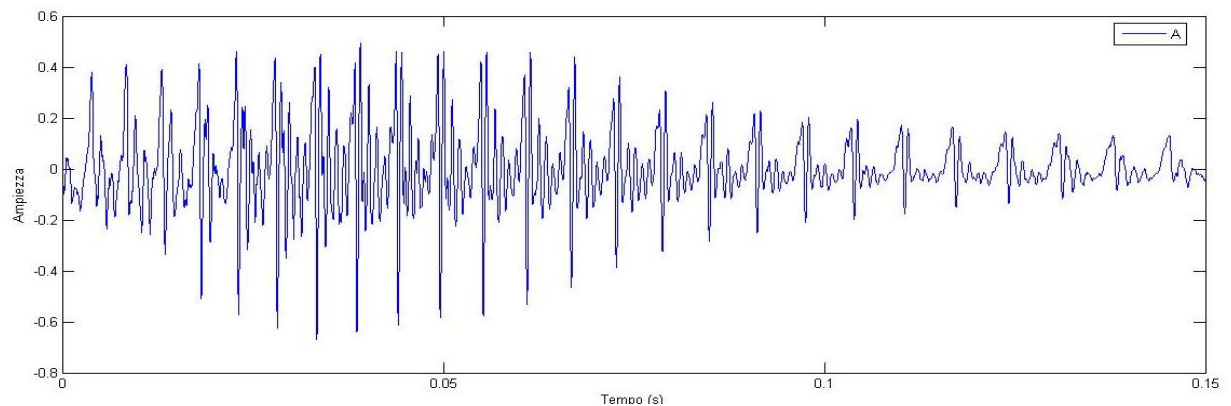


Grafico 4: forma d'onda della vocale A.

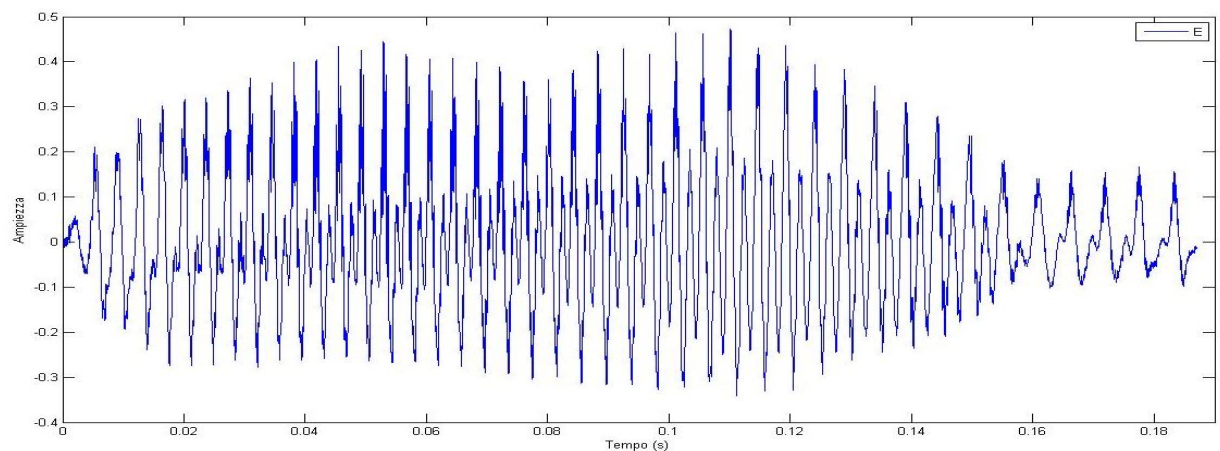


Grafico 5: forma d'onda della vocale E.



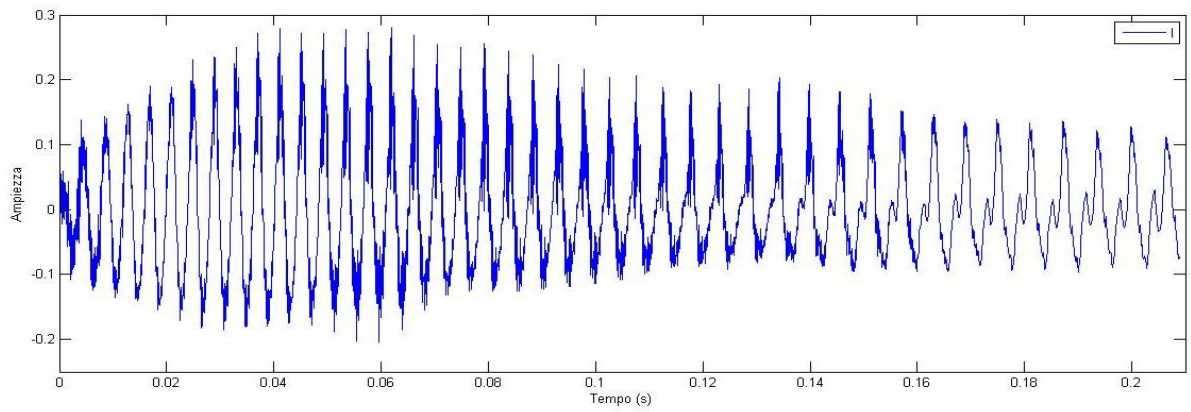


Grafico 6: forma d'onda della vocale I.

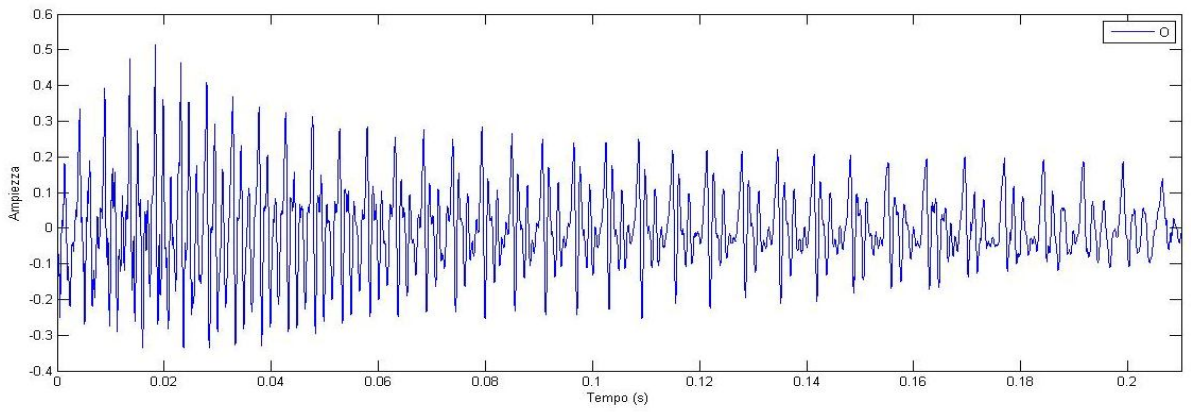


Grafico 7: forma d'onda della vocale O.

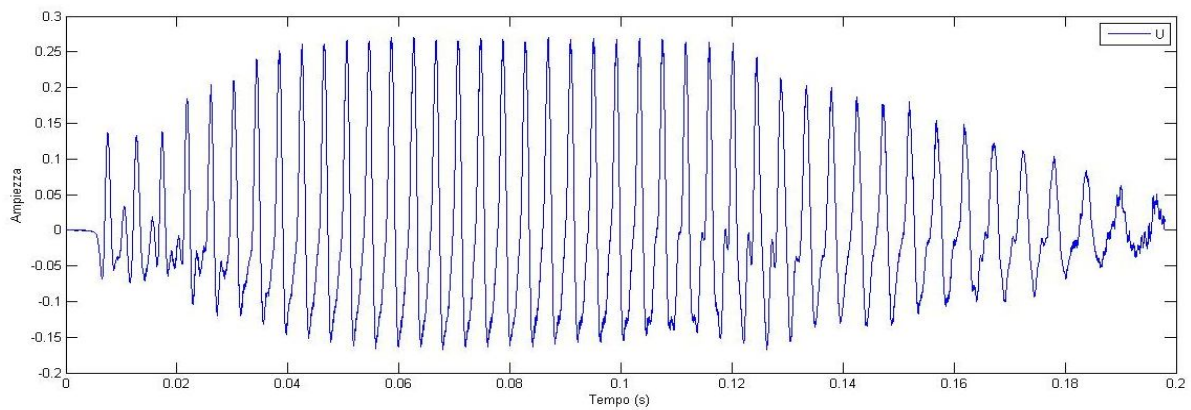


Grafico 8: forma d'onda della vocale U.

Le parole salto e cinque sono state invece scelte casualmente ed estratte da una lista di undici parole bisillabiche in formato stereo *.wav*, per mezzo del programma suddetto.

La forma d'onda di tali parole è riportata nel grafico 9.

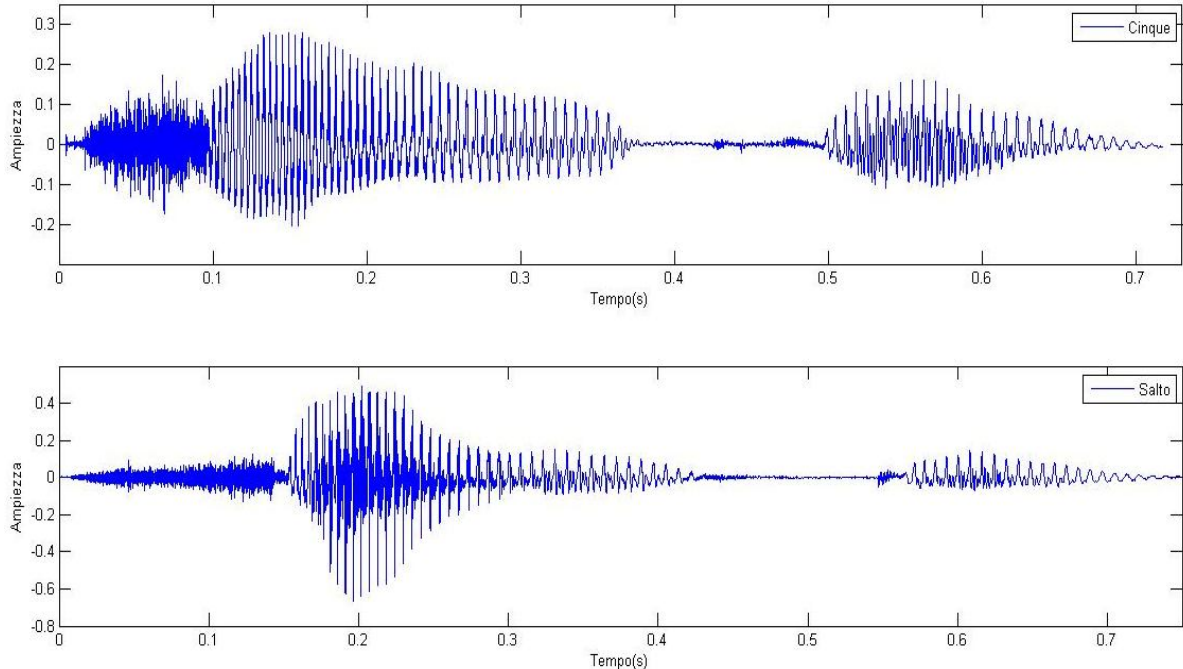


Grafico 9: forma d'onda delle parole "cinque" (in alto) e "salto" (in basso).

I segnali stereo sono tipicamente costituiti da due canali: in questo caso il primo contiene la base di rumore mentre il secondo il segnale vocale stesso. Dal momento che l'obiettivo del presente studio non è quello di ridurre il rumore, bensì quello di valutare il comportamento del *modulation index* e delle formanti delle vocali in funzione della variazione della soglia dell'AGC e del rapporto di compressione, ciascun segnale vocale è stato scomposto in due segnali mono e successivamente depurato dalla componente di rumore.

Le due parole e le cinque vocali sono state poi importate nell'ambiente MATLAB e quindi convertite dal formato *.wav* a *.mat*, rendendole in questo modo utilizzabili dai simulatori implementati.

Dopo la fase di simulazione, tutti i segnali registrati in uscita dal simulatore sono stati normalizzati alla stessa ampiezza di *root-mean-square* (RMS), prima che venisse eseguita

l'analisi spettrale, per eliminare le differenze di ampiezza tra i segnali originali e quelli in uscita dal simulatore, che risultano evidentemente amplificati.

#### **4. Descrizione delle *features* estratte**

##### **4.1. Modulation Index**

Il segnale di speech è caratterizzato da uno spettro variabile nel tempo: tali variazioni contengono l'informazione essenziale per l'identificazione di fonemi, sillabe, parole e frasi.

A questo proposito, nel presente studio è stata focalizzata l'attenzione sull'involuppo temporale: le modulazioni temporali del segnale di speech possono essere descritte in termini di un parametro denominato "modulation index". Un eventuale decremento di tale parametro corrisponde ad una diminuzione dell'energia spettrale in una determinata banda di frequenza, che si traduce in una riduzione delle fluttuazioni temporali del segnale di speech. Da tutto ciò deriva una forte perdita di intelligibilità.

Ai fini del presente studio è quindi necessario studiare ed analizzare il modo in cui il simulatore trasferisce le fluttuazioni dell'involuppo temporale del segnale vocale originale.

Per ricavare il modulation index, che contiene tutte le informazioni richieste, si deve prima di tutto estrarre l'involuppo temporale (ossia la curva che interpola tutti i picchi positivi del segnale) tramite la trasformata di Hilbert, e ricavarne poi lo spettro in frequenza, tramite la trasformata di Fourier (FFT). A questo punto si potrà fare un confronto tra lo spettro dell'involuppo del segnale originale e quello dei segnali elaborati e andare a verificare come sono state trasferite le fluttuazioni temporali e se queste potrebbero portare ad una perdita di intelligibilità.

##### **4.2. Formanti e *vowels space***

Il termine formante indica picchi dello spettro sonoro della voce ed indicano quindi le frequenze di risonanza del tratto vocale.

In prima approssimazione, una specifica vocale è identificata dalla prima e seconda formante (F1 ed F2), mentre le altre due (F3 ed F4) definiscono sfumature linguistiche e personali. Nel piano F1-F2, ogni vocale occupa quindi uno specifico punto, le cui coordinate variano a seconda del tipo di linguaggio adottato. In studi precedenti, si è

riscontrato un aumento di F1 ed F2 in condizione di clear speech rispetto ad un parlato di tipo discorsivo. L'aumento delle formanti di ogni vocale si traduce perciò in un aumento del cosiddetto *vowels space*, ossia l'area del poligono ottenuto congiungendo i punti nel piano F1-F2 corrispondenti a ciascuna vocale.

Nel presente studio, assumendo il segnale originale come punto di riferimento di clear speech, si eseguirà un confronto tra le formanti di suddetto segnale e quelle ricavate dai segnali elaborati, per analizzare il loro comportamento in questa situazione e verificare un'eventuale perdita di intelligibilità.

### **5. Protocollo di *speech processing***

Nella fase di elaborazione dello *speech*, i segnali vocali sono stati processati in modo da poter osservare il comportamento delle *features* stabilite in funzione della variazione della soglia e del rapporto di compressione dell'AGC. Ai fini di una semplice analisi ed interpretazione dei risultati, sono stati analizzati solo due valori di soglia (65 e 75dB) e di rapporto di compressione (2:1 e 4:1), che sono comunque quelli più comunemente usati nelle protesi acustiche moderne. Sono state perciò testate sedici diverse configurazioni (illustrate in tabella 2), nelle quali si osservano valori diversi di soglia e rapporto di compressione (RC) per i canali a bassa e ad alta frequenza. I cosiddetti canali a bassa frequenza sono i primi sette, i cui filtri passa-banda permettono l'elaborazione delle frequenze dello *speech* inferiori a 2kHz, mentre i canali ad alta frequenza sono deputati al trattamento delle frequenze superiori a 2kHz.

Osservando la tabella, si può vedere che le prime quattro configurazioni mantengono la stessa soglia e rapporto di compressione sia per le frequenze gravi sia per quelle acute, la 5 e la 6 e la 9 e la 10 sono caratterizzate dalla stessa soglia ma diversi rapporti di compressione per le basse e alte frequenze mentre le rimanenti otto possono essere definite configurazioni "ibride", in cui entrambi i parametri di interesse si differenziano nei due tipi di canali.

N°configurazione	CANALI A BASSA FREQUENZA		CANALI AD ALTA FREQUENZA	
	SOGLIA [dB]	RC	SOGLIA [dB]	RC
1	65	2:1	65	2:1
2	65	4:1	65	4:1
3	75	2:1	75	2:1
4	75	4:1	75	4:1
5	65	2:1	65	4:1
6	65	4:1	65	2:1
7	65	2:1	75	2:1
8	65	2:1	75	4:1
9	75	2:1	75	4:1
10	75	4:1	75	2:1
11	75	2:1	65	2:1
12	75	4:1	65	4:1
13	65	4:1	75	2:1
14	65	4:1	75	4:1
15	75	2:1	65	4:1
16	75	4:1	65	2:1

Tabella 2: elenco delle sedici configurazioni di AGC studiate.

Per poter eseguire una prima stima del funzionamento del simulatore di apparecchio acustico implementato e stabilire una *baseline* di riferimento per le simulazioni dei segnali “soprasoglia”, ciascuna vocale ed entrambe le parole sono state testate in condizioni standard, ossia tali da non attivare mai i sistemi compressori (AGC). In questo modo è inoltre possibile valutare l’effetto globale della funzione di trasferimento del microfono e dei sedici filtri sul segnale in uscita dal simulatore, pronto per essere trasdotto dal ricevitore e inviato al timpano di un ipotetico portatore di apparecchio acustico.

La stima dell’effetto della soglia dell’AGC sul *modulation index* e sul *vowels space* è stata eseguita fornendo al segnale originale un’amplificazione tale da far sì che questo si trovi, durante l’intera durata della simulazione, sempre al di sopra della soglia di compressione

minore (65dB) e che manifesti solamente alcuni picchi che superano la soglia maggiore (75dB). In questa maniera gli effetti dovuti ad un diverso rapporto di compressione vengono minimizzati, evidenziando contemporaneamente le conseguenze di una diversa soglia di compressione sull'intelligibilità del segnale in uscita.

Per valutare invece le conseguenze della variazione del rapporto di compressione, è necessario che il segnale in ingresso al simulatore si trovi sempre al di sopra della soglia di 75dB, in modo da mantenere attivi tutti i sistemi compressor durante l'intera durata della stimolazione. Un segnale sempre nettamente superiore a 75dB consente di trascurare le conseguenze di una diversa soglia di compressione e pone l'accento sugli effetti derivanti dall'utilizzo di diversi rapporti di compressione.

L'amplificazione fornita nelle diverse condizioni di simulazione è stata calcolata per ciascun segnale in base alla forma d'onda e all'ampiezza dello stesso, così da soddisfare i requisiti delle diverse simulazioni.

I segnali così amplificati sono stati quindi processati dal simulatore nelle sedici diverse configurazioni, raggruppati secondo le due diverse condizioni di simulazione e successivamente sottoposti ad una analisi spettrale.

### 5.1. Estrazione del *modulation index*

Per l'estrazione ed il plottaggio del *modulation index* di ciascun segnale derivante dall'elaborazione della parola "cinque", è stato utilizzato un codice MATLAB in grado di eseguire le operazioni successivamente descritte.

#### Calcolo del *root-mean-square* (rms)

$$rms_w = \sqrt{\frac{\sum_{j=1}^M w_j^2}{M}} \quad j = 1 \dots M;$$

$$rms_{x_i} = \sqrt{\frac{\sum_{j=1}^N x_{ij}^2}{N}} \quad j = 1 \dots N; \quad i = 1 \dots 16;$$

dove  $w$  è il segnale originale, costituito da  $M$  campioni e  $x_i$  è il segnale elaborato per la configurazione  $i$ , costituito da  $N$  campioni.  $M$  ed  $N$  sono numeri diversi in quanto

dipendono dalla diversa frequenza di campionamento dei segnali che, per quelli originali, è pari a 48kHz, mentre per quelli processati 20kHz.

#### Normalizzazione dei segnali elaborati rispetto all'rms del segnale originale.

Questo procedimento, eseguito per tutte le sedici configurazioni, è necessario per eliminare le differenze in ampiezza tra i due tipi di segnali.

$$xn_i = \frac{rms_w}{rms_{x_i}} \cdot x_i$$

#### Inviluppo temporale

Viene calcolato per mezzo della Trasformata di Hilbert, una particolare rappresentazione che, contrariamente ad altre trasformate (Fourier, Laplace, Z,...) non realizza un cambiamento del dominio di definizione, ma il suo risultato è un segnale complesso ancora in funzione del tempo. L'operazione di trasformazione di Hilbert si ottiene come uscita dal cosiddetto *filtro di Hilbert*, caratterizzato dalla funzione di trasferimento:

$$H_H(x) = \begin{cases} -i & \text{per } x > 0 \\ 0 & \text{per } x = 0 \\ i & \text{per } x < 0 \end{cases}$$

Si hanno così 16 segnali complessi  $H_i$ , che rappresentano l'inviluppo temporale del segnale di speech, di cui viene ricavato il modulo per il calcolo dello spettro:  $h_i = |H_i|$

#### Analisi spettrale

Infine, per poter eseguire una stima spettrale dell'inviluppo, è necessario applicare la Trasformata di Fourier ad  $h_i$ , il cui risultato verrà alla fine plottato e permetterà uno studio qualitativo del *modulation index*.

La trasformata di Fourier discreta (DFT) è definita dalla formula

$$Y_q = \sum_{k=0}^{NFFT-1} h_{ik} e^{-j \frac{2\pi}{NFFT} kq}$$

per  $q = 0 \dots NFFT - 1$ .

Come valore di *NFFT* si è scelta la prima potenza di 2 maggiore di *M* ed *N* (a seconda del segnale che si sta plottando), per migliorare la risoluzione in frequenza dello spettro. Essendo *NFFT* maggiore del numero di campioni del segnale, i dati vengono aggiustati con la tecnica dello *zero – padding*.

## 5.2. Estrazione delle formanti e costruzione del *vowels space*

Per l'analisi dello spazio vocalico, analogamente a come si è proceduto per l'estrazione del *modulation index*, sono state estratte le formanti F1 ed F2 da ciascuna vocale originale, “di controllo” e da ogni segnale vocalico derivante dalle sedici configurazioni dei due diversi tipi di simulazione.

Per l'estrazione delle formanti è stato utilizzato un complesso algoritmo in codice MATLAB, di cui viene riportata solamente la stringa di richiamo della funzione:

```
[fund, ffreq] = fundforms(signal, sampleFreq)
```

La funzione *fundforms* riceve in ingresso una stringa numerica contenente il segnale della vocale e la rispettiva frequenza di campionamento e dà in uscita il valore della frequenza fondamentale (*fund*) e un vettore colonna di tre elementi (*ffreq*) contenente le prime tre formanti della vocale analizzata, da cui sono quindi state estratte F1 ed F2.

Entrambi i valori di F1 ed F2 per ciascuna vocale e condizione di lavoro sono stati poi tabulati e plottati in diversi grafici per costruire ed analizzare il *vowels space*.



## CAPITOLO 5

### Risultati

In questa sezione verranno riportati i risultati ottenuti nella fase di elaborazione dei segnali processati dal simulatore di apparecchio acustico. Prima di tutto verranno presentati i risultati relativi alle due parole e, successivamente, quelli relativi alle vocali.

Si rammenta che sono stati eseguiti due tipi di simulazioni, una in cui il segnale originale si trova nettamente al di sopra di entrambe le soglie dell'AGC e una in cui il segnale è sempre maggiore della soglia di 65dB e presenta solo alcuni picchi al di sopra di 75dB.

Per dimostrare l'effettivo funzionamento dei sistemi compressor, per ciascuna parola e modalità di simulazione, verranno messi a confronto, su scala normalizzata, i valori di rms dei segnali elaborati nelle sedici diverse configurazioni. Naturalmente, il valore 1 indica l'assenza di compressione (come per i segnali non elaborati) e l'entità della compressione cresce al diminuire del valore di rms.

Dopodiché si illustreranno gli indici di modulazione dei segnali elaborati a confronto con quello del segnale originale, che ci si aspetta presenti un andamento qualitativo migliore e valori superiori ai segnali elaborati.

Nell'ultima sezione si riportano infine i risultati ottenuti per le cinque vocali della lingua italiana che sono state analizzate.

#### **1. Parole “salto” e “cinque”: *modulation index***

##### **1.1. AGC da 65dB e 75dB sempre in funzione**

Nei primi due grafici riportati, si osserva l'entità della compressione fornita dalle diverse configurazioni, caratterizzate da diverse soglie e rapporti di compressione per i canali a bassa e alta frequenza. Naturalmente, la massima compressione ottenibile con i simulatori a nostra disposizione è quella fornita dalla configurazione con la minima soglia (65dB) e il massimo rapporto di compressione (4:1) su tutti i canali. Al contrario, quella minima si ottiene impostando una soglia di 75dB e un basso rapporto di compressione, pari a 2:1, su tutti i canali. Per quanto riguarda le altre configurazioni, si osservano valori di compressione intermedia, che aumenta in maniera graduale al diminuire della soglia e all'aumentare del rapporto di compressione degli AGC nei diversi canali.

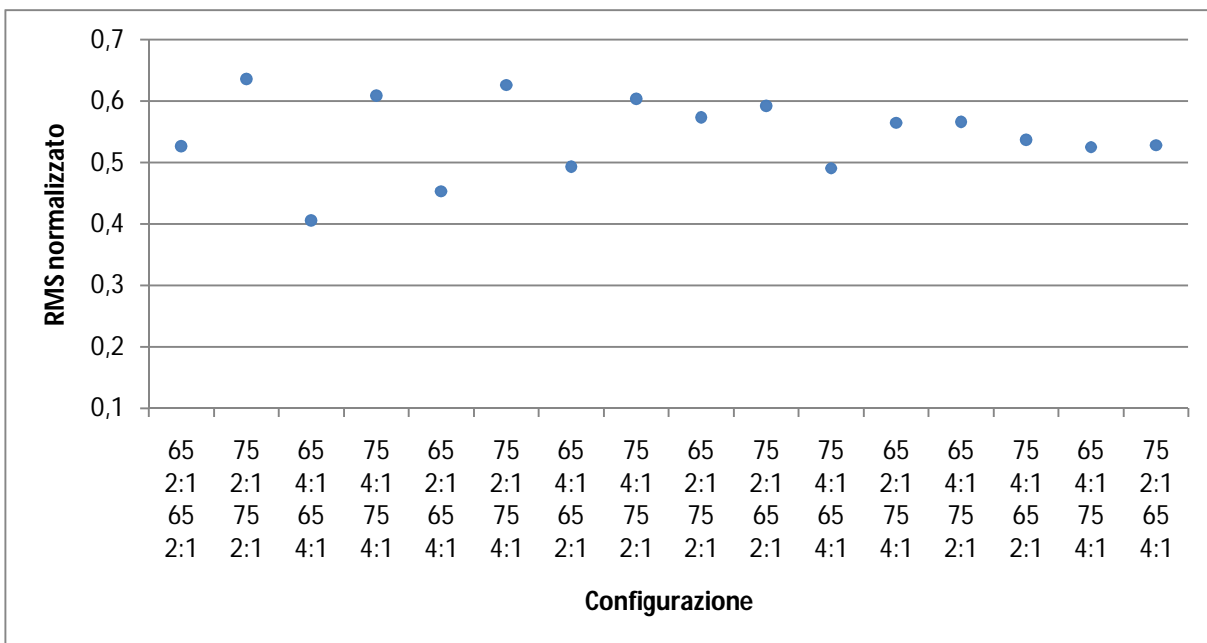


Grafico 100: entità della compressione (su scala normalizzata) subita dalla parola “salto” in funzione delle sedici diverse configurazioni del simulatore di apparecchio acustico digitale, con tutti gli AGC in funzione. Sull’asse delle ascisse sono riportate le sedici configurazioni studiate: prendendo a titolo esemplificativo la prima a destra, i primi due valori (75 e 2:1) rappresentano la soglia in dB e il rapporto di compressione degli AGC dei canali a bassa frequenza (canali 1-7), mentre gli ultimi due la soglia e il rapporto degli AGC dei canali ad alta frequenza (canali 8-16).

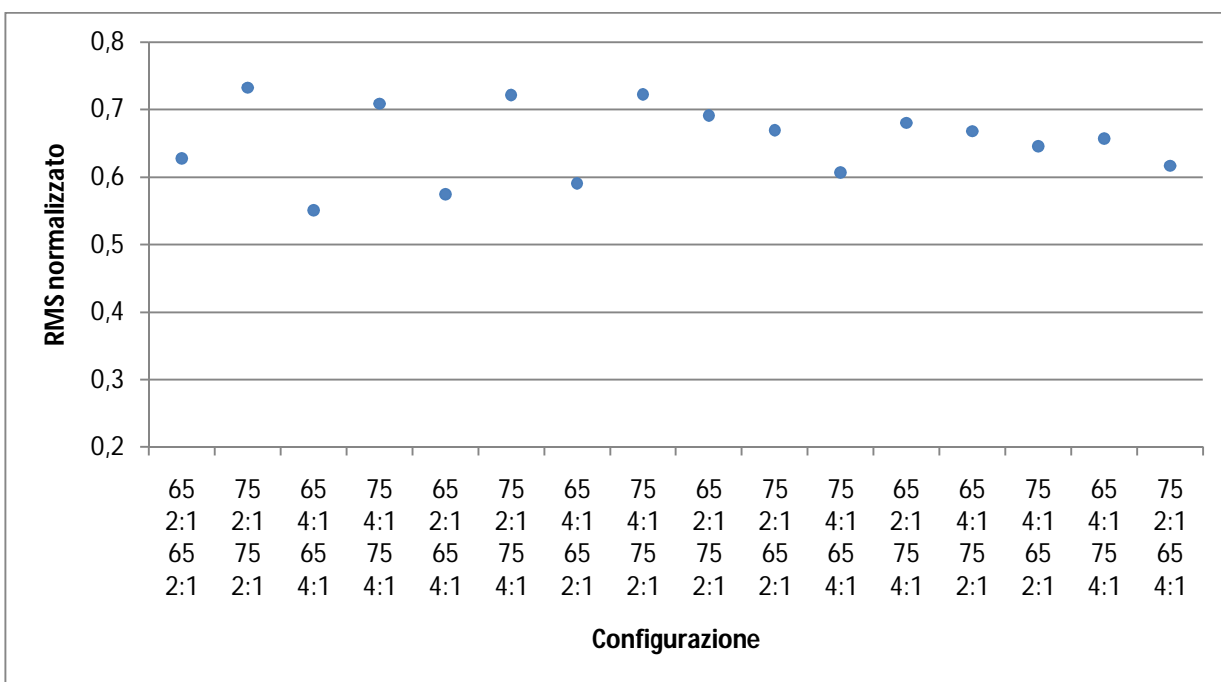


Grafico 11: entità della compressione (su scala normalizzata) subita dalla parola “cinque” in funzione delle diverse configurazioni del simulatore di apparecchio acustico digitale, con tutti gli AGC in funzione.

Nei grafici 12 e 13 sono stati posti a confronto gli indici di modulazione delle parole “salto” e “cinque” non elaborate con quelli ottenuti per le diverse configurazioni del simulatore, in una situazione di attivazione persistente di tutti gli AGC presenti. In questo modo è possibile studiare l’andamento di suddetto indice in funzione del rapporto di compressione.

Come ci si aspettava, si osserva una graduale diminuzione dell’indice di modulazione dei segnali processati rispetto al segnale originale non elaborato, che si potrebbe riflettere pertanto in una perdita progressiva di intelligibilità.

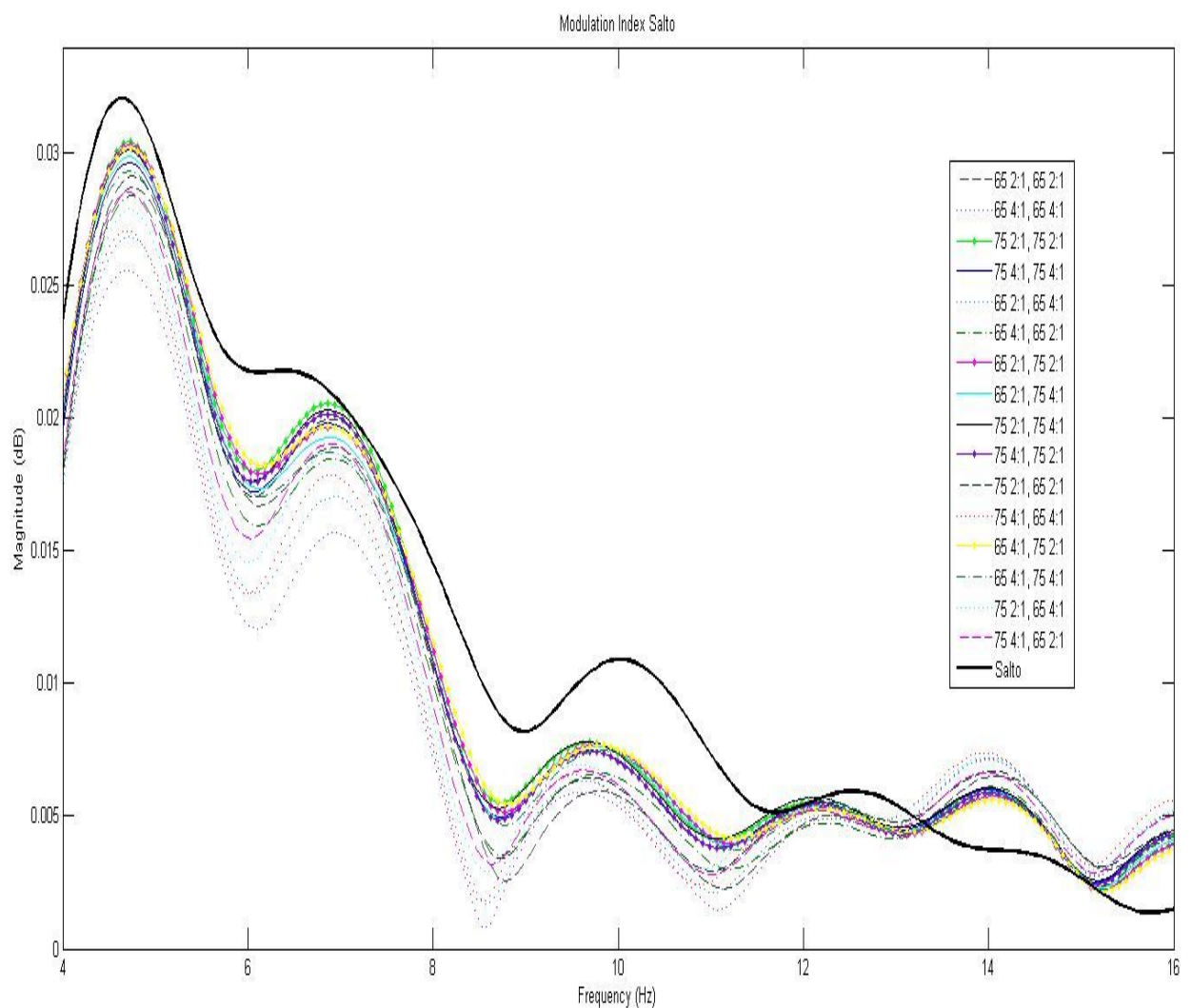


Grafico 12: spettro in frequenza dell'involucro temporale della parola "salto" con tutti gli AGC in funzione: confronto tra il segnale non elaborato (in nero) e quelli ottenuti nelle sedici diverse configurazioni del simulatore.

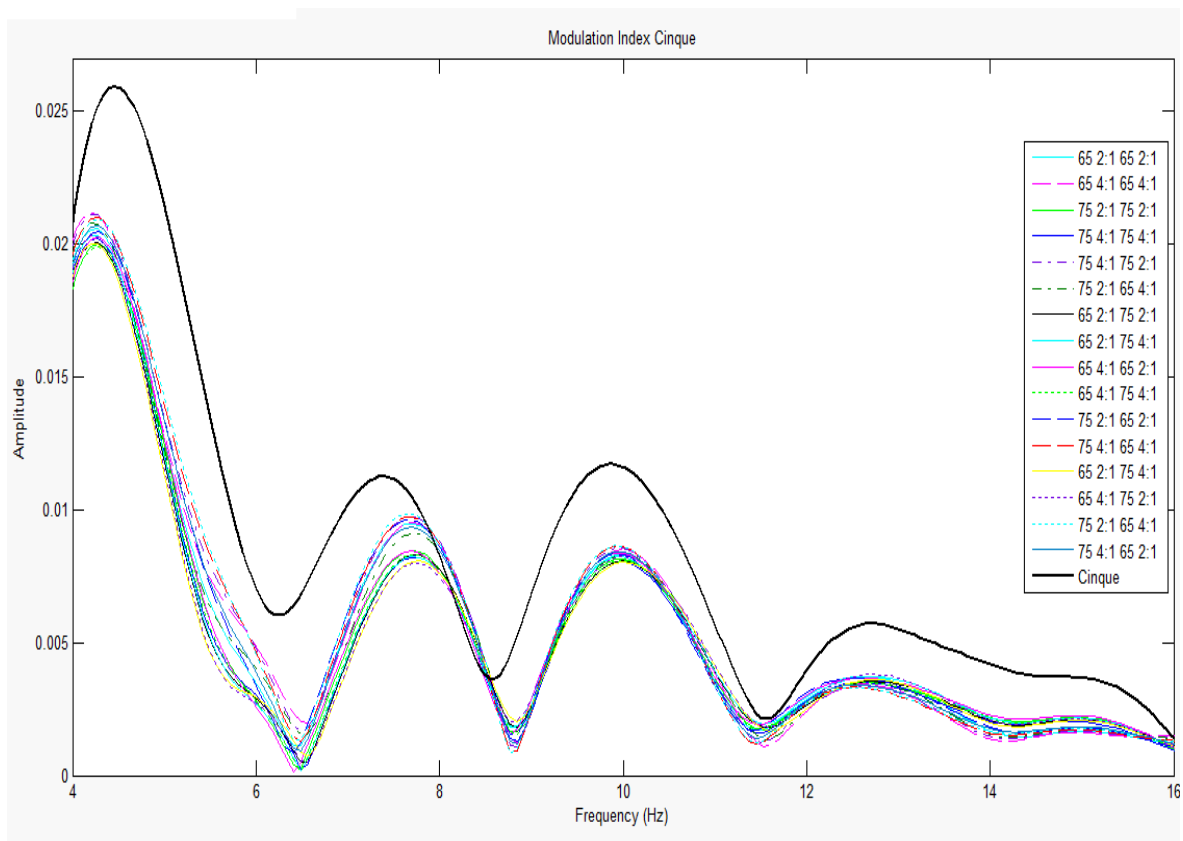


Grafico 13: spettro in frequenza dell'involuppo temporale della parola "cinque" con tutti gli AGC in funzione: confronto tra il segnale non elaborato (in nero) e quelli ottenuti nelle sedici diverse configurazioni del simulatore.

Da un'analisi qualitativa dei grafici dell'indice di modulazione ottenuti, è possibile fare alcune prime considerazioni globali.

Per entrambe le parole si osserva che i segnali elaborati mantengono indicativamente lo stesso andamento del segnale originale, nonostante siano comunque presenti leggere traslazioni delle frequenze centrali dei picchi e, ovviamente, perdite del *modulation index* (o *modulation depth*).

Lo spettro dell'involuppo temporale della parola "salto" è caratterizzato da un primo picco di ampiezza elevata a 4.65Hz, seguito da un *plateau* nel range di frequenze 6-7Hz, che prosegue con andamento decrescente approssimativamente lineare fino a 9Hz, dove si incontra il primo vero *dip* a 9Hz, seguito da un picco di bassa ampiezza e qualità a 10 Hz. Il segnale decresce poi con una successione di leggeri picchi e avvallamenti. I segnali elaborati, caratterizzati anch'essi da un primo picco elevato, presentano, all'aumentare della frequenza, un andamento con pendenza maggiore, una coppia *dip - peak* al posto del *plateau* originale seguito da un profondo avvallamento con frequenza centrale leggermente

inferiore rispetto a quella originale e variabile. Questo slittamento “all’indietro” delle frequenze centrali si rileva anche per i picchi e gli avvallamenti successivi. Infine, degno di nota, è il diverso andamento tra segnale originale ed elaborato che si osserva per le frequenze superiori a 12Hz, in particolare si ha la generazione di una successione picco – avvallamento – picco a 14, 15 e 16Hz.

Per la parola “cinque” si osserva, nello spettro dell’involuppo temporale, un primo picco molto accentuato ad una frequenza di circa 4.45Hz e picchi relativi dei segnali elaborati ad ampiezza minore e frequenza centrale di 4.2Hz. Si incontra poi un primo avvallamento a circa 6.25Hz, corrispondente ad avvallamenti più profondi a 6.5Hz dei segnali elaborati. All’aumentare della frequenza, si osservano due picchi, di ampiezza e qualità inferiori al primo, a 7.4 e 9.9Hz, rispettivamente, separati da un secondo avvallamento, leggermente più profondo, a 8.6Hz; nei segnali elaborati, i due picchi si trovano rispettivamente a 7.7Hz e 10Hz, mentre l’avvallamento, nettamente più profondo dell’originale, si trova a 8.8Hz. Dopo l’ultimo *dip* a 11.5Hz, lo spettro del segnale originale e di quelli elaborati presenta un andamento più dolce.

## 1.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente

Dal grafico 14 si osservano tre livelli di compressione per la parola “salto”: la compressione minima (livello più alto) è ottenuta per le configurazioni che presentano alte soglie degli AGC, il che deriva direttamente dalla condizione di attivazione intermittente dei sistemi compressor a soglia maggiore. Il livello di compressione intermedio si riferisce invece alle configurazioni caratterizzate da una soglia di 65dB e un rapporto di compressione di 2:1 negli AGC dei canali a bassa frequenza, mentre il terzo livello, relativo alla compressione massima ottenibile, si riferisce alle configurazioni con soglia minima (65dB) e rapporto massimo (4:1) degli AGC dei primi sette canali.

Un comportamento analogo è visibile anche per la parola “cinque” (grafico 15).

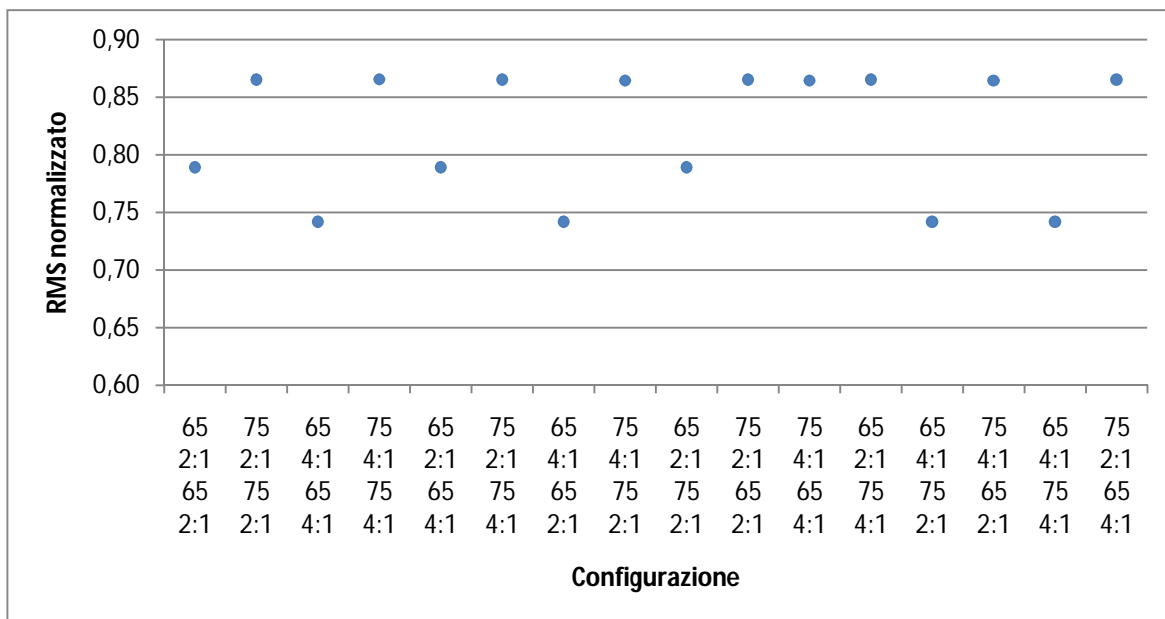


Grafico 14: entità della compressione (su scala normalizzata) subita dalla parola “salto” in funzione delle sedici diverse configurazioni del simulatore di apparecchio acustico digitale, con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente.

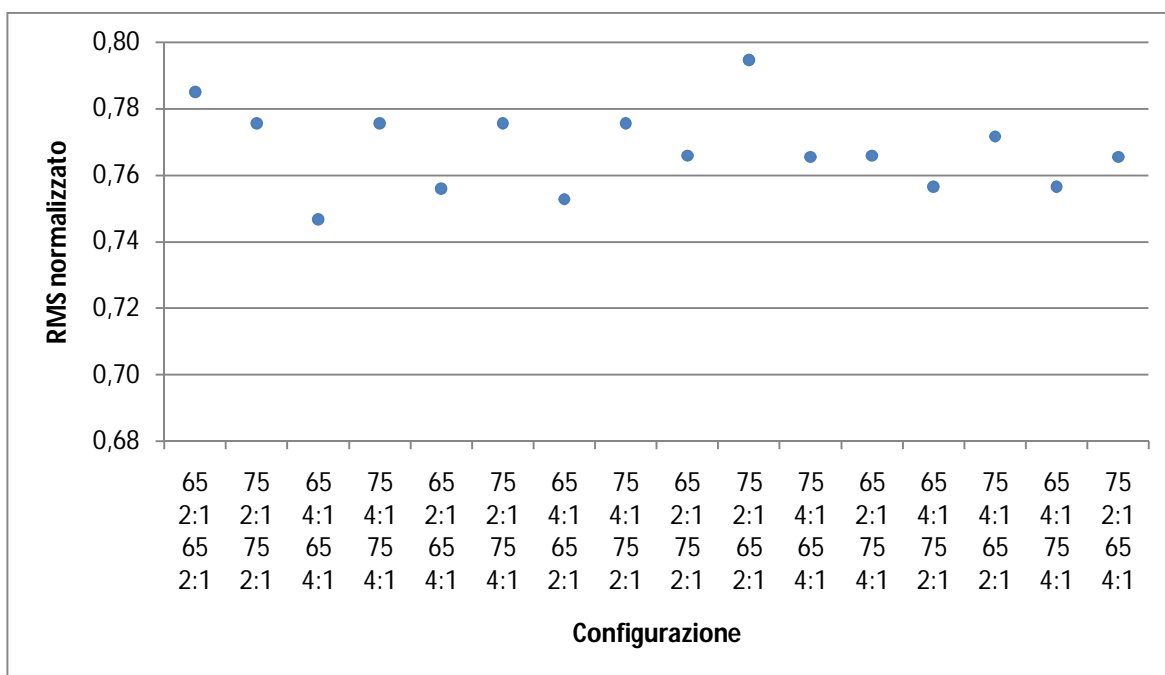


Grafico 15: entità della compressione (su scala normalizzata) subita dalla parola “cinque” in funzione delle sedici diverse configurazioni del simulatore di apparecchio acustico digitale, con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente.

Analogamente alla precedente condizione di lavoro, sono stati riportati, nei grafici 16 e 17, gli spettri degli involucri temporali delle parole “salto” e “cinque”, rispettivamente. Per l’analisi qualitativa del segnale si rimanda al paragrafo precedente, in quanto l’andamento

è analogo, ad eccezione di “salto” per cui non si osserva più la successione accentuata di picchi e avvallamenti per frequenze superiori a 12Hz.

Ciò che invece si nota è la possibilità di raggruppare le configurazioni in tre o quattro categorie differenti, data la totale sovrapposizione dei segnali ottenuti.

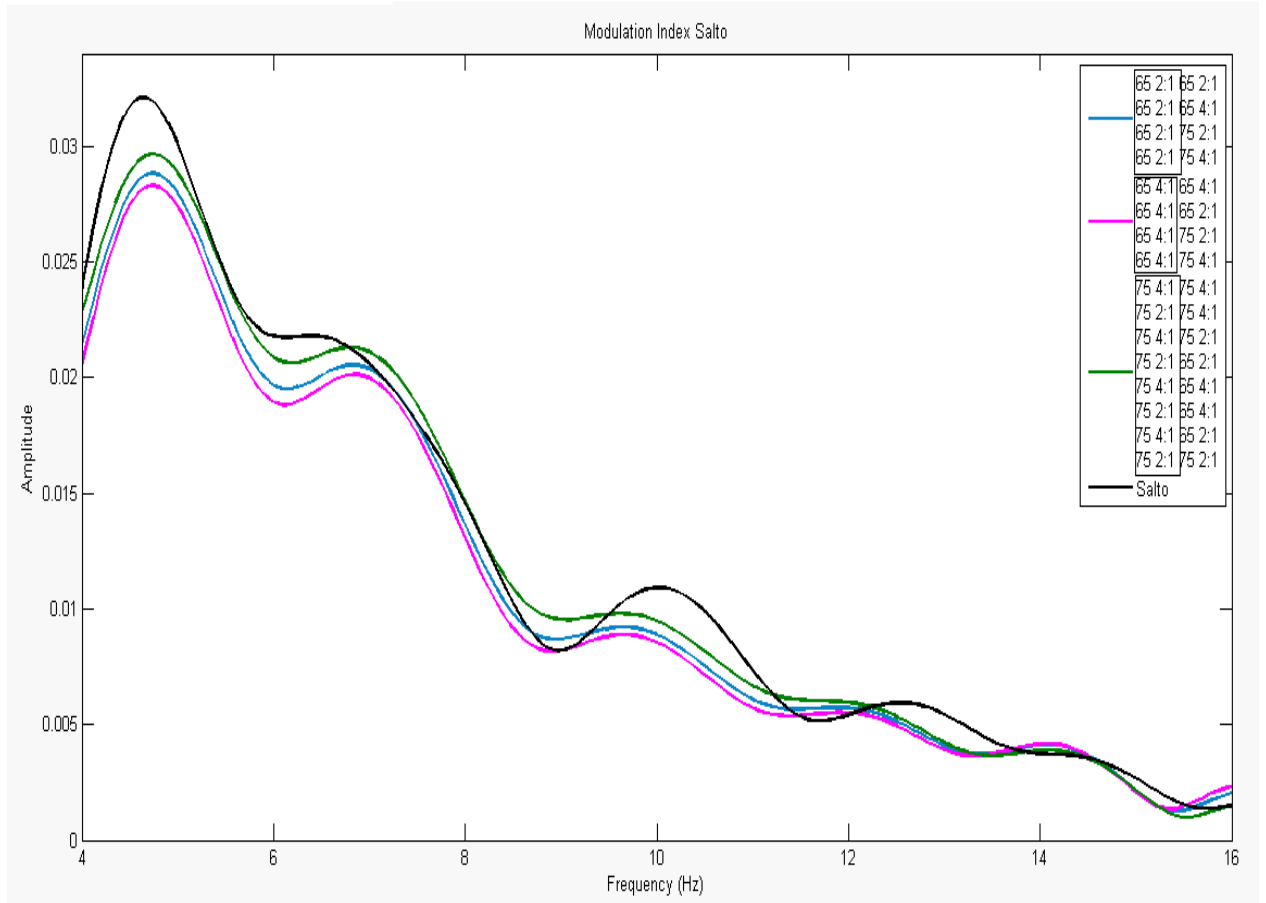


Grafico 16: spettro in frequenza dell'involuppo temporale della parola "salto" con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente: confronto tra il segnale originale (in nero) e tre tipi di configurazioni, rivelatesi equivalenti.

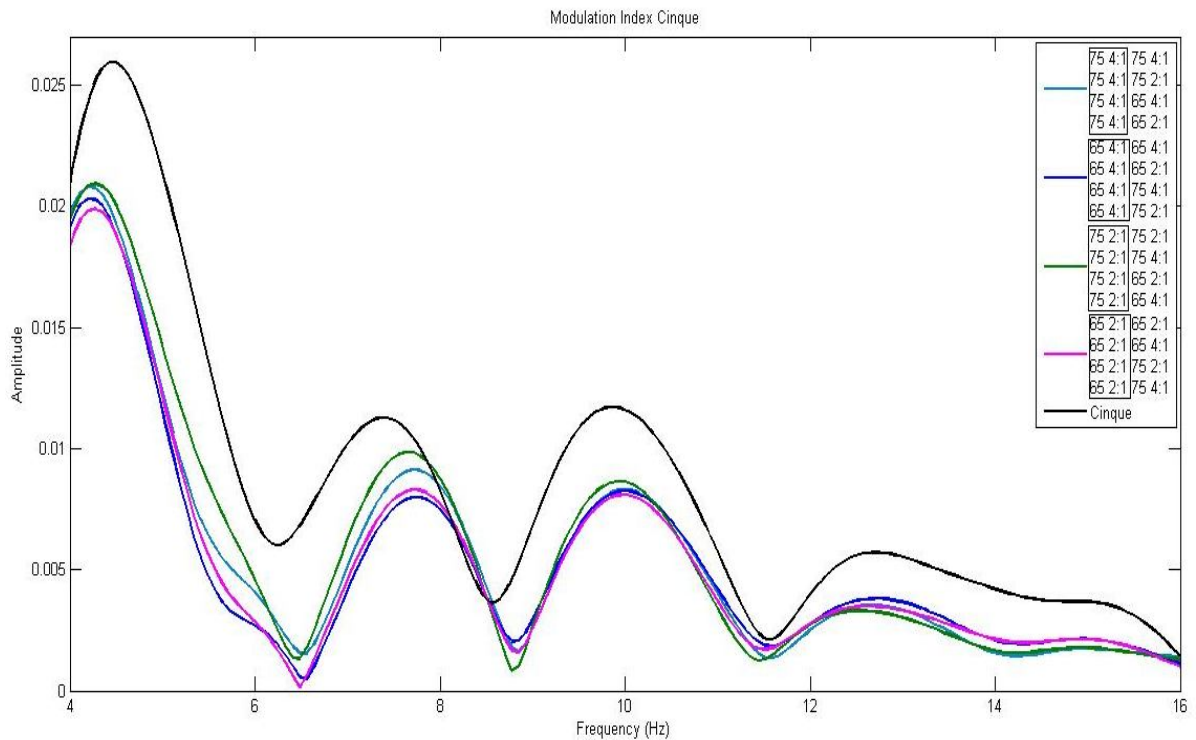


Grafico 17: spettro in frequenza dell'involucro temporale della parola "cinque" con tutti gli AGC da 65dB funzionanti e quelli da 75dB attivati solo occasionalmente: confronto tra il segnale originale (in nero) e quattro tipi di configurazioni, rivelatesi equivalenti.

## 2. Vocali singole: *vowels space*

Si ricorda che i segnali di controllo, di intensità tale da non attivare gli AGC, permettono di osservare l'effetto del microfono e dei filtri del simulatore sul contenuto spettrale dei segnali. Mentre le parole intere non subiscono variazioni spettrali significative, le singole vocali mostrano invece un forte ed eterogeneo cambiamento di entrambe le formanti.

### 2.1. AGC da 65dB e 75dB sempre in funzione

Il grafico 18 riporta lo spazio delle vocali di tutte le sedici configurazioni del simulatore sovrapposte a quella di controllo (linea nera spessa).



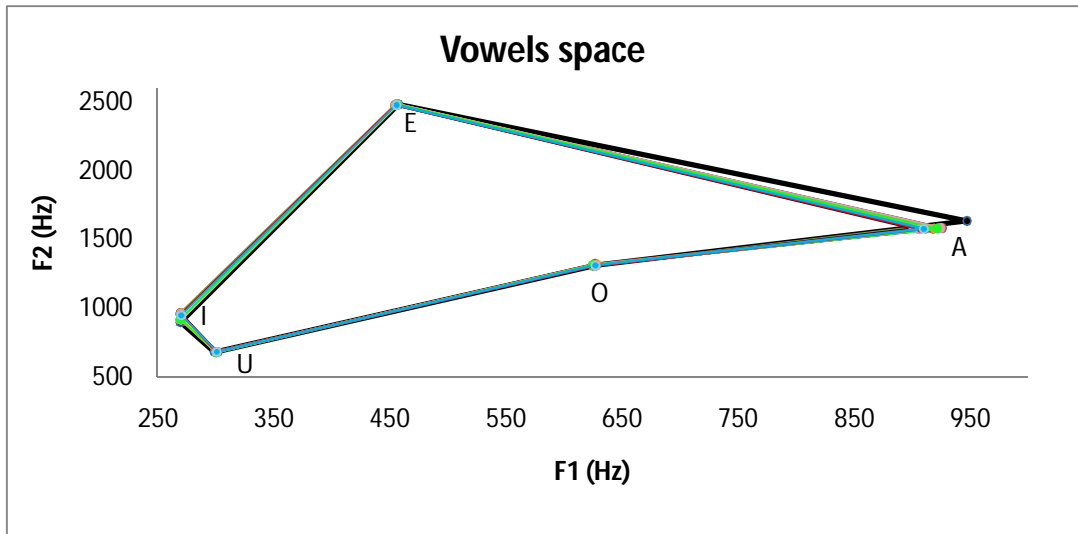


Grafico 18: *vowels spaces* ottenuti dalle sedici configurazioni del simulatore in condizioni di attivazione continua di tutti gli AGC, sovrapposti al *vowels space* di “controllo” (linea e indicatori neri).

## 2.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente

Analogamente alla prima condizione di lavoro, è stato costruito il *vowels space* per le sedici configurazioni anche nella condizione di attivazione intermittente degli AGC da 75dB, risultando però completamente sovrapponibile al *vowels space* delle vocali di controllo (grafico 19).

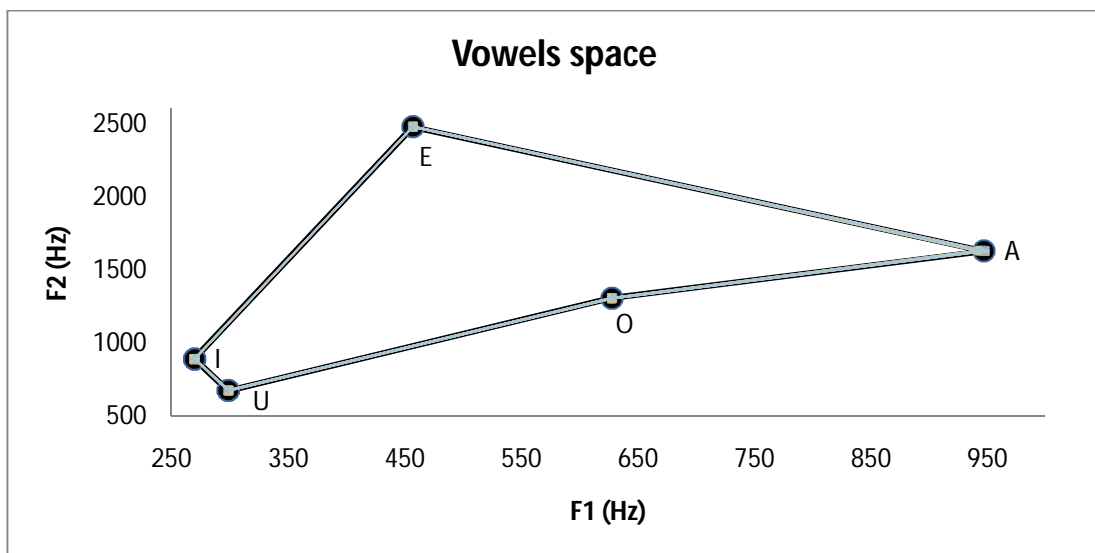


Grafico 19: *vowels spaces* ottenuti dalle sedici configurazioni del simulatore in condizioni di attivazione continua degli AGC da 65dB e funzione intermittente degli AGC con soglia, sovrapposti al *vowels space* di “controllo”.

## CAPITOLO 6

### Discussione

#### 1. Parole “salto” e “cinque”: *modulation index*

##### 1.1. AGC da 65dB e 75dB sempre in funzione

In tale condizione, i segnali in ingresso presentano un'intensità tale da mantenere sempre in funzione tutti gli AGC presenti nel simulatore, sia quelli a bassa soglia di compressione sia quelli ad alta. Così facendo, gli effetti della soglia risultano trascurabili rispetto al rapporto di compressione.

Analizzando lo spettro dell'involuppo temporale della parola “salto”, si osserva una progressiva perdita del *modulation index* mano a mano che aumenta il rapporto di compressione e diminuisce la soglia nei canali ad alta frequenza. Di seguito verranno elencate le sedici configurazioni in ordine di perdita dell'indice di modulazione, a partire da quella con perdita maggiore, quindi più “lontana” dall'originale, non elaborata:

- 1) LF: 65 4:1 HF: 65 4:1
- 2) LF: 75 4:1 HF: 65 4:1
- 3) LF: 65 2:1 HF: 65 4:1
- 4) LF: 75 2:1 HF: 65 4:1
- 5) LF: 65 4:1 HF: 65 2:1
- 6) LF: 75 4:1 HF: 65 2:1
- 7) LF: 65 2:1 HF: 65 2:1
- 8) LF: 75 2:1 HF: 65 2:1
- 9) LF: 65 4:1 HF: 75 4:1
- 10) LF: 65 2:1 HF: 75 4:1
- 11) LF: 75 4:1 HF: 75 4:1
- 12) LF: 75 2:1 HF: 75 4:1
- 13) LF: 65 4:1 HF: 75 2:1
- 14) LF: 65 2:1 HF: 75 2:1
- 15) LF: 75 4:1 HF: 75 2:1
- 16) LF: 75 2:1 HF: 75 2:1

dove la sigla LF (low frequency) sta ad indicare i primi sette canali del simulatore, mentre HF (high frequency) rappresenta gli ultimi nove.

Salta subito all'occhio che le sedici configurazioni seguono un ordine ben preciso: le prime quattro sono infatti caratterizzate da una soglia di 65dB e un rapporto di compressione di 4:1 nei canali ad alta frequenza. Ma anche i canali a bassa frequenza sono soggetti ad un ordine preciso: quello con perdita maggiore ha soglia 65dB e RC (rapporto di compressione) 4:1, seguito da quello a soglia maggiore e uguale RC e così via. Lo stesso ordine è ripetuto per le seguenti quattro configurazioni, mentre per le ultime otto si ha una leggera variazione nell'ordine dei canali a bassa frequenza: la perdita nel *modulation index* aumenta infatti prima all'aumentare del rapporto di compressione e poi al diminuire della soglia.

Per quanto riguarda invece la parola "cinque", si elencano anche in questo caso le sedici configurazioni in ordine di perdita nell'indice di modulazione, a partire sempre da quella che ha generato la perdita più significativa:

- 1) LF: 65 4:1 HF: 65 4:1
- 2) LF: 65 4:1 HF: 65 2:1
- 3) LF: 65 4:1 HF: 75 4:1
- 4) LF: 65 4:1 HF: 75 2:1
- 5) LF: 65 2:1 HF: 65 4:1
- 6) LF: 65 2:1 HF: 65 2:1
- 7) LF: 65 2:1 HF: 75 2:1
- 8) LF: 65 2:1 HF: 75 4:1
- 9) LF: 75 4:1 HF: 65 4:1
- 10) LF: 75 4:1 HF: 65 2:1
- 11) LF: 75 4:1 HF: 75 4:1
- 12) LF: 75 4:1 HF: 75 2:1
- 13) LF: 75 2:1 HF: 65 2:1
- 14) LF: 75 2:1 HF: 65 4:1
- 15) LF: 75 2:1 HF: 75 4:1
- 16) LF: 75 2:1 HF: 75 2:1

Analogamente a quanto osservato per la parola “salto”, anche per “cinque” si nota un ordine ben preciso delle configurazioni del simulatore. Esse sono prima di tutto raggruppabili sulla base della soglia e del rapporto di compressione dei canali a bassa frequenza: come atteso, quelli con perdita minore nel *modulation index* hanno soglia alta ed RC basso (75 e 2:1), seguiti da quelli a parità di soglia e maggior RC, succeduti a loro volta dai canali con soglia pari a 65dB e RC di 4:1 e così via. Anche in questo caso si osserva, all’interno di ciascun raggruppamento, un ulteriore ordine per i canali ad alta frequenza: per le configurazioni 1-4 e 9-12, l’indice di modulazione diminuisce prima all’aumentare del rapporto di compressione e poi al diminuire della soglia, mentre per gli altri due raggruppamenti (5-8 e 13-16) si nota semplicemente una diminuzione più accentuata per i canali a soglia minore.

#### 1.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente

In questa condizione di lavoro, i segnali in ingresso presentano un’intensità tale da mantenere sempre in funzione tutti gli AGC a soglia minore, ma da attivare quelli a soglia maggiore ad intermittenza. Così facendo, si può studiare l’andamento dell’indice di modulazione in funzione della soglia di compressione.

L’analisi spettrale in questa condizione di funzionamento si è rivelata più semplice e meno laboriosa: si è riscontrata infatti una totale sovrapposizione del *modulation index* proveniente da specifiche categorie di configurazioni.

Per la parola “salto”, lo spettro dell’involuppo temporale più simile a quello originale è stato ottenuto indistintamente dalle otto configurazioni con soglia di compressione 75dB nei canali a bassa frequenza, come è chiaramente visibile dal grafico 16 (linea verde).

Un’ulteriore diminuzione del *modulation index* si riscontra in ugual misura impostando la soglia di compressione a 65dB e l’RC a 2:1 nei canali a bassa frequenza, senza riguardo dei parametri dell’AGC negli altri canali (linea turchese).

Naturalmente, le configurazioni caratterizzate dalla massima compressione, cioè da una soglia pari a 65dB e un RC di 4:1, sempre nei canali a bassa frequenza (linea viola), sono quelle che più si discostano dallo spettro dell’involuppo temporale del segnale originale.

Comportamento analogo si riscontra per la parola “cinque” (grafico 17), ad eccezione del fatto che i risultati osservati sono classificabili in quattro categorie anziché in tre: due di queste sono equivalenti a quelle già viste per “salto” (linee blu e viola, relative rispettivamente alle configurazioni con parametri di compressione 65dB e 4:1 e 65dB e 2:1 nei canali a bassa frequenza), mentre le altre due (linee verde e turchese) permettono di distinguere anche l’effetto del rapporto di compressione a parità di soglia di 75dB.

Quest’ultima analisi costituisce un’ulteriore conferma della diminuzione dell’indice di modulazione all’aumentare della soglia di compressione dell’AGC, a prescindere dal rapporto di compressione.

In definitiva, facendo ancora riferimento ai grafici relativi all’entità della compressione globale (grafici 10, 11, 14, 15), si può affermare che, ad un aumento della compressione, esprimibile attraverso la diminuzione del valore di rms del segnale, corrisponde sempre una diminuzione del *modulation index*, da cui deriverà, come documentato in letteratura, una diminuzione dell’intelligibilità del segnale di speech.

## **2. Vocali singole: *vowels space***

### **2.1. AGC da 65dB e 75dB sempre in funzione**

In primo luogo, si è riscontrata una significativa ed eterogenea variazione delle formanti del segnale di controllo rispetto a quelle del segnale originale: nello specifico, le vocali *a*, *e* ed *o* sono caratterizzate da una crescita più o meno accentuata di entrambe le formanti, mentre *u* ed *i* mostrano un aumento di F1 e una diminuzione di F2. Queste variazioni non uniformi si traducono inoltre in due forme diverse del *vowels space* delle vocali originali e di quelle di controllo, elaborate dal simulatore implementato.

Tali variazioni delle formanti, con molta probabilità, sono dovute alla risposta in frequenza intrinseca del microfono combinata con quella dei filtri del modello. Si è pertanto deciso di confrontare i *vowels spaces* dei segnali elaborati con quello di controllo, piuttosto che con quello originale, in modo da avere un punto di riferimento più semplice e preciso da accostare.

Osservando il grafico 18, si nota una scarsa variabilità del *vowels space* tra una configurazione del simulatore e l’altra. In particolare, per le vocali *e*, *o* e *u* non è visibile alcuno scostamento dai valori di controllo di F1 ed F2 (linea nera spessa), mentre *a* ed *i*

mostrano solamente differenze poco significative. Globalmente è tuttavia rilevabile una leggera diminuzione dello spazio delle vocali soggette a compressione rispetto al *vowels space* di controllo, il che è coerente con gli studi presentati in letteratura: di fatto, la restrizione del *vowels space* può generare una perdita di intelligibilità, fenomeno che va di pari passo, tra l'altro, anche con la compressione.

Questa possibile perdita di intelligibilità potrebbe essere comunque talmente lieve da non essere nemmeno percepita dall'orecchio un soggetto ipoacusico, caratterizzato da una risoluzione temporale e in frequenza minori di quelle di un normoudente.

## 2.2. AGC da 65dB sempre in funzione e AGC da 75dB intermittente

Nella condizione di lavoro ad attivazione intermittente degli AGC a soglia più alta (grafico 19), non si osserva invece alcuno scostamento dal *vowels space* di controllo, il che sta ad indicare che, in una situazione di segnali ad intensità intermedia tra 65 e 75dB, l'entità della compressione e la compressione stessa non influiscono minimamente sulle formanti delle vocali.

Le variazioni più sostanziali sono invece quelle fornite dall'azione del microfono e dei filtri del simulatore stesso, che modificano significativamente le formanti delle vocali e fissano addirittura una nuova forma del *vowels space*.

## CONCLUSIONI

Nel presente studio è stato in primo luogo ricostruito, per mezzo della piattaforma Simulink, un modello di apparecchio acustico digitale a sedici canali, utilizzato poi per la simulazione e analisi di segnali vocali di diverso tipo: due parole (salto e cinque) e le cinque vocali della lingua italiana.

Attraverso l'uso di suddette simulazioni e un'analisi spettrale eseguita con il software MATLAB, è stata studiata l'influenza dell'entità della compressione, esprimibile attraverso la soglia e il rapporto di compressione degli AGC presenti nel modello, su due parametri ritenuti responsabili dell'intelligibilità del clear speech: il *modulation index* e il *vowels space*.

I primi risultati mostrano una diminuzione progressiva dell'indice di modulazione all'aumentare della compressione fornita dal simulatore. Tutti i ricercatori nell'ambito del clear speech sono unanimemente d'accordo sulla perdita di intelligibilità legata alla diminuzione del *modulation index*, il che ci permette di affermare l'esistenza di un peggioramento della qualità del clear speech, che aumenta in parallelo alla compressione, ossia alla diminuzione della soglia e l'aumento del rapporto di compressione degli AGC.

Il secondo tipo di analisi è stato effettuato sulla base di studi precedenti che affermano un rapporto diretto tra l'espansione del *vowels space* e il miglioramento della qualità del clear speech. I risultati ottenuti mostrano tuttavia solo una leggera riduzione dello spazio delle vocali in presenza della compressione di segnali di speech ad altissima intensità rispetto ad una condizione di assenza dell'azione degli AGC. Da ciò potrebbe derivare solo una leggera perdita di intelligibilità, che risulterebbe del tutto trascurabile rispetto ad una perdita globale legata ad altri fattori più rilevanti, come la diminuzione dell'energia spettrale in corrispondenza delle frequenze critiche dell'involuppo temporale del segnale di speech.

Ai fini dell'implementazione di un apparecchio acustico digitale reale, potrebbe essere utile considerare i risultati ottenuti nel presente studio per stabilire un compromesso tra la curva frequenza – guadagno e l'intelligibilità. Nel processo di programmazione dei filtri digitali di ciascun canale di una protesi acustica bisogna infatti soddisfare le esigenze reali dei pazienti sulla base del loro audiogramma, in modo da colmare opportunamente le

perdite uditive per ogni banda di frequenza critica. Se tuttavia ciò comporta l'impiego di basse soglie e alti rapporti di compressione degli AGC, bisogna dunque tenere presente che, più questi parametri sono restrittivi e rinforzano la compressione, minore sarà l'intelligibilità risultante del segnale di speech elaborato. Da ciò deriva un deterioramento delle prestazioni globali della protesi, il cui fine ultimo, si ricorda, è quello di migliorare la comprensione del linguaggio e la comunicazione di soggetti ipoacusici e con difficoltà percettive.

Concludendo, nonostante sia stato compiuto un leggero passo in avanti nell'ambito della ricerca su *clear speech* e *speech processing*, rimangono tuttavia aperte ancora molte porte in questo campo, che potranno portare ad un ulteriore sviluppo e perfezionamento degli apparecchi acustici digitali moderni.



## **RINGRAZIAMENTI**

*Un sentito ringraziamento  
ai miei genitori e a mio fratello  
che mi hanno permesso di studiare  
e sopportato sempre le mie crisi di panico.*

*Un grazie particolare  
a Serena, Arianna, Massimo, Alberto e Gianluca  
che mi hanno accompagnata in questi cinque anni  
di studio, fatica, gioie, viaggi e complicità.*

*Dovuti ringraziamenti soprattutto  
alla dott.ssa Tognola e all'ing. Paglialonga  
per la loro pazienza e comprensione.*

*In ricordo di mio zio Roberto,  
il miglior ingegnere mai esistito.*

## **BIBLIOGRAFIA**

Chong KS, Chang JS: A 16-channel low-power nonuniform spaced filter bank core for digital hearing aids. *IEEE transactions on circuits and systems – II: Express briefs*, vol. 53 no. 09, September 2006.

Clerici M: *L'audioprotesista*.

Drullman R, Festen JM, Plomp R: Effect of reducing temporal modulations on speech reception. *Journal of Acoustical Society of America* 95(5): 2670-2680.

Drullman R, Festen JM, Plomp R: Effect of temporal envelope smearing on speech reception. *Journal of Acoustical Society of America* 95(2): 1053-1064.

Drullman R: Temporal envelope and fine structure cues for speech intelligibility. *Journal of Acoustical Society of America* 97(1): 585-592.

Edwards BW: *Signal processing techniques for a DSP hearing aid*. IEEE 1998.

Ferguson SH: Talker differences in clear and conversational speech: vowel intelligibility for normal-hearing listeners. *Journal of Acoustical Society of America* 116(4): 2365-2373.

Ferguson SH, Kewley-Port D: Talker differences in clear and conversational speech: acoustic characteristics of vowels. *Journal of Speech, Language and Hearing Research* 50: 1241-1255.

Kates JM: A test suite for hearing aid evaluation. *Journal of Rehabilitation Research and Development* 27(3): 255-278, 1990.

Kates JM: A time-domain digital simulation of hearing aid response. *Journal of Rehabilitation Research and Development* 27(3): 279-294, 1990.

Katz J et al.: *Handbook of clinical audiology*.

Lewitt H, Neumann A, Mills R, Schwander T: A digital master hearing aid. *Journal of Rehabilitation Research and Development* 23(1): 79-87.

Lewitt H: Digital hearing aids: a tutorial review. *Journal of Rehabilitation Research and Development* 24(4): 7-20.

Moore BCJ: Design and evaluation of a two-channel compression hearing aid. *Journal of Rehabilitation Research and Development* 24(4): 181-192.

Smiljanić R, Bradlow AR: Speaking and hearing clearly: talker and listener factors in speaking style changes. *Language and Linguistic Compass* 3(1): 236-264.

Smiljanić R, Bradlow AR: Temporal organization of English clear and conversational speech. *Acoustical Society of America*. 3171-3182

Sorianello P: Per una rappresentazione dei segmenti vocalici: il caso del senese.

Ventura JC, Morellini L: A multi-band hearing aid emulation using real-time digital signal processing. Elsevier Science Publishers B.V. (North-Holland)

Vicen-Bueno R, Gil-Pita R, Utrilla-Manso M, Álvarez-Pérez L: A hearing aid simulator to test adaptive signal processing algorithms. IEEE 2007.

[www.mathworks.com](http://www.mathworks.com)