

# **POLITECNICO DI TORINO**

## **Corso di Laurea Magistrale in Ingegneria Biomedica**

Tesi di Laurea Magistrale

**Diagnosi precoce del morbo di Parkinson tramite analisi del segnale  
vocale rilevato con microfono in aria e microfono a contatto**



### **Relatore**

Prof. Alessio Carullo

### **Correlatore**

Prof. Alberto Vallan

### **Candidato**

Francesco Bozzi

237739

Anno accademico 2018-2019

*Alla mia famiglia.*

# Abstract

Il Parkinson è una malattia neurodegenerativa con evoluzione lenta ma progressiva, che coinvolge principalmente il sistema motorio con evidenti sintomi quali tremori, rigidità e blocchi nel cammino. Tuttavia, oltre a questi effetti, si riscontrano difficoltà in ambito comunicativo, in quanto il paziente perde il pieno controllo dell'apparato fonatorio, mostrando affievolimento della voce, articolazione indistinta e ripetizioni incontrollate.

Negli ultimi anni si è diffusa l'analisi oggettiva della qualità del segnale vocale grazie all'impiego di dispositivi indossabili in grado di monitorare i soggetti durante le proprie attività quotidiane. Ciò ha permesso l'identificazione di tecniche finalizzate alla diagnosi di patologie a carico dell'apparato fonatorio. Durante l'attività di tesi, la stessa metodologia è stata applicata a pazienti affetti dalla malattia di Parkinson con l'intento di individuare modelli per una diagnosi precoce della patologia. La prima parte del lavoro svolto si è focalizzata sull'acquisizione di campioni vocali da soggetti affetti dalla malattia di Parkinson. Nel corso di tale fase sono stati sottoposti a monitoraggio 22 pazienti presso l'ambulatorio Parkinson della Città della Salute di Torino. Le registrazioni sono state effettuate in un ambiente non isolato acusticamente, tramite un microfono in aria ed un microfono a contatto, seguendo un protocollo ben definito: ripetizione di tre vocali sostenute 'a', ripetizione dei fonemi pa/ta/ka, lettura di un brano foneticamente bilanciato (circa un minuto) ed eloquio spontaneo (circa un minuto). Nella seconda parte del lavoro, i segnali acquisiti sono stati elaborati tramite script Matlab appositamente sviluppati per estrarre indicatori dello stato vocale dei soggetti monitorati. I parametri estratti permettono di valutare la stabilità in frequenza (*jitter*) e in ampiezza (*shimmer*) delle vocali sostenute, in aggiunta alla qualità del segnale sia per la vocale 'a' sia per l'eloquio spontaneo tramite parametri specifici, quali HNR (*Harmonics to Noise Ratio*) e CPPS (*Cepstral Peak Prominence Smoothed*). Per le vocali sostenute acquisite tramite microfono in aria e per l'eloquio spontaneo acquisito con entrambi i microfoni, è stata effettuata una

selezione dei parametri più significativi attraverso l'analisi della matrice di correlazione. I parametri così individuati sono stati analizzati attraverso l'applicativo Matlab 'Classification Learner' al fine di identificare due classificatori (regressione logistica e *coarse tree*) in grado di discriminare soggetti sani da parkinsoniani. Per la vocale sostenuta acquisita con il microfono in aria si è ottenuta la migliore accuratezza con il modello del *coarse tree* (92%) considerando i parametri PPQ, shimmer e CPPS skewness, paragonabile a quella ottenuta con il modello di regressione logistica (88%) con i parametri jitter, PPQ, APQ e CPPS 5°perc. Le accuratezze ottenute per l'eloquio spontaneo con i due microfoni risultano simili tra di loro: 95% con il modello del *coarse tree* (CPPS moda, CPPS range, CPPS kurtosis) e 79% con il modello di regressione logistica (HNR, CPPS media, CPPS kurtosis) utilizzando il microfono in aria, e 92% con il modello del *coarse tree* (F0, *voiced/unvoiced ratio*, CPPS kurtosis) e 76% con il modello di regressione logistica (*voiced/unvoiced ratio*, CPPS range) utilizzando il microfono a contatto.

Questa prima parte del lavoro di tesi si pone quindi alla base di uno studio il cui obiettivo futuro è migliorare modelli di analisi identificati attraverso l'incremento del database dei soggetti monitorati.

La seconda parte del lavoro di tesi si è focalizzata sulla valutazione dell'efficacia di un percorso logopedico specifico per soggetti affetti dalla malattia di Parkinson. Sono stati monitorati presso la onlus "Amici Parkinsoniani Piemonte" 6 soggetti (5 uomini e 1 donna) seguendo lo stesso protocollo e con la stessa strumentazione utilizzata per la prima parte. Lo studio si è basato sul confronto dei segnali acquisiti a due mesi di distanza, ovvero prima e dopo il corso logopedico. Dall'analisi dei risultati relativi alle vocali sostenute, si è riscontrato, a due mesi di distanza, un miglioramento dei parametri di perturbazione per tre dei soggetti analizzati. Analizzando invece, sia per il microfono in aria che il microfono a contatto, le distribuzioni del CPPS del parlato libero la maggioranza dei soggetti in esame ha mantenuto lo stesso tipo di andamento, non rilevando variazioni apprezzabili. Tali confronti sono stati effettuati al termine del corso di logopedia, dunque un'analisi a breve

termine, riscontrando solo un lieve miglioramento nei soggetti della stabilità della fonazione, mantenendo invece stabile il proprio eloquio. Avendo tale corso insegnato ai pazienti esercizi da svolgere durante la propria vita quotidiana, sarebbe interessante un'analisi a lungo termine effettuando futuri monitoraggi.

# Sommario

Abstract .....	2
Indice delle Figure.....	8
Indice delle Tabelle.....	10
1 Introduzione.....	11
1.1 Malattia di Parkinson .....	11
1.2 Diagnosi e valutazione della malattia .....	12
1.3 Variazioni fonetiche nel soggetto parkinsoniano.....	14
2 Stato dell'arte .....	16
2.1 Segnale vocale .....	16
2.2 Parametri del segnale vocale .....	17
2.3 Analisi basate sullo spettro e sul cepstrum.....	18
3 Monitoraggio di soggetti affetti da Parkinson .....	21
3.1 Soggetti monitorati .....	21
3.2 Catena di acquisizione dati.....	21
3.3 Protocollo per le registrazioni .....	23
3.4 Pre-processing dei dati .....	23
3.5 Parametri estratti .....	24
3.5.1 Parametri di variazione in frequenza .....	25
3.5.2 Parametri di variazione in ampiezza.....	27
3.5.3 Parametri ricavati dall'autocorrelazione del segnale .....	29
3.5.4 Distribuzione del CPPS.....	29
3.6 Elaborazione dei dati.....	32
4 Risultati.....	33
4.1 Vocale sostenuta 'a' con microfono in aria.....	33
4.1.1 Feature selection.....	33

4.1.2	Creazione classificatore.....	36
4.1.2.1	Regressione logistica .....	37
4.1.2.2	Coarse tree .....	42
4.2	Eloquio libero con microfono in aria .....	46
4.2.1	Feature selection .....	46
4.2.2	Creazione classificatore.....	48
4.2.2.1	Regressione logistica .....	48
4.2.2.2	Coarse tree .....	51
4.3	Eloquio libero con microfono a contatto.....	54
4.3.1	Feature selection .....	54
4.3.2	Creazione classificatore.....	55
4.3.2.1	Regressione logistica .....	56
4.3.2.2	Coarse tree .....	59
5	Valutazione effetti corso logopedia.....	62
5.1	Vocale sostenuta ‘a’ con microfono in aria.....	62
5.2	Eloquio libero con microfono in aria .....	69
5.3	Eloquio libero con microfono a contatto.....	75
6	Conclusioni.....	82
	Bibliografia.....	84
	Appendice A .....	86

# Indice delle Figure

Figura 1 Rilascio di dopamina in un neurone sano e uno affetto da Parkinson [3].....	12
Figura 2 Foglio punteggiato UPDRS.....	13
Figura 3 Segnale glottico e segnale vocale di un soggetto parkinsoniano per vocale 'a' sostenuta .....	17
Figura 4 Esempio di Cepstrum.....	18
Figura 5 Esempio di CPPS con regressione lineare.....	19
Figura 6 (a) Microfono in aria, (b) microfono a contatto e registratore .....	22
Figura 7 Rappresentazione di jitter e shimmer nel segnale vocale.....	25
Figura 8 esempio distribuzione CPPS per soggetti sani e patologici: (a) vocale sostenuta, (b) lettura, (c) eloquio libero.....	31
Figura 9 Esempio di Parallel Coordinates Plot.....	36
Figura 10 Scelta del taglio del modello di regressione lineare per la vocale a.....	38
Figura 11 Curva ROC del modello a regressione lineare della vocale a.....	39
Figura 12 Probabilità delle vocali a con il modello di regressione logistica.....	40
Figura 13 Probabilità delle vocali 'a' con il modello di regressione logistica per ogni singolo soggetto .....	40
Figura 14 Probabilità delle tre 'a' mediate per ogni singolo paziente .....	41
Figura 15 Decision tree vocale 'a' .....	43
Figura 16 Divisione dello spazio delle features con l'albero decisionale per la vocale 'a'.....	44
Figura 17 Divisione dello spazio delle features con l'albero decisionale per la media delle vocali 'a'.....	45
Figura 18 Esempio di Parallel Coordinates Plot.....	48
Figura 19 Scelta del taglio del modello di regressione lineare per eloquio libero con MIPRO.....	50
Figura 20 Curva ROC del modello a regressione lineare dell'eloquio libero con MIPRO.....	50
Figura 21 Probabilità eloquio libero con MIPRO per modello di regressione logistica .....	51

Figura 22 Decision tree free speech MIPRO .....	52
Figura 23 Divisione dello spazio delle features con l'albero decisionale per free speech MIPRO.....	53
Figura 24 Esempio di Parallel Coordinates Plot.....	56
Figura 25 Scelta del taglio del modello di regressione lineare per eloquio libero con PIEZO .....	58
Figura 26 Curva ROC del modello a regressione lineare dell'eloquio libero con PIEZO .....	58
Figura 27 Probabilità eloquio libero con PIEZO per modello di regressione logistica .....	59
Figura 28 Grafo del free speech PIEZO .....	60
Figura 29 Divisione dello spazio delle features con l'albero decisionale per free speech PIEZO .....	61
Figura 30 Confronto PPQ pre-post corso di logopedia .....	65
Figura 31 Confronto vF0 pre-post corso di logopedia.....	66
Figura 32 Confronto vAm pre-post corso di logopedia.....	67
Figura 33 Distribuzioni CPPS vocale 'a' MIPRO .....	68
Figura 34 Confronto CPPS deviazione standard per MIPRO .....	69
Figura 35 Distribuzioni CPPS free speech MIPRO.....	70
Figura 36 Confronto HNR free speech MIPRO .....	73
Figura 37 Confronto CPPS mean free speech MIPRO.....	74
Figura 38 Confronto CPPS std free speech MIPRO.....	74
Figura 39 Distribuzioni CPPS free speech PIEZO .....	76
Figura 40 Confronto HNR free speech PIEZO .....	79
Figura 41 Confronto CPPS mean free speech PIEZO .....	80
Figura 42 Confronto CPPS std free speech PIEZO .....	80

## Indice delle Tabele

Tabella 1 Soggetti monitorati.....	22
Tabella 2 Riepilogo soggetti sani e soggetti parkinsoniani analizzati .....	24
Tabella 3 Matrice di correlazione dei parametri della vocale 'a'.....	35
Tabella 4 Confusion matrix regressione logistica vocale a.....	39
Tabella 5 Confusion matrix regressione logistica per le 'a' mediate.....	42
Tabella 6 Confusion matrix coarse tree vocale a .....	43
Tabella 7 Confusion matrix regressione logistica per le 'a' mediate.....	45
Tabella 8 Matrice di correlazione con indici di Pearson per parametri dell'eloquio libero con microfono in aria .....	47
Tabella 9 Confusion matrix regressione logistica eloquio libero MIPRO.....	49
Tabella 10 Confusion matrix coarse tree free speech MIPRO .....	52
Tabella 11 Matrice di correlazione con indici di Pearson per parametri dell'eloquio libero con microfono a contatto .....	55
Tabella 12 Confusion matrix eloquio libero PIEZO regressione logistica .....	57
Tabella 13 Confusion matrix coarse tree free speech PIEZO.....	60
Tabella 14 Matrice di indici Pearson per vocale 'a' con MIPRO.....	63
Tabella 15 Indici Pearson pre-post logopedia free speech MIPRO .....	71
Tabella 16 Differenze parametri free speech MIPRO .....	75
Tabella 17 Indici Pearson pre-post logopedia free speech PIEZO.....	77
Tabella 18 Differenze parametri free speech PIEZO.....	81

# 1 Introduzione

Il Capitolo 1 si apre con una descrizione della malattia di Parkinson (PD: parkinson disease), come è comunemente indicato, utile a contestualizzare l'attuale lavoro di tesi. La neuropatologia è descritta ponendo l'accento su sintomi motori e non motori associati alla PD. Al fine di completare le conoscenze di base, il capitolo include anche una descrizione della diagnosi clinica e della valutazione della PD.

## 1.1 Malattia di Parkinson

Il morbo di Parkinson (PD) è un disturbo progressivo neurodegenerativo, multi sistemico con sintomi motori e non motori [1], che fa parte di un gruppo di patologie definite "Disordini del Movimento".

Tradizionalmente, il PD è descritto in relazione a quattro effetti:

- rigidità
- bradicinesia
- instabilità posturale
- tremore a riposo.

La malattia è diffusa in tutto il mondo e si riscontra in entrambi i generi, con una lieve minoranza in quello femminile. L'età di esordio è tra i 57/60 anni, ma una piccola percentuale dei pazienti presenta un esordio del morbo tra i 22 ed i 40 anni. Al di sopra dei 60 anni colpisce l'1÷2% della popolazione, mentre la percentuale aumenta al 3÷5% quando l'età è superiore agli 86 [2].

Le strutture coinvolte nel morbo di Parkinson sono situate in aree profonde del cervello, note come gangli della base, che contribuiscono alla corretta esecuzione dei movimenti. La malattia di Parkinson si presenta quando la produzione di dopamina nel cervello cala considerevolmente. La riduzione dei livelli di dopamina è dovuta alla degenerazione di neuroni, in un'area chiamata substantia nigra.

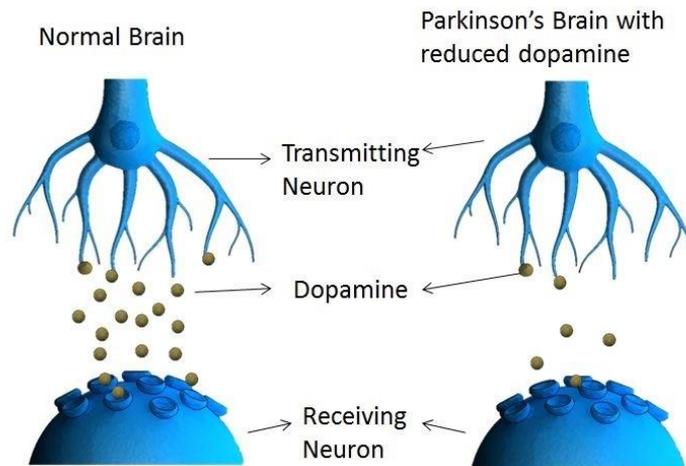


Figura 1 Rilascio di dopamina in un neurone sano e uno affetto da Parkinson [3]

## 1.2 Diagnosi e valutazione della malattia

Non esiste un test diagnostico specifico per identificare il PD. La diagnosi è effettuata sulla base di sintomi clinici che devono soddisfare i criteri della Brain Bank della Parkinson's Disease Society (PDS) per il PD [4].

Le scale più utilizzate per classificare i sintomi clinici sono la scala di *Hoehn e Yahr* e la *Unified Parkinson's Disease Rating Scale* (UPDRS).

La scala di Hoehn e Yahr è lo strumento più comunemente impiegato per valutare la gravità generale del PD, esso utilizza un semplice approccio definendo uno stadio che va dallo 0 (nessun segno di malattia) al 5 (grave disabilità). I pazienti sono solitamente classificati in PD in stadio precoce o tardivo in base alla gravità dei sintomi motori.

L'UPDRS è composto da quattro sottosezioni: Parte I (*Mentation, Behaviour e Mood*), Parte II (attività della vita quotidiana), Parte III (esame motorio), Parte IV (complicanze del trattamento). Le sottosezioni II e III sono ampiamente utilizzate clinicamente e includono rispettivamente l'auto-valutazione del paziente nelle attività quotidiane ed un punteggio clinico dei segni motori. Ogni item è valutato da 0 (normale) a 4 (grave disabilità).

La sezione 1 della Parte III può essere utilizzata come una misura "globale" della valutazione della comunicazione. Comprende dizione, volume, intelligibilità ed espressione. È evidente che a causa degli innumerevoli parametri che stabiliscono il

grado di avanzamento della malattia, per questo studio solo tale sezione sarà utilizzata come indice della gravità del sistema vocale.

**FOGLIO DEL PUNTEGGIO MDS UPDRS**

1.A	Fonte primaria d'informazione	<input type="checkbox"/> Paziente <input type="checkbox"/> Caregiver <input type="checkbox"/> Paziente + caregiver	3.3b	Rigidità del braccio destro	
			3.3c	Rigidità del braccio sinistro	
<b>Parte I</b>			3.3d	Rigidità della gamba destra	
1.1	Deterioramento cognitivo		3.3e	Rigidità della gamba sinistra	
1.2	Allucinazioni e psicosi		3.4a	Movimenti ripetuti dita mano destra	
1.3	Umore depresso		3.4b	Movimenti ripetuti dita mano sinistra	
1.4	Umore ansioso		3.5a	Movimenti della mano destra	
1.5	Apatia		3.5b	Movimenti della mano sinistra	
1.6	Caratteristiche della Sindrome da DDS		3.6a	Movimenti prono-supinazione mano destra	
1.6a	Chi compila il questionario	<input type="checkbox"/> Paziente <input type="checkbox"/> Caregiver <input type="checkbox"/> Paziente + caregiver	3.6b	Movimenti prono-supinazione mano sinistra	
			3.7a	Movimenti ripetuti dita piede destro	
1.7	Disturbi del sonno		3.7b	Movimenti ripetuti dita piede sinistro	
1.8	Sonnolenza diurna		3.8a	Agilità gamba destra	
1.9	Dolore e altre sensazioni		3.8b	Agilità gamba sinistra	
1.10	Problemi urinari		3.9	Alzarsi dalla sedia	
1.11	Problemi di costipazione		3.10	Marcia	
1.12	Sensazione di mancamento nell'assumere la posizione eretta		3.11	Blocco della marcia	
1.13	Affaticabilità		3.12	Stabilità posturale	
<b>Parte II</b>			3.13	Postura	
2.1	Eloquio		3.14	Spontaneità globale dei movimenti	
2.2	Salivazione e perdita di saliva		3.15a	Tremore posturale - mano destra	
2.3	Masticazione e deglutizione		3.15b	Tremore posturale - mano sinistra	
2.4	Attività correlate al mangiare		3.16a	Tremore cinetico - mano destra	
2.5	Vestirsi		3.16b	Tremore cinetico - mano sinistra	
2.6	Igiene personale		3.17a	Ampiezza del tremore a riposo - braccio destro	
2.7	Scrittura		3.17b	Ampiezza del tremore a riposo - braccio sinistro	
2.8	Passatempi e altre attività		3.17c	Ampiezza del tremore a riposo - gamba destra	
2.9	Girarsi nel letto		3.17d	Ampiezza del tremore a riposo - gamba sinistra	
2.10	Tremore		3.17e	Ampiezza del tremore a riposo - labbra/mascella	
2.11	Uscire dal letto, dall'auto, da una poltrona		3.18	Continuità del tremore a riposo	
2.12	Camminare ed equilibrio			La discinesia è presente	<input type="checkbox"/> No <input type="checkbox"/> Si
2.13	Blocco motorio (Freezing)			I movimenti hanno interferito con la valutazione?	<input type="checkbox"/> No <input type="checkbox"/> Si
3a	Il paziente è in terapia?	<input type="checkbox"/> No <input type="checkbox"/> Si		Scala di Hoehn e Yahr	
3b	Stato clinico del paziente	<input type="checkbox"/> Off <input type="checkbox"/> On	<b>Parte IV</b>		
3c	Il paziente è in Levodopa?	<input type="checkbox"/> No <input type="checkbox"/> Si	4.1	Tempo trascorso con discinesie	
3.C1	Se è così, indicare i minuti trascorsi dall'ultima somministrazione:		4.2	Impatto funzionale delle discinesie	
<b>Parte III</b>			4.3	Tempo trascorso in fase OFF	
3.1	Eloquio		4.4	Impatto funzionale delle fluttuazioni	
3.2	Mimica facciale		4.5	Complessità delle fluttuazioni motorie	
3.3a	Rigidità del collo		4.6	Distonia dolorosa in fase OFF	

1 Luglio, 2008

Copyright © 2008 Movement Disorder Society. Tutti i diritti riservati.

Questo documento non può essere copiato, distribuito o comunque utilizzato interamente o in parte senza il previo consenso della Movement Disorder Society: Official Working Document for large Clinimetric Program

**Figura 2 Foglio punteggio UPDRS**

### 1.3 Variazioni fonetiche nel soggetto parkinsoniano

È abitudine consolidata associare al morbo di Parkinson quasi esclusivamente sintomi motori. Tuttavia, tale malattia comporta anche sintomi precoci minori come la variazione della voce che diventa debole e monotona. Molti pazienti hanno descritto le difficoltà di linguaggio come il sintomo più debilitante, e simili ostacoli limitano ai minimi termini la comunicazione, conseguenza che può degenerare nel completo isolamento sociale. Nella malattia di Parkinson le alterazioni della voce sono dovute ad un ridotto coordinamento di alcuni muscoli ed in particolare di quelli che controllano:

- la respirazione
- l'articolazione (pronuncia)
- la prosodia (ritmo, intonazione e velocità dell'eloquio)
- la fonazione (emissione della voce) [5]

I sintomi che possono manifestarsi come conseguenza di tali alterazioni sono:

- **Affievolimento della voce:**  
La voce all'inizio di una frase è forte, ma si affievolisce progressivamente.
- **Monotonia:**  
La voce rimane costantemente allo stesso livello, non variando mai e mancando di espressività.
- **Cambiamento qualitativo:**  
Il suono della voce risulta tremolante, flebile o più acuto, a volte stridulo.
- **Indebolimento del volume:**  
Con l'incedere del tempo la riduzione del volume può arrivare al punto di non udibilità della voce.
- **Rapidità vocale:**  
Le sillabe o le intere parole scorrono senza pausa fino ad ammassarsi, con una lenta accelerazione delle parole verso la fine di una frase.
- **Articolazione indistinta:**  
I suoi delle sillabe finali non sono distinguibili e la fine delle parole talvolta è omessa.

- Involontaria esitazione prima di parlare:  
È difficile iniziare a parlare e poi mantenere la voce salda dall'inizio alla fine di una conversazione o di una frase.
- Ripetizioni incontrollate:  
Le parole, le frasi e i periodi sono ripetuti involontariamente e senza controllo [6].

## 2 Stato dell'arte

Questo capitolo presenta una panoramica sui parametri utilizzati negli ultimi decenni nell'analisi del segnale vocale per valutare lo stato dell'apparato fonatorio e il carico vocale.

### 2.1 Segnale vocale

La fonazione è la produzione di un segnale acustico da una sorgente, che consiste essenzialmente nella vibrazione delle corde vocali. In condizioni normali, e da un punto di vista fisico, quello glottico è un segnale quasi periodico, somma algebrica di una serie di segnali sinusoidali, chiamati componenti spettrali [7]. Ogni componente è caratterizzato da una propria frequenza, intensità e fase. Se il segnale è periodico con periodo fondamentale  $T_0$ , le componenti, definite anche armoniche, hanno frequenze che sono multipli interi della frequenza fondamentale  $F_0 = \frac{1}{T_0}$ . Il termine "quasi" suggerisce che le caratteristiche della frequenza e dell'ampiezza del segnale sono variabili nel tempo. Potrebbero verificarsi perturbazioni a breve termine tra un periodo e l'altro, o a lungo termine, ovvero tra l'inizio e la fine di un eloquio. Queste variazioni di frequenza della fondamentale e dell'ampiezza possono caratterizzare una voce patologica se superiori a determinate soglie. Le variazioni di  $F_0$  oscillano intorno ad un valore medio, che è caratteristico per ogni individuo e varia in base all'età, al genere e al tipo di attività vocale:

- 255 ÷ 440 Hz per i bambini
- 175 ÷ 245 Hz per le donne
- 105 ÷ 160 Hz per gli uomini [7].

Una volta uscito dalle corde vocali, il segnale della glottide è sottoposto all'azione filtrante del tratto vocale, che lo trasforma nel segnale vocale. Tuttavia, la frequenza fondamentale  $F_0$  resta immutata, ad essere modificate sono solo le ampiezze delle varie componenti spettrali. La Figura 3 consente di confrontare le caratteristiche del segnale della glottide, acquisite con un microfono a contatto, e quelle del corrispondente segnale vocale acquisito con microfono in aria.

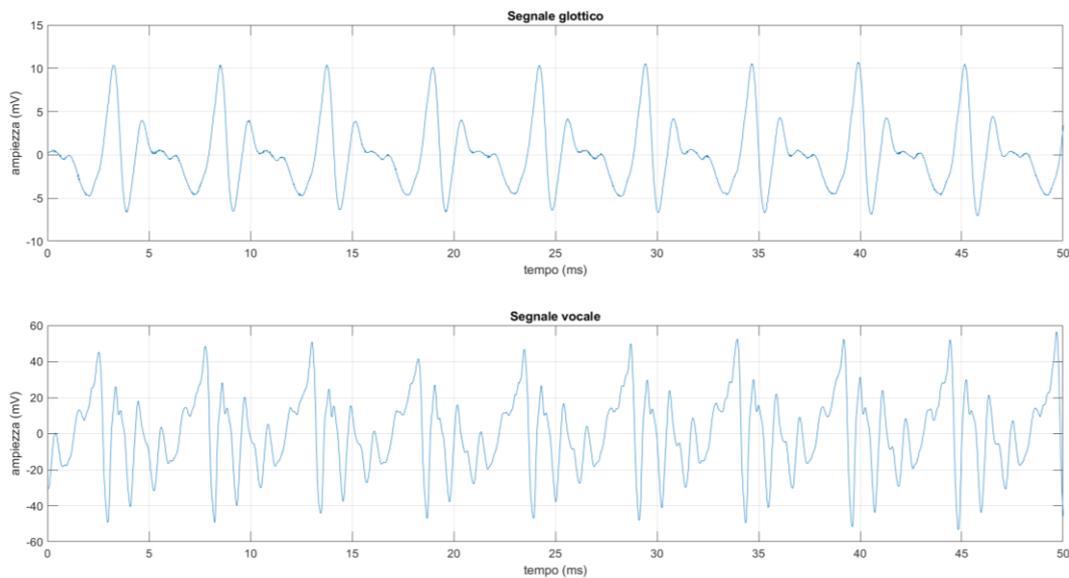


Figura 3 Segnale glottico e segnale vocale di un soggetto parkinsoniano per vocale 'a' sostenuta

## 2.2 Parametri del segnale vocale

Negli ultimi anni, le misure acustiche unite a tecniche di analisi e di elaborazione di segnali hanno assunto un'importanza fondamentale, essendo poco invasive, relativamente a basso costo, in tempo reale e facili da applicare. L'analisi principale è quella della valutazione delle caratteristiche nel dominio del tempo e della frequenza, specialmente la frequenza fondamentale  $F_0$ , e le perturbazioni ad essa legate, in frequenza come il *jitter* ed in ampiezza come lo *shimmer*, ed il rapporto tra componenti armoniche e non armoniche (HNR) [8]. Nell'analisi acustica della produzione di una vocale sostenuta è stato osservato che maggiori variazioni in frequenza ed ampiezza dalla  $F_0$  e un incremento dell' HNR sono indicatori di una instabilità del sistema fonatorio [9]. Si sono mostrate, infatti, significative differenze di tali parametri tra pazienti sani e pazienti affetti dalla malattia di Parkinson, specialmente l'HNR ha confermato il risultato che la voce in questi soggetti diventi più grave con l'avanzamento della patologia [10]. In recenti studi si è inoltre mostrato come durante un eloquio libero o una lettura, soggetti parkinsoniani presentano un'alterazione del ritmo e della velocità di parole, per cui parametri come *speech rate*

*variation* e *rhythm variation* risultano essere strettamente correlati con la malattia e rilevanti per identificare pazienti affetti da tale morbo [11].

### 2.3 Analisi basate sullo spettro e sul cepstrum

I parametri descritti in precedenza indicano le perturbazioni del segnale utili a livello clinico per la valutazione oggettiva della qualità della voce, e hanno una validità limitata all'accuratezza nel determinare la  $F_0$ , per cui piccoli errori nel suo calcolo possono essere rilevanti per l'analisi dei parametri da essa derivanti. Inoltre, l'analisi ne risulta utile solo su dati provenienti da vocali sostenute e non su di un eloquio libero. Per sopperire a tali limitazioni si utilizza un parametro che non prende in considerazione la frequenza fondamentale, ovvero il CPPS (*Cepstral Peak Prominence Smoothed*) che permette l'analisi anche del parlato continuo [12].

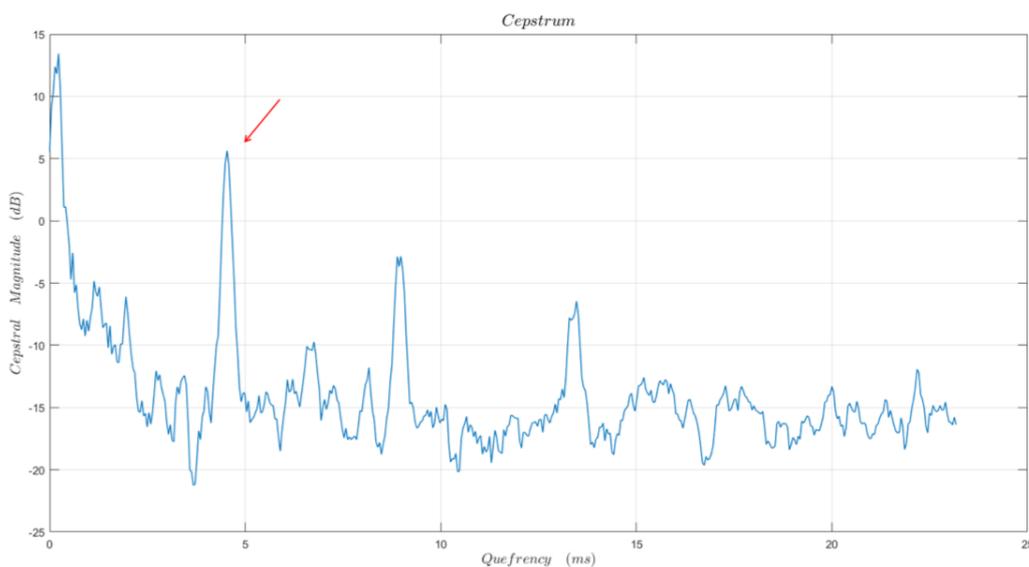
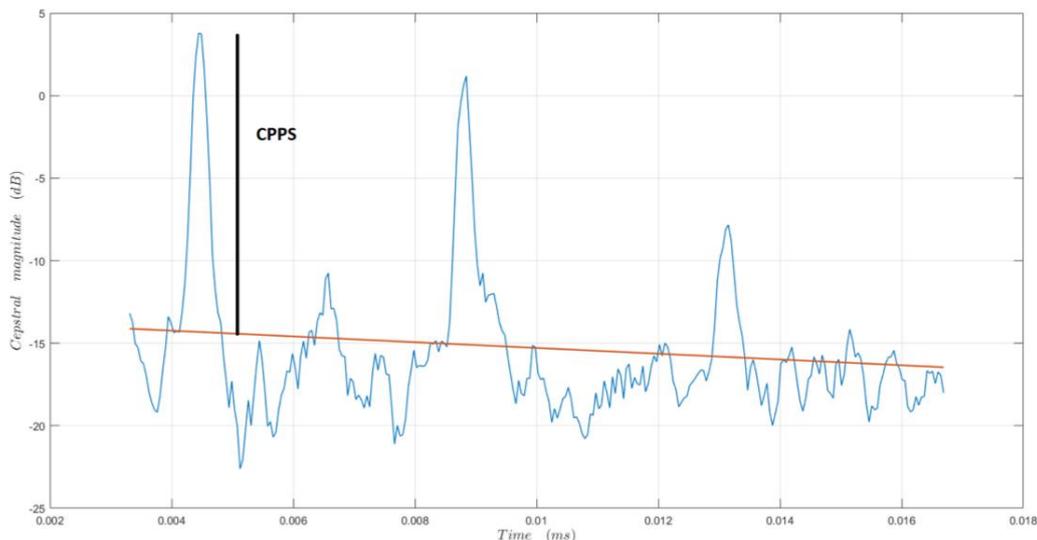


Figura 4 Esempio di Cepstrum

Il metodo della densità spettrale di potenza consente di ottenere informazioni importanti sulla voce, ma la ricerca si è mossa negli ultimi anni verso un nuovo approccio importante basato sulla misurazione del cepstrum. Nel 1963 Borget et al. [13] ha definito il concetto di cepstrum come “la trasformata di Fourier del logaritmo della trasformata di Fourier del segnale, a volte chiamato *spettro dello spettro*”. Dunque il cepstrum è definito nel dominio del tempo, per cui i termini *cepstrum* e *quefreny* sono l’anagramma delle parole *spectrum* e *frequency*.

In particolare, il cepstrum mostra il picco di energia di ogni componente di frequenza. Il segnale periodico evidenzia una struttura armonica ben definita e il cepstrum contiene un picco in corrispondenza della frequenza fondamentale, come è possibile vedere dalla Figura 4. L'ampiezza di questo picco è definito come CP (*Cepstral Peak*) e si trova tra 3 ms e 16 ms ( $60 \div 300$  Hz). Questo parametro non dipende solo dalla composizione armonica del segnale, ma da tutta la sua energia, per cui è influenzato anche dal rumore. Ecco perché è stato definito un parametro alternativo, il CPP (*Cepstral Peak Prominence*), che è la differenza tra il picco ed il corrispondente valore di quella quefrequency ottenuta con la regressione lineare. Infine, è stato definito il parametro CPPS (*Cepstral Peak Prominence Smoothed*) ottenuto con lo *smoothing* del segnale per avere un valore più significativo del picco (un esempio è mostrato in Figura 5) [14].



**Figura 5** Esempio di CPPS con regressione lineare

Considerando una finestra mobile di frame lungo il segnale con segmenti di overlap si ottiene quindi un diverso valore di CPPS per ognuno di esso. Con tali valori calcolati è possibile costruire una distribuzione da cui estrarre diversi parametri statistici come la media, mediana, moda, range e deviazione tipo. [15]. L'ampia presenza di studi sul CPPS ha permesso la nascita di diversi software come praat, speech tool e ADSV, per calcolare i parametri statistici al fine di discriminare soggetti sani e soggetti affetti da patologie all'apparato vocale. I risultati di differenti studi hanno dimostrato che il 5°

percentile è il miglior discriminante per segnali acquisiti con microfoni in aria e la deviazione tipo per quelli con microfono a contatto in vocali sostenute 'a' [16], mentre il 95° percentile ha mostrato migliori risultati nella discriminazione dell'eloquio libero [17].

### **3 Monitoraggio di soggetti affetti da Parkinson**

Questo capitolo è dedicato alle modalità di monitoraggio dei soggetti affetti da Parkinson, dalla raccolta dei dati alla procedura e agli strumenti con cui sono stati raccolti, includendo le fasi di pre-processing e processing.

#### **3.1 Soggetti monitorati**

In questo studio sono stati monitorati presso l'ambulatorio Parkinson della Città della Salute di Torino 22 pazienti, 14 uomini e 8 donne con età media di 68 anni, tutti affetti dalla malattia di Parkinson, come riportato dalla successiva tabella (Tabella 1), con differente grado di avanzamento della malattia.

#### **3.2 Catena di acquisizione dati**

I segnali vocali dei soggetti monitorati in questo studio sono stati acquisiti attraverso due differenti tipi di microfono:

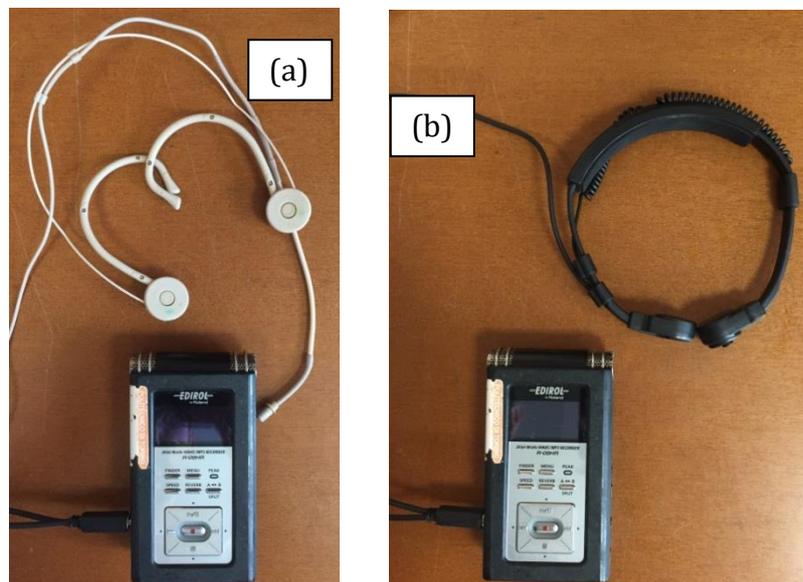
- un microfono in aria del tipo a guancia (MIPRO MU55-HR) posto ad una distanza di circa 2.5 cm dalle labbra del soggetto;
- un microfono a contatto di tipo piezoelettrico (HX505-1-1), che è un collare il cui elemento sensibile deve essere posizionato in prossimità della fossetta del giugulo per captare le vibrazioni delle corde vocali.

Entrambi i microfoni sono stati collegati ad un registratore portatile (EDIROL ROLAND R09-HR) con ingresso stereo a due canali. Il canale destro è utilizzato per acquisizione con microfono in aria, il canale sinistro per l'acquisizione con microfono a contatto. Tutte le registrazioni sono state effettuate simultaneamente con una frequenza di campionamento  $f_s = 44100$  Sa/s e con una risoluzione di 16 bit. La strumentazione utilizzata è mostrata nella Figura 6.

## Monitoraggio di soggetti affetti da Parkinson

<i>Soggetto</i>	<i>Sesso</i>	<i>Età</i>	<i>Durata malattia</i>	<i>UPDRS Speech</i>	<i>HY</i>
1	M	80	3	1	2
2	M	64	1	1	2
3	F	89	7	3	4
4	M	71	3	1	2
5	F	54	6	2	2
6	M	68	11	2	2
7	F	79	8	2	4
8	M	49	19	2	2
9	M	41	2	0	1
10	M	69	8	1	2
11	M	43	5	1	2
12	F	86	3	1	2
13	M	75	1	1	2
14	M	49	6	2	2
15	F	75	17	3	4
16	M	83	6	2	2
17	F	67	7	2	2
18	M	67	3	1	2
19	M	65	6	1	2
20	F	75	2	0	1
21	M	70	6	1	2
22	F	78	6	2	2

**Tabella 1 Soggetti monitorati**



**Figura 6 (a) Microfono in aria, (b) microfono a contatto e registratore**

### 3.3 Protocollo per le registrazioni

Le acquisizioni dei segnali vocali, avvenute simultaneamente con entrambi i microfoni descritti in precedenza, hanno previsto un protocollo ben definito. Tale protocollo è stato strutturato in 4 parti distinte della durata complessiva di circa 5 minuti:

1. fonazione della vocale 'a' mantenuta per più tempo possibile (3 ÷ 10 secondi), con un volume di confort. È stato richiesto ai soggetti coinvolti di ripetere tre volte l'emissione della vocale lasciando il tempo necessario per riprendere fiato;
2. ripetizione dei fonemi pa/ta/ka per circa 10 volte il più velocemente possibile, senza riprendere fiato tra una ripetizione e la successiva;
3. lettura di un brano foneticamente bilanciato, riportato in Appendice A, della durata di circa 1 minuto;
4. eloquio su un tema libero della durata di circa 1 minuto.

Le acquisizioni sono state effettuate in un ambiente non controllato acusticamente, dove il un rumore di fondo medio non è mai stato inferiore a 50 dB.

### 3.4 Pre-processing dei dati

Una volta ottenute le registrazioni dei soggetti monitorati presso l'ambulatorio Parkinson, i file in formato .wav sono stati trasferiti su un PC, dove tramite l'applicativo Audacity (ver. 2.3.0) l'intera registrazione è stata suddivisa nelle singole parti previste dal protocollo:

- *vocale sostenuta 'a'*: per ogni registrazione sono state separate le tre vocali in singoli file denominati A1, A2, A3, eliminando la parte iniziale e finale della registrazione per evitare instabilità del segnale;
- *pa/ta/ka*: sono state estratte dal file completo la ripetizione dei fonemi;
- *lettura*: la registrazione della lettura è stata tagliata dall'inizio del brano 'Bulka' fino alla parola 'sanguisuga', al fine di bilanciare il tempo di analisi con quello dell'eloquio libero e ridurre i tempi computazionali;

- *eloquio libero*: è stato tagliato dalla registrazione la parte inerente all’eloquio libero dall’inizio del parlato fino ad un massimo di 1 minuto per conformarsi con la durata di analisi del file della lettura e ridurre i tempi computazionali.

Tale procedimento è stato impiegato anche su registrazioni effettuate in precedenti lavori di tesi, utilizzando gli stessi microfoni, su persone diagnosticate sane all’apparato fonatorio, al fine di costituire il gruppo di controllo. I soggetti scelti sono stati selezionati per cercare di rendere il data-set di confronto quanto più bilanciato possibile, come si può notare dalla Tabella 2 dove è riportata anche l’età media dei gruppi considerati.

	Vocale A			Reading			Free Speech											
	Mipro		Piezo	Mipro		Piezo	Mipro		Piezo									
	M	F	età	M	F	età	M	F	età									
Pazienti	14	8	60	14	7	67	14	8	60	12	7	65	14	8	60	12	7	65
Sani	14	8	41	13	8	23	14	8	41	11	8	32	14	8	41	11	8	32

**Tabella 2 Riepilogo soggetti sani e soggetti parkinsoniani analizzati**

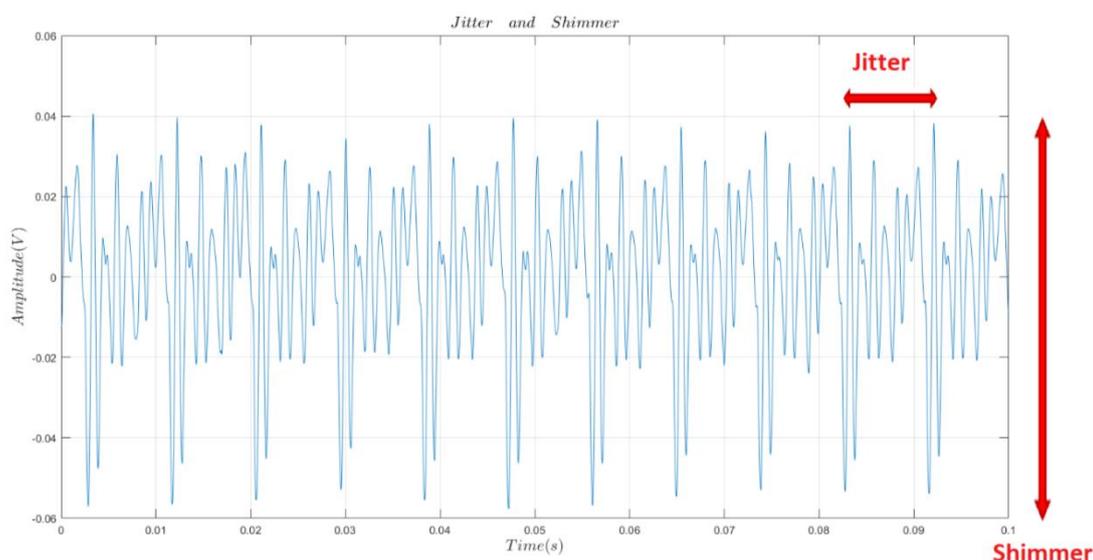
Come è possibile notare, il numero di soggetti monitorati risulta inferiore con il microfono a contatto, poiché trattandosi di microfono a girocollo ne risulta problematico l’utilizzo in caso di soggetti con collo troppo piccolo o troppo grande, o nel caso il cui il soggetto da monitorare sia impossibilitato ad indossarlo (es. utilizzo di ossigenatori).

I singoli file .wav ottenuti sono stati mantenuti in formato stereo.

### 3.5 Parametri estratti

I file segmentati delle tre ripetizioni della vocale \a\ sostenuta per i pazienti affetti dal morbo di Parkinson e per i soggetti sani, sono stati successivamente elaborati al fine

di estrarre parametri acustici significativi per la discriminazione dei due gruppi monitorati. Oltre alla frequenza fondamentale  $F_0$ , sono stati estratti 5 parametri relativi alla variazione della frequenza (*Jita*, *Jitt*, *RAP*, *PPQ*,  $vF_0$ ) e 4 relativi alla variazione in ampiezza (*Shim*, *ShdB*, *APQ*,  $vAm$ ). In aggiunta ai parametri dipendenti dalla frequenza fondamentale, significativi solo per l'analisi della vocale sostenuta, sono stati estratti parametri utili anche per la discriminazione dei file di lettura ed eloquio libero. In particolare, sono stati calcolati 3 parametri analizzando l'autocorrelazione del segnale (*HNR*,  $V/uV$ ,  $V/S$ ), e le distribuzioni del CPPS con 9 statistiche descrittive (*media*, *mediana*, *moda*, *standard deviation*, *range*, *5° percentile*, *95° percentile*, *skewness*, *kurtosis*)



**Figura 7** Rappresentazione di jitter e shimmer nel segnale vocale

### 3.5.1 Parametri di variazione in frequenza

Dopo il calcolo della frequenza fondamentale  $F_0$  sono stati estratti i seguenti parametri di variazione del periodo:

- **Jita** ( $\mu s$ ) è il **jitter assoluto** e descrive la variazione del periodo fondamentale  $T_0$ , calcolando la media in valore assoluto tra due consecutivi valori del periodo.

$$Jita = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^{(i)} - T_0^{(i+1)}| \quad (1)$$

Con  $T_0^{(i)}$  i periodi estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero dei periodi.

- **Jitt (%)** è il **jitter locale** e descrive come il *Jita* la variazione del periodo fondamentale  $T_0$ , calcolando il rapporto della media in valore assoluto tra due consecutivi valori del periodo rispetto alla media del periodo fondamentale dell'intero segnale.

$$Jitt = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_0^{(i)} - T_0^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}} \quad (2)$$

Con  $T_0^{(i)}$  i periodi estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero dei periodi.

- **RAP (%)** è il **Relative Average Perturbation** di tre periodo, ovvero il periodo fondamentale e i suoi 2 consecutivi, muovendosi con passo di uno.

$$RAP = \frac{\frac{1}{N-2} \sum_{i=2}^{N-1} \left| \frac{T_0^{(i-1)} + T_0^{(i)} + T_0^{(i+1)}}{3} - T_0^{(i)} \right|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}} \quad (3)$$

Con  $T_0^{(i)}$  i periodi estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero dei periodi. Questo parametro è simile al jitter ma invece di calcolare la media tra due periodi successivi effettua uno smoothing su 3 periodi sottraendo quello centrale.

- **PPQ (%)** è il *Pitch Period Perturbation Quotient* di 5 periodi consecutivi, ovvero la media relativa al periodo fondamentale con un fattore di smoothing pari a 5.

$$PPQ = \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{1}{5} \sum_{r=0}^4 T_0^{(i+r)} - T_0^{(i+2)} \right|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}} \quad (4)$$

Con  $T_0^{(i)}$  i periodi estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero dei periodi.

- **vF<sub>0</sub> (%)** è *Fundamental Frequency Variation* relativo alla variabilità della standard deviation della frequenza fondamentale  $F_0$  rispetto alla media della stessa

$$vF_0 = \frac{\sigma}{F_0} \times 100 = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (F_0 - F_0^{(i)})^2}}{\frac{1}{N} \sum_{i=1}^N f_0^{(i)}} \times 100 \quad (5)$$

Con  $F_0$  la media della frequenza fondamentale,  $\sigma$  la deviazione tipo della  $f_0$ , e  $f_0^{(i)}$  la singola frequenza estratta.

### 3.5.2 Parametri di variazione in ampiezza

Così come per la variazioni in frequenza, sono stati estratti parametri legati all'instabilità in ampiezza per ogni ciclo consecutivo della fondamentale  $F_0$ :

- **ShdB (dB)** è lo *shimmer assoluto* che descrive la variabilità in valore assoluto dell'ampiezza picco per picco in ogni periodo consecutivo

$$ShdB = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| 20 \log \frac{A^{(i+1)}}{A^{(i)}} \right| \quad (6)$$

Dove  $A^{(i)}$  i valori picco-picco estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero di impulsi.

- **Shim** (%) è lo *shimmer locale* che descrive le variazioni relative del valore picco-picco in periodi consecutivi

$$Shim = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A^{(i)} - A^{(i+1)}|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \quad (7)$$

Dove  $A^{(i)}$  i valori picco-picco estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero di impulsi. Sia lo Shim che il ShdB sono relative allo stesso tipo di variazione in ampiezza ma misurate in maniera differenti.

- **APQ** (%) è l'*Amplitude Perturbation Quotient* che descrivere la variazione relativa tra 11 periodi consecutivi con passo di 1:

$$APQ = \frac{\frac{1}{N-10} \sum_{i=1}^{N-10} \left| \frac{1}{11} \sum_{r=0}^{10} A^{(i+r)} - A^{(i+5)} \right|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \quad (8)$$

Dove  $A^{(i)}$  i valori picco-picco estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero di impulsi. L'APQ è il parametro meno sensibile all'errore di variazione ma non riesce ad analizzare una variazione a breve termine dell'ampiezza del segnale vocale.

- **vAm** (%) è il *Peak Amplitude Variation* e fornisce una variabilità relativa delle variazioni dell'ampiezza picco-picco (a breve-lungo termine) all'interno del campione vocale analizzato:

$$vAm = \frac{\sigma}{A_0} \times 100 = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (A_0 - A^{(i)})^2}}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \times 100 \quad (9)$$

Dove  $A^{(i)}$  i valori picco-picco estratti dal segnale vocale e  $i = 1, 2, \dots, N$  il numero di impulsi,  $A_0$  è il valor medio dei valori di picco-picco estratti [18].

### 3.5.3 Parametri ricavati dall'autocorrelazione del segnale

- **HNR (dB)** è l'*Harmonics to noise ratio* ovvero la misura del rumore additivo sul segnale di interesse:

$$HNR = 10 \log_{10} \left[ \frac{AC_v(T)}{AC_v(0) - AC_v(T)} \right] \quad (10)$$

Dove  $AC_v(0)$  è l'autocorrelazione al ritardo 0 (la potenza) e  $AC_v(T)$  è il valore di autocorrelazione al ritardo T.

- **V/uV (%)** è il *Voiced/unvoiced ratio*, ovvero il rapporto tra i frame sonori e quelli non sonori all'interno del singolo file estratto. Questa misura è eseguita considerando la porzione di frame con un HNR positivo (fonemi sonori) rispetto alla lunghezza delle parti non armoniche (fonemi silenti).
- **V/S (%)** è il *Voiced/silence ratio*, ovvero il rapporto tra parte armonica e non armonica del segnale rispetto all'intera durata del file (considerando anche le pause).

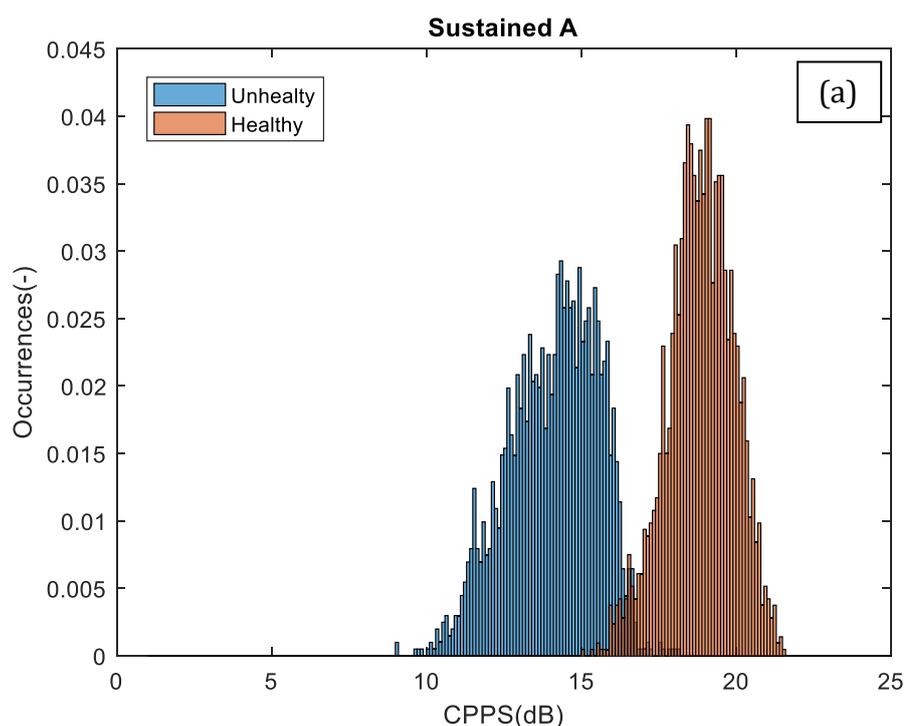
### 3.5.4 Distribuzione del CPPS

Il Cepstral Peak Prominence Smoothed è una misura dell'ampiezza del picco del cepstrum normalizzata per l'ampiezza complessiva. Dato il segnale vocale  $y$ , è possibile definire:

$$y_{FFT} = 20\log|FFT(y)|$$

$$y_{FFT2} = 20\log|FFT(y_{FFT})|$$

Dove la  $y_{FFT2}$  è definita come il cepstrum del segnale  $y$  [19]. Generalmente il Cepstrum è calcolato per diverse finestre in cui il segnale è diviso, eseguendo uno smoothing nel dominio del tempo e nel dominio della frequenza. Il valore del CPPS è la differenza tra l'ampiezza del picco e il valore corrispondente nella regressione lineare calcolata sul cepstrum. È calcolato quindi un singolo valore di CPPS per ogni finestra di analisi, per cui è possibile costruire una sua distribuzione (come mostrato in Figura 8) estraendo 9 parametri statistici: *media*, *mediana*, *moda*, *standard deviation*, *range*, *5° percentile*, *95° percentile*, *skewness* e *kurtosis*.



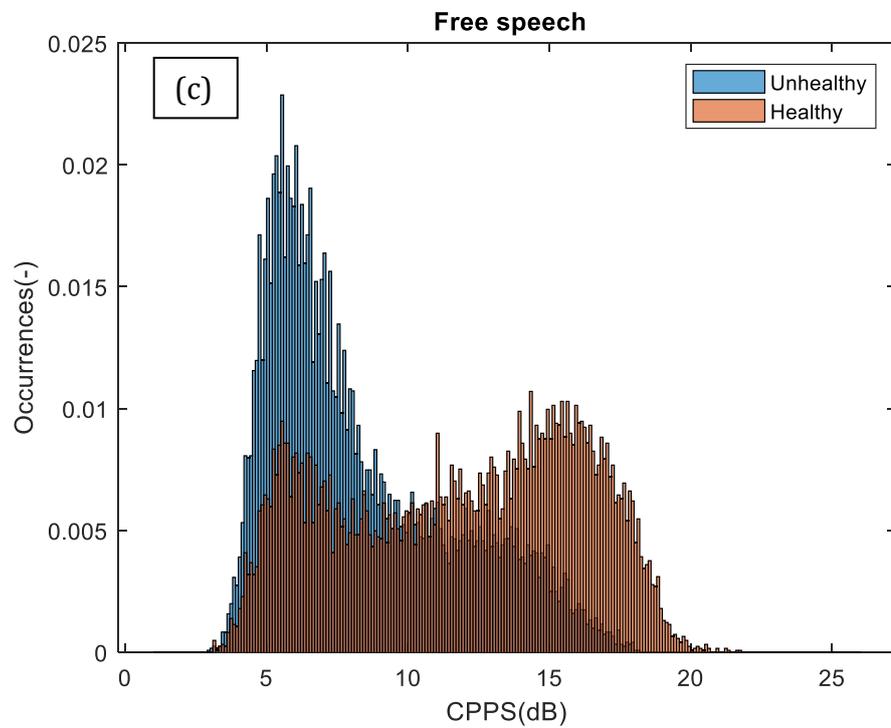
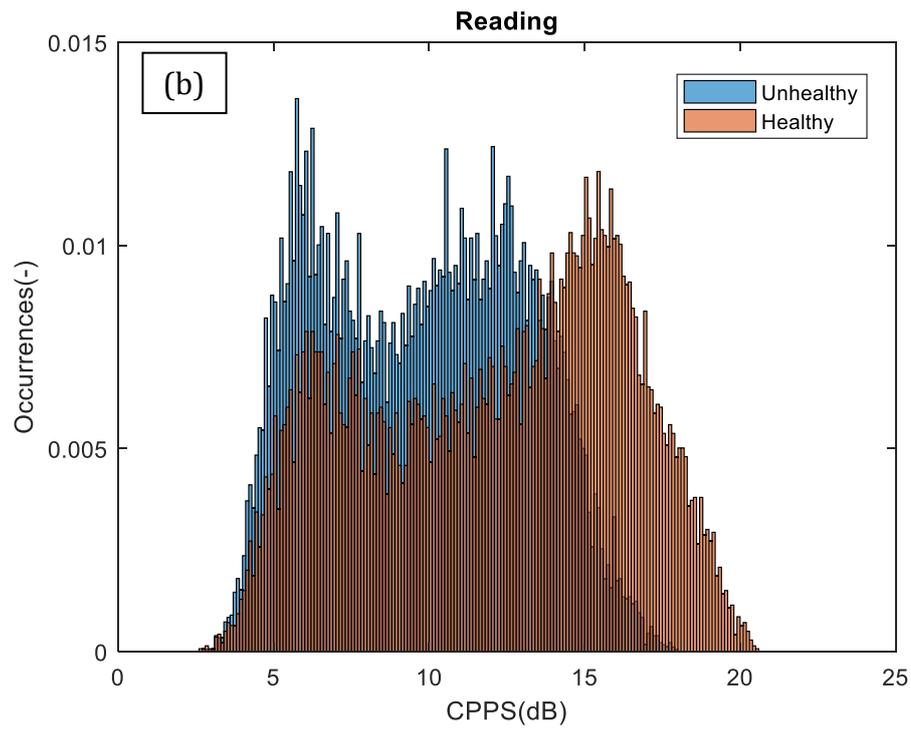


Figura 8 esempio distribuzione CPPS per soggetti sani e patologici: (a) vocale sostenuta, (b) lettura, (c) eloquio libero

### 3.6 Elaborazione dei dati

Per permettere l'elaborazione dei segnali acquisiti e l'estrazione dei dati descritti nel capitolo precedente, sono stati sviluppati due appositi script per l'ambiente di calcolo numerico Matlab® R2018b.

1. Il primo script è utilizzato soltanto per l'elaborazione dei file contenenti la vocale sostenuta 'a'. È stato sviluppato per permettere l'estrazione della frequenza fondamentale  $F_0$  e il calcolo dei parametri relativi alla variazione in frequenza ed in ampiezza descritti nei paragrafi precedenti. Inoltre tale script permette di calcolare il valore di  $F_0$ , SPL e di HNR per ogni frame di analisi, per potersi ricostruire eventualmente in futuro un grafico a dispersione per analizzare la distribuzione di tali valori. I frame di analisi sono calcolati su finestre di 1024 campioni. Prima di effettuare l'analisi, i segnali sono stati sovracampionati di un fattore 8 per aumentare la risoluzione temporale e dell'ampiezza.
2. Il secondo script è utilizzato per l'elaborazione dei file sia contenenti la vocale sostenuta 'a', sia quelli contenenti la lettura e l'eloquio libero. Lo script sviluppato permette di calcolare i valori di CPPS e di analizzare le distribuzioni ad essi associati. I segnali registrati, con frequenza di campionamento pari a 44100 Hz, vengono sottocampionati con fattore pari a 2 per ottenere una frequenza di campionamento pari a 22050 Hz conforme al calcolo del CPPS e delle sue distribuzioni [19], considerando così frame di 1024 campioni (~22 ms), con passo di sovrapposizione pari a 88 campioni (~2 ms). All'interno dello script vi è anche l'estrazione dei parametri di HNR, V/uV e V/S.

## 4 Risultati

In questo capitolo vengono presentati i risultati dei modelli ottenuti per la discriminazione di soggetti sani e soggetti affetti da malattia di Parkinson per la vocale sostenuta registrata con microfono in aria e per l'eloquio libero registrato con entrambi i microfoni. In primis, vengono proposti i metodi di scelta dei parametri descrittivi utili alla classificazione dei pazienti, successivamente i modelli di classificazione utilizzati per la discriminazione e i risultati ottenuti in termini di accuratezza, sensibilità e specificità.

### 4.1 Vocale sostenuta 'a' con microfono in aria

Come già specificato nel capitolo precedente, per l'analisi della vocale sostenuta 'a' sono stati ricavati 21 parametri descrittivi, 5 di variazione in frequenza: *Jitt*, *Jita*, *RAP*, *PPQ*,  $vF_0$ ,  $F_0$  *media*; 4 di variazione in ampiezza: *Shim*, *ShdB*, *APQ*,  $vAm$ ; 2 ricavati dall'autocorrelazione del segnale: *HNR*,  $V/U_v$ , e 9 statistiche descrittive della distribuzione del CPPS: *media*, *mediana*, *moda*, *standard deviation*, *range*, *5 percentile*, *95°percentile*, *skweness*, *kurtosis*.

#### 4.1.1 Feature selection

Per creare un classificatore che meglio potesse discriminare i soggetti in analisi, una volta ottenuti tutti i parametri è stata effettuata una selezione degli stessi, allo scopo di creare un subset minimo di features non correlate tra di loro, comparandole l'una con le altre attraverso l'indice di correlazione di *Pearson* ' $\rho$ '. Questo coefficiente è la misura della correlazione lineare tra due variabili X e Y:

$$\rho(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y}$$

Dove  $cov(X, Y)$  è la convoluzione tra X e Y, mentre  $\sigma_X$  e  $\sigma_Y$  le deviazioni tipo di X e Y. Il risultato è un valore tra +1 e -1, con +1 correlazione lineare positiva, -1 correlazione lineare negativa, 0 non correlazione.

Nella Tabella 3 sono mostrati i coefficienti di correlazione di Pearson ottenuti comparando ogni parametro della vocale sostenuta 'a' con tutti i restanti parametri, tenendo in considerazione la loro significatività ai fini statistici andando a considerare solo quelli che risultano avere un  $pValue > 0.05$ , non verificando quindi l'ipotesi nulla. Tali coefficienti infatti, non sono riportati. Il parametro  $V/S$ , irrilevante per l'analisi della vocale 'a', non è stato riportato poiché il suo valore risulta essere sempre pari a  $\sim 100\%$ , avendo opportunamente segmentato in precedenza il file. I coefficienti sulla diagonale sono evidenziati perché rappresentano la correlazione del parametro con se stesso, per cui non restituisce alcuna informazione. I coefficienti evidenziati in azzurro, invece, rappresentano quelli maggiormente correlati tra di loro, i seguenti in modo particolare:

- *Jitt, Jita, Shim, ShdB;*
- *Shim, ShdB, HNR;*
- *CPPS media, CPPS mediana, CPPS moda, CPPS 5°percentile, CPPS 95°percentile;*
- *CPPS standard deviation, CPPS range.*

Al fine di effettuare una prima scrematura dei parametri ridondanti per le successive analisi, ci si è focalizzati sul subset di parametri: *Jitt, PPQ,  $vF_0$ , APQ,  $vAm$ , HNR,  $V/Uv$ , CPPS media, CPPS range, CPPS skewness, CPPS kurtosis.*

La matrice è stata ottenuta per il microfono in aria Mipro.

## Risultati

MIPRO	<i>Jitt</i>	<i>Jita</i>	<i>RAP</i>	<i>PPQ</i>	<i>vF<sub>0</sub></i>	<i>Shim</i>	<i>ShdB</i>	<i>APQ</i>	<i>vAm</i>	<i>F<sub>0</sub></i>	<i>HNR</i>	<i>V/Uv</i>	<i>CPPS Media</i>	<i>CPPS Median</i>	<i>CPPS Moda</i>	<i>CPPS Std</i>	<i>CPPS Range</i>	<i>CPPS 5°prc</i>	<i>CPPS 95°prc</i>	<i>CPPS Skew</i>	<i>CPPS Kurt</i>	
<i>Jitt</i>	1	0,99	1,00	0,72	0,65	0,82	0,87	0,77	0,36		-0,65	-0,57	-0,41	-0,39	-0,42	0,62	0,52	-0,58	-0,18			
<i>Jita</i>		1	0,98	0,71	0,64	0,79	0,85	0,77	0,38		-0,65	-0,55	-0,36	-0,34	-0,37	0,63	0,56	-0,54				
<i>RAP</i>			1	0,71	0,65	0,81	0,86	0,76	0,35		-0,64	-0,55	-0,41	-0,38	-0,42	0,60	0,50	-0,56	-0,18			
<i>PPQ</i>				1	0,54	0,66	0,68	0,66			-0,58	-0,42			-0,24	0,66	0,42	-0,38				
<i>vF<sub>0</sub></i>					1	0,53	0,56	0,52	0,28	0,20	-0,45	-0,41	-0,34	-0,30	-0,32	0,51	0,46	-0,48		-0,22		
<i>Shim</i>						1	0,99	0,97	0,43		-0,82	-0,52	-0,37	-0,34	-0,34	0,63	0,49	-0,54				
<i>ShdB</i>							1	0,97	0,43		-0,82	-0,55	-0,39	-0,36	-0,36	0,65	0,52	-0,56				
<i>APQ</i>								1	0,42		-0,82	-0,55	-0,33	-0,31	-0,30	0,66	0,52	-0,53				
<i>vAm</i>									1		-0,44		-0,27	-0,24	-0,19	0,45	0,48	-0,42		-0,24		
<i>F<sub>0</sub></i>										1		-0,23	-0,51	-0,50	-0,48			-0,46	-0,47		-0,20	
<i>HNR</i>											1	0,37	0,20			-0,56	-0,48	0,39		0,22		
<i>V/Uv</i>												1	0,39	0,38	0,38	-0,34	-0,19	0,44	0,25			
<i>CPPS Media</i>													1	1,00	0,93	-0,31		0,90	0,90	-0,23	0,21	
<i>CPPS Median</i>														1	0,94	-0,28		0,87	0,90	-0,30	0,23	
<i>CPPS Moda</i>															1	-0,28		0,81	0,82	-0,31	0,22	
<i>CPPS Std</i>																1	0,82	-0,66		-0,26		
<i>CPPS Range</i>																	1	-0,44	0,19	-0,42	0,45	
<i>CPPS 5°prc</i>																		1	0,67			
<i>CPPS 95°prc</i>																			1	-0,22	0,19	
<i>CPPS Skew</i>																				1	-0,58	
<i>CPPS Kurt</i>																						1

Tabella 3 Matrice di correlazione dei parametri della vocale 'a'

### 4.1.2 Creazione classificatore

Considerando solo uno dei parametri di ogni set delle features ridondanti presentate nel paragrafo precedente, si è utilizzata l'applicazione di Matlab 'Classification Learning' per la creazione di un classificatore. Grazie all'analisi del *Parallel Coordinates Plot* è stato possibile verificare quali fossero i parametri utili per separare le classi, in modo da poterlo addestrare solo con le features più utili.

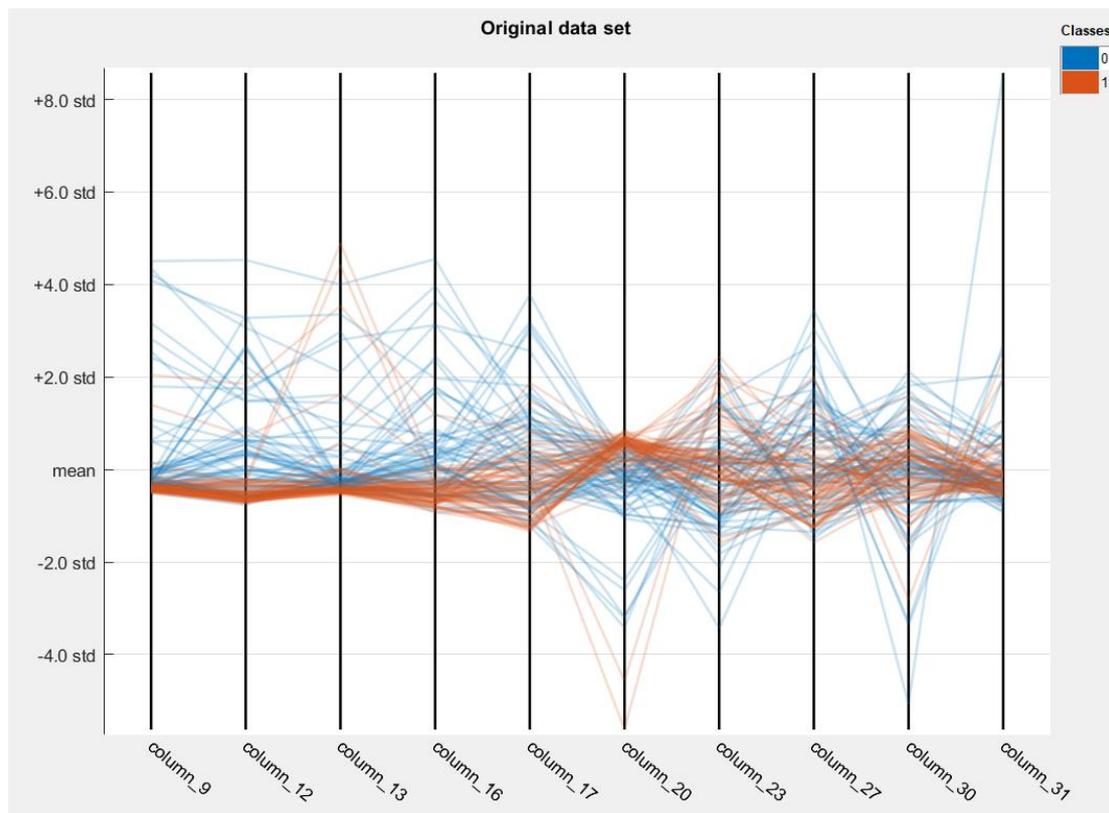


Figura 9 Esempio di Parallel Coordinates Plot

Lo scopo di quest'analisi statistica è stata quella di scegliere quali parametri meglio potessero discriminare soggetti sani da soggetti affetti dalla malattia di Parkinson. Per questo motivo si è scelta di creare due modelli di classificatore: regressione logistica e *coarse tree*. La scelta è caduta su questi due modelli perché:

- **Regressione logistica:** è un modello di regressione non lineare in presenza di una variabile dipendente di tipo dicotomica, utile per stabilire la probabilità con cui un certo dato appartenga o meno ad uno dei due valori della variabile dipendente. L'impiego è finalizzato a classificare un dato che può assumere

come unici valori 0 o 1, restituendo la probabilità che esso appartenga alla classe 1 con la seguente funzione:

$$P(1) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}$$

Con P che indica la probabilità, X il vettore delle variabili indipendenti,  $\beta$  il vettore dei parametri.

- **Coarse tree:** è un modello di classificazione del tipo albero decisionale. Scomponere un set di dati in sottoinsiemi sempre più piccoli, mentre è sviluppato un nuovo albero decisionale. Il risultato è un albero con nodi decisionali e nodi a fogli, in cui il nodo decisionale ha due o più rami per la decisione, mentre quello a foglia rappresenta la classificazione finale. Il nodo decisionale più in alto corrisponde al miglior predittore chiamato nodo radice [20]. Nel caso della tipologia 'coarse', l'albero presenta al massimo 4 nodi, utile nel caso di piccoli dataset per evitare il problema dell'overfitting.

Per la costruzione dei modelli di classificatore per la vocale 'a' sostenuta, si è scelto di utilizzare le tre vocali singolarmente in modo da poter creare un classificatore più robusto avendo come input un dataset non da 44 elementi (22 soggetti affetti da Parkinson e 22 soggetti sani), ma da 132-

#### 4.1.2.1 *Regressione logistica*

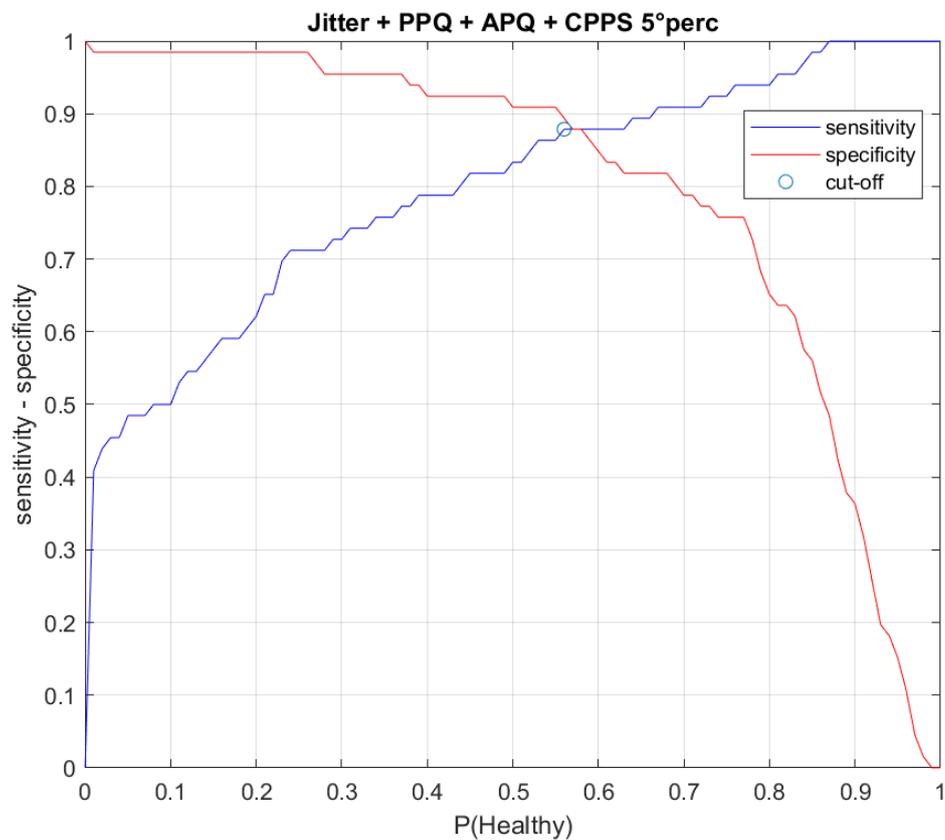
Per la creazione del modello di regressione lineare, oltre all'aver analizzato il *Parallel Coordinates Plot*, per la scelta dei parametri ci si è anche aiutati andando a verificare il *p-value* delle *features*. Se tale valore risulta infatti inferiore a 0.05, l'ipotesi nulla non è rifiutata e il parametro risulta significativamente irrilevante alla creazione del modello.

Il modello di regressione logistica che meglio è riuscito a discriminare la due classi è risultato il seguente:

$$P(\text{Healthy}) = \frac{e^{-3.99+3.45 \cdot \text{Jitt}-3.66 \cdot \text{PPQ}-0.93 \cdot \text{APQ}+0.75 \cdot \text{CPPS}_{5pr}}}{1 + e^{-3.99+3.45 \cdot \text{Jitt}-3.66 \cdot \text{PPQ}-0.93 \cdot \text{APQ}+0.75 \cdot \text{CPPS}_{5pr}}}$$

Ovvero un modello a quattro parametri: *Jitt*, *PPQ*, *APQ* e *CPPS 5°percentile*.

Una volta creato il modello, si è provveduto a scegliere la soglia ottimale per la distinzione nelle due classi analizzando i grafici di sensibilità (veri positivi) e specificità (veri negativi), concedendo priorità alla sensibilità.



**Figura 10** Scelta del taglio del modello di regressione lineare per la vocale a

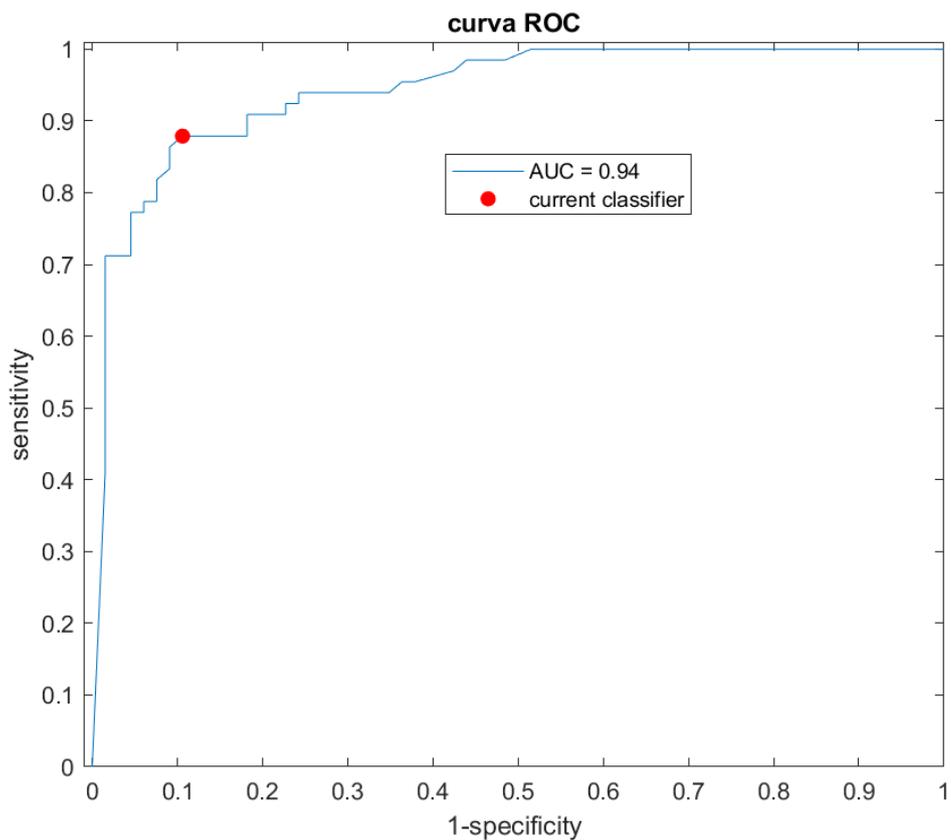


Figura 11 Curva ROC del modello a regressione lineare della vocale a

La scelta è ricaduta su una soglia per la distinzione pari a 0.56 tra le due classi, in modo da ottenere una sensibilità migliore lasciando inalterata la specificità.. Come è possibile notare dalla Figura 12, il classificatore così costruito presenta una confusion matrix del tipo (Tabella 4):

		Classe predetta	
		0	1
Classe reale	0	58	8
	1	7	59

Tabella 4 Confusion matrix regressione logistica vocale a

Con:

- 0 classe Parkinson
- 1 classe Sani

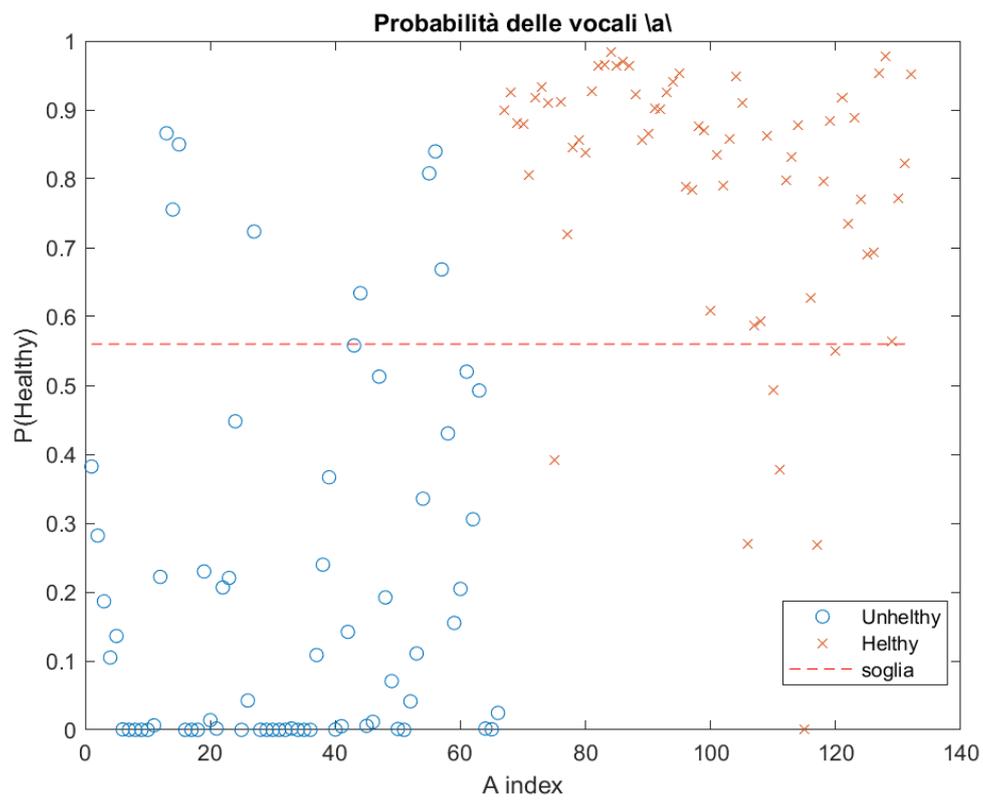


Figura 12 Probabilità delle vocali a con il modello di regressione logistica

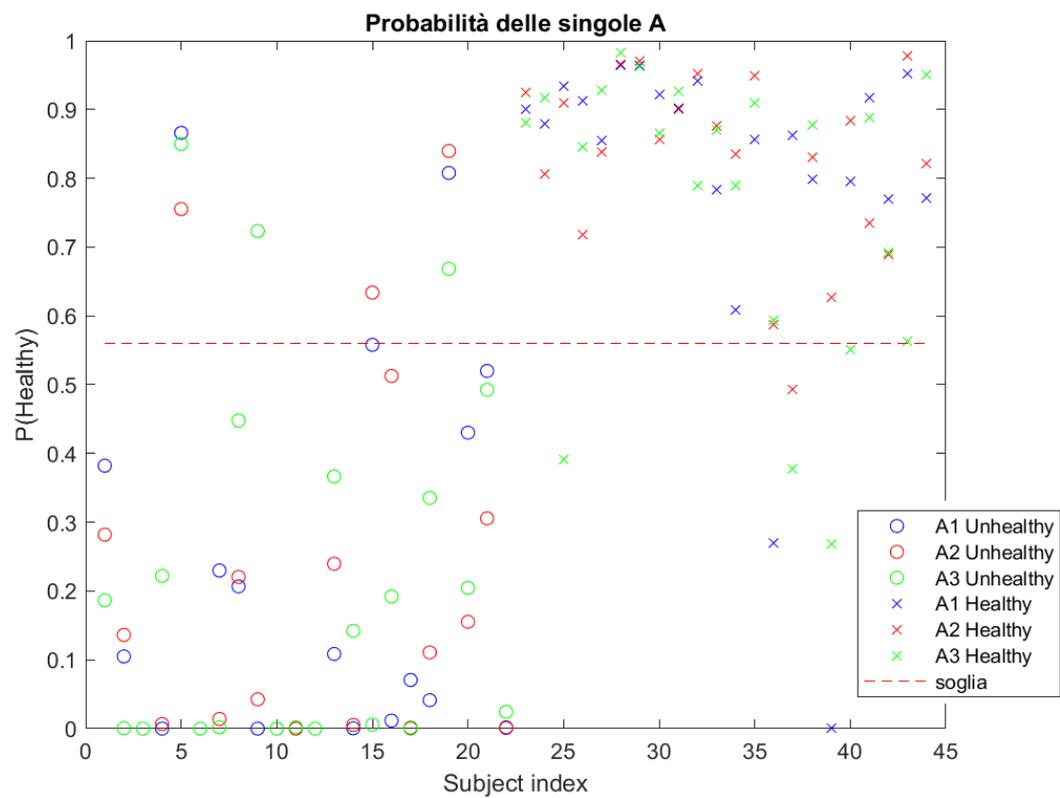


Figura 13 Probabilità delle vocali 'a' con il modello di regressione logistica per ogni singolo soggetto

Ottenendo così ottime prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 88.6%
- *Sensitivity* = 87.8%
- *Specificity* = 89.4%

Infine il classificato creato è stato testato considerando le un'unica 'a' per soggetto, in modo che ogni paziente avesse un singolo responso. Per ottenere i nuovi parametri delle 'a', sono stati considerati i valor medi delle singole 'a' registrate da ogni soggetto monitorato.

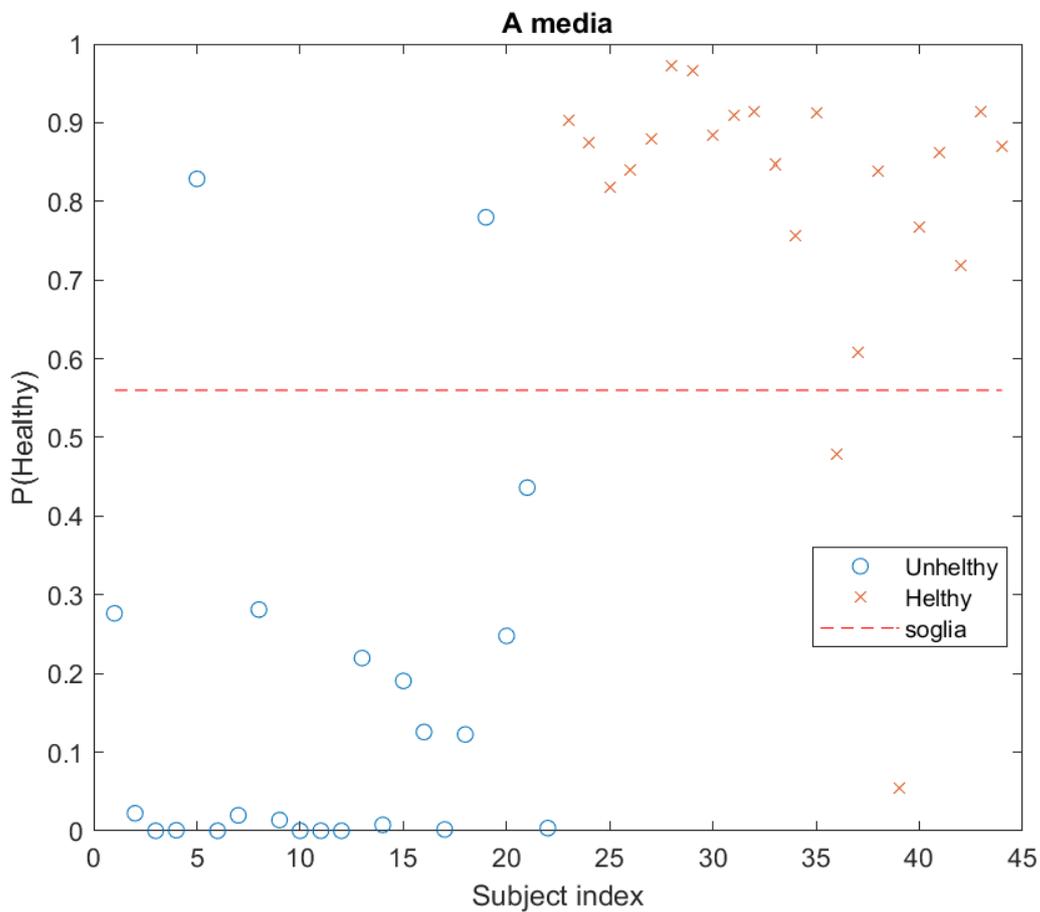


Figura 14 Probabilità delle tre 'a' mediate per ogni singolo paziente

Come si nota dalla Figura 14, le prestazioni del modello risultano essere pressoché invariate:

		Classe predetta	
		0	1
Classe reale	0	20	2
	1	2	20

Tabella 5 Confusion matrix regressione logistica per le 'a' mediate

Ottenendo così prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 90.9%
- *Sensitivity* = 90.9%
- *Specificity* = 90.9%

#### 4.1.2.2 *Coarse tree*

Il classificatore con il modello del *coarse tree*, a differenza della regressione logistica, non restituisce una probabilità di appartenenza alla classe, ma è del tipo *in/out*, ovvero appartiene ad una determinata classe o meno. Questo perché il modello divide lo spazio in  $n$  dimensioni con  $n$  pari al numero di parametri rilevanti (Figura 16), che meglio riesce a discriminare il dataset nelle due classi di appartenenza.

In seguito (Figura 15) è riportato l'albero di decisione per la distinzioni nelle due classi.

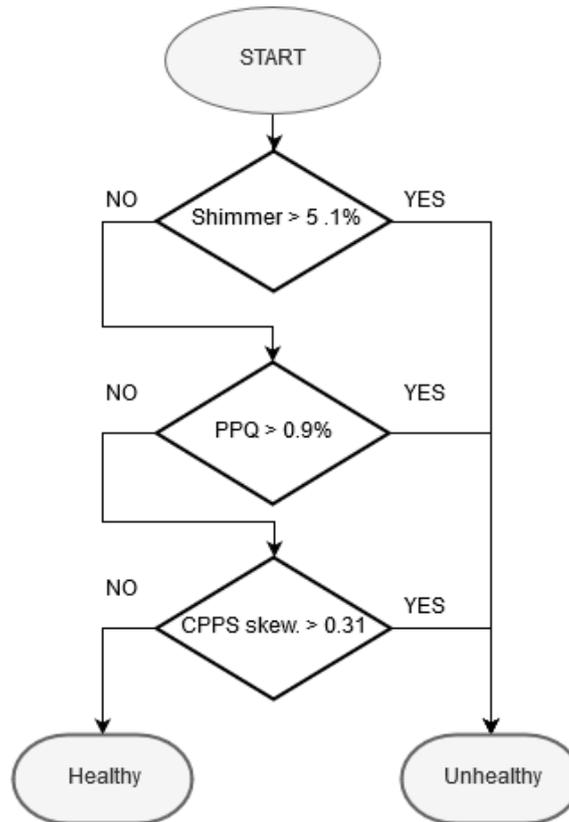


Figura 15 Decision tree vocale 'a'

Come si nota dal grafo, i parametri scelti per la classificazione sono risultati: *Shim*, *PPQ* e *CPPS skewness*.

Il classificatore così costruito presenta una confusion matrix del tipo (Tabella 6):

		Classe predetta	
		0	1
Classe reale	0	60	6
	1	4	62

Tabella 6 Confusion matrix coarse tree vocale a

Con:

- 0 classe Parkinson
- 1 classe Sani

Ottenendo così ottime prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 92.4%
- *Sensitivity* = 90.9%
- *Specificity* = 93.9%

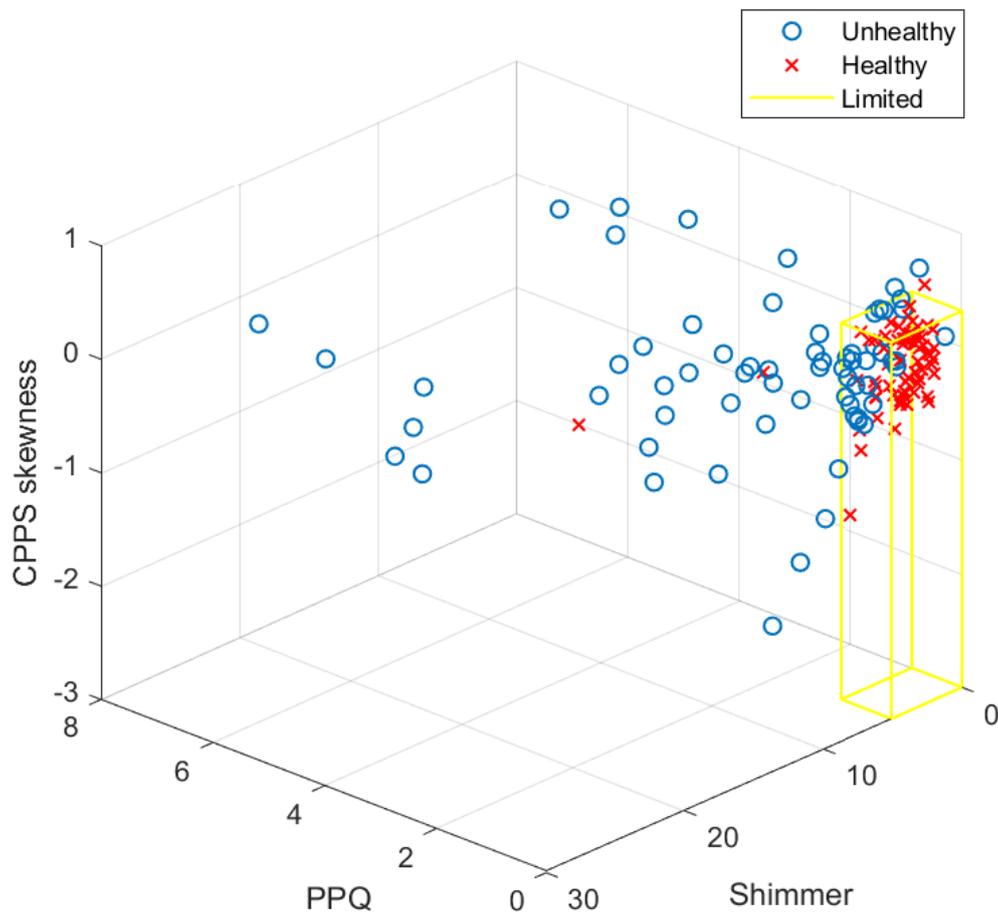
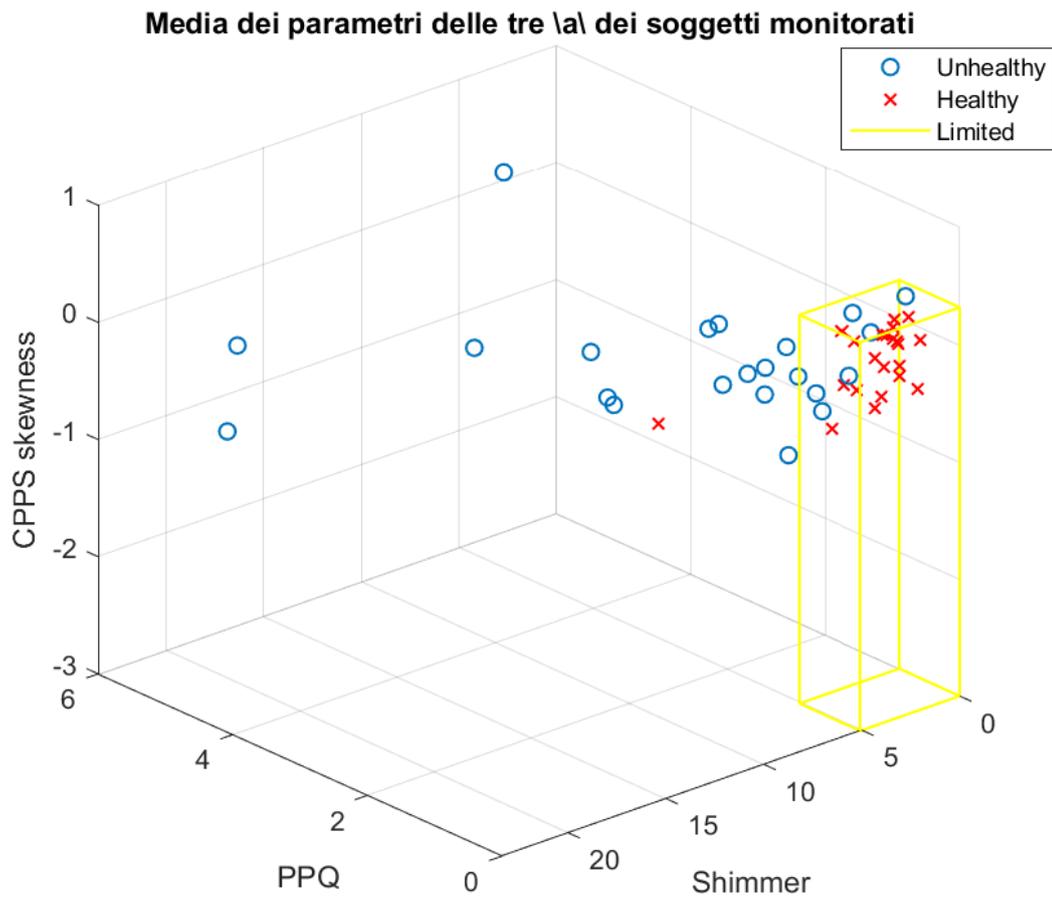


Figura 16 Divisione dello spazio delle features con l'albero decisionale per la vocale 'a'

Anche in questo caso il classificatore creato è stato testato considerando le medie dei parametri delle tre \a\ di ogni soggetto monitorato per ottenere un'unica discriminante tra le due classi di appartenenza per ogni singolo paziente.



**Figura 17** Divisione dello spazio delle features con l'albero decisionale per la media delle vocali 'a'

Le prestazioni del modello risultano essere pressoché invariate:

		Classe predetta	
		0	1
Classe reale	0	19	3
	1	2	20

**Tabella 7** Confusion matrix regressione logistica per le 'a' mediate

Ottenendo così prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 88.6%
- *Sensitivity* = 86.3%
- *Specificity* = 90.9%

## 4.2 Eloquio libero con microfono in aria

Per l'analisi dell'eloquio libero registrato con il microfono in aria, come descritto nei paragrafi precedenti, sono stati calcolati 13 parametri, 1 di variazione in frequenza:  $F_0$  *media*; 2 ricavati dall'autocorrelazione del segnale *HNR*,  $V/U_v$ ; e 9 parametri statistici relativi al CPPS: *media*, *mediana*, *moda*, *standard deviation*, *range*, *5 percentile*, *95°percentile*, *skewness*, *kurtosis*.

Come già specificato parametri descrittivi di variazione in frequenza e in ampiezza come: *Jitt*, *Jita*, *RAP*, *PPQ*,  $vF_0$ , *Shim*, *ShdB*, *APQ*,  $vAm$ ; non sono stati calcolati a differenza della vocale 'a' sostenuta perché irrilevanti per la discriminazione di patologie all'apparato fonatorio.

### 4.2.1 Feature selection

Come per la vocale 'a' la prima discriminazione effettuata sui parametri descrittivi è stata effettuata analizzando l'indice di correlazione di *Pearson* ' $\rho$ '. Nella Tabella 8 riportata in seguito, sono presentati i coefficienti di correlazione di *Pearson* ottenuti comparando ogni parametro ricavato dai file dell'eloquio libero con tutti i restanti parametri.

MIPRO	$F_0$	HNR	V/Uv	Media	Median	Moda	CPPS Std	CPPS Range	CPPS 5°pr	CPPS 95°pr	CPPS Skew	CPPS Kurt
$F_0$	1						-0,39					0,41
HNR		1	0,70	0,33	0,30				0,41		-0,40	0,35
V/Uv			1	0,33	0,32				0,35		-0,36	
CPPS Media				1	1,00	0,88		0,67	0,87	0,92	-0,95	0,57
CPPS Median					1	0,91		0,69	0,85	0,92	-0,96	0,57
CPPS Moda						1		0,70	0,70	0,80	-0,87	0,48
CPPS Std							1	0,38	-0,42	0,36		-0,67
CPPS Range								1	0,36	0,76	-0,63	
CPPS 5°prc									1	0,68	-0,82	0,75
CPPS 95°prc										1	-0,81	
CPPS Skew											1	-0,67
CPPS Kurt												1

Tabella 8 Matrice di correlazione con indici di Pearson per parametri dell'eloquio libero con microfono in aria

I coefficienti sulla diagonale sono evidenziati perché rappresentano la correlazione del parametro con se stesso, per cui non restituisce nessuna informazione. I coefficienti evidenziati in azzurro invece rappresentano i coefficienti maggiormente correlati tra di loro, in particolare i seguenti:

- *CPPS media, CPPS mediana, CPPS moda, CPPS 5°percentile, CPPS 95°percentile, CPPS skewness.*

Al fine di effettuare una prima scrematura dei parametri ridondanti per le successive analisi ci si è focalizzati sul subset di parametri:  $F_0$ , HNR, V/Uv, CPPS media, CPPS standard deviation, CPPS range, CPPS kurtosis.

La matrice è stata ottenuta per il microfono in aria MIPRO. Inoltre si può notare che alcuni coefficienti non sono stati riportati. Tali valori omessi non risultavano statisticamente significativi, avendo un  $pValue > 0.05$ , non verificando quindi l'ipotesi nulla.

## 4.2.2 Creazione classificatore

Considerando quindi solo uno dei parametri del set delle features ridondanti, è stata utilizzata l'applicazione di Matlab 'Classification Learning' per la creazione dei due modelli di classificatori: *Regressione logistica* e *Coarse tree*. Grazie all'analisi del *Parallel Coordinates Plot* è stato possibile verificare quali fossero i parametri utili per separare le classi, in modo da poterlo addestrare solo con le features più utili.

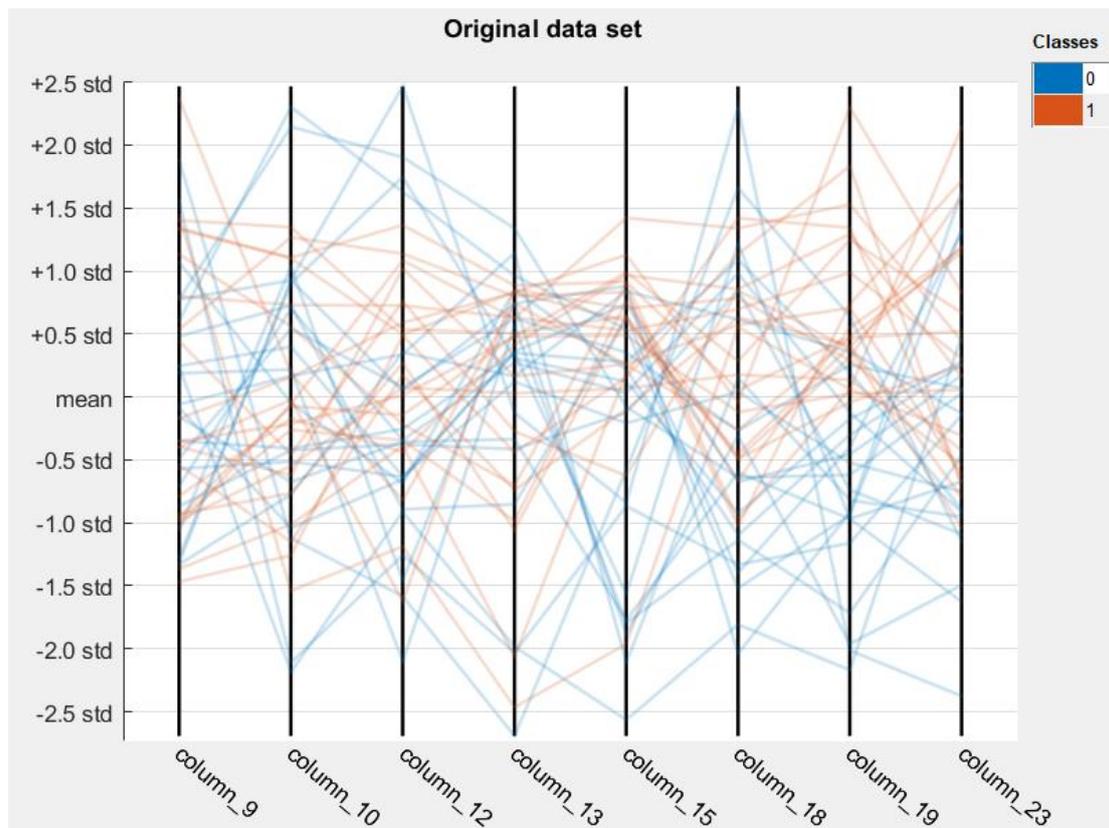


Figura 18 Esempio di Parallel Coordinates Plot

Lo scopo di quest'analisi statistica è stata quella di scegliere quali parametri meglio potessero discriminare soggetti sani da soggetti affetti dalla malattia di Parkinson.

### 4.2.2.1 *Regressione logistica*

Nel processo di creazione del modello di regressione lineare, oltre all'aver analizzato il *Parallel Coordinates Plot*, per la scelta dei parametri ci si è affidati alla verifica del *p-value* delle *features*. Se tale valore, infatti, risulta inferiore a 0.05, l'ipotesi nulla non è rifiutata e il parametro risulta significativamente irrilevante alla creazione del modello.

Il modello di regressione logistica che meglio è riuscito a discriminare la due classi è risultato il seguente:

$$P(\text{Healthy}) = \frac{e^{-4.74-0.42 \cdot \text{HNR}+0.88 \cdot \text{CPPS mean}+1.52 \cdot \text{CPPS kurtosis}}}{1 + e^{-4.74-0.42 \cdot \text{HNR}+0.88 \cdot \text{CPPS mean}+1.52 \cdot \text{CPPS kurtosis}}}$$

Ovvero un modello a tre parametri: *HNR*, *CPPS media* e *CPPS kurtosis*.

Una volta creato il modello, è stata scelta la soglia ottimale per la distinzione nelle due classi analizzando i grafici di sensibilità (veri positivi) e specificità (veri negativi), concedendo priorità alla sensibilità.

La scelta è ricaduta su una soglia per la distinzione pari a 0.53 tra le due classi in modo da ottenere una sensibilità migliore lasciando inalterata la specificità, come è possibile notare dalla Figura 19. Il classificatore così costruito presenta una confusion matrix del tipo (Tabella 9):

		Classe predetta	
		0	1
Classe reale	0	16	6
	1	3	19

Tabella 9 Confusion matrix regressione logistica eloquio libero MIPRO

Con:

- 0 classe Parkinson
- 1 classe Sani

Ottenendo così prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 79.5%
- *Sensitivity* = 72.7%
- *Specificity* = 86.3%

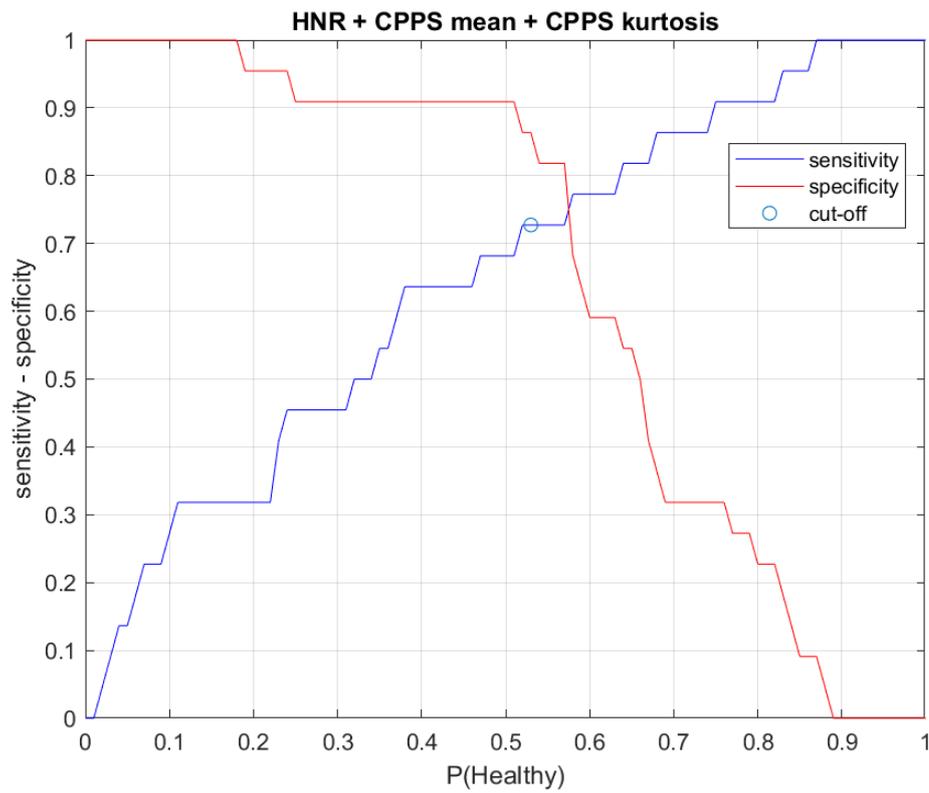


Figura 19 Scelta del taglio del modello di regressione lineare per eloquio libro con MIPRO

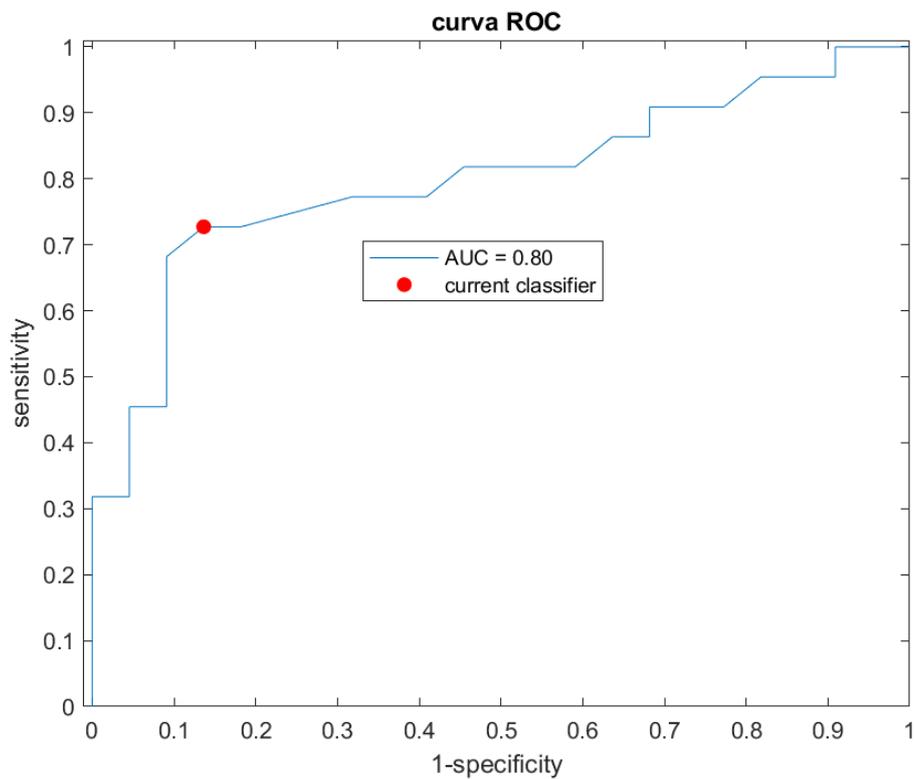


Figura 20 Curva ROC del modello a regressione lineare dell'eloquio libero con MIPRO

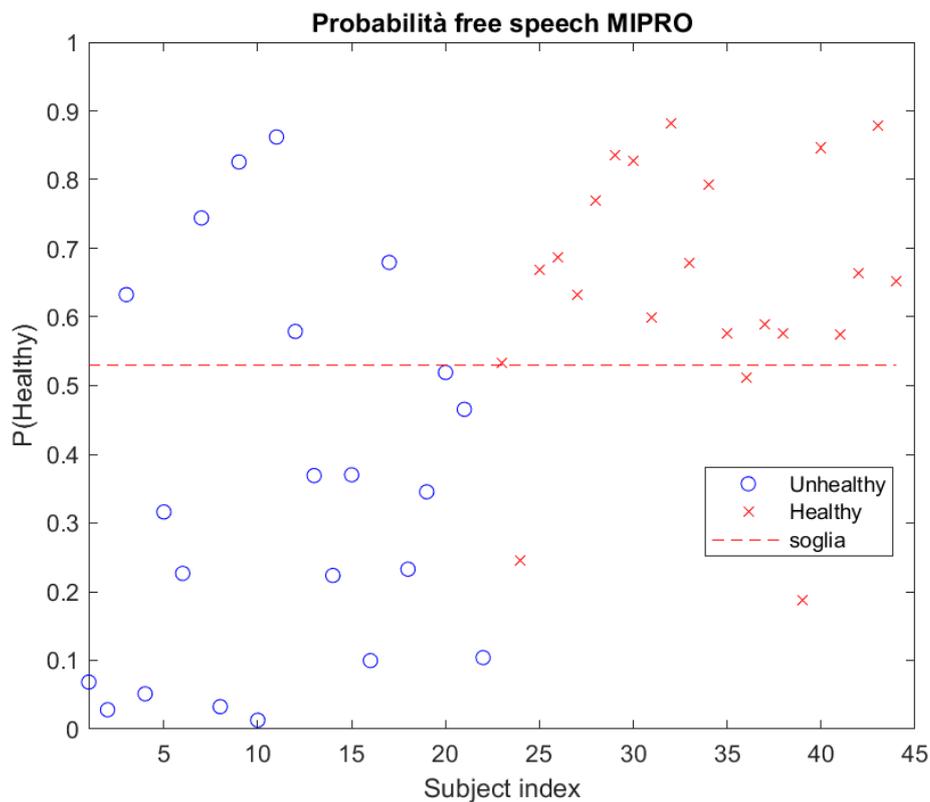


Figura 21 Probabilità eloquio libero con MIPRO per modello di regressione logistica

#### 4.2.2.2 *Coarse tree*

Il classificatore con il modello del *coarse tree*, a differenza della regressione logistica non restituisce una probabilità di appartenenza alla classe ma è del tipo *in/out*, ovvero appartiene ad una determinata classe o meno. Questo perché il modello divide lo spazio in  $n$  dimensioni con  $n$  pari al numero di parametri rilevanti (Figura 23), che meglio riesce a discriminare il dataset nelle due classi di appartenenza. In seguito (Figura 22) è riportato l'albero di decisione per la distinzioni nelle due classi. Come si nota dal grafico, i parametri scelti per la classificazione sono risultati: *CPPS range*, *CPPS kurtosis*, *CPPS moda*.

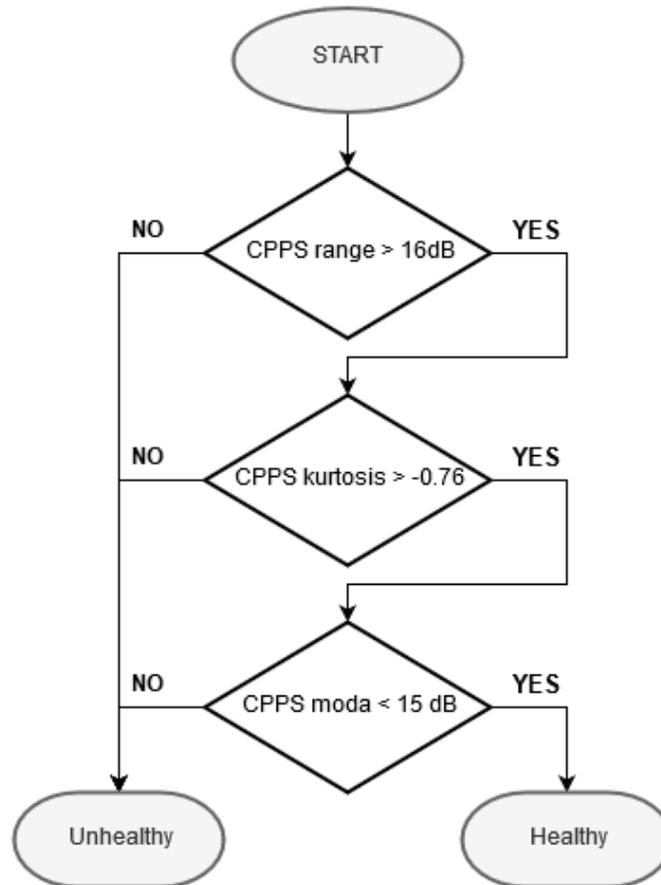


Figura 22 Decision tree free speech MIPRO

Il classificatore così costruito presenta una confusion matrix del tipo (Tabella 10):

		Classe predetta	
		0	1
Classe reale	0	21	1
	1	2	20

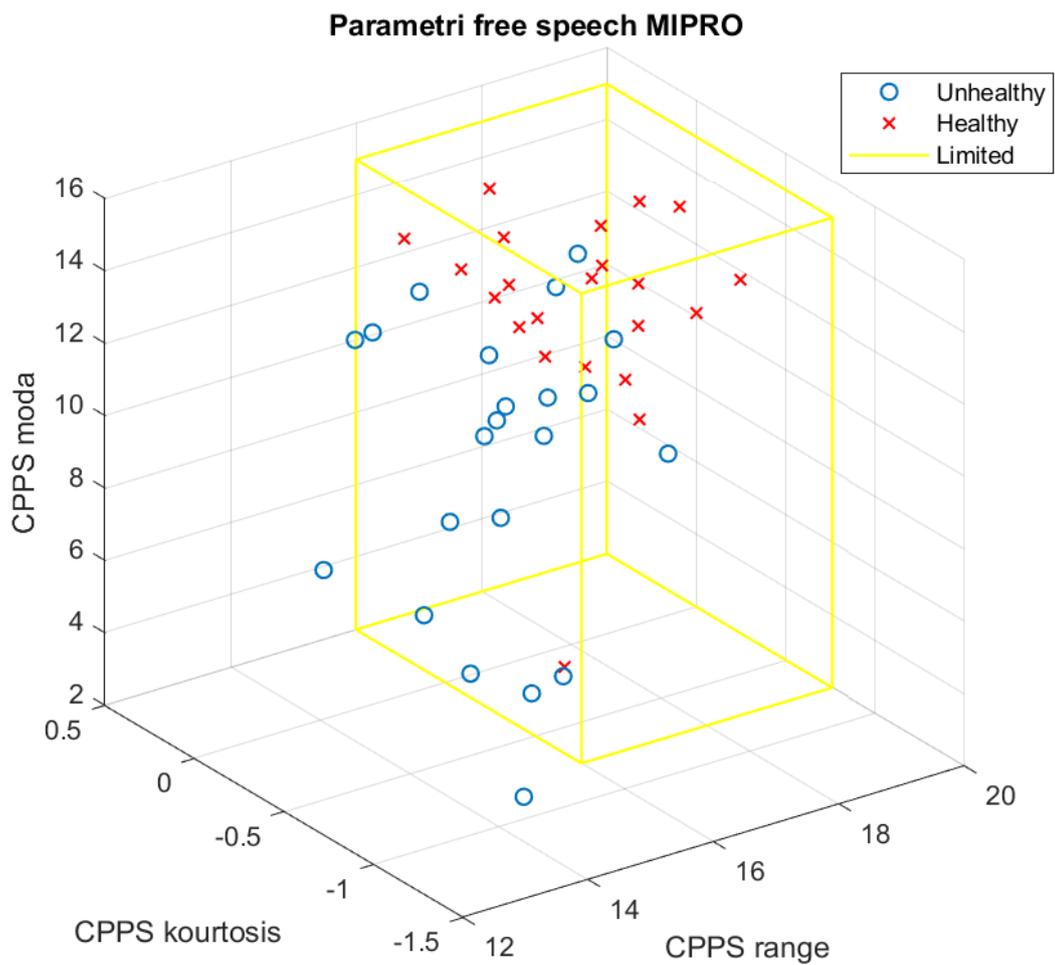
Tabella 10 Confusion matrix coarse tree free speech MIPRO

Con:

- 0 classe Parkinson
- 1 classe Sani

Ottenendo così ottime prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 93.1%
- *Sensitivity* = 95.4%
- *Specificity* = 90.9%



**Figura 23** Divisione dello spazio delle features con l'albero decisionale per free speech MIPRO

### 4.3 Eloquio libero con microfono a contatto

Per l'analisi dell'eloquio libero registrato con il microfono a contatto, così come per i file registrati con il microfono in aria, sono stati calcolati 13 parametri, 1 di variazione in frequenza:  $\Delta F_0$ ; 2 di ricavate dall'auto correlazione del segnale *HNR*, *V/Uv*; e 9 parametri statistici relativi al CPPS: *media*, *mediana*, *moda*, *standard deviation*, *range*, *5°percentile*, *95°percentile*, *skewness*, *kurtosis*.

Come già specificato parametri descrittivi di variazione in frequenza e in ampiezza come: *Jitt*, *Jita*, *RAP*, *PPQ*,  $vF_0$ , *Shim*, *ShdB*, *APQ*, *vAm*; non sono stati calcolati a differenza della vocale 'a' sostenuta perché irrilevanti per la discriminazione di patologie all'apparato fonatorio.

#### 4.3.1 Feature selection

Come per la vocale 'a' e per il microfono in aria, la prima discriminazione effettuata sui parametri descrittivi è stata effettuata analizzando l'indice di correlazione di *Pearson* ' $\rho$ '. Nella Tabella 11 riportata in seguito, sono presentati i coefficienti di correlazione di *Pearson* ottenuti comparando ogni parametro ricavato dai file dell'eloquio libero registrato con il microfono a contatto con tutti i restanti parametri.

I coefficienti sulla diagonale sono evidenziati perché rappresentano la correlazione del parametro con se stesso, per cui non restituisce nessuna informazione. I coefficienti evidenziati in azzurro invece rappresentano i coefficienti maggiormente correlati tra di loro, in particolare i seguenti:

- *CPPS media*, *CPPS mediana*, *CPPS moda*, *CPPS 5°percentile*, *CPPS 95°percentile*, *CPPS skewness*;
- *CPPS skewness*, *CPPS kurtosis*.

PIEZO	$F_0$	HNR	V/U <sub>v</sub>	CPPS Media	CPPS Median	CPPS Moda	CPPS Std	CPPS Range	CPPS 5°pr	CPPS 95°pr	CPPS Skew	CPPS Kurt
$F_0$	1						-0,41		0,34		-0,44	0,51
HNR		1	0,73	0,53	0,51	0,38	-0,45		0,60	0,36	-0,59	0,63
V/U <sub>v</sub>			1	0,39	0,35				0,46		-0,35	0,39
CPPS Media				1	0,99	0,82		0,53	0,88	0,90	-0,94	0,76
CPPS Median					1	0,83		0,55	0,84	0,88	-0,95	0,72
CPPS Moda						1		0,45	0,61	0,84	-0,75	0,49
CPPS Std							1		-0,65		0,44	-0,70
CPPS Range								1	0,35	0,65	-0,40	
CPPS 5°pr									1	0,69	-0,83	0,89
CPPS 95°pr										1	-0,73	0,52
CPPS Skew											1	-0,83
CPPS Kurt												1

Tabella 11 Matrice di correlazione con indici di Pearson per parametri dell'eloquio libero con microfono a contatto

Al fine di effettuare una prima scrematura dei parametri ridondanti per le successive analisi ci si è focalizzati sul subset di parametri:  $\Delta F_0$ , HNR, V/S, V/U<sub>v</sub>, CPPS media, CPPS standard deviation, CPPS range, CPPS kurtosis.

La matrice è stata ottenuta per il microfono a contatto. Inoltre si può notare che alcuni coefficienti non sono stati riportati. Tali valori omessi non risultavano statisticamente significativi, avendo un  $pValue > 0.05$ , non verificando quindi l'ipotesi nulla.

### 4.3.2 Creazione classificatore

Considerando quindi solo uno dei parametri del set delle features ridondanti, si è utilizzata l'applicazione di Matlab 'Classification Learning' per la creazione dei due modelli di classificatori: *Regressione logistica* e *Coarse tree*. Grazie all'analisi del *Parallel Coordinates Plot* è stato possibile verificare quali fossero i parametri utili per separare le classi, in modo da poterlo addestrare solo con le features più utili.

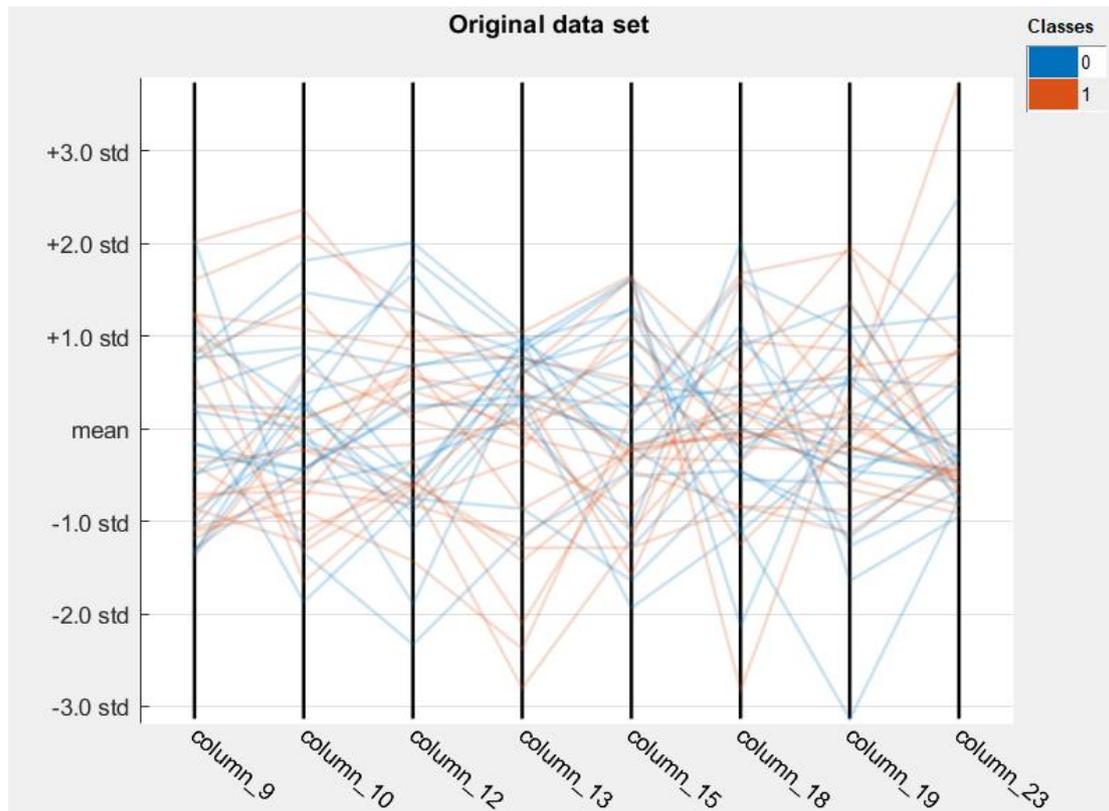


Figura 24 Esempio di Parallel Coordinates Plot

Lo scopo di quest'analisi statistica è stato quello di scegliere quali parametri meglio potessero discriminare soggetti sani da soggetti affetti dalla malattia di Parkinson. La creazione dei classificatori per l'eloquio libero con microfono a contatto è stato effettuata su un numero di soggetti pari a 38 (19 parkinsoniani e 19 sani), inferiore al numero di soggetti per le precedenti analisi con il microfono in aria. In quanto a causa della presenza di ossigenatori, non è stato possibile per 3 pazienti indossare il collare con i microfoni piezoelettrici.

#### 4.3.2.1 *Regressione logistica*

Nella creazione del modello di regressione lineare, oltre all'aver analizzato il *Parallel Coordinates Plot*, per la scelta dei parametri è stato verificato il *p-value* delle *features*. Se tale valore risulta infatti inferiore a 0.05, l'ipotesi nulla non è rifiutata e il parametro risulta significativamente irrilevante alla creazione del modello.

Il modello di regressione logistica che meglio è riuscito a discriminare la due classi per l'eloquio libero registrato con microfono a contatto è risultato il seguente:

$$P(\text{Healthy}) = \frac{e^{2.31-0.19 \cdot \text{Voiced} \setminus \text{Unvoiced} + 1.00 \cdot \text{CPPS range}}}{1 + e^{2.31-0.19 \cdot \text{Voiced} \setminus \text{Unvoiced} + 1.00 \cdot \text{CPPS range}}}$$

Ovvero un modello a due parametri: *Voiced\Unvoiced* e *CPPS range*.

Una volta creato il modello, si è provveduto a scegliere la soglia ottimale per la distinzione nelle due classi, analizzando i grafici di sensibilità (veri positivi) e specificità (veri negativi), dando priorità alla sensibilità.

Si è scelti una soglia per la distinzione pari a 0.50 tra le due classi in modo da ottenere una sensibilità migliore lasciando inalterata la specificità. Come è possibile notare dalla Figura 25. Il classificatore così costruito presenta una confusion matrix del tipo (Tabella 12):

		Classe predetta	
		0	1
Classe reale	0	15	4
	1	5	14

Tabella 12 Confusion matrix eloquio libero PIEZO regressione logistica

Con:

- 0 classe Parkinson
- 1 classe Sani

Ottenendo così prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 76.3%
- *Sensitivity* = 78.9%
- *Specificity* = 73.6%

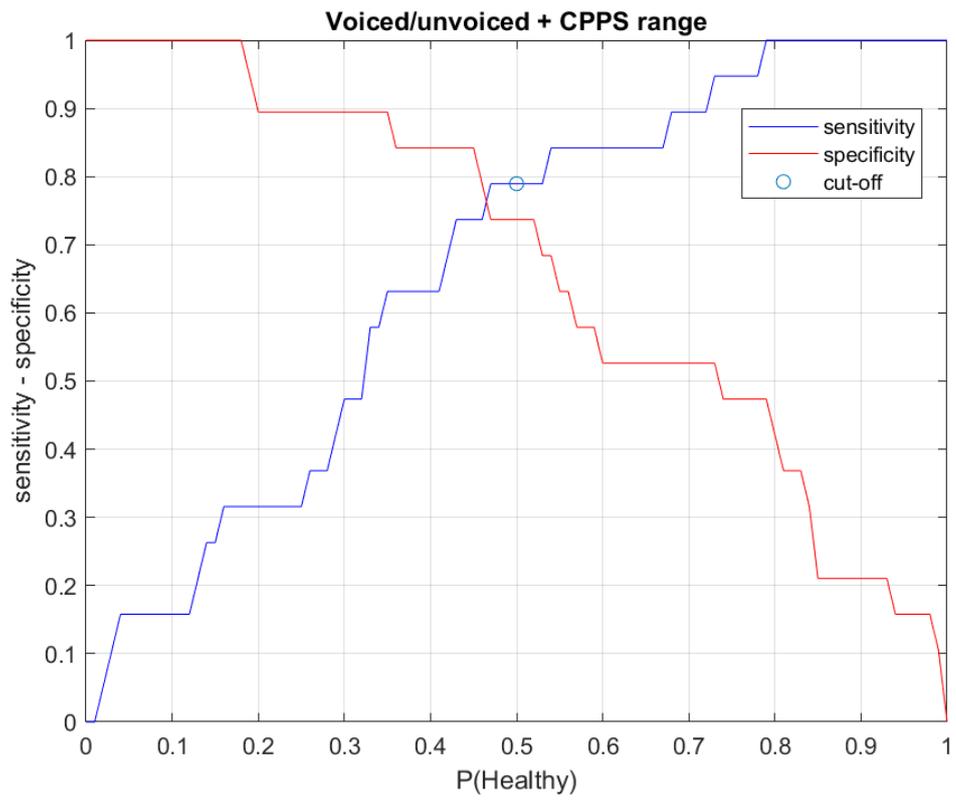


Figura 25 Scelta del taglio del modello di regressione lineare per eloquio libro con PIEZO

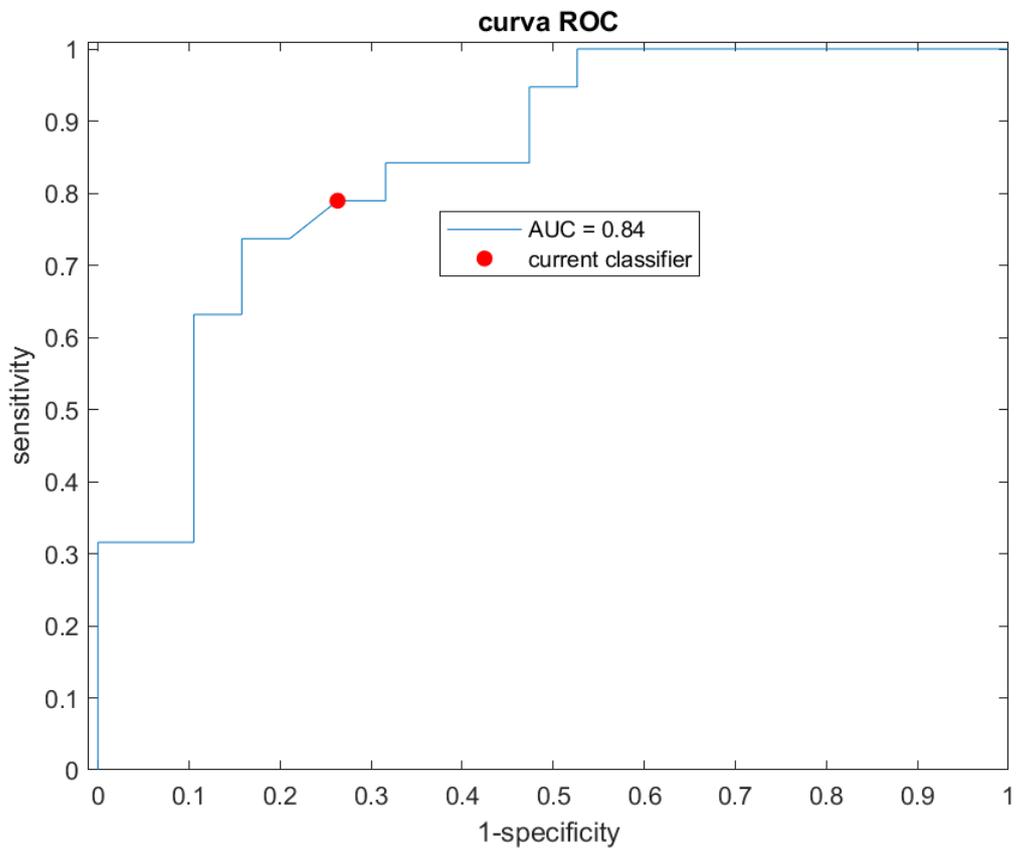


Figura 26 Curva ROC del modello a regressione lineare dell'eloquio libero con PIEZO

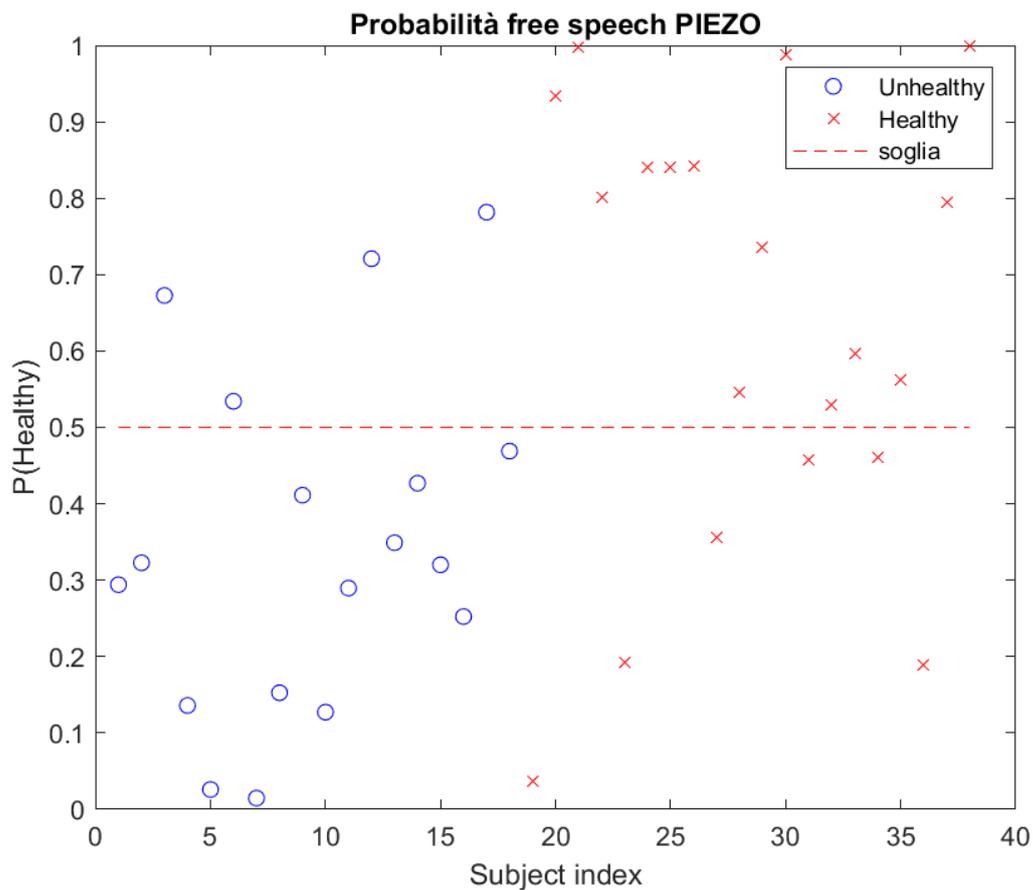


Figura 27 Probabilità eloquio libero con PIEZO per modello di regressione logistica

#### 4.3.2.2 Coarse tree

Il classificatore con il modello del *coarse tree*, a differenza della regressione logistica, non restituisce una probabilità di appartenenza alla classe ma è del tipo *in/out*, ovvero appartiene ad una determinata classe o meno. Questo perché il modello divide lo spazio in  $n$  dimensioni con  $n$  pari al numero di parametri rilevanti (Figura 29), che meglio riesce a discriminare il dataset nelle due classi di appartenenza. In seguito (Figura 28) è riportato l'albero di decisione per la distinzioni nelle due classi. Come si nota dal grafo, i parametri scelti per la classificazione sono risultati:  $\Delta F_0$ , *Voiced/Unvoiced*, *CPPS media*.

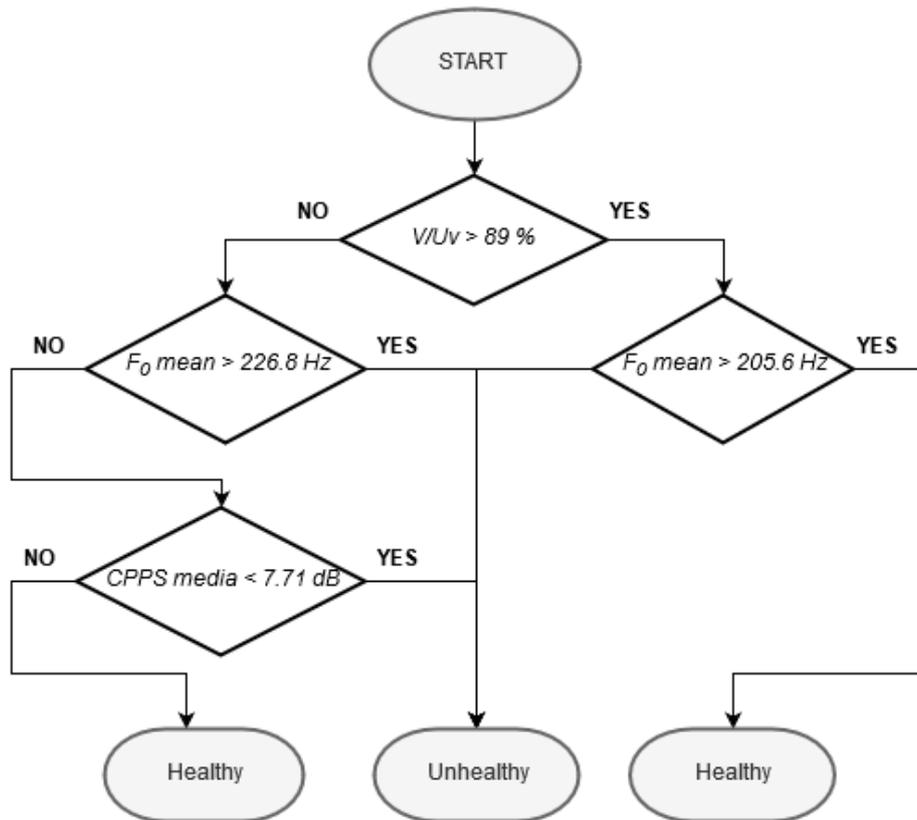


Figura 28 Grafo del free speech PIEZO

Il classificatore così costruito presenta una confusion matrix del tipo (Tabella 13):

		Classe predetta	
		0	1
Classe reale	0	18	1
	1	3	16

Tabella 13 Confusion matrix coarse tree free speech PIEZO

Con:

- 0 classe Parkinson
- 1 classe Sani

Ottenendo così ottime prestazioni in termini di accuratezza, sensibilità e specificità:

- *Accuracy* = 89.4%
- *Sensitivity* = 94.7%
- *Specificity* = 84.2%

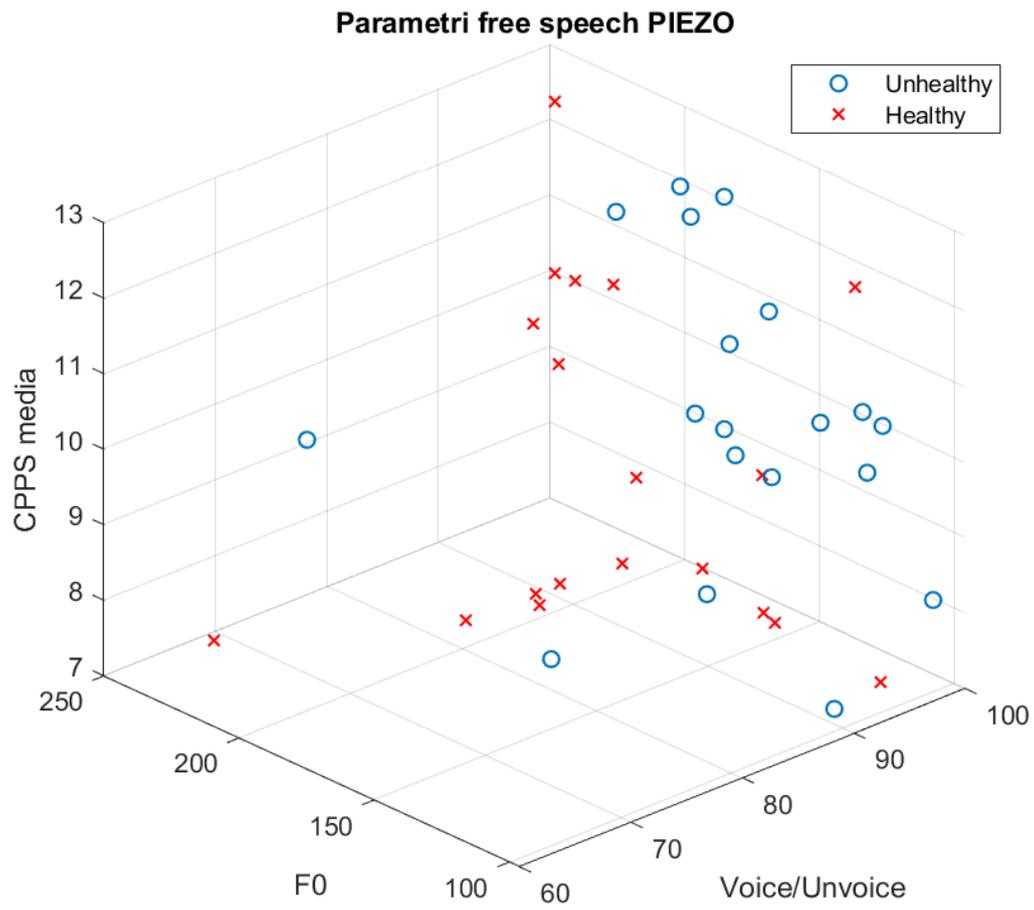


Figura 29 Divisione dello spazio delle features con l'albero decisionale per free speech PIEZO

## 5 Valutazione effetti corso logopedia

In questo lavoro di tesi è stato portato in parallelo anche un secondo studio, relativo alla valutazione dell'efficacia di un corso di logopedia sui parametri vocali di soggetti affetti dalla malattia di Parkinson.

In collaborazione con l'Associazione Amici Parkinsoniani Piemonte sono stati monitorati 6 soggetti (5 uomini e 1 donna). La registrazione del segnale vocale è avvenuta in due differenti momenti. La prima ad inizio di un corso di logopedia specifico per i soggetti, la seconda a due mesi dall'inizio dello stesso corso.

Le registrazioni sono state effettuate seguendo il protocollo ed utilizzando la strumentazione descritti rispettivamente nei capitoli 3.3 e **Errore. L'origine iferimento non è stata trovata..**

Così come per la creazione del classificatore, la discriminazione soggetti sani e soggetti affetti da malattia di Parkinson è basata sull'analisi dei parametri estratti dalla vocale 'a' sostenuta registrata con microfono in aria e l'eloquio libero registrato con entrambi i microfoni.

### 5.1 Voce sostenuta 'a' con microfono in aria

Per la scelta dei parametri da analizzare al fine di verificare eventuali variazioni nei suoi valori dopo il trattamento logopedico, anche in questo caso è stato fatto ricorso alla matrice di correlazione ricavata tramite indici di *Pearson* ' $\rho$ '. Nella Tabella 14 riportata in seguito, sono presentati i coefficienti di correlazione di *Pearson* ottenuti comparando ogni parametro ricavato dai file dell'eloquio libero con tutti i restanti parametri. I coefficienti sulla diagonale sono evidenziati perché rappresentano la correlazione del parametro con se stesso, per cui non restituisce nessuna informazione.

Valutazione effetti corso logopedia

MIPRO	<i>Jitt</i>	<i>Jita</i>	<i>RAP</i>	<i>PPQ</i>	<i>vF<sub>0</sub></i>	<i>Shim</i>	<i>ShdB</i>	<i>APQ</i>	<i>vAm</i>	<i>F<sub>0</sub></i>	<i>HNR</i>	<i>V/Uv</i>	<i>CPPS Media</i>	<i>CPPS Median</i>	<i>CPPS Moda</i>	<i>CPPS Std</i>	<i>CPPS Range</i>	<i>CPPS 5°prc</i>	<i>CPPS 95°prc</i>	<i>CPPS Skew</i>	<i>CPPS Kurt</i>	
<i>Jitt</i>	1	0,97	1,00	0,97	0,51	0,91	0,91	0,93	0,56		-0,77	-0,90	-0,56	-0,54	-0,50			-0,64				
<i>Jita</i>		1	0,98	0,94	0,52	0,84	0,85	0,86	0,57		-0,74	-0,90	-0,51	-0,48		0,52		-0,65				
<i>RAP</i>			1	0,96	0,51	0,91	0,92	0,92	0,55		-0,77	-0,88	-0,55	-0,53	-0,48			-0,63				
<i>PPQ</i>				1	0,57	0,87	0,87	0,91	0,54		-0,75	-0,91	-0,53	-0,51	-0,48			-0,60				
<i>vF<sub>0</sub></i>					1							-0,64										
<i>Shim</i>						1	1	0,98	0,54		-0,87	-0,71	-0,49	-0,48	-0,47			-0,55				
<i>ShdB</i>							1	0,97	0,56		-0,89	-0,70	-0,47	-0,45				-0,54				
<i>APQ</i>								1	0,59		-0,85	-0,79	-0,52	-0,51	-0,49			-0,56				
<i>vAm</i>									1	0,48	-0,53	-0,52				0,52	0,48				-0,51	
<i>F<sub>0</sub></i>										1												0,48
<i>HNR</i>											1	0,60				-0,47		0,49				
<i>V/Uv</i>												1	0,53	0,50	0,46			0,58	0,46			
<i>CPPS Media</i>													1	1,00	0,99			0,92	0,95	-0,48	0,55	
<i>CPPS Median</i>														1	0,99			0,89	0,96	-0,52	0,54	
<i>CPPS Moda</i>															1			0,86	0,97	-0,52	0,56	
<i>CPPS Std</i>																1	0,83	-0,49				
<i>CPPS Range</i>																	1		0,48			
<i>CPPS 5°prc</i>																		1	0,77		0,54	
<i>CPPS 95°prc</i>																			1	-0,48	0,51	
<i>CPPS Skew</i>																				1	-0,56	
<i>CPPS Kurt</i>																						1

Tabella 14 Matrice di indici Pearson per vocale 'a' con MIPRO

I coefficienti evidenziati in azzurro, invece, rappresentano quelli maggiormente correlati tra di loro, in particolare i seguenti:

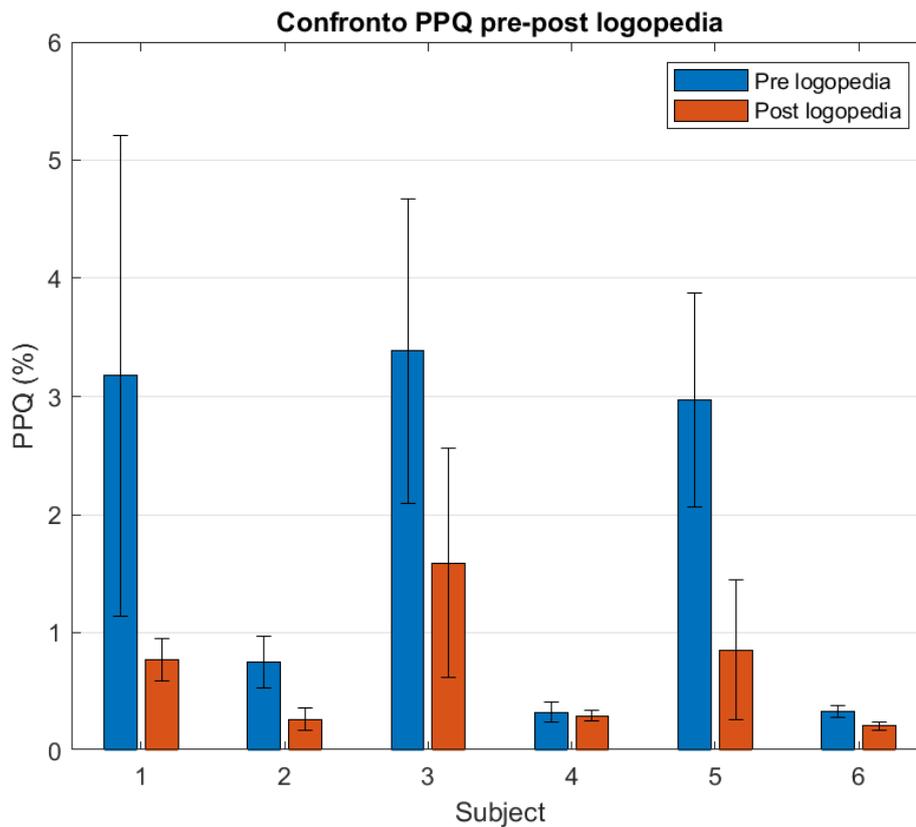
- *Jitt, Jita, RAP, PPQ, Shim, ShdB, APQ;*
- *Shim, ShdB, APQ, HNR;*
- *CPPS media, CPPS mediana, CPPS moda, CPPS 5°percentile, CPPS 95°percentile;*
- *CPPS standard deviation, CPPS range.*

Ci si è focalizzati quindi sul subset di parametri: *PPQ, vF<sub>0</sub>, vAm, CPPS 95°percentile, CPPS range, CPPS skewness, CPPS kurtosis.*

La matrice è stata ottenuta per il microfono in aria MIPRO. Inoltre si può notare che alcuni coefficienti non sono stati riportati. Tali valori omessi non risultavano statisticamente significativi, avendo un *pValue* > 0.05, non verificando quindi l'ipotesi nulla.

Dopo aver scelto i parametri analizzando la matrice di correlazione, sono stati confrontati i risultati ottenuti prima e dopo il corso di logopedia attraverso differenti istogrammi. Ad ognuno di essi è associato un parametro e per ogni soggetto sono presenti due barre, una in blu corrispondente alla media dei valori delle tre 'a' sostenute, ed in rosso il corrispettivo valore della media dei parametri dopo 2 mesi. Ogni barra presenta anche la sua deviazione tipo. Una bassa deviazione tipo indica che i singoli valori tendono ad essere simili al valor medio, mentre una alta *standard deviation* indica che i tre valori hanno un range molto ampio.

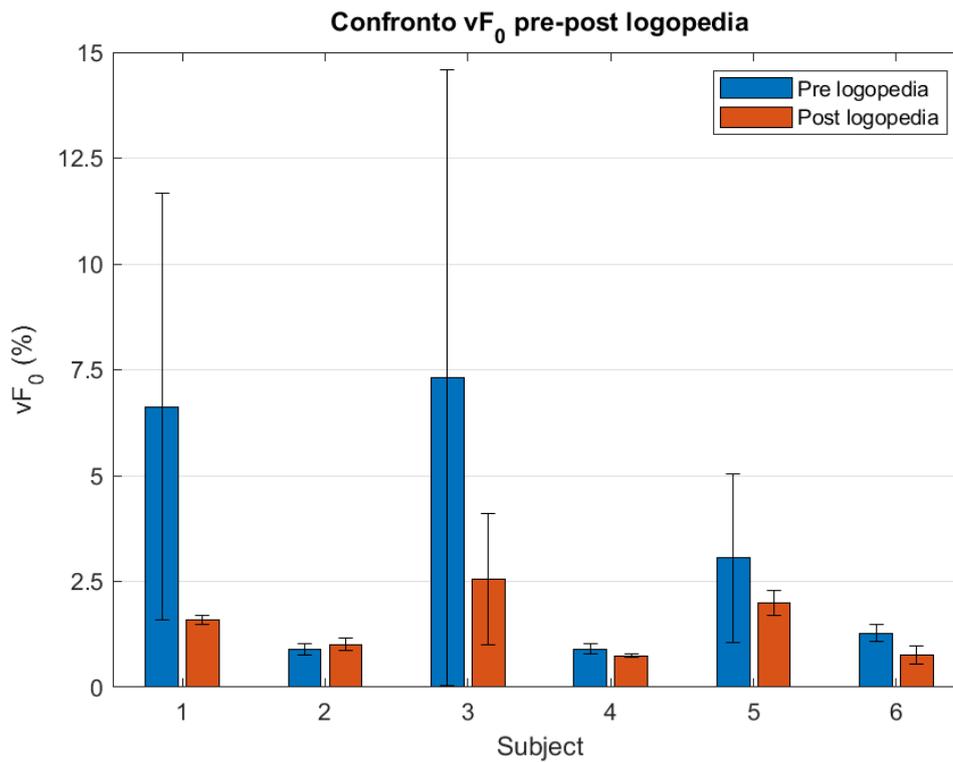
Come si può notare dalla Figura 30, in cui è rappresentato il confronto del PPQ (%), tutti i soggetti hanno riscontrato un miglioramento in termini di stabilità in frequenza, dimostrando, quindi, che il corso ha dato la capacità ai soggetti di ottenere un controllo della stabilità in frequenza della propria voce.



**Figura 30 Confronto PPQ pre-post corso di logopedia**

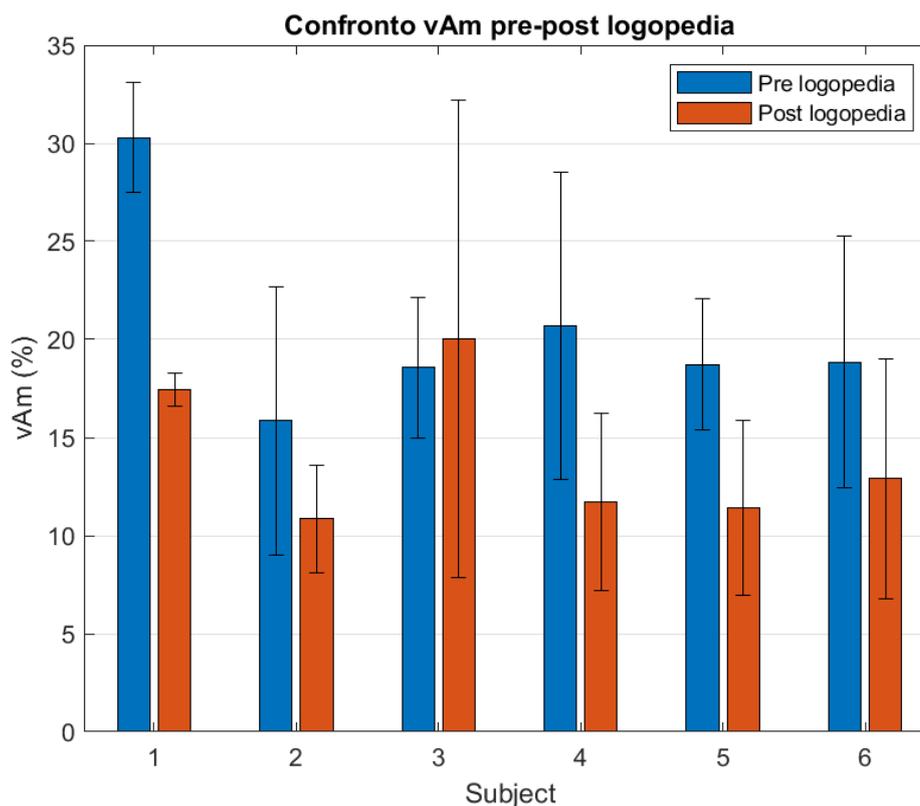
Specialmente per il primo, terzo e quinto soggetto si è riscontrato un netto miglioramento della stabilità in frequenza, passando da un valore di ( $\sim 3\%$ ) ad uno di ( $\sim 0.5\%$ ). Il secondo soggetto ha riscontrato un lieve miglioramento passando da un valore di ( $\sim 0.8\%$ ) ad uno di ( $\sim 0.2\%$ ), mentre il terzo e il sesto presentano un valore pressoché invariato di tale parametro. Si può affermare, dunque, che tale parametro risulta rilevante in termini di discriminazione tra pre e post corso di logopedia.

In Figura 31 è presente il confronto di  $\nu F_0(\%)$ . Anche in questo caso, come per il PPQ, si può notare un miglioramento nella stabilità della frequenza fondamentale dopo che i soggetti hanno effettuato il corso di logopedia. Anche in questo caso i soggetti che hanno beneficiato maggiormente sono stati il primo ed il terzo, passando da un valore di ( $\sim 7.5\%$ ) ad uno di ( $\sim 2\%$ ). Il quinto soggetto ha riscontrato un lieve miglioramento passando da un valore di ( $\sim 3\%$ ) ad uno di ( $\sim 2\%$ ), mentre il secondo, il quarto ed il sesto presentano un valore pressoché simile nelle due registrazioni.



**Figura 31 Confronto  $vF_0$  pre-post corso di logopedia**

Nella Figura 32 è riportato il confronto pre-post corso del  $vAm(\%)$ . Analizzando il bar diagram, è possibile riscontrare che anche la stabilità in ampiezza risulta un parametro discriminante post corso. Quasi tutti i soggetti presentano un miglioramento stabile passando da un valore di ( $\sim 18\%$ ) ad uno di ( $\sim 12\%$ ). Specialmente il primo soggetto, che così come per la stabilità in frequenza risulta chi ha riscontrato maggiori miglioramenti nel tempo, passando da un valore di ( $\sim 30\%$ ) ad uno di ( $\sim 17\%$ ). L'unica eccezione è costituita dal terzo soggetto che in caso di miglioramenti nella stabilità in frequenza, non ne presenta nella stabilità in ampiezza.



**Figura 32 Confronto vAm pre-post corso di logopedia**

In seguito (Figura 33) è riportata la distribuzione del CPPS complessiva delle tre ‘a’ sostenute. Per ogni soggetto sono riportati i grafici pre corso di logopedia e due mesi dopo lo stesso corso. I due grafici delle distribuzioni sono sovrapposti per osservare eventuale cambiamenti. Si nota uno spostamento verso sinistra della media del CPPS per il primo, terzo, quarto e sesto paziente, passando da un valore di ( $\sim 14dB\%$ ) ad uno di ( $\sim 12dB$ ) mentre uno spostamento verso destra della media del CPPS per il secondo e quinto soggetto passando da un valore di ( $\sim 12dB$ ) ad uno di ( $\sim 15dB$ ). Un netto miglioramento pre e post corso, visibile anche solo osservando i grafici, è nel restringimento delle distribuzioni, cioè può essere verificato attraverso il bar diagram della deviazione tipo del CPPS (Figura 34). Per il primo, quarto, quinto e sesto paziente tale valore risulta calare di ( $\sim 0.5dB$ ) dopo il corso di logopedia, rispetto al secondo e terzo soggetto per cui il corso risulta influente in termini di variazione di dispersione del CPPS.

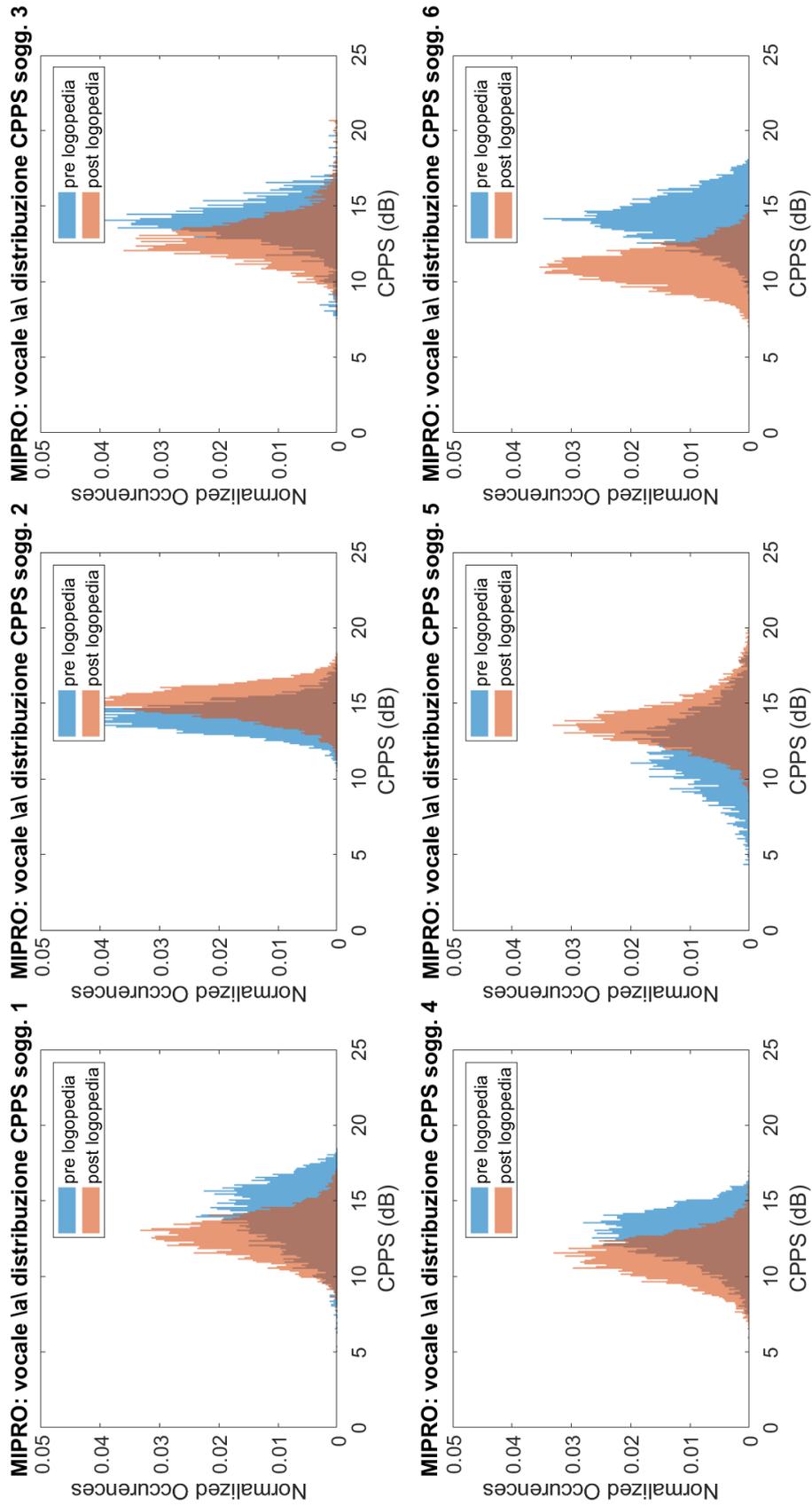
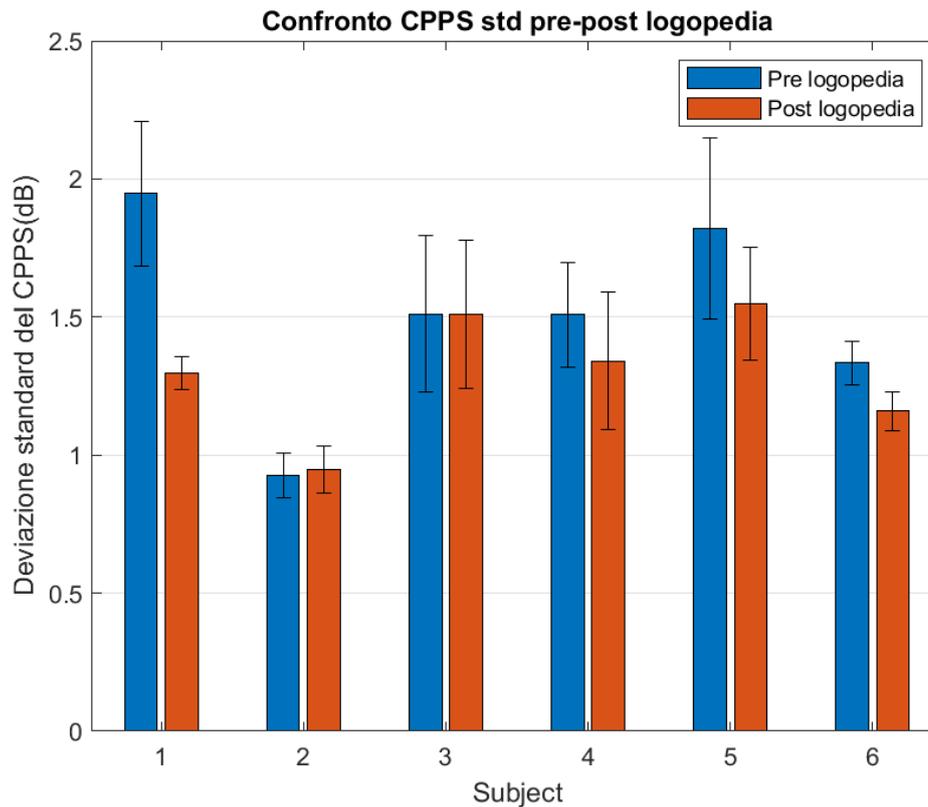


Figura 33 Distribuzioni CPPS vocale 'a' MIPRO



**Figura 34 Confronto CPPS deviazione standard per MIPRO**

Infatti le distribuzioni di questi ultimi due soggetti risultano pressoché simili, ad eccezione di uno shift di qualche dB della media rispettivamente verso destra e verso sinistra.

## 5.2 Eloquio libero con microfono in aria

L'eloquio libero è stato analizzato attraverso la sola analisi dei parametri del CPPS. Nella Figura 35 sono presenti i grafici delle rispettive distribuzioni del CPPS dei soggetti analizzati per quanto riguarda l'eloquio registrato tramite microfono in aria prima del corso di logopedia e a due mesi dal suo inizio. A differenza della vocale 'a' sostenuta, analizzando la sovrapposizione delle distribuzioni prima e dopo il corso di logopedia per l'eloquio libero, non si nota quasi nessuna differenza, e ciò è dovuto al *CPPS vocal print*. Le distribuzioni tra il prima e il dopo presentano un leggero spostamento verso dB più alti dopo il corso, che rappresenta un miglioramento nella salute dell'apparato vocale.

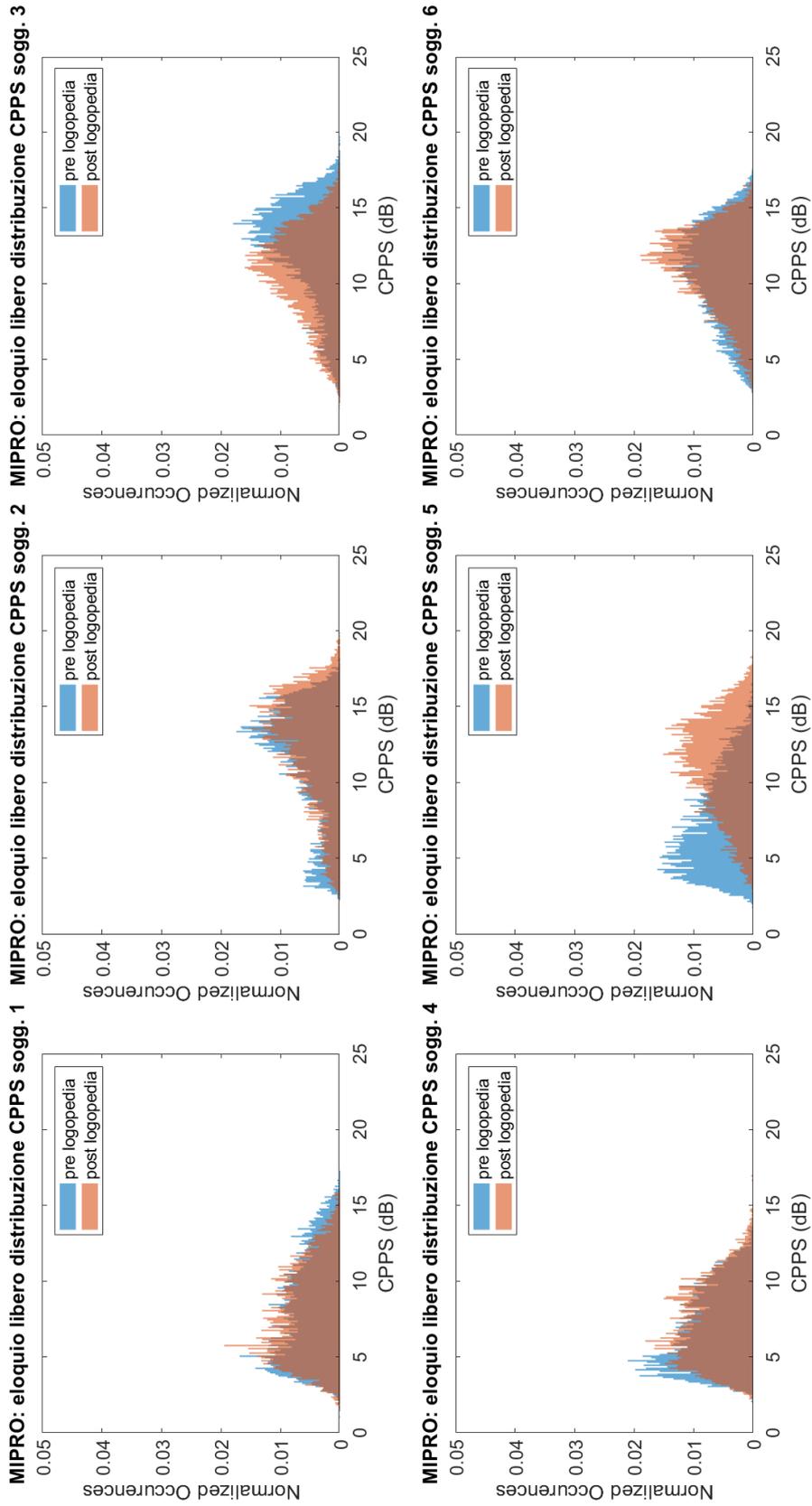


Figura 35 Distribuzioni CPPS free speech MIPRO

Specialmente il quinto paziente presenta una notevole diminuzione di occorrenze a bassi dB, risultando quindi il soggetto che ha maggiormente ? del corso, presentando parametri nettamente migliorativi sia nella distribuzione del CPPS dell'eloquio libero, sia per la variazione in frequenza e ampiezza nella vocale sostenuta. L'unico soggetto ad avere un più basso spostamento verso dB risulta essere il terzo soggetto, anche se inizialmente sembrava ottenere dei miglioramenti in stabilità vocale nella vocale 'a' sostenuta.

Per la scelta dei parametri da analizzare per verificare eventuali variazioni nei suoi valori dopo il trattamento logopedico, anche in questo caso si è ricorso alla matrice di correlazione ricavata tramite indici di *Pearson 'ρ'*. Nella Tabella 15 riportata in seguito, sono presentati i coefficienti di correlazione di *Pearson* ottenuti comparando ogni parametro ricavato dai file dell'eloquio libero con tutti i restanti parametri.

MIPRO	$F_0$	HNR	V/Uv	CPPS Media	CPPS Median	CPPS Moda	CPPS Std	CPPS Range	CPPS 5°pr	CPPS 95°pr	CPPS Skew	CPPS Kurt
$F_0$	1											
HNR		1	0,63									
V/Uv			1									
CPPS Media				1	1,00	0,90		0,71	0,80	0,93	-0,97	0,51
CPPS Median					1	0,91		0,71	0,77	0,92	-0,98	0,50
CPPS Moda						1		0,55	0,76	0,76	-0,92	0,65
CPPS Std							1	0,62		0,66		
CPPS Range								1		0,81	-0,64	
CPPS 5°prc									1	0,58	-0,77	0,70
CPPS 95°prc										1	-0,85	
CPPS Skew											1	-0,53
CPPS Kurt												1

Tabella 15 Indici Pearson pre-post logopedia free speech MIPRO

I coefficienti sulla diagonale sono evidenziati perché rappresentano la correlazione del parametro con se stesso, per cui non restituisce nessuna informazione. I coefficienti evidenziati in azzurro invece rappresentano i coefficienti maggiormente correlati tra di loro, in particolare i seguenti:

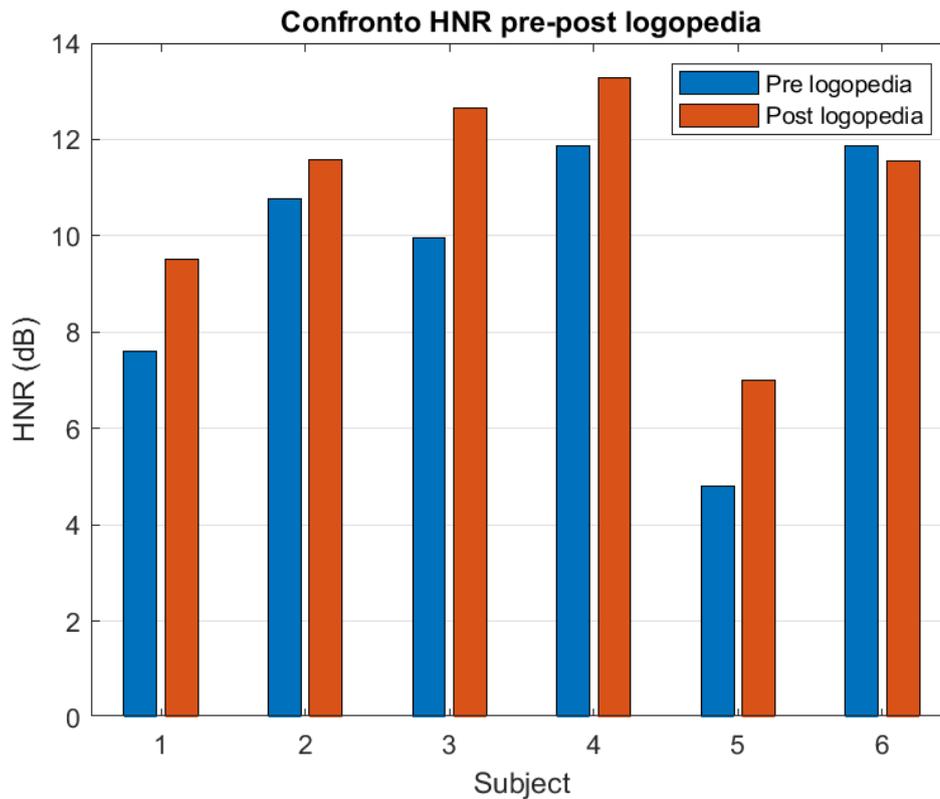
- *CPPS media, CPPS mediana, CPPS moda, CPPS 5°percentile, CPPS 95°percentile, CPPS skewness;*

Ci si è focalizzati quindi sul subset di parametri: : *HNR, CPPS media, CPPS standard deviation, CPPS kurtosis.*

La matrice è stata ottenuta per il microfono in aria MIPRO. Inoltre si può notare che alcuni coefficienti non sono stati riportati. Tali valori omessi non risultavano statisticamente significativi, avendo un *pValue* > 0.05, non verificando quindi l'ipotesi nulla.

Dopo aver scelto i parametri analizzando la matrice di correlazione, sono stati confrontati i risultati ottenuti prima e dopo il corso di logopedia, attraverso differenti istogrammi. Ad ognuno di esso è associato un parametro e per ogni soggetto sono presenti due barre, una in blu corrispondente al valore ottenuto durante la prima registrazione, ed una in rosso per il corrispettivo valore a 2 mesi dal corso di logopedia.

In Figura 36 è presentato il confronto del' HNR prima e dopo il corso di logopedia. Analizzando il diagramma si nota che tale parametro non è discriminante per valutare il corso, in quanto il suo valore risulta maggiore nella seconda registrazione ad eccezione del sesto soggetto in cui tale valore risulta stabile. Uno dei motivi di tale risultato potrebbe essere causato dal fatto che le registrazioni sono state acquisite tramite microfono in aria in una stanza non insonorizzata, per cui tra le due diverse registrazioni il rumore di fondo potrebbe risultare differente.



**Figura 36 Confronto HNR free speech MIPRO**

In Figura 37 è rappresentata la media del CPPS (dB). Si può notare, nonostante le distribuzioni nelle due registrazioni restino quasi invariate, un lieve spostamento a dB più elevati da parte delle distribuzioni, ad eccezione del terzo soggetto per cui la distribuzione sembra spostarsi verso dB più bassi. In Figura 38 è presentato il parametro *CPPS std*, e si nota come tale parametro resti pressoché invariato, a dimostrazione che il corso di logopedia ha portato sì miglioramenti in stabilità di frequenza e ampiezza dell'apparato fonatorio, ma non relativi vantaggi per quanto riguarda la fonazione in un parlato quotidiano.

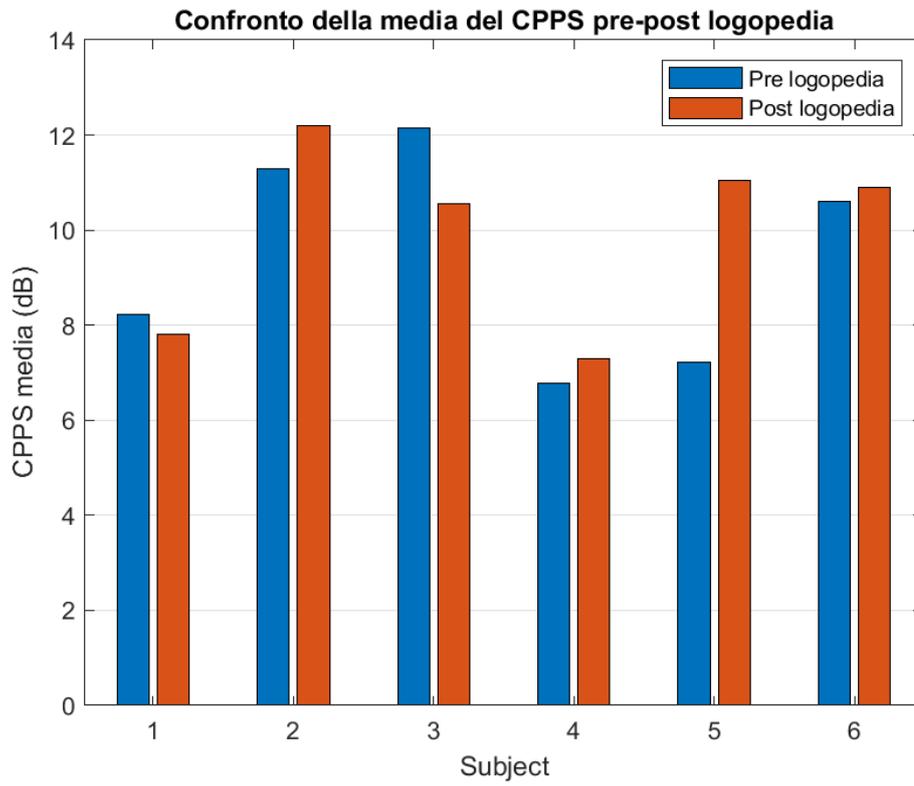


Figura 37 Confronto CPPS mean free speech MIPRO

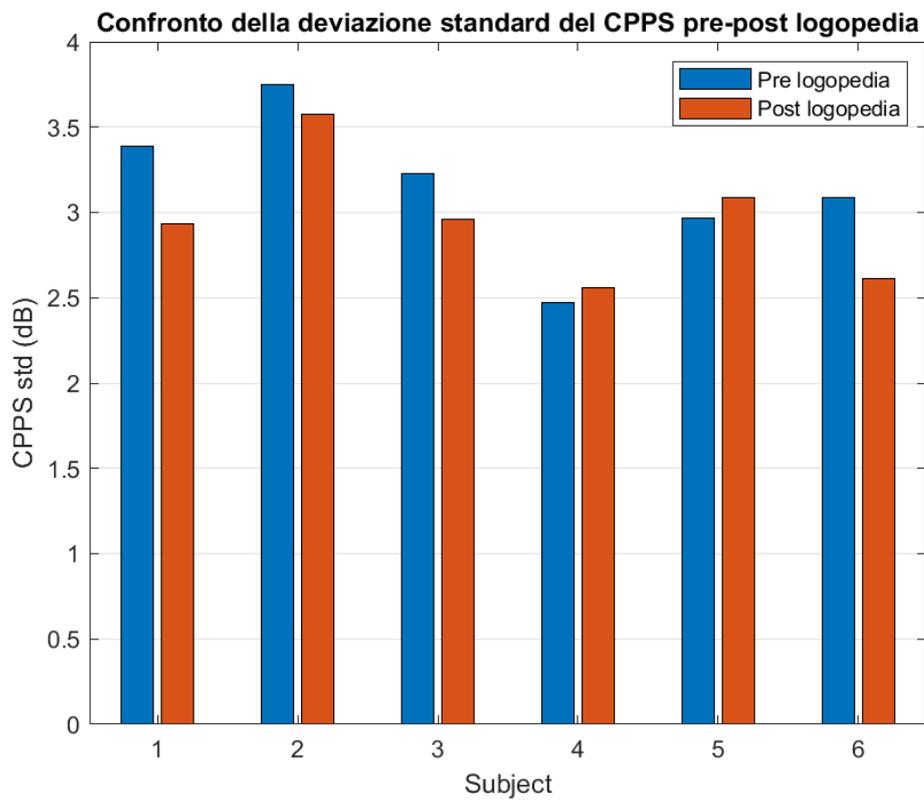


Figura 38 Confronto CPPS std free speech MIPRO

In seguito è riportata la variazione relativa ai singoli parametri dei sei soggetti monitorati tramite microfono in aria a distanza da due mesi da una registrazione e l'altra (Tabella 16).

<b><i>Eloquio libero MIPRO</i></b>		<b><i>HNR</i></b>	<b><i>CPPS media</i></b>	<b><i>CPPS std</i></b>	<b><i>CPPS kurtosis</i></b>
<b><i>Soggetto 1</i></b>	<i>Pre</i>	7.60 dB	8.23 dB	3.39 dB	-0.89
	<i>Post</i>	9.50 dB	7.82 dB	2.93 dB	-0.71
<b><i>Soggetto 2</i></b>	<i>Pre</i>	10.75 dB	11.29 dB	3.75 dB	-0.54
	<i>Post</i>	11.58 dB	12.19 dB	3.57 dB	-0.35
<b><i>Soggetto 3</i></b>	<i>Pre</i>	9.97 dB	12.15 dB	3.23 dB	-0.11
	<i>Post</i>	12.64 dB	10.56 dB	2.96 dB	-0.39
<b><i>Soggetto 4</i></b>	<i>Pre</i>	11.86 dB	6.77 dB	2.47 dB	-0.87
	<i>Post</i>	13.28 dB	7.29 dB	2.56 dB	-0.52
<b><i>Soggetto 5</i></b>	<i>Pre</i>	4.81 dB	7.23 dB	2.97 dB	-0.57
	<i>Post</i>	6.99 dB	11.05 dB	3.06 dB	-0.57
<b><i>Soggetto 6</i></b>	<i>Pre</i>	11.88 dB	10.60 dB	3.08 dB	-0.68
	<i>Post</i>	11.56 dB	10.89 dB	2.61 dB	-0.22

**Tabella 16** Differenze parametri free speech MIPRO

### **5.3 Eloquio libero con microfono a contatto**

Così come per l'eloquio libero acquisito con microfono in aria, anche quello registrato tramite microfono a contatto è stato analizzato attraverso la sola analisi dei parametri del CPPS. Nella Figura 39 sono presenti i grafici delle rispettive distribuzioni del CPPS dei soggetti analizzati per quanto riguarda l'eloquio registrato tramite microfono a contatto prima del corso di logopedia e a due mesi dal suo inizio. Analizzando la sovrapposizione delle distribuzioni prima e dopo il corso di logopedia per l'eloquio libero, non si notano rilevanti differenze rispetto a quelle ottenute tramite microfono in aria.

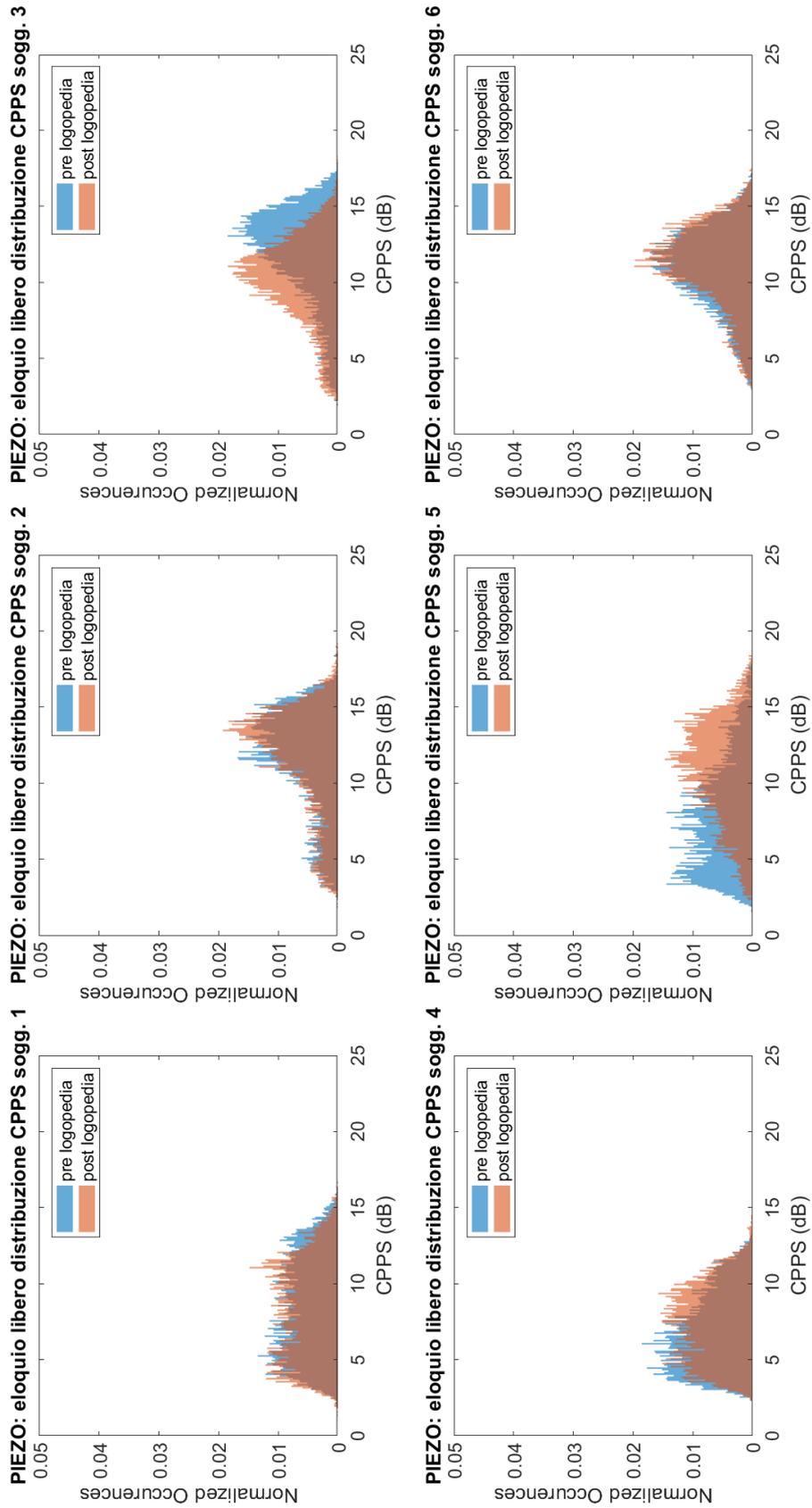


Figura 39 Distribuzioni CPPS free speech PIEZO

Anche per il microfono a contatto, le distribuzioni tra il prima e il dopo presentano un leggero spostamento verso dB più alti dopo il corso, che è rappresentativo di un miglioramento nella salute dell'apparato vocale. Il quinto paziente in particolare, presenta una notevole diminuzione di occorrenze a bassi dB, risultando quindi il soggetto che ha maggiormente ? del corso, presentando parametri nettamente migliorativi sia nella distribuzione del CPPS dell'eloquio libero, sia per la variazione in frequenza e ampiezza nella vocale sostenuta. L'unico soggetto ad avere uno spostamento verso dB più bassi risulta essere il terzo, anche se inizialmente sembrava ottenere dei miglioramenti in stabilità vocale nella vocale 'a' sostenuta.

Nella scelta dei parametri da analizzare per verificare eventuali variazioni nei suoi valori dopo il trattamento logopedico, anche in questo caso è stato fatto ricorso alla matrice di correlazione ricavata tramite indici di *Pearson 'ρ'*. Nella Tabella 17 riportata in seguito, sono presentati i coefficienti di correlazione di *Pearson* ottenuti comparando ogni parametro ricavato dai file dell'eloqui libero con tutti i restanti parametri.

MIPRO	$F_0$	HNR	V/Uv	CPPS Media	CPPS Median	CPPS Moda	CPPS Std	CPPS Range	CPPS 5°pr	CPPS 95°pr	CPPS Skew	CPPS Kurt
$F_0$	1											
HNR		1	0,73				-0,52					
V/Uv			1				-0,66					
CPPS Media				1	1,00	0,86		0,72	0,79	0,88	-0,93	0,64
CPPS Median					1	0,89		0,70	0,76	0,86	-0,95	0,62
CPPS Moda						1		0,58	0,63	0,72	-0,86	0,52
CPPS Std							1	0,68		0,67		
CPPS Range								1		0,90	-0,53	
CPPS 5°prc									1	0,50	-0,72	0,81
CPPS 95°prc										1	-0,71	
CPPS Skew											1	-0,61
CPPS Kurt												1

Tabella 17 Indici Pearson pre-post logopedia free speech PIEZO

I coefficienti sulla diagonale sono evidenziati perché rappresentano la correlazione del parametro con se stesso, per cui non restituisce nessuna informazione. I coefficienti evidenziati in azzurro invece rappresentano i coefficienti maggiormente correlati tra di loro, in particolare i seguenti:

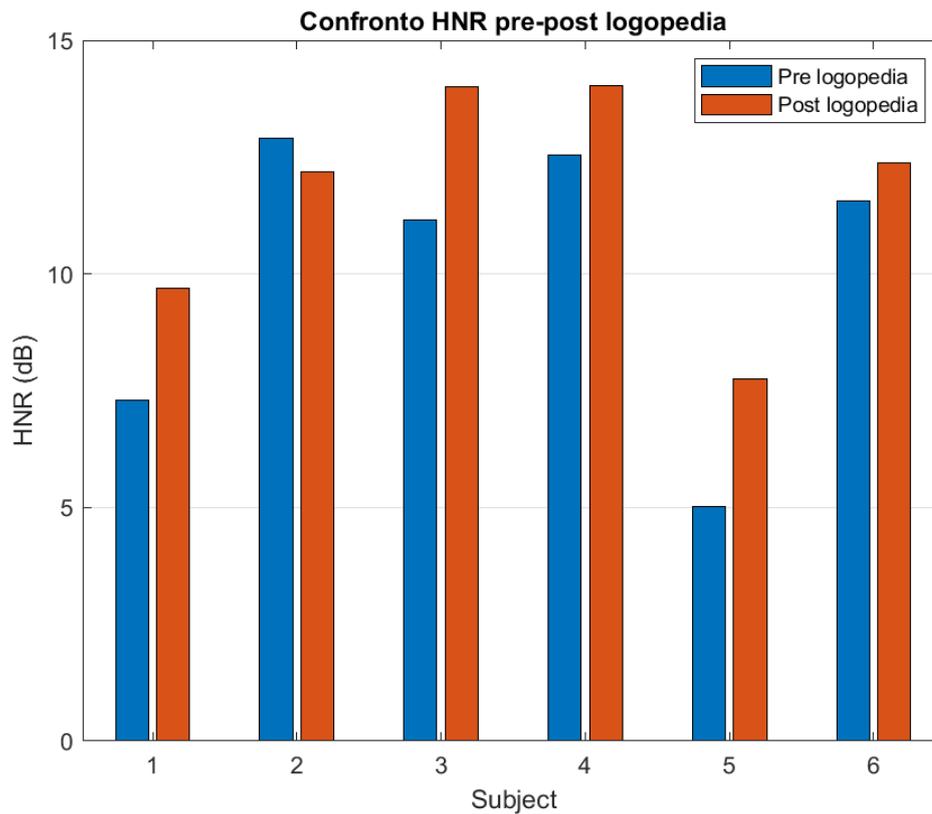
- *CPPS media, CPPS mediana, CPPS moda, CPPS 95°percentile, CPPS skewness;*
- *CPPS 5°percentile, CPPS kurtosis.*

Ci si è focalizzati quindi sul subset di parametri: : *HNR, CPPS media, CPPS standard deviation, CPPS kurtosis.*

La matrice è stata ottenuta per il microfono a contatto. Inoltre si può notare che alcuni coefficienti non sono stati riportati. Tali valori omessi non risultavano statisticamente significativi, avendo un *pValue* > 0.05, non verificando quindi l'ipotesi nulla.

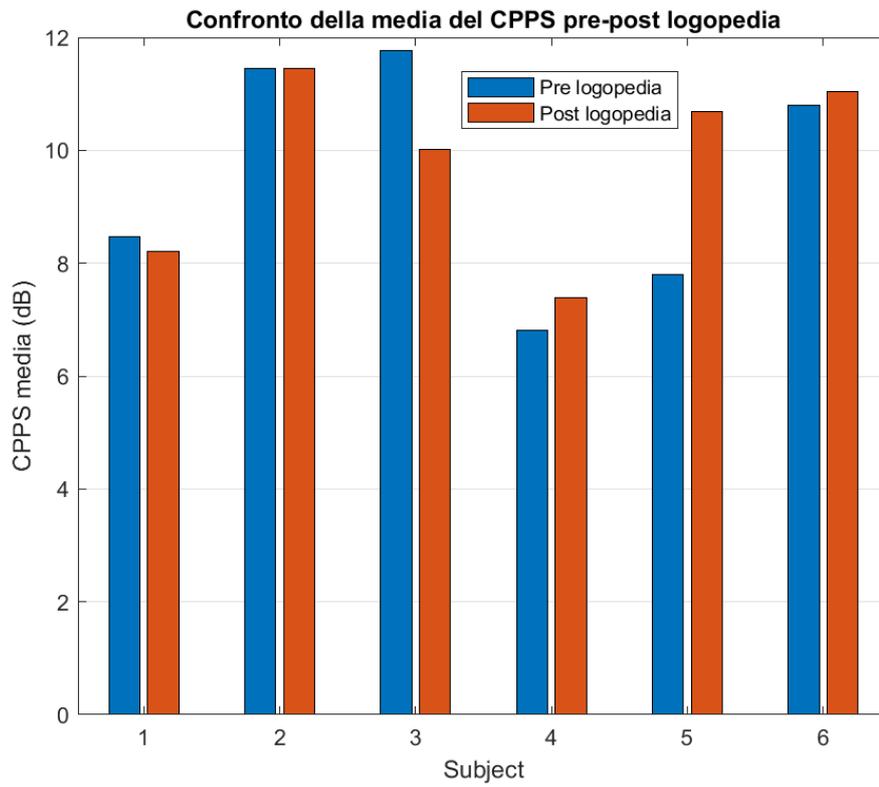
Dopo aver scelto i parametri analizzando la matrice di correlazione, sono stati confrontati i risultati ottenuti prima e dopo il corso di logopedia attraverso differenti istogrammi. Ad ognuno di esso è associato un parametro e per ogni soggetto sono presenti due barre, una in blu corrispondente al valore ottenuto alla prima registrazione, ed una in rosso per il corrispettivo valore a 2 mesi dal corso di logopedia.

In Figura 40 è presentato il confronto del' HNR prima e dopo il corso di logopedia. Analizzando il bar diagram si nota che tale parametro risulta avere dei valori molto simili ai risultati ottenuti tramite registrazioni con microfono in aria, ovvero nessun paziente presenta miglioramenti per quanto riguarda l'eloquio, a differenza della vocale sostenuta. Tutti i soggetti presentano un incremento in dB del rapporto segnale rumore, ad eccezione del secondo soggetto che presenta una lieve diminuzione.

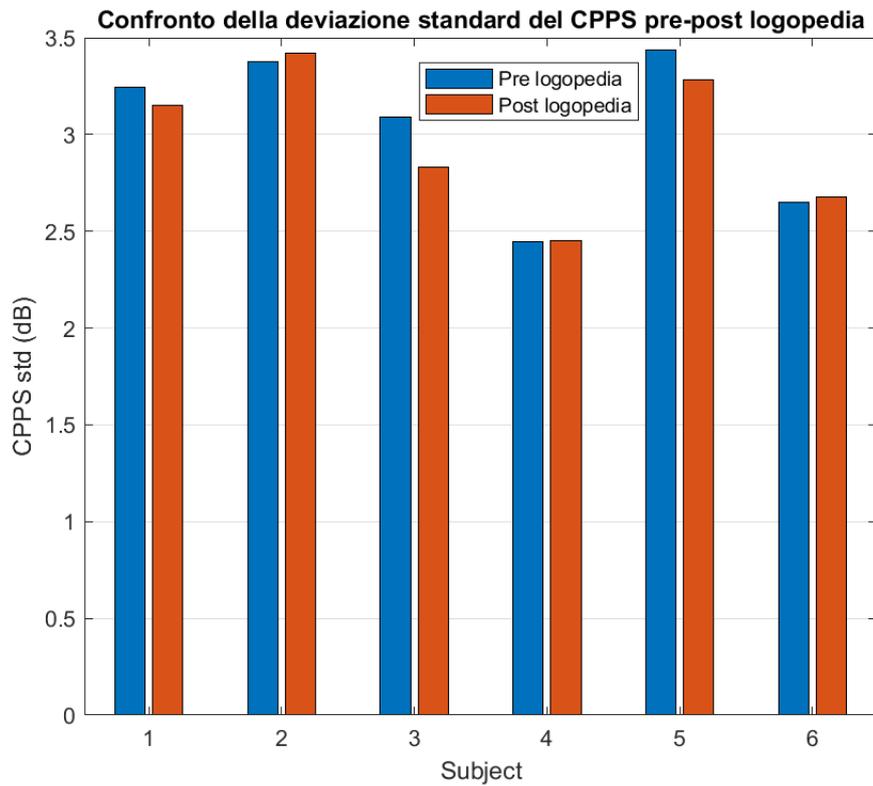


**Figura 40 Confronto HNR free speech PIEZO**

In Figura 41 è rappresentata la media del CPPS (dB). Si può notare, nonostante le distribuzioni nelle due registrazioni restino quasi invariate, un lieve spostamento a dB più elevati da parte delle distribuzioni, ad eccezione del terzo soggetto per cui la distribuzione sembra spostarsi verso dB più bassi. In Figura 42 è presentato il parametro *CPPS std*, e si nota come tale parametro resti pressoché invariato, a dimostrazione che il corso di logopedia ha portato sì miglioramenti in stabilità di frequenza e ampiezza dell'apparato fonatorio, ma non relativi vantaggi per quanto riguarda la fonazione in un parlato quotidiano.



**Figura 41 Confronto CPPS mean free speech PIEZO**



**Figura 42 Confronto CPPS std free speech PIEZO**

In seguito è riportata la variazione relativa ai singoli parametri dei sei soggetti monitorati tramite microfono a contatto prima e dopo i due mesi di corso di logopedia (Tabella 18).

<b>Eloquio libero PIEZO</b>		<b>HNR</b>	<b>CPPS media</b>	<b>CPPS std</b>	<b>CPPS kurtosis</b>
<b>Soggetto 1</b>	<i>Pre</i>	7.60 dB	8.28 dB	3.25 dB	-1.06
	<i>Post</i>	9.69 dB	8.20 dB	3.15 dB	-1.00
<b>Soggetto 2</b>	<i>Pre</i>	12.91 dB	12.11 dB	3.37 dB	-0.34
	<i>Post</i>	12.18 dB	12.35 dB	3.42 dB	-0.24
<b>Soggetto 3</b>	<i>Pre</i>	11.15 dB	12.41 dB	3.09 dB	0.22
	<i>Post</i>	14.00 dB	10.36 dB	2.83 dB	-0.02
<b>Soggetto 4</b>	<i>Pre</i>	12.55 dB	6.50 dB	2.44 dB	-0.77
	<i>Post</i>	14.02 dB	7.44 dB	2.45 dB	-0.85
<b>Soggetto 5</b>	<i>Pre</i>	5.01 dB	7.44 dB	3.44 dB	-0.45
	<i>Post</i>	7.74 dB	11.07 dB	3.28 dB	-0.51
<b>Soggetto 6</b>	<i>Pre</i>	11.57 dB	11.06 dB	2.65 dB	-0.16
	<i>Post</i>	12.38 dB	11.40 dB	2.68 dB	0.02

**Tabella 18 Differenze parametri free speech PIEZO**

## 6 Conclusioni

Lo scopo principale di questo studio è stato quello di creare un supporto per la diagnosi del morbo di Parkinson tramite il monitoraggio del segnale vocale e l'analisi dei suoi parametri descrittivi. Un gruppo di 22 pazienti affetti dal morbo di Parkinson è stato monitorato durante la consueta visita annuale presso l'ambulatorio Parkinson della Città della Salute di Torino.

Ai pazienti è stato chiesto di seguire un protocollo ben preciso: ripetere per tre volte la vocale 'a', ripetere i fonemi pa/ta/ka, leggere un brano e parlare di un tema libero. Durante tale operazione sono stati acquisiti il segnale vocale e la vibrazione delle corde vocali tramite due microfoni, uno in aria ed uno a contatto.

Nella fase di elaborazione sono stati analizzati i parametri di CPPS e della sua stima statistica, e di HNR per l'eloquio libero e la lettura, mentre per la vocale sostenuta, sono stati calcolati anche parametri di variazione in ampiezza e frequenza.

Infine, sono stati costruiti specifici classificatori in grado di discriminare soggetti affetti dalla malattia di Parkinson da soggetti sani. Il dataset è stato composto da 44 pazienti (22 soggetti monitorati e 22 sani già presenti nel database), utilizzando le features precedentemente estratte. I classificatori sono stati creati analizzando la vocale 'a' registrata tramite microfono in aria e l'eloquio libero con entrambi i microfoni. La scelta si è focalizzata sul modello della regressione logistica e del coarse tree. Il modello a regressione logistica per la 'a' sostenuta che ha riportato la migliore capacità di discriminazione tra soggetto sano e parkinsoniano è stato quello a quattro parametri: Jitter, PPQ, APQ, CPPS 5° percentile, con accuratezza dell'88% considerando un taglio a  $P(0.56)$ . Per il modello del coarse tree, considerando i parametri PPQ, shimmer e CPPS skewness, si è ottenuta un'accuratezza del 92%. Le accuratezze ottenute con i classificatori analizzando l'eloquio libero risultano avere nel caso del modello a regressione logistica un'accuratezza meno elevata: 79% con il modello di regressione logistica (HNR, CPPS media, CPPS kurtosis) utilizzando il microfono in aria, mentre 76% (*voiced/unvoiced ratio*, CPPS range) impiegando il microfono a contatto. Le prestazioni del modello coarse tree risultano, invece, stabili: 95% di accuratezza (CPPS moda, CPPS range, CPPS kurtosis) tramite l'uso del microfono in aria, e 92% di accuratezza utilizzando il microfono a contatto.

Questi risultati sono promettenti per una futura discriminazione supportata da un numero più ampio di soggetti monitorati. Eventualmente si potrebbe portare avanti il solo studio dei parametri analizzati con registrazioni tramite microfono a contatto che risultano essere affetti da un rumore di fondo molto più basso rispetto alle registrazioni effettuate con microfono in aria.

Per questo lavoro di tesi è stato effettuato inoltre un secondo studio, coinvolgendo 6 soggetti affetti dalla malattia del Parkinson, monitorati in due differenti occasioni. La prima ad inizio di un corso di logopedia specifico, e la seconda a due mesi di distanza, a corso ormai terminato. Anche in questo caso le registrazioni hanno seguito il protocollo descritto in precedenza, utilizzando la stessa strumentazione. Lo scopo è stato verificare se grazie ad esercizi specifici per la fonazione è possibile contrastare l'avanzamento della patologia dal punto di vista vocale.

Dall'analisi dei risultati relativi alla vocale 'a' sostenuta, si è riscontrato che, a fine corso, variazioni in ampiezza e frequenza della fondamentale sono migliorati, specialmente per il terzo soggetto e in modo moderato per il primo e per il sesto, mentre analizzando la sovrapposizione delle distribuzioni del CPPS del parlato libero, non è comparsa alcuna rilevante variazione.

L'analisi è stata effettuata a corso di logopedia ultimato riscontrando, dunque, nel breve termine un lieve miglioramento della stabilità della fonazione, mantenendo invece stabile il proprio eloquio nonostante l'avanzamento della malattia. Interessante sarebbe poter effettuare futuri monitoraggi così da verificare eventuali miglioramenti a lungo termine, avendo tale corso insegnato tecniche che ogni soggetto potrebbe e dovrebbe sfruttare durante la propria vita quotidiana.

## Bibliografia

- [1] S. Przedborski, Etiology and Pathogenesis of Parkinson's Disease, Jankovic J.
- [2] H. Bergman e D. G., Pathophysiology of Parkinson's Disease: From Clinical Neurology to Basic Neuroscience and Back. Movement Disorders, 2002.
- [3] [Online]. Available: <http://www.agescimarche.it>.
- [4] S. Fahn e R. Elton, Recent Developments in Parkinson's Disease, Fahn S, Goldstein M, Calne D, 1987.
- [5] «Parkinson.it,» [Online]. Available: <https://www.parkinson.it/problemi-di-parola.html>.
- [6] «Gondola Parkinson,» [Online]. Available: <https://www.gondola-parkinson.com/it/morbo-di-parkinson/>.
- [7] W. De Colle, Voce Computer, analisi acustica digitale del segnale verbale (Il sistema CSL-MDVP), Omega Edizioni, 2001.
- [8] Y. Maryn, P. Corthals, P. Van Cauwenberge, N. Roy e M. De Bodt, «Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels,» *Journal of voice*, n. 24, pp. 540-555, 2010.
- [9] F. Jimenez, J. Gamboa e A. Nieto, «Acoustic voice analysis in untreated patient with Parkinson's disease,» *Parkinson Relat Disord.*, n. 3, pp. 111-116, 1997.
- [10] A. Tsanas, M. Little, P. McSharry e E. Hunter, «Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease,» *IEEE Trans. Biomed. Eng.*, vol. 59, n. 5, pp. 1264-1271, 2010.
- [11] S. Skodda, H. Rinsche e U. Schlegel, «Progression of dysprosody in Parkinson's disease over time,» *Mov. Disord.*, vol. 24, pp. 716-722, 2009.
- [12] A. Castellana, F. Casassa e G. E. Puglisi, «Nuovi parametri acustici utili nella diagnostica e nella prevenzione di patologie vocali,» in *42° Convegno Nazionale*, Firenze, 2015.

- [13] Y. D. Haman-Ackah, M. D. Deirde e G. S. Goding, «The relationship between cepstral peak prominence and selected parameters of dysphonia,» *Journal of voice*, vol. 16, n. (1), pp. 20-27, 2002.
- [14] J. Hillendrand, R. A. Cleveland e R. L. Erickson, «Acoustic correlates of breath vocal quality.,» *Journal of Speech, Language and Hearing research*, vol. 37, n. (4), pp. 769-778, 1994.
- [15] Y. D. Heman-Ackah, D. D. Michael e G. S. Goding, «The relationship between cepstral peak prominence and selected parameters of dysphonia,» *Journal of voice*, vol. 1, n. 16, pp. 20-27, 2002.
- [16] G. Fava, N. P. Paolillo, G. Oliveira e M. Behlau, «Cross cultural adaptation of the itaian version of voice activitu participation profile,» *CoDAS*, vol. 26, pp. 252-255, 2014.
- [17] A. Castellana, A. Carullo, A. Astolfi, M. S. Bisetti e J. Colombini, «Vocal health assessment by means od Cepstral Peak Prominence Smoothed distribution in continuous speech,» *IEEE Instrumentation and Measurement Society*, 2018.
- [18] L. Park, «Multi Dimensional Voice Program (MDVP) Model 5105,» Kay PENTAX, NJ USA.
- [19] A. Castellana, A. Carullo, S. Corbellini e A. Astolfi, «Discriminating Pathological Voice From Healthy Voice Using Cepstal Peak Prominence Smoothed Distribution in Sustained Vowel,» *IEEE Transactions on Instrumentation and Measurement*, vol. 67, pp. 646-654, 2018.
- [20] E. Ficarra, *Bioinformatics*, Torino: Politecnico di Torino, 2018.

## Appendice A

### Brano foneticamente bilanciato

Avevo un bulldog che si chiamava Bulka. Era tutto nero salvo una macchia bianca all'estremità delle zampe anteriori. Nei cani di questa razza, la mandibola è sempre prominente, così i denti superiori vengono a collocarsi dietro a quelli inferiori. Ma quella di Bulka era tanto grossa che tra gli uni e gli altri denti rimaneva molto spazio. Aveva il muso largo, grandi occhi neri e brillanti e i canini sempre scoperti, perfettamente bianchi. Somigliava a un grugno. Bulka era assai forte. E se afferrava qualcosa tra i denti non c'era verso che mollasse la sua preda. Stretti i canini nella carne dell'avversario, serrava la mascella e rimaneva sospeso come un cencio ad un chiodo: attaccato come una sanguisuga. Un giorno che era stato lanciato contro un orso, gli afferrò tra i denti un orecchio. L'orso cercava di colpirlo con una zampa, scuoteva la testa, ma non se ne poteva sbarazzare: finì per rovesciare il testone in terra per schiacciarvi il cane. Su quest'ultimo, però, perché lasciasse la presa, dovemmo gettare una secchia di acqua gelata. Lo avevo avuto da ragazzo e gli davo da mangiare io stesso. Quando dovetti partire a prestar servizio ne Caucaso, decisi di non prenderlo con me e cercai di andarmene senza che lo sapesse. Ordinai che lo tenessero rinchiuso. Ero giunto alla prima tappa, stavo per ripartire con i cavalli freschi, quando ad un tratto notai una palla nera e brillante che avanzava velocissima sulla strada. Era Bulka col suo collare di rame al collo. Correva a perdifiato; si gettò su di me, mi leccò la mano e poi, la lingua ciondoloni, si stese all'ombra sotto la vettura. Seppi più tardi che aveva rotto un vetro per seguirmi; era saltato dalla finestra: aveva percorso venti chilometri d'estate, sotto un sole bruciante.