

UNIVERSITA' DEGLI STUDI DI SALERNO

Dipartimento di Fisica "E.R. Caianiello"



DOTTORATO DI RICERCA IN SCIENZE E TECNOLOGIE DELL'INFORMAZIONE,
DEI SISTEMI COMPLESSI E DELL'AMBIENTE

OSCILLAZIONI MECCANICHE IN SISTEMI BIOLOGICI

Germana Buccheri

Tutor

Dott.ssa M. Falanga

Coordinatore

Ch.mo R. Scarpa

INDICE

Introduzione:	I
Capitolo1: L'apparato fonatorio	1
1.1 Descrizione anatomica.....	1
1.2 La fonazione.....	6
1.2.1 Il flusso glottale.....	6
1.2.2 I muscoli della laringe.....	7
1.2.3 Descrizione e modello tratto voce.....	8
1.2.4 I diversi tipi di voce.....	11
Capitolo2: Il segnale vocale	13
2.1 Sintesi del flusso glottale	14
2.2 La sorgente glottale	15
2.2.1 Il modello a due masse simmetrico.....	15
2.2.2 Le equazioni di pressione nella glottide	19
2.3 I tratti vocale e nasale.....	21
2.4 I limiti della teoria classica.....	21

Capitolo3: Risultati sperimentali.....	23
3.1 Data set.....	24
3.2 Tecniche di estrazione del pitch.....	24
3.2.1 Spettrogramma e densità spettrale di potenza.....	25
3.2.2 Codifica predittiva lineare (LPC).....	26
3.2.3 Mutua informazione.....	31
3.2.4 Intertempi.....	31
3.3 Il ruolo della trachea nell'emissione dei suoni vocalici.....	34
3.3.1 Il transiente d'attacco.....	34
3.3.2 Il confronto con il suono emesso da una canna d'organo...35	35
3.4 Analisi delle componenti principali.....	36
3.5 Analisi delle componenti indipendenti.....	38
3.6 Modello fenomenologico del vocal tract.....	49
Capitolo4: Le oscillazioni meccaniche del cuore.....	51
4.1 L'apparato cardiocircolatorio.....	51
4.1.1 Il sangue e il cuore.....	51
4.1.2 L'attività meccanica del cuore.....	53
4.2 Il ciclo cardiaco.....	54
4.3 Data set.....	58
4.3.1 Analisi spettrale dei segnali.....	58

4.4 Tecniche di analisi applicate al segnale cardiaco.....	60
4.4.1 Spettrogramma.....	60
4.4.2 Analisi delle componenti indipendenti.....	61
4.4.3 La trasformata di Hilbert.....	62
Conclusioni.....	66
Appendici.....	68
Appendice A: Analisi delle Componenti Indipendenti (ICA).....	69
Appendice B: La Trasformata di Hilbert.....	71
Bibliografia.....	73
Produzione scientifica.....	77

Introduzione

La comprensione dei sistemi non lineari passa attraverso lo studio delle dinamiche oscillatorie ad essi associati e può considerarsi un argomento che richiede notevoli sforzi per essere completamente capito.

Il comportamento dei sistemi naturali spesso è associato a fenomeni di vibrazioni meccaniche, apparentemente diversi tra loro per forma e funzione ma che in realtà generano oscillazioni globalmente organizzate. Gli aspetti chiave della fenomenologia sono cicli limite, auto oscillazioni, relazioni di scala, dissipazione.

Nella presente tesi si affrontano lo studio e la caratterizzazione di autoscillazioni in alcuni sistemi biologici ponendo l'attenzione sull'apparato fonatorio e marginalmente sul sistema cardiaco.

Per molti anni le cosiddette scienze dure non si sono interessate allo studio dei sistemi biologici che invece sono caratterizzati fortemente da dinamiche oscillatorie che permettono una loro più approfondita comprensione e modellazione.

I processi fisici coinvolti nella produzione del suono in strumenti musicali auto-sostenuti e nell'apparato fonatorio (l'esempio più semplice è la produzione di suoni vocalici), le oscillazioni del condotto vulcanico, le oscillazioni della Terra, il battito cardiaco, costituiscono alcuni dei possibili esempi fisici in cui si osservano vibrazioni non lineari. Questi sistemi mostrano oscillazioni non lineari organizzate imputabili all'interazione fluido-solido o fluido-fluido. La sincronizzazione tra le scale temporali del flusso e i tempi caratteristici associati alle diverse parti rappresenta il meccanismo di innesco essenziale alla generazione di auto-oscillazioni. Le informazioni sulle caratteristiche dei meccanismi di sorgente di questi sistemi verranno esplorate dallo studio dei dati sperimentali ed osservazionali con l'obiettivo di costruire un modello fisico esplicativo della fenomenologia osservata. Il lavoro di ricerca si concentrerà sullo studio dell'emissione dei suoni vocalici attraverso l'apparato fonatorio e del funzionamento del cuore che sono associati all'interazione non lineare tra la fase fluida (es. flusso d'aria, sangue) e la struttura solida vibrante (trachea, tessuto muscolare). Tale interazione coinvolge meccanismi di sincronizzazione tra i tempi caratteristici della dinamica dei fluidi, le vibrazioni del condotto (trachea) e il miocardio che si risolve nella generazione di autoscillazioni non lineari globalmente organizzate.

La comprensione di questi elementi non lineari, permette di stabilire il comportamento dell'emissione dei suoni vocalici e del funzionamento del cuore fornendo nel primo caso elementi di riconoscibilità di un singolo individuo o di una classe di individui e nel secondo caso permette di controllare il funzionamento del cuore. In particolare questa fase di ricerca iniziale potrebbe fornire informazioni per effettuare diagnosi sulla attività cardiaca con una metodologia innovativa.

Si è investigato il meccanismo delle autoscillazioni, migliorandone le osservazioni, sia attraverso l'impiego di strumentazione quali vibrometri laser doppler e sia attraverso metodologie informatiche avanzate al fine della modellazione teorica. Sono state effettuate misure simultanee di campi di vibrazione meccanica e di campi acustici per quanto riguarda lo studio dell'apparato fonatorio e sono state effettuate misure di vibrazione meccanica per l'apparato cardiaco cercando di

capire meglio il meccanismo di sincronizzazione delle oscillazioni coerenti fluido-solido e la geometria delle strutture solide vibranti. Le vibrazioni delle strutture solide sono state misurate attraverso vibrometri laser mentre, le misure di campo acustico sono state effettuate da un sistema di registrazione e da una catena microfonica. L'utilizzo congiunto della vibrometria laser per la misura delle vibrazioni meccaniche e dell'apparato fonico per le registrazioni acustiche mira alla comprensione di molti processi che caratterizzano la dinamica dei sistemi in esame, quali:

innescamento delle vibrazioni in strutture solide come risultato dell'accoppiamento non lineare tra un campo fluidodinamico e meccanico;

fenomeni di amplificazione e risonanza.

Ad esempio per l'apparato fonatorio, il primo punto viene affrontato andando a misurare in camera anecoica il campo meccanico ed acustico in diversi operatori che emettono suoni vocalici per capire, ad esempio, le differenze tra uomini e donne nella generazione delle autoscillazioni. Le vibrazioni osservate si innescano difatti da una opportuna sincronizzazione tra le scale temporali del flusso e i tempi caratteristici propri della parte solida.

Relativamente al punto due lo scopo è guardare al campo acustico generato da più strutture solide che oscillano simultaneamente. Questi meccanismi possono portare a fenomeni di amplificazione e risonanza.

I sistemi d'interesse sono caratterizzati, come già detto, da una serie di processi fortemente accoppiati che risultano essere non lineari e operanti su scale spaziali e temporali diverse, su differenti ordini di grandezza. La modellazione fisico-matematica della totalità di questi sistemi complessi richiede la comprensione di come l'informazione di eventi alla microscala influenzi il comportamento del sistema alla macroscale. Quest'aspetto viene evidenziato nello studio dei sistemi biologici in esame. Gli esseri viventi hanno una struttura temporale e i fenomeni bioperiodici esistono a tutti i livelli di organizzazione, dagli eucarioti unicellulari ai pluricellulari tra cui l'uomo. In essi convivono un'ampia serie di fenomeni oscillatori, su diverse scale temporali che variano da pochi millisecondi (moti delle proteine, attività di scarica neuronale), a secondi (battito cardiaco e respirazione), ad ore (ritmo sonno-veglia), a giorni (ciclo ovarico), mesi ed anni (variazioni metaboliche stagionali, o lo stesso ciclo vitale).

Un tipico esempio di più fenomeni naturali governati da dinamiche oscillatorie che vivono su scale temporali diverse è il segnale delle oscillazioni meccaniche indotte lungo la trachea di un operatore che emette un suono vocalico. La sua analisi spettrale mette in evidenza in una prima parte la presenza di fenomeni caratterizzati da basse frequenze e in una seconda parte la presenza di fenomeni a frequenze più alte. I fenomeni naturali descritti dai valori di bassa frequenza sono quelli che caratterizzano il processo di respirazione e il battito cardiaco, mentre le frequenze più alte sono associate all'emissione dei suoni vocalici.

Lo studio di questa tipologia di segnale è affrontato nel seguente lavoro di ricerca, la sua analisi dettagliata attraverso tecniche standard ed innovative ci hanno permesso di trarre importanti conclusioni sul funzionamento del cuore e dell'apparato fonatorio.

Capitolo 1

L'apparato fonatorio¹

Il termine fonazione (derivato dal greco, "voce, suono") indica il processo fisiologico in base al quale si produce un suono o un rumore per mezzo degli organi vocali. Concorrono nella fonazione: un meccanismo di vibrazione, una forza, prodotta dal flusso d'aria, che dai polmoni, per mezzo dei bronchi e della trachea, raggiunge e attraversa la glottide, mettendo in vibrazione le corde vocali vere; un apparato di risonanza, costituito da tutte le cavità situate al di sopra della glottide (la porzione superiore della laringe, la faringe e le cavità nasali e orali). Il sistema nervoso interviene nella fonazione svolgendo un'azione di coordinamento fra le varie componenti e una di controllo per mezzo dell'udito.

La fonetica articolatoria si occupa in primo luogo di descrivere l'anatomia degli organi preposti alla produzione del suono, vale a dire dell'apparato fonatorio. In secondo luogo, essa si occupa di studiare la fisiologia di tali organi, cioè il processo mediante il quale essi producono i foni: tale processo è detto fonazione. Infine, la fonetica articolatoria fornisce i criteri e i termini per la classificazione dei suoni. L'apparato fonatorio è l'insieme delle strutture anatomiche che l'uomo utilizza per parlare. Si noti che esso è formato per la maggior parte da organi che svolgono principalmente altre funzioni, come quelle digerenti e respiratorie.

1.1 Descrizione anatomica

Vengono descritte le caratteristiche anatomiche degli organi coinvolti nella fonazione, facendo riferimento ai seguenti piani geometrici (Fig.1):

Trasversale: individua le metà superiore ed inferiore del corpo;

Coronale o frontale: individua le metà anteriore e posteriore;

Sagittale o mediano: individua le metà destra e sinistra.

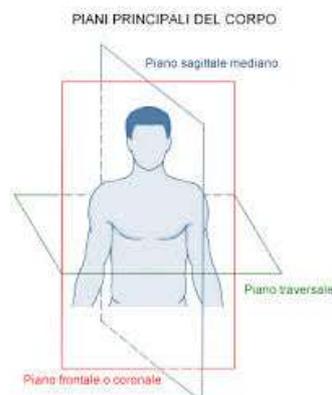


Fig.1: I piani principali del corpo umano.

¹ L'anatomia dell'apparato fonatorio e il processo di fonazione è descritto in questo capitolo considerando la struttura e i contenuti presenti nella tesi del dott. E. Marchetto dal titolo "INVERSIONE DI UN MODELLO FISICO DELL'APPARATO FONATORIO MEDIANTE PROGRAMMAZIONE DINAMICA E RETI RBF", per fornire ulteriori dettagli e una maggiore comprensione del fenomeno in esame.

L'apparato fonatorio è l'insieme delle strutture anatomiche che l'uomo utilizza per parlare. È formato da organi che, pur appartenendo primariamente agli apparati respiratorio e digerente, svolgono anche funzioni fonatorie. Risalendo lungo i vari tratti verso le labbra si trovano:

- i bronchi
- la trachea
- la laringe
- la faringe, che si divide in laringofaringe, orofaringe e rinofaringe
- le cavità nasale ed orale.

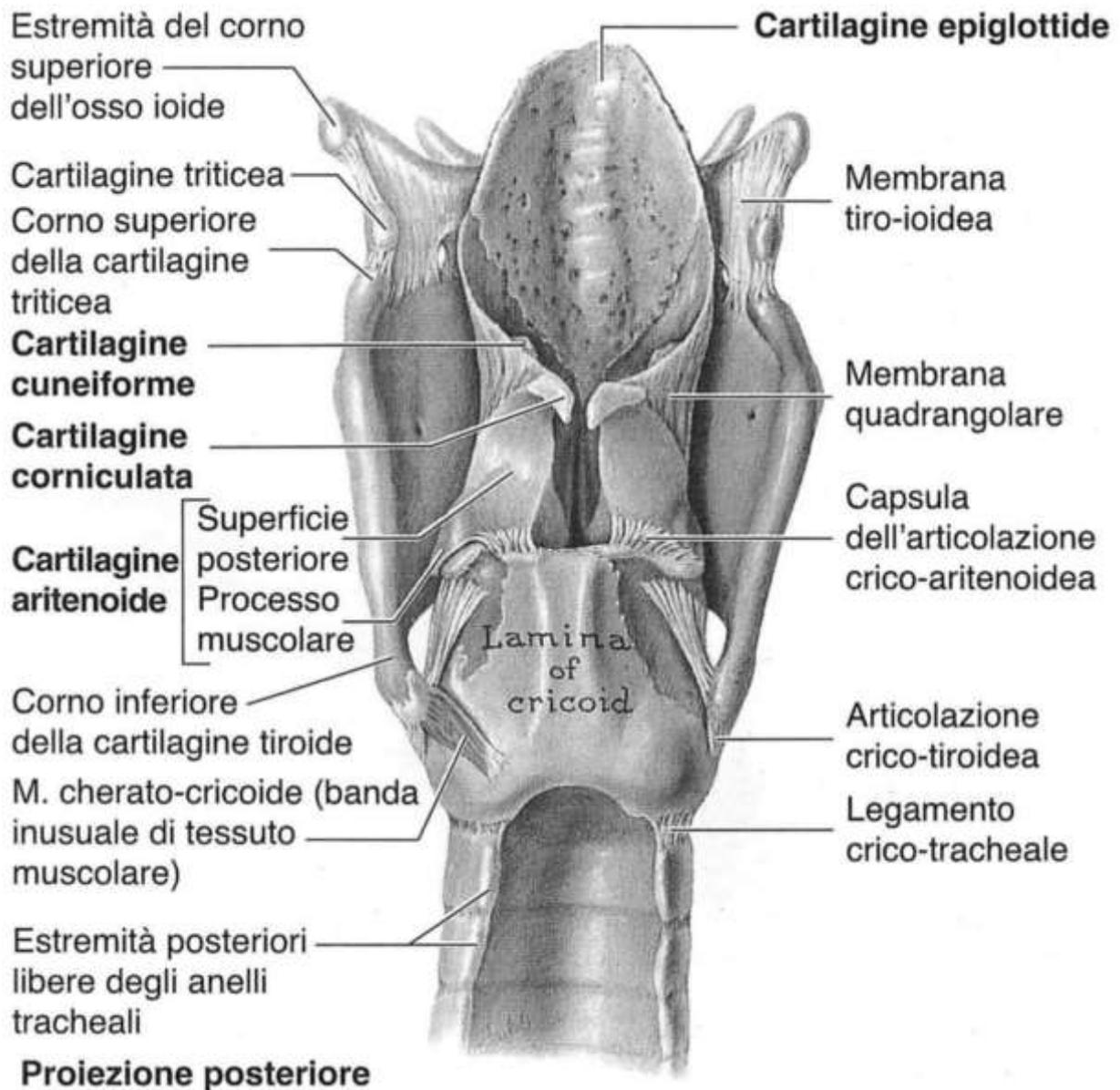


Fig. 2: Vista posteriore sul piano frontale della laringe.

La laringe è un condotto che inizia dietro la lingua, facendo seguito alla faringe, e continua nella trachea. Permette il passaggio dell'aria inspirata ed espirata e consente l'emissione dei suoni

(fonazione). Essa è provvista di un dispositivo di chiusura che, durante la deglutizione, impedisce che il bolo alimentare passi nelle vie respiratorie.

Nella laringe si trovano una serie di cartilagini e muscoli, che svolgono un ruolo fondamentale per quanto riguarda la fonazione. L'unico osso presente nella laringe è quello ioide che fornisce supporto alla cartilagine triticea che, a sua volta, è collegata alla cartilagine tiroide; quest'ultima, come si osserva nella figura 3, è la principale costituente della laringe. Quest'organo è costituito da nove cartilagini di cui tre impari e tre pari, di seguito descritte.

Tiroide: è la maggiore delle cartilagini; la parte inferiore delle due lamine che la costituiscono sono unite al centro per formare la prominenza laringea detta "pomo d'Adamo". Al di sopra di tale prominenza le due lamine si dischiudono; proseguendo lungo i loro bordi, uniti all'osso ioide dalla membrana tiro-ioidea, si giunge al corno superiore. In basso, invece, i corni inferiori si articolano con la cartilagine cricoide formando le articolazioni crico-tiroidee. Esse permettono la rotazione e lo scivolamento della cartilagine tiroidea, permettendo alle corde vocali di variare la propria lunghezza.

Cricoid: questa cartilagine, forma un anello completo, circonda la trachea e la raccorda alla laringe attraverso i legamenti crico-tiroideo e crico-tracheale.

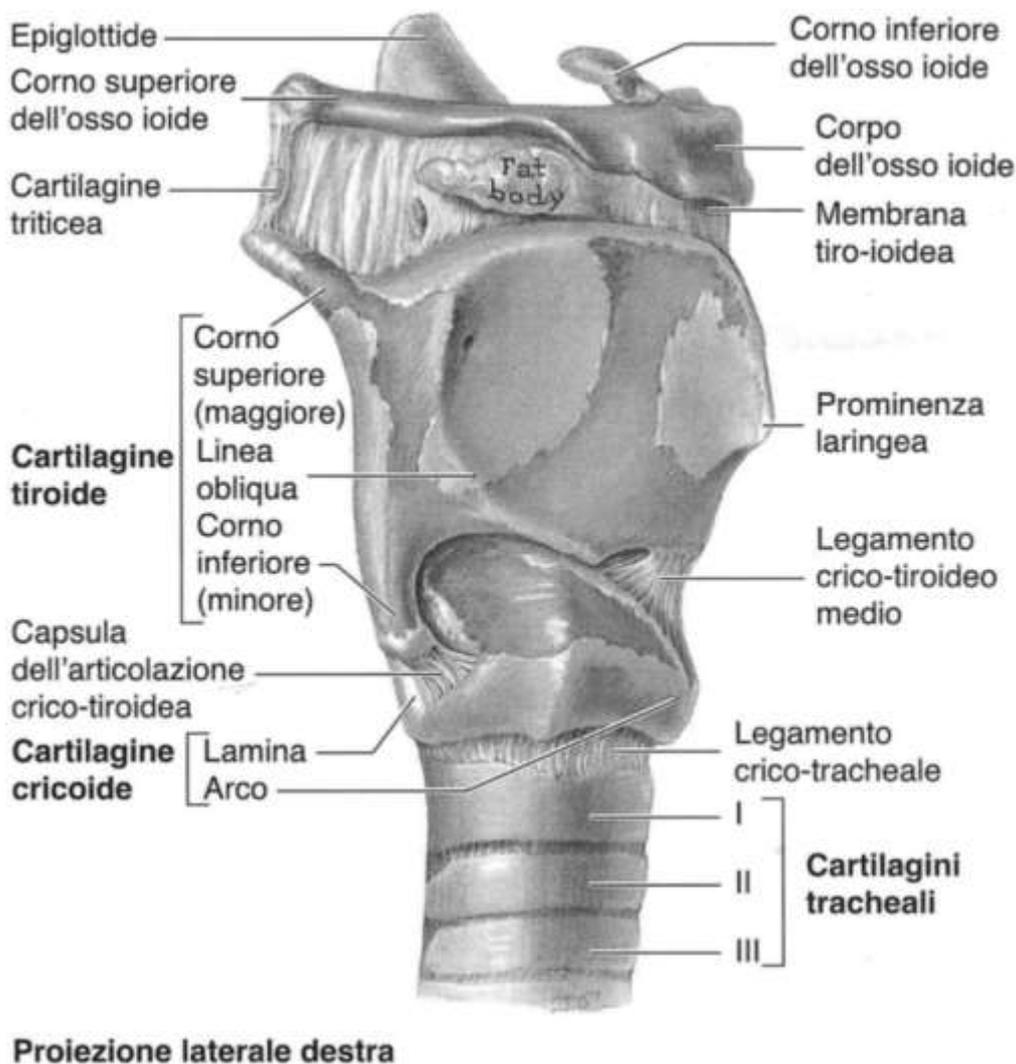


Fig. 3: Vista da destra sul piano mediano della laringe.

Aritenoidi: piccole cartilagini pari che trovano la loro collocazione nella cartilagine tiroide e che si articolano con l'estremità superiore della cartilagine cricoide, formando le articolazioni crico-aritenoidee. In Fig. 2 è evidente la posizione delle cartilagini aritenoidi e la loro funzione di supporto alle cartilagini corniculate e cuneiformi. L'importanza delle cartilagini aritenoidi si apprezza osservando la Fig. 4; dove è chiaro che ciascuna corda vocale (o piega) nasce proprio dal bordo superiore delle cartilagini aritenoidi ed è facile intuirne il ruolo importante per la fonazione.

Epiglottide: è posta dietro l'osso ioide e sopra la cartilagine tiroidea. Ha la forma di una foglia che rimane libera all'interno della laringe e il suo picciuolo, tramite il legamento tiroepiglottico, è legato alla cartilagine tiroide.

Corniculate e cuneiformi: sono piccole porzioni di tessuto cartilagineo poste nella parte posteriore delle pieghe ariepiglottiche, ben visibili in Fig.2. Le cartilagini corniculate si articolano sulle cartilagini cuneiformi, mentre queste ultime non si articolano con alcuna altra cartilagine.

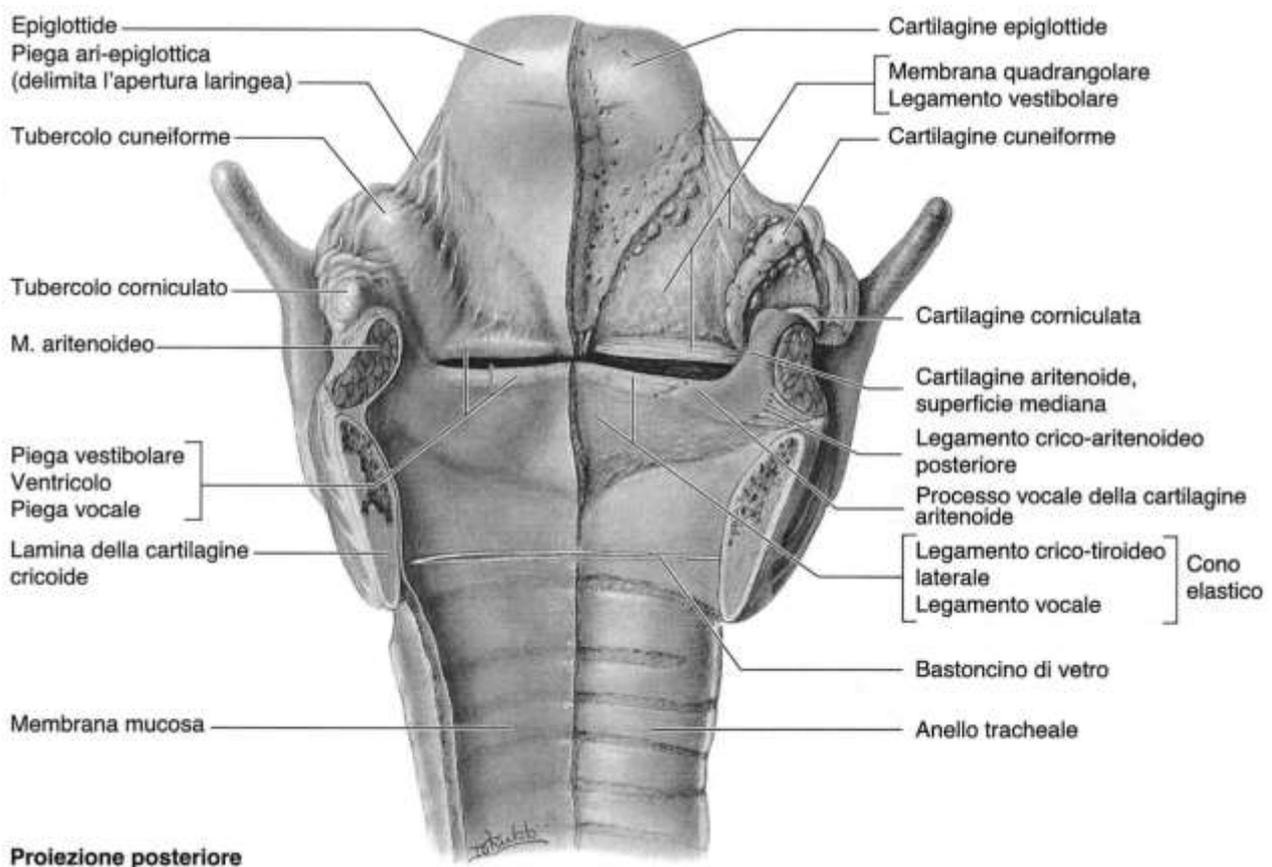


Fig. 4: Vista anteriore dell'interno sul piano coronale della laringe.

Vengono descritte dettagliatamente le corde vocali che si trovano nella cavità laringea. Ciascuna membrana quadrangolare (Fig. 3), congiunge i margini laterali delle cartilagini aritenoide e l'epiglottide; i loro bordi inferiori, invece, sono liberi e costituiscono le pieghe vestibolari (o corde vocali false). Le "false" corde vocali che si trovano immediatamente sopra le "vere" corde vocali (Fig. 4) e impediscono ai solidi e ai liquidi ingeriti di entrare nella trachea durante la deglutizione.

La corda vocale vera è costituita:

1. da una sistema legamentoso (legamento vocale), vera struttura portante,

2. da una formazione longitudinale deformabile (legamento elastico), alla quale si deve la capacità di sopportare uno stress tensionale (ogni incremento di frequenza è dovuto a un aumento di lunghezza cordale) e di scaricare il trauma che le corde ricevono nel momento in cui si avvicinano, 3. dal muscolo tiroaritenoidio (Fig.5).

Le corde vocali possono essere paragonate a due elastici, rivestiti da muscolo e mucosa, posizionati su un piano orizzontale e connessi anteriormente all'angolo dietro formato dalla cartilagine tiroide e posteriormente a due piccole cartilagini, dette aritenoidi, che ruotano su se stesse e sono capaci di slittare medialmente l'una verso l'altra. In anatomia i punti nei quali le corde vocali sono vincolate al sistema di sostegno laringeo vengono chiamate commessure. Presso la commessura anteriore esse sono vicinissime e i loro estremi convergono. Le corde vocali dall'avanti all'indietro divergono, disegnando durante l'inspirazione una V a vertice anteriore (Fig.5)

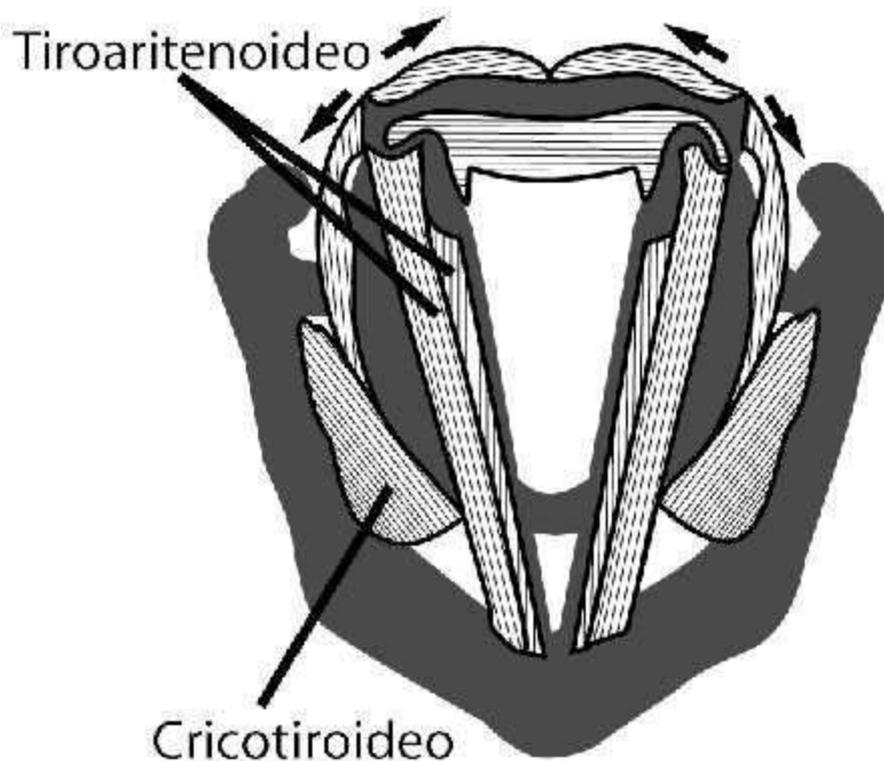


Fig.5: le corde vocali.

È proprio questa situazione anatomica che rende possibile, a seguito della rotazione della aritenoidi sul proprio asse, lo spostamento medio-laterale della corda.

Le corde vocali sono al massimo della loro apertura durante la respirazione e permettono all'aria di transitare, mentre, quando ci prepariamo a fonare le corde vengono avvicinate alla linea mediana per la contrattura del muscolo tiroaritenoidio. È questo improvviso ostacolo, trovato dall'aria in uscita dalla trachea, che permette la nascita della vibrazione e la produzione dei suoni. Quando la fonazione termina, la contrattura di un sistema muscolare posizionato posteriormente alle aritenoidi (muscoli dei postici), produce una rotazione inversa delle cartilagini che riapre la glottide.

Il muscolo tiroaritenoidio, oltre a permettere la rotazione della aritenoidi e, con questa, l'avvicinamento delle corde vocali tra loro, contraendosi accorcia e compatta le corde stesse, aumentandone la massa.

All'esterno della scatola laringea è posizionato il muscolo cricotiroideo, che contraendosi fa basculare la cartilagine tiroide su una struttura sottostante (cartilagine cricoide) alla quale essa è articolata. Tale movimento, simile all'abbassamento della visiera di un elmo del quale la cartilagine cricoide sia la mentoniera, determina un allungamento della corda, un aumento della sua tensione longitudinale e una diminuzione della sua massa.

La contrazione combinata dei due sistemi muscolari: tiroaritenoidico (accorciatore) e cricotiroideo (allungatore) modula la dimensione delle corde vocali, producendo le diverse frequenze della voce e le diverse voce quality (maggiore attività del tiroaritenoidico: tono più grave, timbro glottico più pieno; maggiore attività del cricotiroideo: tono più acuto, alleggerimento del timbro glottico) [Magnani, 2010].

Il termine voce quality, descritto nella sez. 1.2.2, tradotto letteralmente come "qualità della voce" indica i differenti tipi di voce ottenuti, ad esempio, a causa del particolare stato d'animo o stato emotivo di chi parla oppure a causa di alcune patologie.

Nel seguito, non si farà riferimento alle corde vocali ma alla glottide; questo termine identifica tutto l'apparato vocale della laringe: le corde vocali, i loro processi (cioè i legamenti con le cartilagini) e la rima della glottide, ovvero l'apertura tra le corde vocali stesse.

1.2 La fonazione

1.2.1 Il flusso glottale

In questo lavoro di ricerca si prende in considerazione solo la produzione di suoni vocalici. Alcuni suoni basati sull'uso degli organi articolatori (ad esempio le consonanti occlusive o fricative, quali [t] [b] [q] ed [ʃ] [ʧ]) non sono legati al treno di impulsi uscente dalle corde vocali, ma nascono dalla presenza di una pressione subglottale che si ottiene da una leggera compressione dei polmoni.

Altri suoni, ad esempio i nasali [m] [n], sono prodotti attraverso il flusso glottale; ma non sono considerati vocalici perché l'aria interagisce anche con ostacoli diversi dalle corde vocali.

Si descrivono i diversi stadi del processo di fonazione considerando i soli suoni vocalici. I muscoli della glottide dispongono le corde vocali con forma e dimensione adatti alla fonazione da realizzare; nel seguito consideriamo una posizione della glottide che corrisponde ad una voce quality descritta in letteratura con il termine "modale".

Le corde vocali, quando si trovano nella posizione di equilibrio, presentano tra di esse un'area che coincide con la sezione intraglottale della laringe; le stesse corde potrebbero essere in parziale contatto tra di esse, relativamente al tipo di fonazione che si produce.

I polmoni, attraverso alcuni muscoli che agiscono sul diaframma, vengono compressi e forniscono una certa pressione subglottale, che rimane costante durante la fonazione e che agisce sulle corde vocali. L'aumentando della pressione provoca un allontanamento (abduzione) delle corde vocali l'una dall'altra, vincendo anche l'eventuale adesione tra le due corde; l'area della glottide aumenta.

Nell'area sopraglottale si verifica un "getto" d'aria coincidente con l'impulso glottale, che provoca una diminuzione della pressione intraglottale e nel momento di flusso massimo del getto si ha la conclusione della fase di apertura e l'inizio della fase di chiusura della glottide.

Se diminuisce la pressione intraglottale, torna a prevalere l'azione dei muscoli portando nuovamente al valore nominale l'area della glottide. Si conclude così la fase di chiusura.

Se il flusso d'aria proveniente dai polmoni è sufficiente per sostenere le fasi di apertura e chiusura della glottide si ottiene l'emissione di un suono vocalico. In letteratura sono presenti descrizioni

accurate, a cui si rifanno i modelli fisici proposti [Titze, 1988] che considerano lo spessore della singola corda vocale e, lungo tale dimensione, evidenziano il comportamento del tessuto. In questo modo si possono descrivere i diversi tipi di deformazioni della corda vocale durante la produzione dell'impulso e gli effetti sulle caratteristiche dell'impulso stesso.

1.2.2 I muscoli della laringe

Si illustrano in dettaglio i muscoli laringei, data la loro importanza per la qualità della fonazione. In riferimento alla Fig. 6 vengono citati i seguenti muscoli:

crico-tiroideo: unisce la parte anteriore e laterale della cartilagine cricoide con il margine inferiore della cartilagine tiroide; la sua azione non incide direttamente sulle corde vocali, ma interessa l'intera laringe.

crico-aritenoideo posteriore: nasce nella superficie posteriore della cartilagine cricoide e raggiunge la cartilagine aritenoide.

crico-aritenoideo laterale: come l'omonimo posteriore, unisce la cartilagine cricoide con la aritenoide; ha origine dall'arco della cricoide.

tiro-aritenoideo: unisce la aritenoide con la parte posteriore della cartilagine tiroide.

aritenoidei, trasverso ed obliquo: vanno da una cartilagine aritenoide a quella sul lato opposto.

vocale: E' una parte delle corde vocali; rilascia la parte posteriore delle corde stesse mentre aumenta la tensione della parte anteriore.

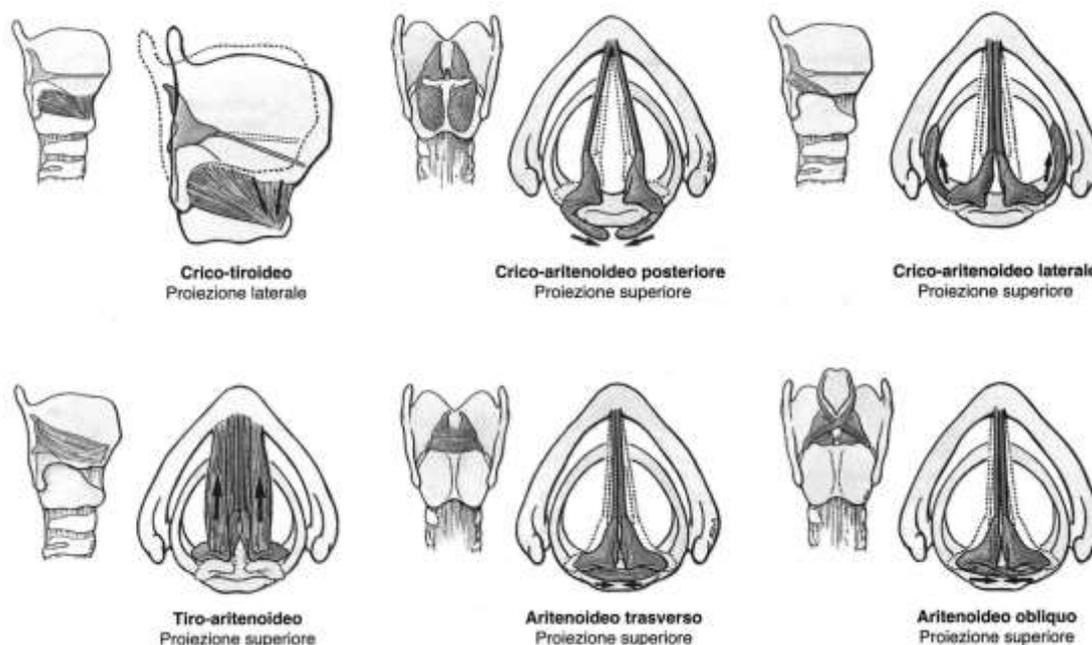


Fig. 6: I muscoli laringei.

I muscoli descritti si classificano in base alla loro funzione. Il muscolo crico-tiroideo è chiamato tensore perché aumenta la tensione delle corde vocali; il suo antagonista è il muscolo tiro-aritenoideo, detto rilasciatore. Il muscolo crico-aritenoideo laterale è chiamato adduttore perché tende ad avvicinare le corde vocali; al contrario sono detti muscoli abduttori il crico-aritenoideo posteriore e l'aritenoide trasverso.

Il posizionamento delle corde vocali durante il processo di fonazione è controllato dai muscoli crico-aritenoide laterale ed aritenoide trasverso. Infatti, se il muscolo aritenoide trasverso non si contrae, le cartilagini aritenoidi non sono abbastanza addotte e l'aria che passa attraverso la glottide non riesce a mettere in vibrazione le corde vocali. Questo tipo di fonazione corrisponde alla voce quality detta whispery.

I muscoli crico-tiroideo, crico-aritenoideo laterale e tiro-aritenoideo sono correlati alle variazioni di voce quality.

1.2.3 Descrizione e modello del tratto vocale

Il tratto vocale è costituito dagli organi che la colonna d'aria proveniente dalla trachea deve attraversare prima di uscire all'esterno: nella sua parte più interna troviamo la laringe, mentre il suo confine esterno è segnato dalle due labbra (ed eventualmente dalle narici del naso). Mentre la trachea non è modificabile nel suo volume interno, il tratto vocale è uno spazio in cui possono essere creati in modo volontario o involontario modifiche al flusso d'aria. Molte delle superfici del tratto vocale sono infatti mobili: perciò partendo dal medesimo impulso glottale si ha la diversificazione delle vocali variando la conformazione della lingua, delle labbra e della cavità orale in genere. La comunicazione verbale si realizza mediante un'onda sonora (o segnale acustico) trasmessa attraverso un mezzo (generalmente l'aria) (*livello di trasmissione*) dal parlante (*livello pneumo-fono-articolatorio*) all'ascoltatore (*livello di ricezione uditiva*). La voce dunque, in accordo con la teoria acustica di Fant [Fant, 1960] e Flanagan [Flanagan J. L, 1965], è il risultato di una modificazione da parte di un filtro complesso e variabile nel tempo di un segnale prodotto da una sorgente (teoria della sorgente-filtro). Il meccanismo fisiologico che porta all'origine del segnale a livello della sorgente è inteso come fonazione, mentre la modificazione dello stesso da parte del filtro come articolazione-risonanza. Il caso più semplice è la produzione dei suoni vocalici in cui la sorgente è costituita dalle corde vocali, mentre il filtro è riferito alle strutture sovraglottiche (Fig.7)

Esiste una vasta letteratura sull'azione esercitata dal tratto orale sugli impulsi glottali; in questa sezione vengono descritti i punti più importanti, evitando le descrizioni anatomiche.

L'azione del tratto vocale superiore per ottenere il segnale vocale partendo dal flusso glottale viene discussa in questa sezione. In letteratura, il tratto vocale è modellato in modo approssimativo ma sufficiente, a descrivere il suo effetto sugli impulsi glottali, con una serie di tubi acustici, di lunghezza e sezione differenti che sono collegati in serie. In questo modello si considerano la presenza di risonanze, armoniche in caso di tubo a sezione costante.

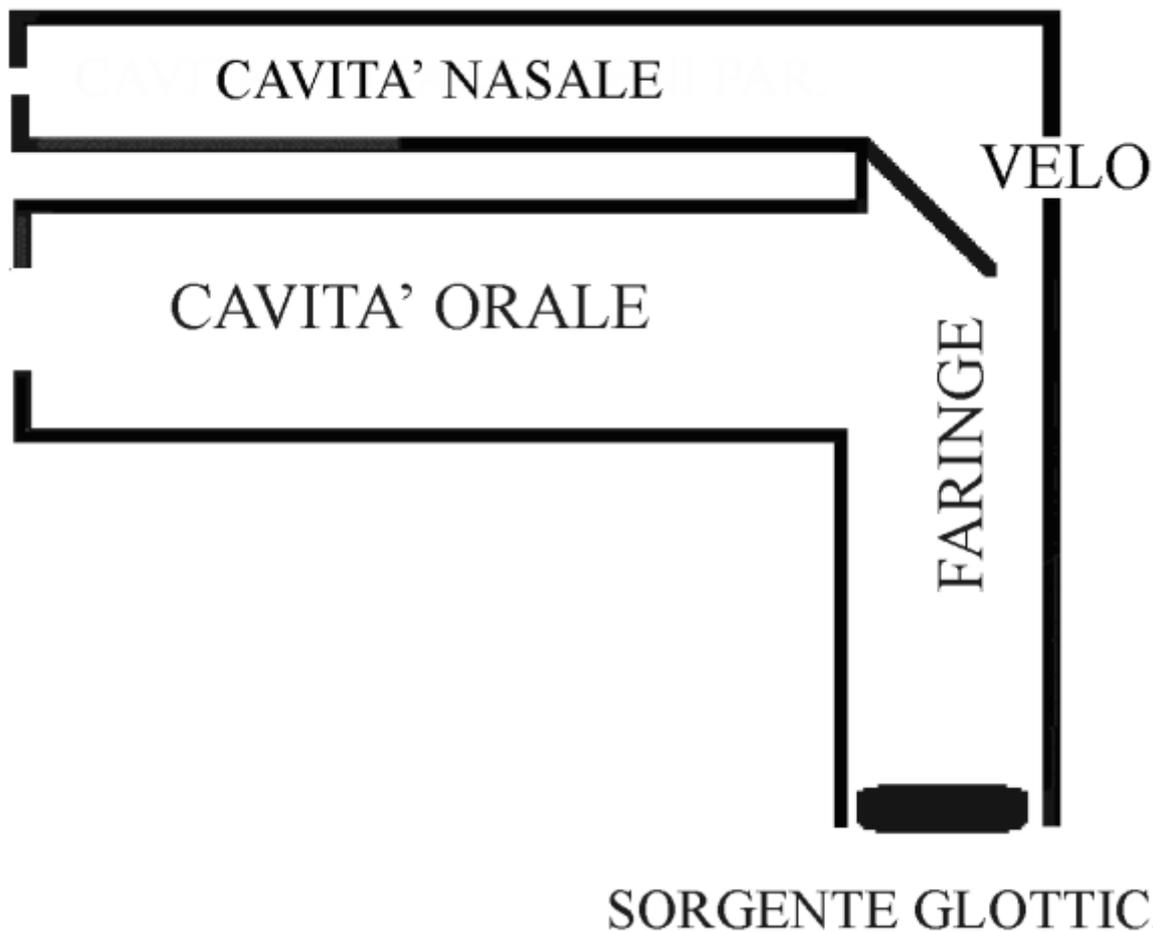


Fig. 7: rappresentazione schematica del sistema di produzione del segnale verbale con sorgente glottica e filtro costituito dal tratto vocale sovraglottico (faringe, cavità orale ed eventuale inserimento in parallelo, per azione del deviatore velo-palatino, delle cavità nasali ed annessi).

Applicando un suono puro al tratto orale, cioè una pressione acustica che varia secondo una funzione sinusoidale ed avente una precisa frequenza f , il suono stesso risulta attenuato dall'attraversamento del tratto orale. Questa attenuazione è proporzionale alla differenza tra f e la frequenza della più vicina risonanza. In altre parole se un suono puro attraversa un tubo acustico rimane inalterato solo se ha una frequenza che coincide con una risonanza; altrimenti tanto è maggiore la differenza rispetto ad una risonanza e tanto più risulta attenuato. Nel caso del tratto orale, ed analogamente nel tratto nasale, le risonanze vengono dette formanti e caratterizzano ciascun suono vocalico; generalmente risultano essere ben più alte della frequenza fondamentale (o frequenza di pitch) f_0 dell'impulso glottale.

È nel tubo di risonanza sopraglottideo, infatti, che la voce acquista il suo particolare timbro, in relazione alla forma ed alle dimensioni delle varie cavità. Nel caso della fonazione dei suoni vocalici è apprezzabile l'importanza degli organi articolatori. Questi suoni si differenziano l'un l'altro essenzialmente per una variazione della posizione della lingua. Gli spostamenti della lingua, durante la produzione dei vari suoni producono un cambiamento della forma e delle dimensioni della cavità orale e delle altre parti del tubo di risonanza. Questo determina una variazione della funzione di trasferimento caratteristica del tubo di risonanza, che si comporta un filtro. Tale funzione presenta dei massimi, cioè delle attenuazioni minime, in corrispondenza di precise zone dello spettro di frequenza, chiamate frequenze formanti o semplicemente formanti.

Nella tabella si riportano le frequenze in Hz delle prime tre formanti delle vocali italiane [Story and Titze, 1996].

Formante	[a]	[æ]	[i]	[o]	[u]	[ɛ]	[ʌ]
F1	754	692	333	540	389	624	707
F2	1195	1873	2332	922	987	1853	1161
F3	2685	2463	2986	2584	2299	2475	2591

Tabella 1.1: Frequenza dei primi tre formanti delle vocali italiane; le ultime due colonne si riferiscono alla versione “aperta” di [æ] e [o].

La prima formante (F1) compresa tra i 300 e gli 800 Hz a seconda della vocale pronunciata, si determina nella cavità orale ed è direttamente correlata con la maggiore o minore apertura mandibolare.

La seconda formante (F2) compresa tra i 900 ed i 2500 Hz viene creata nella regione faringea e dipende dalla posizione linguale (lingua retroposta = abbassamento di F2 e viceversa).

La terza formante (F3) compresa tra i 2300 e i 3000 Hz è determinata dai movimenti del muscolo orbicolare della bocca : una maggiore protrusione labiale comporta un abbassamento di F3.

Ricordiamo che, come spiegato, il suono proveniente dalle corde vocali è ricco di armoniche. Durante il suo attraversamento del tratto vocale, o nasale, molte di queste armoniche vengono attenuate o eliminate mentre le rimanenti danno il solo suono desiderato. Osservando la figura 8 è possibile avere un esempio di come la conformazione del tratto vocale porti a differenti risultati sonori. La posizione della lingua per la vocale neutra [ɜ] è indicata dalla linea nera; si nota come la sezione della cavità vocale sia costante dalla glottide alle labbra. Il risultato è uno spettro sonoro caratterizzato da picchi equidistanti in frequenza.

Cambiare la forma del tratto vocale alzando, abbassando e portando in avanti o all’indietro la lingua conduce a spettri acustici con differenti forme, come accade con le altre vocali mostrate in figura.

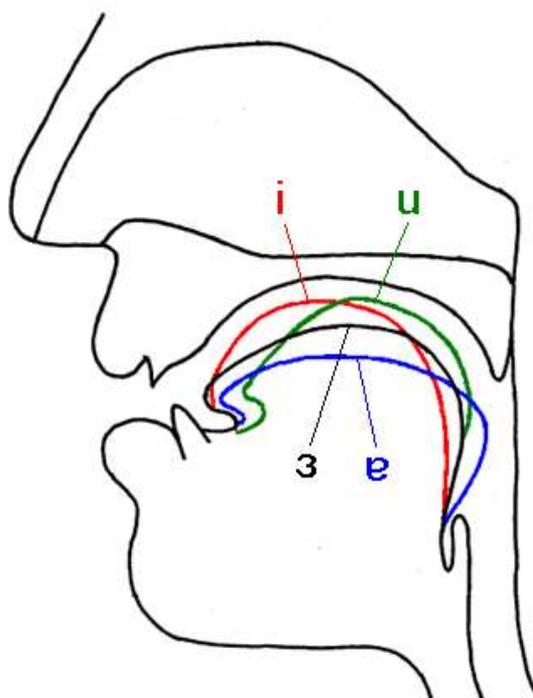


Fig. 8: Posizione della lingua durante l’emissione delle vocali [ɜ], [a], [i] e [u] (da www.ling.mq.edu.au).

1.2.4 I diversi tipi di voce

I diversi tipi di voce vengono indicati in letteratura con il termine “Voice Quality”. Questo termine identifica tutte le caratteristiche della qualità della voce, intese come il “tipo” di voce usato da chi parla. È la glottide che, in base alle varie posizioni assunte dalle corde vocali, modifica la colonna d'aria che giunge dai polmoni per dar luogo a diversi modi di fonazione.

I numerosi studi pubblicati hanno introdotto una serie di termini per descrivere le voice quality, così definiti:

Modal: la voce modale, ovvero il tipo neutro di fonazione, è caratterizzata da un tipo di fonazione nella quale le pliche vocali vibrano in modo periodico (il ciclo di apertura e chiusura è regolare nel tempo);

Soft/Loud: le voci tenue o alta sono legate a voci, rispettivamente, di bassa o alta intensità.

Breathy: questa voce, il cui nome potrebbe essere tradotto come “aspirata”, è caratterizzata da un'eccessiva emissione d'aria durante il processo di fonazione. A volte il tono risulta essere accompagnato da raucedine, mentre in altri casi è preceduto o seguito da un sibilo. La perdita d'aria e il movimento asimmetrico delle corde vocali generano una componente di rumore che si affianca al tono glottale. Si ha un'imperfetta adduzione delle corde vocali durante la fase di chiusura del ciclo fonatorio.

Whispery : la voce sussurrata è distinta per l'assenza di fonazione: durante l'articolazione del parlato si verifica la sola emissione d'aria e le corde vocali non vibrano.

Harsh: questa voce, il cui nome potrebbe essere tradotto come “ruvida”, è caratterizzata dalla presenza di toni vocali a causa dei quali si percepisce una specie di sfrigolio. La laringe spesso è tesa e a volte è sollevata fino quasi a raggiungere la posizione assunta durante la deglutizione. Le cause che portano l'individuo alla produzione di voce ruvida sono l'imitazione, stati d'animo di aggressione od ostilità, il bisogno di farsi sentire in presenza di rumore o un uso improprio dei livelli di pitch.

Hoarse: la caratteristica principale della voce rauca è che presenta aperiodicità alle basse frequenze, presumibilmente legata a una qualche forma di vibrazione irregolare delle corde vocali.

Creaky: in questo caso le corde vocali vibrano lentamente e irregolarmente, perché il flusso d'aria che attraversa la glottide è molto scarso; in queste condizioni solo una parte delle corde vocali riesce a vibrare. Questa voice quality, detta anche “laringalizzata”, è solitamente legata a stati d'animo, quali stanchezza o noia; è altresì nota la sua presenza in alcuni casi nei quali la frequenza della voce raggiunge il suo limite inferiore, ad esempio in seguito ad alcuni toni.

Tremulous: la voce chiamata tremula è caratterizzata da una rapida oscillazione del valore della frequenza di pitch, dovuto allo stato di tensione delle corde vocali, noto come micro-tremor. Essa è associata a condizioni di stress. In situazioni di pressione psicologica i muscoli coinvolti nella determinazione della frequenza fondamentale subiscono delle contrazioni causate dalla tensione procurata dallo stress; ciò interferisce sul controllo e la stabilità del pitch stesso.

Hypernasal: la voce ipernasale si ha quando chi parla non riesce a bloccare il flusso d'aria attraverso la cavità nasale durante l'emissione dei suoni. Si tratta di una voce patologica, presente in soggetti che hanno il palato danneggiato o la paralisi dei muscoli che controllano il movimento del palato molle. Molti dei suoni ottenuti sono nasali o nasalizzati.

Hyponasal: un altro tipo di voce patologica è quella iponasale, caratterizzata dallo scarso o nullo flusso di aria attraverso la cavità nasale durante la fonazione. Un esempio di questo tipo di voce si

ha quando la cavità nasale è ostruita a causa di un forte raffreddore. In questi casi i suoni nasali sono sostituiti da suoni orali.

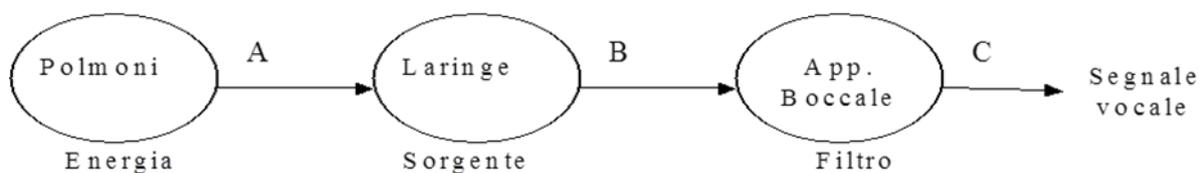
Capitolo 2

Il segnale vocale

La fisica alla base della produzione del parlato è da sempre di grande interesse scientifico. Un complesso studio teorico e sperimentale permette sia la comprensione dei fenomeni fisici coinvolti sia la possibilità di sviluppare modelli capaci di riprodurre i diversi suoni.

Comprendere il meccanismo con cui il cervello organizza la complessa attività del parlato nel tempo è importante per comprendere sia lo speech e sia il funzionamento del cervello. Sicuramente in aggiunta al controllo neurale, un ulteriore livello di organizzazione temporale è fornita da dinamiche oscillatorie, che sono intrinseche del tratto vocale. Alcuni modelli fisici, fin dall'inizio dello studio della produzione della voce, guardano a queste oscillazioni intrinseche ed inglobano questa fenomenologia in un meccanismo di eccitazione persistente non lineare che dà luogo ad oscillazioni auto-sostenute. Infatti, all'inizio del XIX secolo, Airy osservando il comportamento della laringe all'interno del complesso sistema del tratto vocale (polmone-trachea laringe, tratto vocale superiore), introdusse all'interno dell'equazione che lo descriveva un termine di ritardo, tale equazione viene oggi utilizzata per descrivere i modelli epidemici.

Circa un secolo dopo il lavoro Airy, Wegel scrisse una lagrangiana relativa alla laringe la cui variazione fornisce un sistema di equazioni le cui soluzioni producono oscillazioni auto-sostenute. Nel corso degli anni, rilevanti contributi per modellare il tratto vocale e descrivere l'emissione dei suoni ci sono stati forniti dai lavori di numerosi scienziati. In accordo con la teoria acustica di Fant [Fant, 1960] e Flanagan [Flanagan, 1965], il segnale verbale è il risultato di una modificazione da parte di un *filtro complesso* a volumi *variabili* nel tempo, di un segnale prodotto da una *sorgente*. Il meccanismo fisiologico che genera il segnale a livello di sorgente è inteso come *fonazione*, mentre la modificazione dello stesso da parte del filtro come *articolazione- risonanza (teoria lineare della sorgente-filtro)*.



Il caso più semplice è la produzione dei fonemi vocalici in cui la sorgente è costituita dall'attività delle corde vocali, mentre il filtro è riferito alle strutture del tratto vocale sovraglottico. Nella produzione dei fonemi consonantici il processo è più complesso in quanto vi è la comparsa, nel punto di articolazione, di una sorgente generatrice di un segnale aperiodico (rumore) che si sostituisce (consonanti sorde) o si somma al segnale periodico glottico (consonanti sonore). In entrambi i casi il segnale sarà sempre sottoposto alle modificazioni timbriche del filtro articolatorio.

Riferendoci alla produzione delle vocali, la fonazione (attività di sorgente) appartiene in senso stretto alla vibrazione glottica per effetto dell'aria espirata, mentre la risonanza (attività di filtro) al tratto vocale sovraglottico, dalle corde vocali escluse le labbra con aggiunta in parallelo, per mezzo del deviatore velo-palatino, della cavità nasale.

Dal punto di vista acustico, in condizioni di normale vibrazione glottica, il segnale sorgente laringeo è un segnale quasi periodico. Secondo la teoria lineare il segnale è la somma di una serie di segnali sinusoidali, definiti componenti o parziali, ciascuno avente una determinata frequenza, ampiezza e fase (serie o integrale di Fourier). Il termine quasi indica invece che il segnale sorgente, nelle sue caratteristiche di frequenza, ampiezza e forma d'onda, può modificarsi nel tempo.

Tali modificazioni in condizioni di normalità sono molto lente, per cui, considerando un breve intervallo temporale, il segnale stesso in prima approssimazione può dirsi del tutto periodico (analisi spettrale a breve termine). Il segnale sorgente dunque avrà una sua frequenza che corrisponde alla Frequenza Fondamentale f_0 (ed il cui inverso è il Periodo Fondamentale T_0) ed una serie di componenti o parziali dette armoniche, in quanto hanno frequenza che è multiplo intero della fondamentale.

La voce umana come segnale vocale non totalmente armonico, non completamente periodico e non lineare si presta ad essere analizzata con tecniche standard ed innovative che descrivano il segnale vocale sfruttando gli strumenti matematici propri dei Sistemi dinamici.

Nel caso di irregolare oscillazione cordale il segnale sorgente sarà sicuramente aperiodico con componenti che non sono armoniche in quanto la loro frequenza non è multiplo di una fondamentale. In presenza della componente aperiodica si ha la voce percettivamente definita disfonica. La disfonia è uno dei disturbi della voce che va dalla raucedine fino all'afonia. In pratica con il termine disfonia si intende la difficoltà nel produrre una voce "fisiologica" indotta da cause che possono essere organiche o funzionali.

Il modello che descrive il processo fonatorio e le sue successive modifiche che vengono in questo capitolo riportati, secondo la teoria classica, forniscono un ruolo marginale alla parte inferiore dell'apparato fonatorio nella produzione dei suoni. L'attività di ricerca proposta mira ad evidenziare la funzione fondamentale della trachea nello spiegare il fenomeno della fonazione ed a proporre un nuovo modello fisico matematico per la sintesi dei suoni vocalici.

2.1 Sintesi del flusso glottale

In questo capitolo si descrivono, i modelli fisico matematici usati per la sintesi del flusso glottale mediante calcolatore.

Si definisce modello fisico, la rappresentazione schematica del sistema reale; il modello matematico è, invece, la rappresentazione analitica (vale a dire in termini di equazioni) del precedente modello fisico. Un modello fisico può essere a parametri concentrati o distribuiti. Riferendosi al sistema in esame, si può affermare che un modello fisico a parametri concentrati è caratterizzato dalla combinazione di tre soli elementi di base: masse, molle e smorzatori. In generale, un modello fisico a parametri distribuiti è caratterizzato da un modello matematico ove compaiono equazioni differenziali alle derivate parziali, mentre un modello concentrato è caratterizzato da equazioni differenziali ordinarie.

I modelli a parametri concentrati si possono dividere in due categorie: quelli ad elementi finiti e quelli meccanici. I primi si applicano a corpi fisici che possono essere suddivisi in un certo numero, anche molto grande, di parti. Ciascuna di queste è modellata come un sistema fisico semplice e

distinto il cui effetto si manifesta sia sugli elementi adiacenti sia, a livello macroscopico, sul modello completo oggetto della simulazione. Tra le applicazioni dei modelli ad elementi finiti della sintesi vocale ricordiamo quello di Alipour-Haghighi e Titze, [Alipour-Haghighi et al., 1983].

I modelli meccanici permettono una descrizione del sistema da simulare mediante una sua idealizzazione in elementi fisici: masse, molle e smorzatori. Al contrario dei modelli ad elementi finiti, che cercano di descrivere dettagliatamente la dinamica del fenomeno simulato, i modelli meccanici forniscono una descrizione complessivamente precisa del fenomeno. Nel caso in esame il processo fonatorio risulta ben descritto da un punto di vista macroscopico.

Questo tipo di modellazione è più complicato dell'approccio ad elementi finiti che si basa sulla ripetizione di blocchi sempre uguali la cui interazione avviene sempre con la medesima relazione. I modelli meccanici impongono invece di riuscire a descrivere il fenomeno che si vuole simulare attraverso delle relazioni che non sono di banale stesura.

L'attività di ricerca di alcuni scienziati cerca realizzare buoni modelli meccanici della sorgente glottale e del tratto vocale perché, a fronte di un livello di dettaglio più che sufficiente per descrivere i fenomeni d'interesse, essi richiedono una quantità di calcoli sicuramente inferiore rispetto agli equivalenti modelli ad elementi finiti, così come molto minore è il numero di parametri necessari al controllo dei modelli. Quest'ultima considerazione è fondamentale per le applicazioni alla codifica efficiente del parlato, per la quale i modelli esposti in questo capitolo sembrano dare discreti risultati [Marchetto E., 2007].

2.2 La sorgente glottale

2.2.1 Il modello a due masse simmetrico

Molti modelli meccanici della sorgente glottale si possono trovare in letteratura, alcuni dei quali cercano di simulare tutti i modi del sistema con modelli a dieci [Wong et al., 1991] e sedici masse; [Titze, 1973, 1974]. Nel tempo la ricerca [Berry et al., 1994] ha mostrato come il comportamento delle corde vocali è dominato dai primi modi di oscillazione. Questo importante risultato ha messo in evidenza l'inutilità di simulare, per la sintesi del flusso glottale, tutti i modi mediante modelli sedici masse. Modelli con un ridotto numero di masse in movimento riescono a catturare un buon livello di caratteristiche oscillatorie.

In letteratura si trovano vari modelli meccanici con queste caratteristiche [Titze and Story, 2002; Avanzini, 2007], ma in questo lavoro di tesi si è scelto di descrivere quello che garantisce il miglior compromesso tra la complessità della formulazione e l'accuratezza dei risultati. Il modello presentato è quello a due masse di Ishizaka e Flanagan [Ishizaka and Flanagan, 1972] che risulta essere il più utilizzato in ambito scientifico. Inoltre bisogna precisare che esistono due formulazioni del modello a due masse: simmetrica ed asimmetrica. Nella formulazione simmetrica si suppone che le due corde vocali si comportino esattamente nello stesso modo, mentre in quella asimmetrica, si suppone un comportamento delle corde vocali simile ma non identico perciò si modellano quattro masse anziché due. Nel presente lavoro è stata illustrata in dettaglio solo il modello a due masse simmetrico ed in riferimento alle figure 1 e 2 si evince che il fulcro del modello sono le due masse m_1 ed m_2 che simulano il comportamento di una sola delle due corde vocali [Marchetto E.,2007].

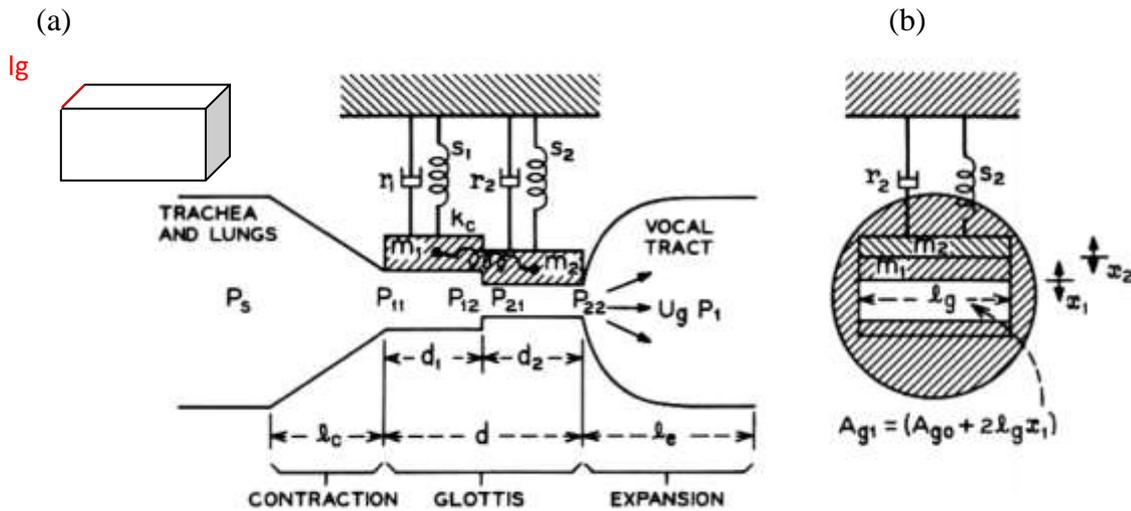


Fig. 1: Vista del modello a due masse sul piano coronale

In Fig.2 è ben rappresentato che le masse possono muoversi solo lungo l'asse x. Entrambe le masse sono accoppiate alla parete della laringe, supposta immobile, mediante i sistemi molla-smorzatore caratterizzati dai seguenti parametri: k_1 , r_1 e k_2 , r_2 . Infine si considera la molla $k_{1,2}$ che vincola le masse tra di loro; dato l'unico grado di libertà delle masse stesse si deduce che $k_{1,2}$ agisce sulla differenza di posizione delle due masse, ma non le "avvicina" come si sarebbe indotti a pensare.

Il modello è descritto analiticamente dalle relazioni matematiche che sono riportate in questo capitolo, partendo dalle seguenti equazioni di differenziali:

$$m_1 \ddot{x}_1 + r_1 \dot{x}_1 + k_1[x_1 - x_{01}] + k_c[x_1 - x_2] = Ld_1 p_1 \quad (2.1)$$

$$m_2 \ddot{x}_2 + r_2 \dot{x}_2 + k_1[x_2 - x_{02}] - k_c[x_1 - x_2] = Ld_2 p_2 \quad (2.2)$$

Lo spostamento delle masse m_1 ed m_2 dalle rispettive posizioni di equilibrio x_{01} e x_{02} è indicato dalle variabili x_1 e x_2 che sono funzioni del tempo mentre sono facilmente riconoscibili gli effetti delle costanti m_i , r_i e k_i . A destra di ciascuna equazione troviamo un valore costante Ld_i , che dipende dalle dimensioni fisiche delle corde vocali, moltiplicato per p_i ; che rappresenta la pressione esercitata sulla superficie Ld_i .

I termini delle equazioni assumono dunque le dimensioni di una forza che varia nel tempo.

Si ricavano le funzioni p_1 e p_2 partendo dalle pressioni che, idealmente, agiscono agli estremi delle masse m_1 ed m_2 rispetto alle dimensioni d_1 e d_2 ; risalendo lungo il tratto vocale, facendo riferimento alla figura 1, si incontra l'estremo uno della massa m_1 , poi l'estremo 2 della massa m_1 ed in seguito i due estremi della massa m_2 . Vengono determinati quattro valori di pressione che risultano essere funzioni del tempo: p_{11} , p_{12} , p_{21} e p_{22} . Questi valori di pressione ottenuti ci permettono di determinare p_1 e p_2 come qui di seguito illustrato:

$$p_1 = 1/2[p_{11} + p_{12}] \quad p_2 = 1/2 [p_{21} + p_{22}] \quad (2.3)$$

I valori istantanei di pressione agli estremi delle due masse vengono determinati tenendo conto della fisica del flusso d'aria attraverso la glottide. In particolare si hanno le seguenti transizioni di pressioni:

$p_s \rightarrow p_{11}$ La pressione subglottale p_s , dovuta alla compressione dei polmoni sotto l'azione dei muscoli, è applicata ad un tratto in cui si suppone esistere un flusso ideale stazionario e al quale si può applicare la legge di Bernoulli che manipolata porta alla seguente relazione:

$$p_s - p_{11} = 0.69 \rho_{\text{air}} \frac{u^2}{A_1^2}$$

dove:

- ρ_{air} è la densità dell'aria;
- u è il flusso d'aria che, come la densità, rimane costante in tutto il modello: si suppone l'incompressibilità del fluido;
- A_1 è l'area tra la massa m_1 e la sua controparte simmetrica, che si ottiene dalle dimensioni e dallo spostamento x_1 .

$p_{11} \rightarrow p_{12}$ Durante l'attraversamento della sezione del modello su cui insiste la parete della massa m_1 si suppone che la caduta di pressione sia governata da perdite legate all'attrito viscoso; tale caduta aumenta linearmente con la distanza:

$$p_{11} - p_{12} = 12\nu d_1 \frac{u l_g^2}{A_1^3}$$

dove:

- ν è il coefficiente di viscosità dell'aria;
- l_g è una delle dimensioni delle masse m_i , comprensibile osservando la Fig. 1.

$p_{12} \rightarrow p_{21}$ Nel passaggio tra la prima e la seconda massa del modello una variazione di energia cinetica determina una caduta di pressione. Il flusso rimane costante, ma la sezione che simula la glottide subisce una brusca variazione che impone un cambiamento altrettanto repentino della pressione.

$$p_{12} - p_{21} = \frac{1}{2} \rho_{\text{air}} u^2 \left(\frac{1}{A_1^2} - \frac{1}{A_2^2} \right)$$

dove u , A_1 ed A_2 sono funzioni del tempo e rappresentano il flusso e l'area delle sezioni trasversali della glottide.

$p_{21} \rightarrow p_{22}$ Questa caduta di pressione è simile alla variazione $p_{11} \rightarrow p_{12}$ precedentemente considerata, anch'essa descrive un modello di flusso in presenza attrito viscoso:

$$p_{21} - p_{22} = 12\nu d_2 \frac{u l_g^2}{A_2^3}$$

$p_{22} \rightarrow p$ Questa parte del modello, simile alla $p_{12} \rightarrow p_{21}$, si basa sulla variazione della sezione entro cui si muove il flusso d'aria; per calcolarla si applica la seconda legge di Newton $\mathbf{F} = m\mathbf{a}$ e si considera come massa m quella dell'aria in movimento. Si ottiene:

$$p_{22} - p = \frac{1}{2} \rho_{\text{air}} \frac{u^2}{A_2^2} \left[2 \frac{A_2}{S} \left(1 - \frac{A_2}{S} \right) \right]$$

dove:

- S è l'area della superficie con cui la massa m_2 insiste sul tratto attraversato dall'aria, pari a $l_g d_2$.

Mettendo insieme le equazioni 2.3 con le cadute di pressione determinate si ottengono due relazioni, che esprimono le pressioni, cioè le forze, che sollecitano le masse m_1 ed m_2 . Le relazioni finali sono:

$$p_1 = p_s - 0.69 \rho_{\text{air}} \frac{u^2}{A_1^2} - 6 v d_1 \frac{u l_g^2}{A_1^3} \quad (2.4)$$

$$p_2 = p + \frac{1}{2} \rho_{\text{air}} \frac{u^2}{A_2^2} \left[2 \frac{A_2}{S} \left(1 - \frac{A_2}{S} \right) \right] + 6 v d_2 \frac{u l_g^2}{A_2^3} \quad (2.5)$$

Queste relazioni mettono in evidenza che la variabile di output del modello, p , rimane implicita.

Per migliorare la descrizione della fase di chiusura della glottide è necessario aggiustare il modello, tenendo conto degli studi di Sondhi e Schroeter [Sondhi and Schroeter, 1987]. Infatti quando le masse m_1 e/o m_2 collidono con le loro controparti simmetriche, vengono aumentate di un certo fattore sia le costanti elastiche k_i sia le costanti di smorzamento r_i per tener conto del maggior effetto dei fenomeni dissipativi.

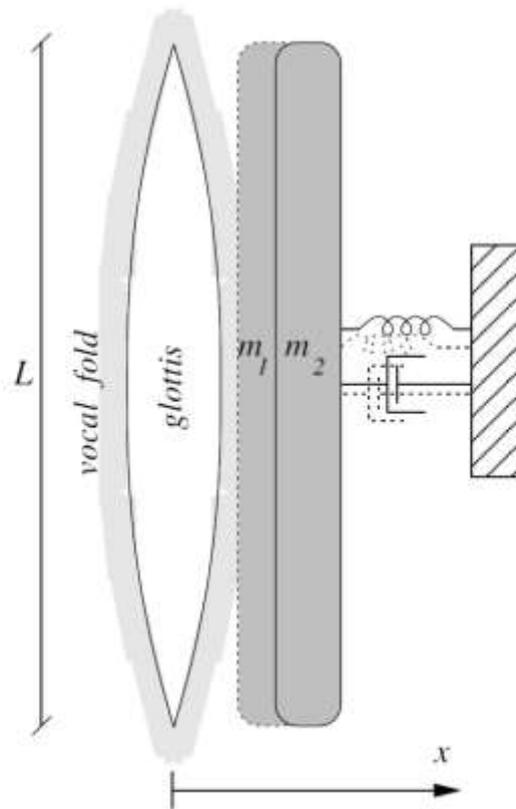


Fig. 2.: Vista dall'alto del modello a due masse sul piano trasversale.

2.2.2 Le equazioni di pressione nella glottide

Il modello a due masse simmetrico esposto nella sez. 2.2.1 è quello descritto nell'articolo di Ishizaka and Flanagan del 1972, il più accreditato della letteratura scientifica relativa a questo argomento. Nel tempo al modello originale sono state apportate delle modifiche e sono state scritte nuove relazioni dovute a studi successivi [Story and Titze, 1995]. Vengono modificate tutte le relazioni che esprimono le cadute di pressione nella glottide, mentre restano invariate le equazioni del moto delle masse.

Le assunzioni fatte per riscrivere le nuove equazioni di pressione, adottate anche in [Avanzini et al., 2006, Marchetto E., 2007], sono di seguito riportate:

- Nella regione subglottale compresa tra i polmoni (trachea) ed il punto in cui l'area intraglottale è minima il flusso può essere descritto utilizzando la legge di Bernoulli. Questo equivale a supporre che il flusso sia stazionario, incomprimibile e privo di viscosità.
- Nel punto di minima area intraglottale il flusso risulta “staccato” dalle pareti della laringe, cioè dalle corde vocali.
- Nella regione che va dal punto in cui l'area intraglottale è minima fino all'uscita della glottide, verso il tratto vocale, il flusso costituisce un “getto” staccato dalle pareti che ha diametro costante ed anche la pressione si assume essere costante.
- Uscito dall'area glottale il getto d'aria subisce un'espansione ed aderisce alle pareti del condotto; generando una pressure recovery descritta dalle relazioni dovute a [Ishizaka and Matsudaira, 1972].

Nella regione in cui si possono applicare le regole del flusso bernoulliano, ovvero al di sotto del punto di minima area intraglottale, la pressione presente ad una certa sezione può essere calcolata con la seguente formula:

$$P(a) = p_s - \frac{1}{2} \rho u^2 \left(\frac{1}{a^2} - \frac{1}{a_m^2} \right) \quad (2.6)$$

dove a è l'area della sezione di interesse, u è il flusso e ρ è la densità dell'aria.

Nel punto di area minima a_m il sistema transita dal regime bernoulliano a quello a "getto"; applicando la 2.6 come condizione al contorno si ha

$$P_m = p_s - \frac{1}{2} \rho u^2 \left(\frac{1}{a_m^2} - \frac{1}{a_s^2} \right) \quad (2.7)$$

La pressione nel punto di area minima, P_m , deve essere identica a quella che si ricava dall'equazione che descrive la pressure recovery così come ricavata nel lavoro di [Ishizaka and Matsudaira, 1972]:

$$P_m = P_i - \frac{1}{2} \rho K_e u^2 \left(\frac{1}{a_m^2} \right) \quad (2.8)$$

dove P_i è la pressione di ingresso al tratto vocale e k_e è un coefficiente di pressione definito come:

$$K_e = \frac{2a_m}{a_i} \left(1 - \frac{a_m}{a_i} \right)$$

Dopo aver uguagliato la 2.7 e la 2.8, risolvendo per u^2 e andando a sostituire nella 2.6 si ricavano le seguenti due equazioni:

$$P = p_s - (p_s - P_i) \frac{a^{-2} - a_s^{-2}}{a_m^{-2}(1 - k_e) - a_s^{-2}} \quad (2.9)$$

valida, al di sotto della glottide, dove si può considerare un regime bernoulliano,

$$P = P_i - (p_s - P_i) \frac{k_e - a_m^{-2}}{a_m^{-2}(1 - k_e) - a_s^{-2}} \quad (2.10)$$

valida, al di sopra della glottide, dove si può considerare un regime a "getto".

Esse permettono di calcolare, data l'area di sezione a , la pressione P . Supponendo piccolo il valore dell'area della sezione (a) si possono riscrivere le due equazioni (2.9) e (2.10) ed ottenere le seguenti equazioni:

$$P = p_s - (p_s - P_i) \left(\frac{a_m}{a} \right)^2 \quad \text{regime bernoulliano}$$

$$P = P_i \quad \text{regime a "getto"}$$

Quando una delle masse m_i e la sua controparte simmetrica collidono si pone:

$P=P_s$ sotto la collisione

$P=0$ nella collisione

$P=P_i$ sopra la collisione

2.3 I tratti vocale e nasale

In questo lavoro di ricerca si è interessati a comprendere le caratteristiche del flusso glottale attraverso la determinazione della pressione acustica che, misurata alla bocca di chi parla, costituisce il risultato finale della fonazione. Per una completa comprensione del fenomeno, le simulazioni del tratto orale e di quello nasale devono essere piuttosto accurate perché le conformazioni delle cavità che raccolgono il flusso glottale possono modificarne alcune caratteristiche.

Gli effetti del tratto vocale sul flusso dell'impulso glottale sono ampiamente discussi nell'attività di ricerca di [Childers and Wong, 1994]. Tali effetti non sono molto importanti ai fini dell'intelligibilità del parlato sintetico, ma si ritiene siano fondamentali per rendere la sintesi il più naturale possibile ovvero non possono essere trascurati se si è interessati al riconoscimento vocale. Vengono descritti brevemente gli effetti del tratto vocale superiore sul flusso glottale. [Marchetto E.,2007].

Skewing: L'interazione della sorgente con il tratto vocale determina uno spostamento della forma d'onda del flusso rispetto all'andamento dell'area della glottide; si può osservare che il flusso raggiunge il suo massimo con un certo ritardo rispetto al massimo dell'area intraglottale. L'andamento del flusso nel tempo subisce una variazione chiamata skewing.

Ripples: Durante la fase di apertura del ciclo di fonazione si possono osservare ripple (ondulazioni) nel segnale del flusso. Essi sono attribuiti all'energia della prima formante che la glottide dissipa durante la fase di apertura (glottal damping).

Sovrapposizione: Quando non si verifica nessun ripple l'energia dissipata dalla glottide viene trasportata al periodo successivo; questo può modificare la derivata del flusso nell'istante di chiusura, che cambia le caratteristiche spettrali dell'impulso.

Costrizioni: La produzione di suoni fricativi può creare degli impedimenti nella zona sopraglottale, che provocano una variazione degli andamenti della caduta di pressione intraglottale e del movimento delle corde vocali; l'effetto che si osserva è un allungamento della fase aperta nella produzione dell'impulso.

Molti scienziati mostrano come la presenza di un tratto vocale superiore incida sul modello fisico e matematico del flusso glottale facilitando l'instaurarsi della necessaria oscillazione [Ishizaka and Flanagan, 1972; Maratea, 2003-2004; Titze and Story, 1997]. L'assenza nel modello delle cavità superiori rappresenta una condizione non verosimile che toglie naturalezza ai suoni riprodotti.

2.4 I limiti della teoria classica

Nel lavoro di ricerca si analizza il segnale vocale studiandone le caratteristiche "elementari" delle vocali italiane. L'analisi del segnale vocale, effettuata fino ad oggi, parte dal presupposto che esso sia armonico e lineare, cioè ben approssimabile con una serie finita di funzioni armoniche e tale che possa essere descritto da equazioni differenziali lineari.

Presupporre che le equazioni che descrivono il segnale siano lineari equivale ad affermare che le ampiezze e le frequenze del segnale non variano nel tempo. Ciò, non è vero neanche per gli strumenti musicali autosostenuti (es. canne d'organo) [De Lauro et al., 2007], ma è decisamente falso per la voce umana, la cui intonazione non è fissa nel tempo. L'apparato produttore della voce, oltre ad essere molto complesso e ancora non del tutto capito, è difficilmente "gestibile" con precisione e soggetto a continue e molteplici variazioni determinate, da più fattori, primo fra tutti quello emozionale, che lo rendono poco controllabile.

Inoltre le vocali sono state sempre trattate come periodiche, quando in realtà sono segnali non completamente periodici ed contaminati dal rumore; un segnale sonoro considerato "*elementare*", come per esempio una vocale, non solo non è periodico, ma possiede una parte relativa all'attacco (*transitorio di attacco*), una dove assume un andamento *quasi* periodico (*quasi-stazionarietà*) e un'altra di decadimento (*transitorio di decadimento*).

Attualmente la descrizione delle peculiarità del suono è fornita sia in termini di aspetti legati alla sua percezione uditiva, sia in termini di osservazioni nei domini del tempo e della frequenza; quindi usando come strumento di valutazione ed analisi o quello fisiologico, comunemente detto "*orecchio*", coadiuvato dalle capacità empiriche di esperti del settore (cantanti, insegnanti di canto, linguisti, ecc.), o quello prettamente tecnico che utilizza strumenti forniti dalla matematica, fisica, ingegneria. Purtroppo, però, non esiste al momento la possibilità di collegare direttamente le sensazioni uditive con precise osservazioni fisiche. Da quanto ci risulta, inoltre, la voce non è mai stata analizzata con i metodi dei *Sistemi dinamici*, i quali studiano il comportamento di fenomeni a "metà strada" tra sistemi lineari con orbite periodiche e rumore, dove riteniamo, proprio, di collocare il segnale vocale per le sue caratteristiche temporali oggettive; precisiamo d'altronde che, attualmente, si ignora ed è ancora oggetto di ricerca, il sistema dinamico che produce il segnale vocale.

Il presente lavoro nasce, quindi, dalla constatazione della carenza di ricerche di tipo sperimentale sulla voce come segnale vocale *non totalmente armonico, non completamente periodico e non lineare*, ed effettuate da persone sufficientemente competenti in entrambi i settori scientifico e musicale, con l'intento di ricavare informazioni ulteriori a quelle possibili tramite le tecniche tradizionali di elaborazione del segnale. L'analisi del segnale vocale è stata effettuata su vari operatori e si è notato che la frequenza del suono varia, ma rimane invariata la riconoscibilità delle vocali. Da qui la ricerca di un modello matematico interpretativo che descriva il segnale vocale e che colleghi le sensazioni uditive con precise osservazioni fisiche sfruttando gli strumenti matematici propri dei *Sistemi dinamici*.

Capitolo 3

RISULTATI SPERIMENTALI

Le vibrazioni meccaniche associate alla trachea ed alla laringe quando vengono emessi suoni vocalici sono state studiate utilizzando due Vibrometri laser Doppler, Polytec OFV-5000 e PDV-100 con acquisizione nella banda di frequenza fino a 20KHz, fondo scala di $\pm 4V$ e con sensibilità pari a 5 mm/s/V. I segnali laser registrati sono stati confrontati con i segnali acustici. Si è osservato una propagazione verso l'alto del segnale associato alle vibrazioni meccaniche della trachea e si è messo in evidenza un ritardo tra le tracce dei segnali registrati, dimostrando che le oscillazioni meccaniche nella trachea iniziano prima che nella laringe e prima delle oscillazioni acustiche. Queste oscillazioni tracheali sono auto-oscillazioni nel tempo e sono associate alla produzione della frequenza di pitch, indicando un'ulteriore instabilità idrodinamica a livello tracheale. Questo porta a nuove intuizioni nel meccanismo che controlla il pitch nel discorso.

Nella attività di ricerca proposta, si è sperimentalmente osservato, allo stesso tempo, il campo acustico e il campo meccanico indotto lungo il torace dal sistema polmone-trachea. I segnali ottenuti sono stati analizzati e sono state stabilite alcune conclusioni sul ruolo che un flusso d'aria non costante all'interno del sistema polmoni-trachea può avere nei modelli di produzione del suono. Essendo interessati alla sintesi del parlato, in particolare a quella dei suoni vocalici, cercando di riprodurre il più fedelmente possibile la voce dell'operatore che parla, non si può non prendere in considerazione la prosodia. La prosodia è l'insieme delle caratteristiche sonore di una lingua che interessano un dominio più ampio del singolo fono. Le principali caratteristiche prosodiche sono il tipo e la posizione degli accenti, la melodia, l'intonazione, la velocità di eloquio, il ritmo, la durata, le pause e la qualità della voce [Selting 1995, Soriano 2006]. Il rapporto tra ciò che si può misurare a livello acustico/fisico e le caratteristiche prosodiche effettivamente percepite dall'apparato uditivo umano è molto complesso ed il principale parametro acustico che corrisponde, alla percezione dell'intonazione è la frequenza di pitch o fondamentale (f_0). A tal proposito particolare attenzione è stata posta sulle tecniche di estrazione della frequenza di pitch e delle formanti dei suoni vocalici. Sono state utilizzate tecniche standard e tecniche innovative, ad es. la frequenza fondamentale è stata valutata dall'analisi dei segnali associati alle vibrazioni meccaniche del tratto trachea/ laringe con tecniche nel dominio della frequenza (es. LPC) e tecniche nel dominio del tempo (es. analisi degli intertempi).

L'effetto di risonanza del tratto vocale sul segnale glottico è quello di aumentare o diminuire l'ampiezza di alcune armoniche; se l'involuppo spettrale del segnale glottico si presenta monotonamente discendente senza minimi o massimi apprezzabili, il segnale verbale nel suo involuppo evidenzia picchi o zone a massima energia, denominate formanti (F_1 , F_2 , F_3 ecc...). Ogni formante è caratterizzata dalla sua frequenza, ampiezza e larghezza di banda, quest'ultima definita come intervallo frequenziale contenente la frequenza della formante e delimitata agli estremi dalle due frequenze che hanno una attenuazione di 3 dB rispetto all'ampiezza della formante. La

larghezza di banda è tanto maggiore quanto più alta è la frequenza della formante. Lo studio delle loro caratteristiche (frequenza, ampiezza e banda) fornisce informazioni su come il soggetto utilizza le proprie cavità sovraglottiche.

3.1 Data set

Tutti gli esperimenti sono stati condotti in camera anecoica (Fig.1) che è un laboratorio strutturato per ridurre il più possibile la riflessione di segnali acustici sulla parete.

Gli esperimenti sono stati condotti su uomini e donne e in situazioni diverse: ad esempio voce modale e sussurrata. Sono state effettuate misure simultanee delle oscillazioni meccaniche (misure di velocità) in due punti del torace lungo il tratto trachea/laringe mediante due Vibrometri laser Doppler, posti ad una distanza l'uno dall'altro di 0.11m, (vedi Fig.1) ed è stata registrata l'emissione acustica con dei microfoni (si riporterà solo la traccia relativa al canale CH3 per comodità). I segnali registrati delle oscillazioni meccaniche dovute all'emissione di suoni vocalici sono molto significativi, infatti ascoltandoli e confrontandoli con quelli acustici, è facile apprezzare una traccia di riconoscimento dell'operatore che li ha pronunciati.



Fig.1 Camera anecoica e configurazione dell'apparato sperimentale. Le oscillazioni meccaniche sono misurate simultaneamente in due punti del torace utilizzando due Vibrometri laser Doppler, (Polytec OFV-5000 e PDV-100) con acquisizione nella banda di frequenza fino a 20KHz, fondo scala di $\pm 4V$ e con sensibilità pari a 5 mm/s/V. Spostando i laser di 0,01 mm è stato investigato il tratto trachea-laringe. Il campo acustico è registrato contemporaneamente alle oscillazioni meccaniche utilizzando due microfoni.

3.2 Tecniche di estrazione del pitch

Per estrarre le frequenze associate al pitch ed alle formanti dei suoni vocalici dai segnali microfonici e dalle oscillazioni meccaniche associate ai segnali laser, sono state utilizzate e confrontate diverse metodologie: 1) spettrogramma e densità spettrale di potenza; 2) Codifica predittiva lineare (LPC) 3) Metodo di mutua informazione 4) Analisi degli intertempi [Falanga, 1999; Falanga, 2003].

In particolare, i segnali vocalici sono stati filtrati utilizzando un filtro passa banda con una finestra tra gli 80-2500 Hz per le seguenti ragioni: a) si è posta l'attenzione sulla gamma delle frequenze che interessano il discorso, b) sono state eliminate le frequenze che caratterizzano il respiro e il ciclo cardiaco nei segnali laser che appartengono all'intervallo 1-30 Hz; c) è stato eliminato il contributo di qualsiasi fonte dovuta alla corrente di rete.

3.2.1 Metodo 1: spettrogramma e densità spettrale di potenza

Si riportano i segnali relativi alle oscillazioni acustiche (CH3) e a quelle meccaniche (CH1, CH4) di una donna di 37 anni per le cinque vocali italiane, emesse nella sequenza [a], [æ], [i], [o], [u], ed i relativi spettrogrammi (Fig.2).

Nello spettrogramma del segnale acustico sono evidenti le formanti dei suoni in esame, in accordo con lo spettrogramma riportato in Ferrero et al., 1978. Lo spettrogramma del segnale laser relativo alle oscillazioni meccaniche della trachea è caratterizzato da una frequenza il cui periodo corrisponde alla frequenza di pitch. Infatti, sovrapponendo il segnale meccanico registrato relativo alla trachea (CH1) e il corrispondente segnale acustico (CH3), si osserva un'apparente modulazione prosodica dell'oscillazione meccanica rispetto a quella acustica (Fig. 3).

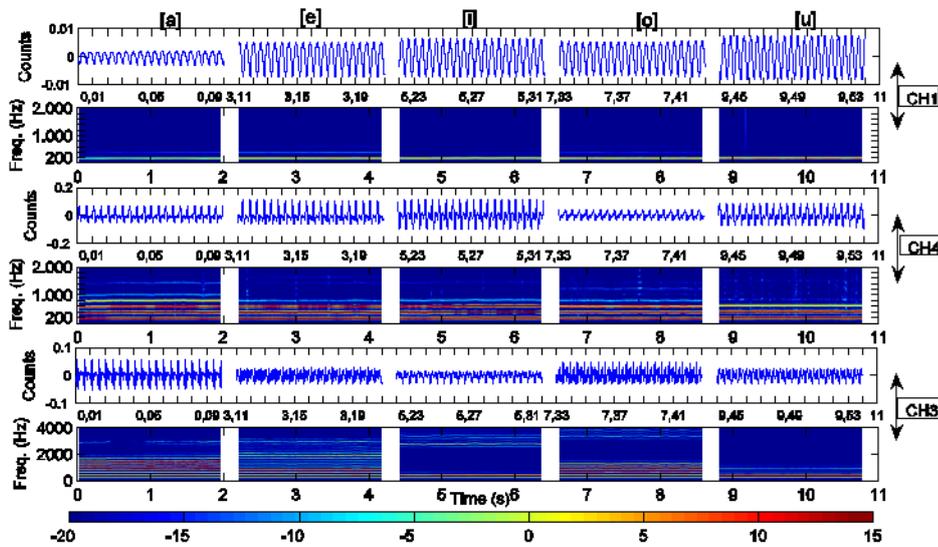


Fig.2 Forme d'onda associate alle vocali italiane pronunciate da una donna di 37 anni acquisite con due laser (CH1, CH4) e un microfono (CH3) ed i relativi spettrogrammi.

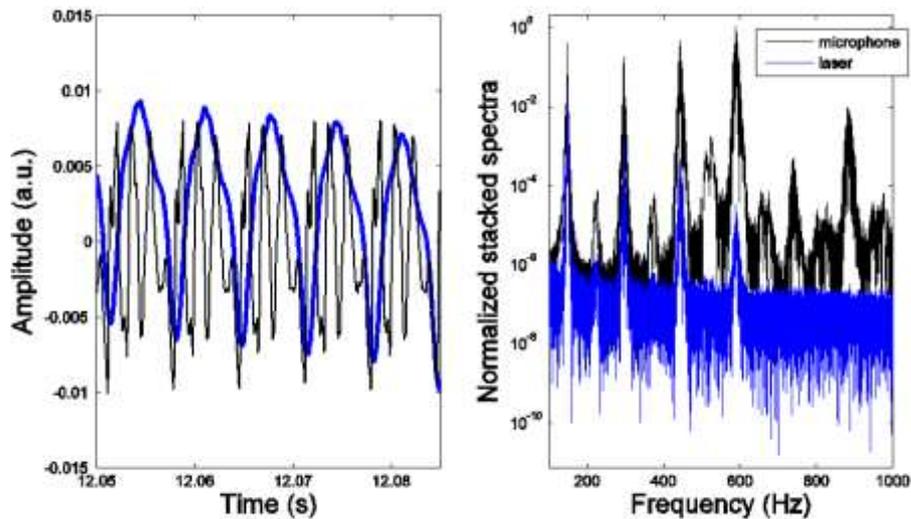


Fig.3. Sovrapposizione del segnale acustico registrato con il microfono (CH3) e segnale laser tracheale (CH1) in unità normalizzate relativamente al suono della vocale [e] emessa da un maschio e il relativo spettro di potenza. Si può notare che CH3 modula CH1 e un ritardo tra le tracce può essere rivelato. Questa caratteristica sarà approfondita in seguito.

3.2.2 Metodo 2: codifica predittiva lineare (LPC)

La metodica di analisi LPC (Linear Predictive Coding) consente la determinazione dell'involuppo spettrale del segnale verbale, e quindi fornisce precise informazioni sulle formanti. I picchi formantici, o la frequenza della formante, non necessariamente coincidono con le componenti armoniche.

Tenendo presente il modello sorgente/filtro della produzione del parlato, partendo dalla teoria dei sistemi e ricordando che la trasformata zeta è l'estensione nel caso discreto della trasformata di Laplace, si ha che, nel dominio delle frequenze, possiamo scrivere:

$$O(z)=I(z)\cdot H(z)$$

dove $O(z)$ è la rappresentazione in trasformata Z del segnale vocale nel dominio della frequenza, $I(z)$ equivalentemente, è la trasformata Z del segnale sorgente, $H(z)$ è la risposta in frequenza del filtro e con $H(z) = \frac{1}{1-\sum_{i=1}^p z^{-i}a_i}$.

La codifica predittiva lineare sfrutta la caratteristica del filtro di sintesi (modulo quadrato della funzione di trasferimento) che presenta un andamento simile all'involuppo del segnale di ingresso. Un codificatore LPC può essere visto come un modello dell'apparato di produzione sonora: l'eccitazione costituisce il segnale sorgente e le modifiche apportate a questo segnale dal tratto vocale superiore sono prodotte dal filtro di sintesi. La qualità del segnale sintetico prodotto con un codificatore LPC è molto bassa: il parlato è innaturale ma comprensibile, ha un timbro metallico e

le caratteristiche dell'operatore risultano alterate. La bassa qualità è dovuta in buona parte alla semplicità del modello.

La tecnica di analisi si basa sull'autocorrelazione e sulla predizione lineare dei campioni, ovvero ogni campione è funzione (lineare) di quello che lo precede. In pratica un campione è uguale alla somma dei campioni precedenti moltiplicati per determinati coefficienti (o pesi), meno un certo errore e . Il numero di coefficienti o pesi costituisce l'ordine dell'LPC.

Supponiamo che ogni campione $x[n]$ di un segnale vocale possa essere predetto utilizzando una combinazione lineare di un numero p di campioni precedenti; $\bar{x}[n]$ che è il segnale predetto e costituirà una stima approssimata del segnale originale:

$$\bar{x}[n] = a_1 x[n-1] + a_2 x[n-2] + \dots + a_p x[n-p]$$

Gli a_i sono detti coefficienti di predizione. La differenza fra il segnale e la sua predizione è detta errore o residuo di predizione, e rappresenta la parte di segnale che non può essere predetta dal suo passato. In generale, detto $e[n]$ l'errore che si compie in questo processo di approssimazione, si avrà che:

$$x[n] = \bar{x}[n] + e[n]$$

Ovvero:

$$x[n] = \sum_{i=1}^p a_i x[n-i] + e[n]$$

I coefficienti di predizione vengono calcolati in modo che la sequenza predetta $\bar{x}[n]$ sia il più simile possibile, con il metodo dei minimi quadrati, alla sequenza $x[n]$, cioè in modo che l'energia del segnale differenza sia più piccola possibile; analiticamente occorre minimizzare la quantità E .

Il valore dei coefficienti a_i può essere calcolato considerando gli scarti quadratici fra i valori di predizione e i valori effettivi:

$$E = \sum_{n=-\infty}^{+\infty} e[n]^2 = \sum_{n=-\infty}^{+\infty} (\bar{x}[n] - x[n])^2$$

Minimizzando E si ricava un sistema di P equazioni lineari (dette equazioni di Yule-Walker) nelle incognite a_i , che permettono quindi il calcolo dei coefficienti di predizione che rendono minima l'energia del residuo.

E' chiaro che, minore è l'energia del segnale residuo, maggiore è l'informazione che è stata rimossa dal segnale e concentrata nei coefficienti a_i . Dalla conoscenza dei coefficienti di predizione e del residuo di predizione è possibile ricostruire il segnale di partenza $x[n]$.

Il metodo che è stato descritto per il calcolo degli a_i è noto col nome di metodo della "Autocorrelazione" ma esistono diversi altri metodi per il calcolo dei parametri LPC (ad esempio il metodo della "Covarianza").

I coefficienti a_i , dunque, caratterizzano il filtro lineare. Il calcolo di questi coefficienti è un calcolo tempo-variabile, di conseguenza viene effettuato su un segnale diviso in finestre e ripetuto in ogni finestra.

La finestra temporale utilizzata per l'analisi dei segnali è stata di 2 sec, è stata scelta in modo che il segnale possa essere considerato stazionario al suo interno. Il calcolo di LPC basato su finestre (frame based) fornisce un metodo di codifica del segnale vocale, cioè definisce una funzione di trasformazione che accetta in ingresso la sequenza dei campioni digitali del segnale e fornisce in uscita una molto meno onerosa tabella di coefficienti.

L'analisi LPC sui segnali sperimentali è stata effettuata con un ordine di predizione pari a 12. Più avanti sarà descritto il motivo per cui è stata fatta questa scelta.

La metodica tuttavia per poter funzionare correttamente presume da un lato che la sorgente glottica sia completamente indipendente (acusticamente) dalla trachea e dalle regioni sovraglottiche, e dall'altro che nel segnale non siano presenti antirisonanze rappresentate nell'involuppo come anti-formanti. La metodica LPC è quindi poco precisa nell'analisi di suoni nasali, di molte consonanti o di voci con parziale accoppiamento acustico fra glottide e trachea come quelle velate o soffiate. Quando ci si appresta all'analisi di predizione lineare in genere è necessario impostare l'ordine dell'LPC che è uguale al doppio del numero di formanti attese più 2 (questa correzione definisce la pendenza media dell'involuppo spettrale). In letteratura è noto che le prime tre formanti sono le più significative, al fine di non perdere utili informazioni sul segnale analizzato si è scelto un ordine di predizione pari a 12 ($5 \times 2 + 2 = 12$) per poter valutare le prime 5 formanti. Con questa tecnica di decodificazione sono stati analizzati diversi segnali vocali. È riportato per un operatore un esempio di confronto tra lo spettro LPC e la densità dello spettro di potenza (P.S.D) per ciascuna vocale. L'analisi LPC è stata effettuata sul segnale acustico e sui segnali laser registrati lungo il tratto tracheale, permettendo di estrarre la frequenza di pitch. Generalmente le donne hanno tratto vocale più corto rispetto a quello degli uomini e dimensioni più piccole delle cavità, perciò le frequenze associate alle formanti di una voce femminile (Tabella a) hanno un valore più grande rispetto a quelle associate ad una voce maschile.

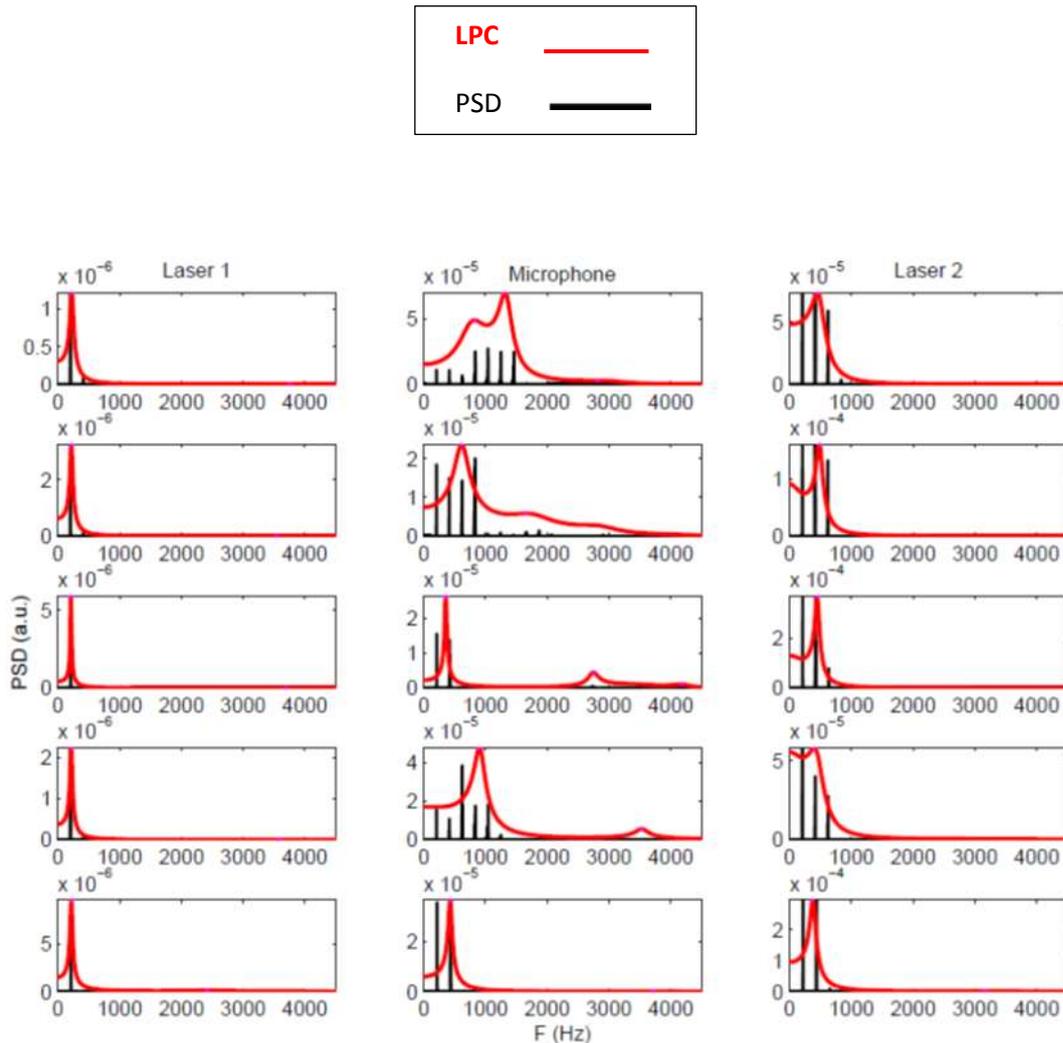


Fig.4: Confronto tra uno spettro LPC e uno spettro di potenza per un operatore che pronuncia le cinque vocali dell'alfabeto italiano. Estrazione della frequenza di pitch tramite LPC dal segnale laser.

Nella tabella sono riportate i valori sperimentali delle formanti e della frequenza di pitch dei suoni vocalici prodotti da una donna di 37 anni e da un uomo di 67 anni. I valori trovati sperimentalmente sono in accordo con i valori riportati in letteratura. E' evidente che la frequenza di pitch e il valore delle formanti nella donna sono tendenzialmente più alti di quelli maschili. Per comprendere la natura della diversità dei suoni emessi dagli esseri umani bisogna considerare due cause generali di variazione: il fattore fonetico ed il fattore organico. La variazione di tipo fonetico dipende dalle differenze nel modo in cui un individuo usa il suo apparato vocale, ad esempio dipende dalla posizione della lingua o dallo stile usato. La variazione di tipo organico riguarda differenze individuali nella forma, nella dimensione, nelle proprietà istologiche e meccaniche degli organi vocali. Le due principali fonti di variazione di tipo organico delle vocali sono connesse al sesso e all'età del parlante, e vengono meglio descritte da misure fisiche quali la frequenza fondamentale (f_0), determinata primariamente da dimensione e massa della laringe, e le frequenze formantiche (F_n , dove $n=1-3$), dove la stessa costrizione genera diversi valori a seconda, principalmente, della

lunghezza e della forma del condotto vocale. Ricordiamo che le differenze tra vocali pronunciate da uomini e donne sono dovute non solo a differenze anatomiche relative alla dimensione della laringe, alla massa delle corde vocali e alle dimensioni del condotto vocale ma anche alla fisiologia della vibrazione delle corde vocali, che nelle donne e nei bambini presuppone un'interazione sorgente-filtro abbastanza consistente, che è responsabile della qualità “soffiata” (breathy).

[a]				[e]			
H _z	Laser 1	Voiced	Laser 2	H _z	Laser 1	Voiced	Laser 2
F0	225			F0	225		
F1		825	450	F1		612	487
F2		1312		F2		1650	
F3		2812		F3		2800	
F4				F4			

[i]				[o]			
H _z	Laser 1	Voiced	Laser 2	H _z	Laser 1	Voiced	Laser 2
F0	212			F0	225		
F1		350	450	F1			400
F2		2787		F2		900	
F3				F3			
F4		4200		F4		3500	

[u]			
H _z	Laser 1	Voiced	Laser 2
F0	225		
F1		437	375
F2			
F3			
F4		3637	

female)

[a]				[e]			
H _z	Laser 1	Voiced	Laser 2	H _z	Laser 1	Voiced	Laser 2
F0	125		106	F0	125		
F1		443	518	F1		500	525
F2	1031	1056		F2		1606	
F3		2168		F3		2356	
F4		3493		F4		3800	

[i]				[o]			
H _z	Laser 1	Voiced	Laser 2	H _z	Laser 1	Voiced	Laser 2
F0	150			F0	137		
F1		356	456	F1		493	531
F2		2000		F2	762	756	
F3		3281		F3		2306	
F4		4037		F4		3425	

[u]			
H _z	Laser 1	Voiced	Laser 2
F0	150		162
F1		325	525
F2		537	
F3		2368	
F4		3962	

male)

Tabella a: Estrazione della frequenza di pitch e delle formanti delle vocali italiane di una donna di 37 anni ed un uomo di 67 anni.

3.2.3 Metodo 3: mutua informazione

Ulteriori informazioni sulla frequenza di pitch può essere ottenuta con metodi basati sulla correlazione. Tenendo conto del comportamento non lineare della fisica alla base del fenomeno in esame, adottiamo l'Average Mutual Information (AMI) per stimare la frequenza di pitch. In particolare, dal confronto tra i segnali acustici e quelli laser, si ha una chiara indicazione della frequenza di pitch e di tutti gli altri contributi in frequenza relativi alla vocale prodotta. In Fig.6, si traccia sia l'AMI: il secondo picco massimo fornisce informazioni sul pitch e si può notare come semplice è la forma del laser rispetto al microfono.

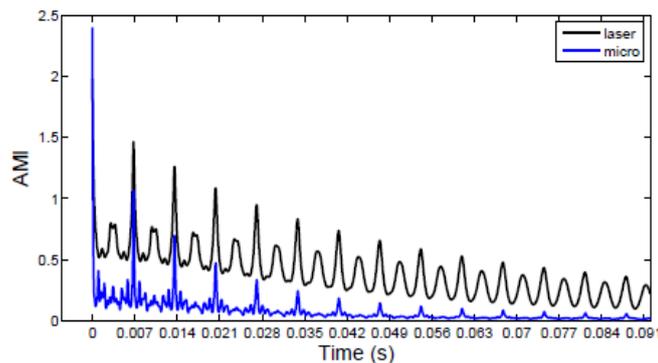


Fig.6: Confronto della funzione AMI relativa ad un segnale laser ed acustico.

3.2.4 Metodo 4: intertempi

Generalmente, il pitch viene estratto nel dominio del tempo come distanza tra due massimi consecutivi del segnale acustico. Per valutare questa distanza temporale è stata utilizzata l'analisi statistica degli intertempi applicata sia ai segnali acustici e sia ai segnali laser. La statistica dei segnali evidenzia che il valor medio della distribuzione corrisponde al valore della frequenza di pitch e si osserva che la distribuzione degli intertempi è più piccata per il segnale laser associato alle oscillazioni del tratto tracheale evidenziando che tale segnale è un segnale di base. La figura (Fig.7) mostra l'analisi effettuata sui segnali registrati di un operatore uomo di 67 anni che emette un suono vocalico e.

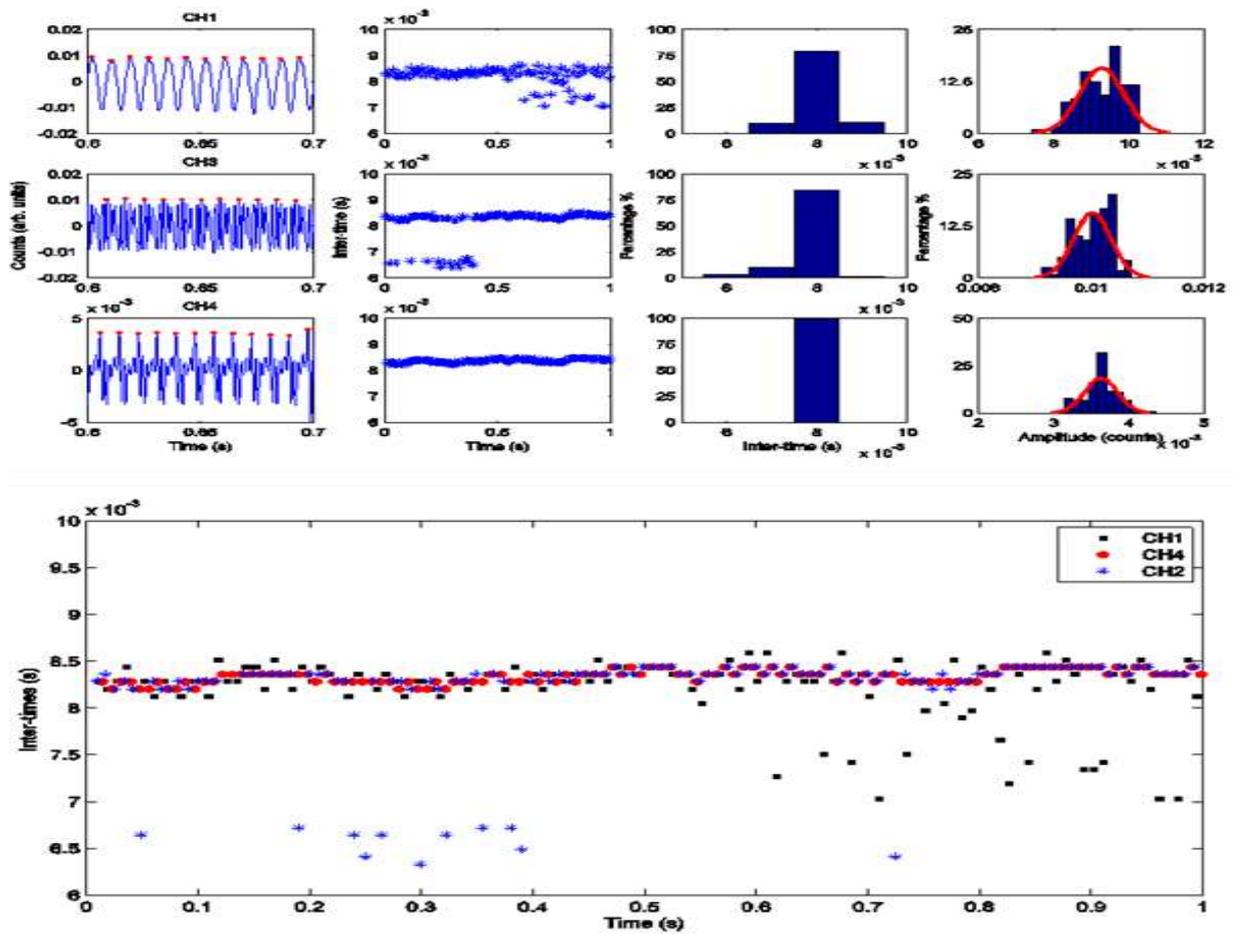


Fig.7 Analisi degli intertempi.

E' stato valutato anche il coefficiente di variabilità C_v , per i segnali riportati in figura, i cui valori ottenuti sono 0,02; 0,4; 0,1 rispettivamente per i segnali registrati sui canali CH1, CH3 e CH4. Il coefficiente di variabilità è un parametro che ci permette di valutare la gaussianità della distribuzione degli intertempi ed è definito:

$$C_v = \frac{\sigma \Delta t}{\overline{\Delta t}}$$

dove $\sigma \Delta t$ è la deviazione standard e $\overline{\Delta t}$ è il valore medio degli intertempi. Sapendo che $C_v = 1$ per un processo di Poisson, $C_v > 1$ per un processo clusterizzato e $C_v = 0$ per un processo periodico e $C_v \rightarrow \infty$ per distribuzioni uniformi.

I tre CV calcolati hanno un valore prossimo a zero. Il coefficiente di variabilità corrispondente al segnale di trachea è un ordine di grandezza inferiore al valore relativo al segnale acustico. Questo indica una maggiore complessità del segnale della glottide rispetto al segnale della trachea, che può essere giustamente considerato come un segnale di base.

Per consolidare questo risultato sono stati calcolati i coefficienti di correlazione tra i segnali del campo vibrazionale associati alle cinque vocali italiane registrati su quattro punti della trachea partendo dal basso e salendo verso la laringe. I segnali risultano ben correlati tra loro e si nota che spostandoci verso la laringe la correlazione tra i segnali diminuisce, questo conferma che il segnale di trachea è un segnale di base caratteristico di ogni singolo individuo. Si riportano i coefficienti di correlazione ottenuti per una donna di 37 anni. Il punto 1 è il punto di misura più basso sulla trachea mentre il punto 4 è il punto più vicino alla glottide.

Punto 1

	a	e	i	o	u
a	1	0.99	0.96	0.98	0.86
e	0.99	1	0.97	0.98	0.88
i	0.96	0.97	1	0.99	0.96
o	0.99	0.98	0.98	1	0.95
u	0.86	0.87	0.95	0.95	1

Punto 2

	a	e	i	o	u
a	1	0.99	0.98	0.99	0.96
e	0.99	1	0.98	0.98	0.93
i	0.98	0.98	1	0.98	0.95
o	0.99	0.98	0.98	1	0.97
u	0.96	0.93	0.95	0.97	1

Punto 3

	a	e	i	o	u
a	1	0.99	0.95	0.99	0.80
e	0.99	1	0.96	0.99	0.79
i	0.96	0.96	1	0.96	0.79
o	0.99	0.99	0.96	1	0.80
u	0.80	0.80	0.79	0.79	1

Punto 4

	a	e	i	o	u
a	1	0.98	0.88	0.80	0.67
e	0.98	1	0.89	0.79	0.61
i	0.88	0.89	1	0.76	0.63
o	0.80	0.79	0.76	1	0.64
u	0.66	0.61	0.63	0.64	1

3.3 Il ruolo della trachea nell'emissione dei suoni vocalici

Nella teoria filtro/sorgente e nel modello a due masse utilizzato in letteratura per studiare e modellare il segnale vocalico, come precedentemente visto nel capitolo 2, la trachea non interviene nella produzione dei suoni. Gli esperimenti effettuati, invece, hanno messo in evidenza che il segnale tracheale presenta già un'informazione sulla frequenza di pitch. Questo farebbe pensare che il suono non sia attribuibile alla sola differenza di pressione attraverso la glottide, ma che la trachea stessa è sorgente del suono vocalico. A sostegno di questa affermazione, del tutto innovativa, sono stati effettuati i seguenti esperimenti: 1) analisi e confronto del transiente d'attacco tra i segnali laser 2) confronto tra il suono emesso da una canna d'organo e il suono vocalico emesso da un uomo.

3.3.1 Il transiente d'attacco

In questa sezione viene analizzato il transiente d'attacco dei segnali laser (CH1, CH4), di un uomo ed una donna che emettono la vocale [e], per dedurre la sorgente segnale vocale. E' evidente un ritardo temporale tra i due segnali registrati ai punti di misura selezionati che rivela un flusso d'aria verso l'alto (cioè da CH1 a CH4). Il ritardo di tempo stimato tra le due misure mediante l'analisi di cross-correlazione è 0.035 s per una donna che corrisponde ad una velocità di circa 30 m/s ricordando che i due laser sono separati di 0.11 m, ed è $4 \cdot 10^{-4}$ s indicando una velocità di propagazione di circa 270 m/s per un uomo. Questi risultati sperimentali si possono leggere come dovuti al fatto che la trachea inizia ad oscillare in modo indipendente e prima dei segnali acustici fino a raggiungere il regime stazionario quando tutto l'apparato vocale globalmente auto-oscilla.

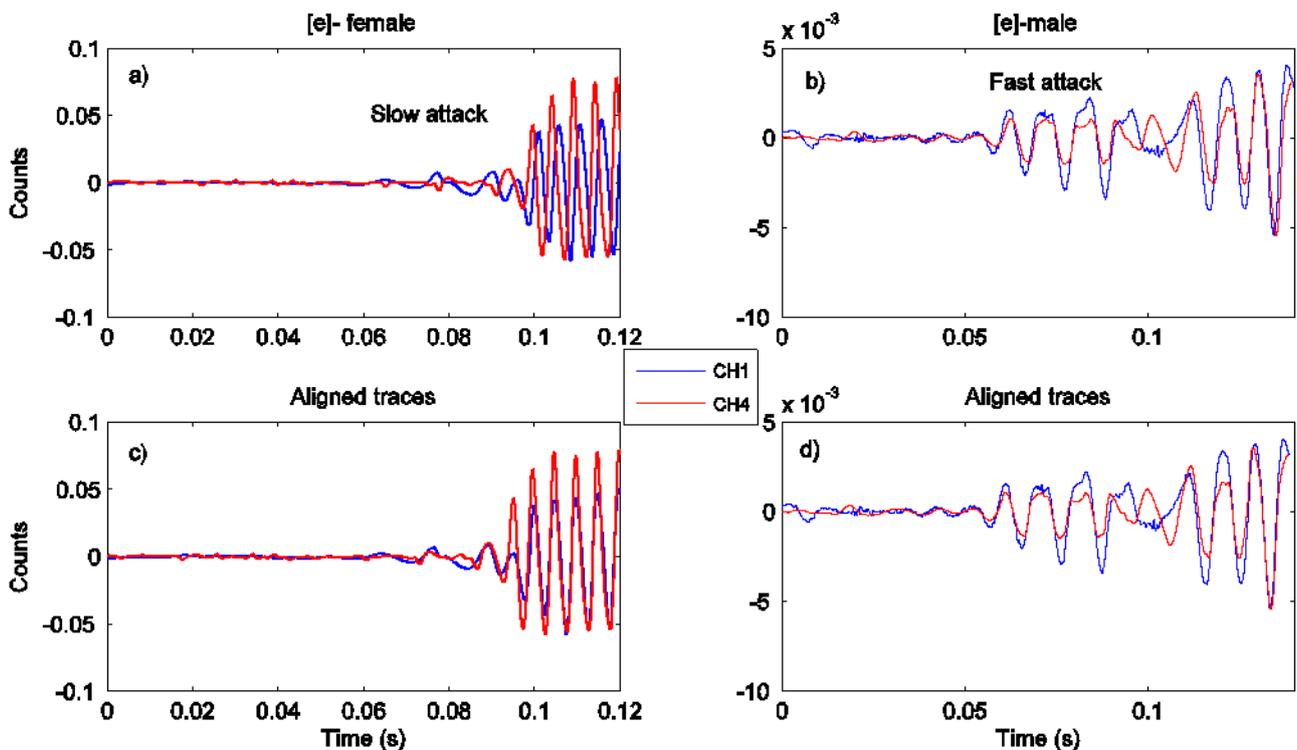


Fig.8 Il segnale d'attacco delle oscillazioni meccaniche ai due punti di misura per un uomo e una donna che emettono la vocale [e].

Questi attacchi ricordano l'andamento del transiente d'attacco in una canna d'organo ed è possibile parlare di attacco lento per le donne e attacco veloce per gli uomini.

In letteratura è noto che in una canna d'organo, all'aumentare della pressione con cui il flusso d'aria attraversa il piede si allunga nel tempo anche il transiente d'attacco (Fig.9), ossia la parte del suono prodotto dallo strumento prima che sia raggiunta la fase stazionaria. Si può distinguere un attacco lento che corrisponde a pressioni di insufflazione più basse e di un attacco veloce che corrisponde a pressioni maggiori [Verge et al., 1994].

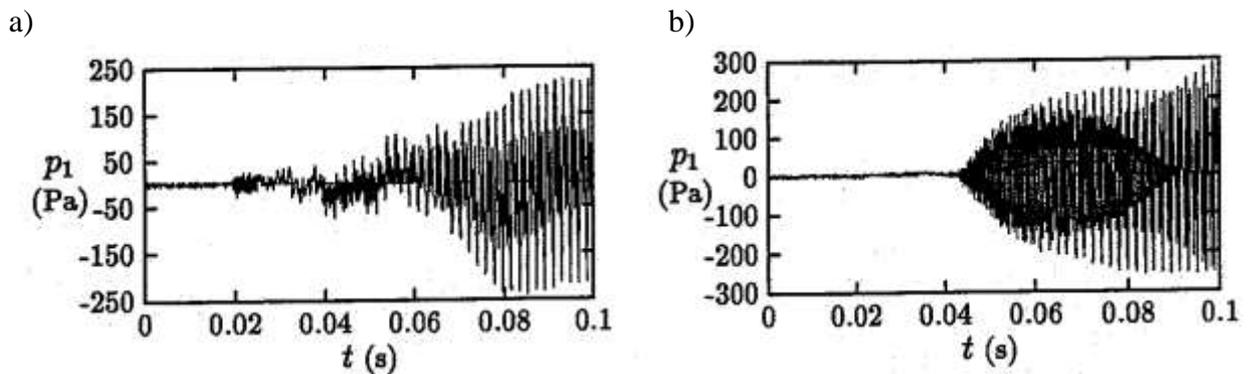


Fig.9 (a) Attacco veloce e (b) attacco lento in una canna d'organo.

L'analogia nasce dal fatto che un uomo ha una capacità polmonare maggiore rispetto a quella di una donna, ciò gli permette di inviare alla canna "trachea" un flusso d'aria a pressione maggiore rispetto a quello di insufflazione di una donna.

3.3.2 Il confronto con il suono emesso da una canna d'organo

Nella teoria classica la trachea permette semplicemente il passaggio dell'aria dai polmoni alla glottide senza svolgere nessun ruolo nella produzione dei suoni. Confrontando il suono emesso da una canna d'organo con quello emesso dall'apparato fonatorio umano, secondo la teoria sorgente/filtro, la trachea svolgerebbe lo stesso ruolo del piede nella canna. Il piede consente solo all'aria prodotta di fluire nel risonatore, ma la misura di vibrazioni acustica e meccanica in una canna d'organo che suona mostrano un comportamento diverso rispetto al comportamento sperimentalmente osservato nella voce. Nella canna quando un getto d'aria nel piede urta contro l'edge si genera un tono. Questo movimento di aria attorno al labbro superiore eccita la colonna d'aria che risuona all'interno del corpo della canna e induce il campo acustico esterno registrato dal microfono come prima arrivo. La vibrazione solo successivamente viene trasferita al piede della canna (CH1) quando nella fase stazionaria viene visualizzata una complessiva auto-oscillazione.

Se la trachea svolgesse la stessa funzione del piede nella canna, si dovrebbe notare che il segnale registrato su CH1 partirebbe dopo quello il segnale acustico, invece le analisi mostrano il contrario. Ciò indica che il ruolo del trachea nella produzione vocali così come nella formazione del pitch deve essere preso in considerazione e non deve essere trascurato.

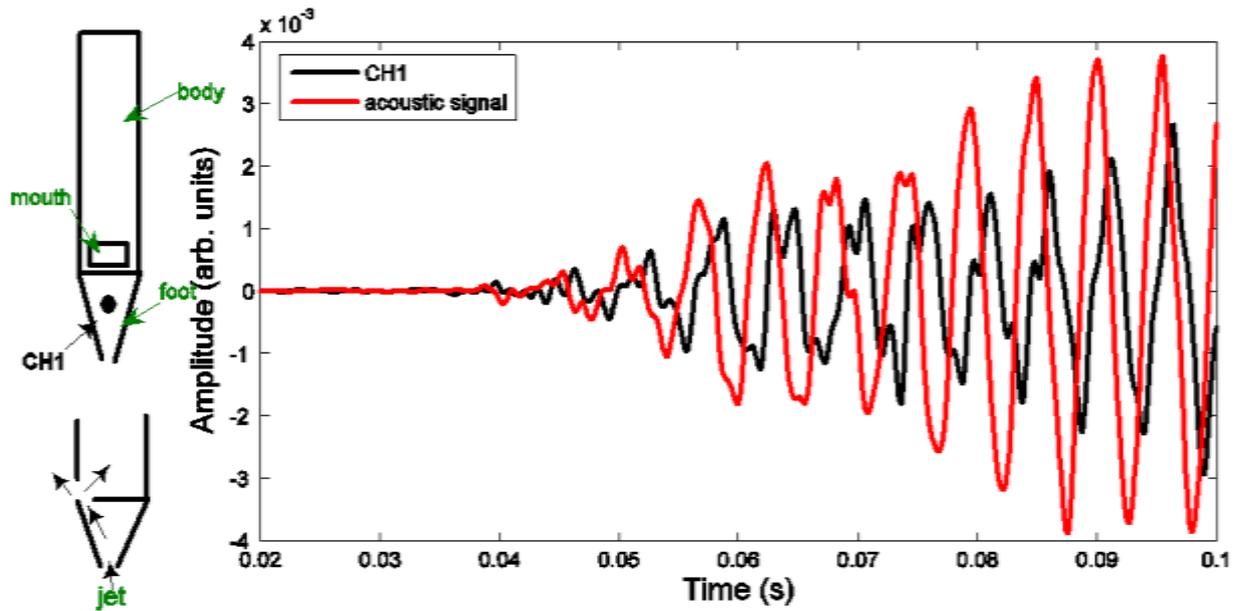


Fig.10 Confronto dell'attacco tra il segnale tracheale e quello acustico per valutare quale tra i due segnali parte prima.

3.4 Analisi delle componenti principali

Una dinamica ad alta dimensionalità è generalmente richiesta per descrivere i sistemi fisici che rappresentano un segnale vocale, mentre praticamente è preferibile riuscire a descriverli con una dinamica a bassa dimensionalità. L'Analisi delle Componenti Principali (PCA) o trasformazione di Karhunen-Loeve (KL) è una tecnica lineare utilizzata per ridurre la dimensione caratteristica del sistema in esame e si basa sul criterio del minimo errore quadratico medio.

L'approccio PCA è una tecnica utilizzata principalmente per la riduzione della dimensionalità dei sistemi ed opera essenzialmente come un metodo di rappresentazione di segnali, esprime cioè un dato segnale come somma di componenti ortonormali e scorrelate. Gli autovalori dello spazio delle fasi che conservano la più significativa quantità di informazioni sono conservati. Precedenti lavori nell'ambito dei segnali vocalici sulla PCA si possono trovare in letteratura [Mokthari et al., 2007].

La pre-analisi effettuata con la PCA è stata eseguita sui quattro segnali acustici registrati emessi da diversi operatori che emettevano suoni vocalici. Questa analisi ha permesso di stabilire, attraverso il contenuto informativo relativo agli autovalori della matrice di covarianza, anche che il numero delle componenti significative sono al più quattro. E' riportato un esempio di analisi PCA sul segnale acustico e sul segnale di oscillazione meccanica registrato su due punti lungo la trachea per una donna di 37 anni che emette la vocale O.

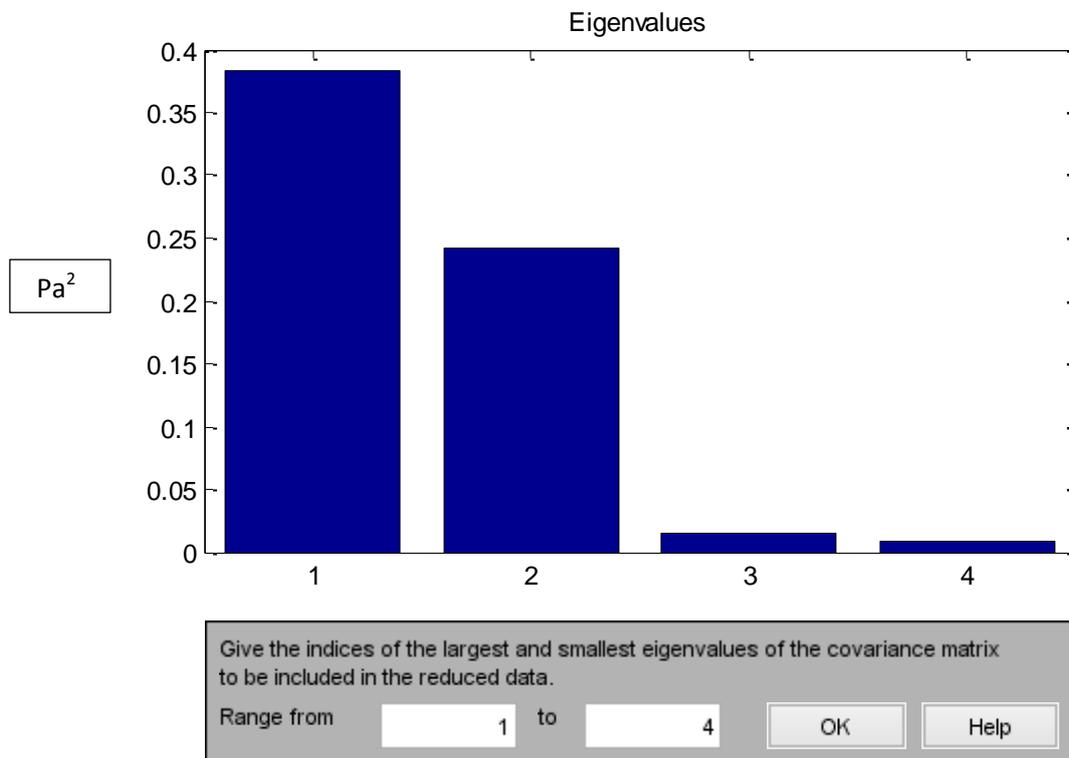


Fig. 11: Analisi del segnale acustico mediante tecnica PCA.

La pre-analisi, effettuata con la PCA è stata effettuata considerando la registrazione di quattro tracce di segnali acustici eseguite contemporaneamente a due a due. Essa rivela quattro autovalori di cui uno dei quali conserva il 64% dell'informazione relativa al sistema in esame.

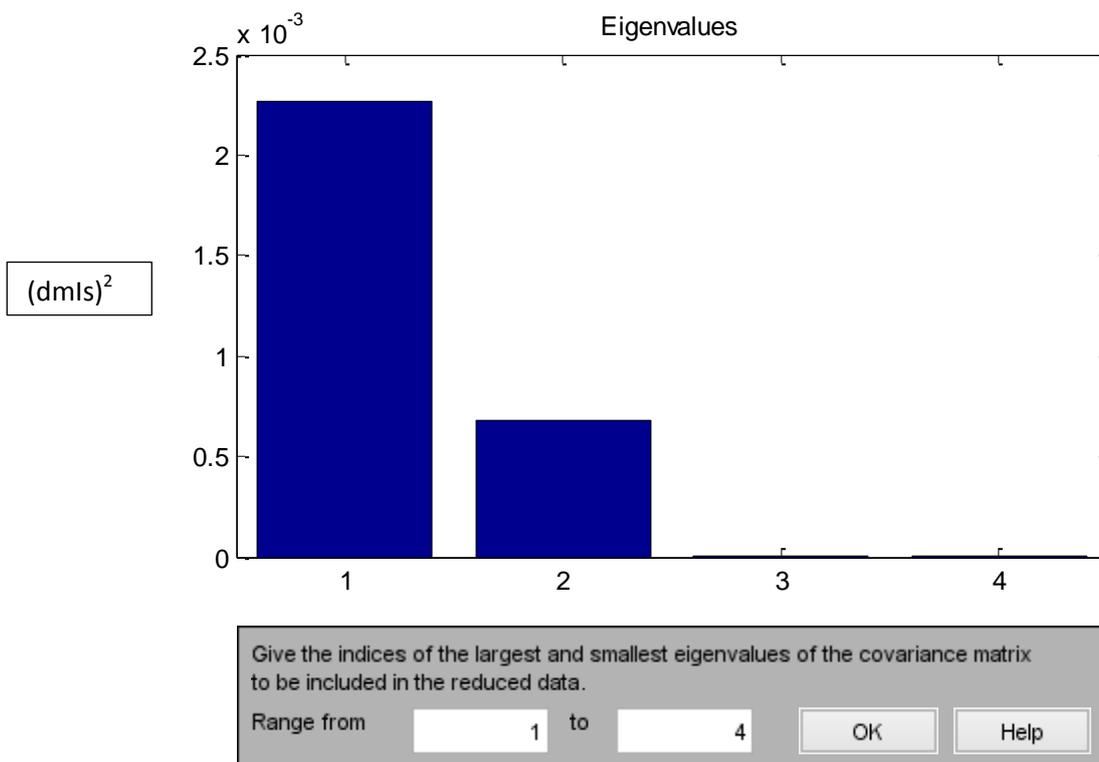


Fig.12: Analisi del segnale relativo all'oscillazione meccanica mediante tecnica PCA.

La pre-analisi, effettuata con la PCA è stata effettuata considerando quattro tracce di segnali laser, rivela due autovalori di cui uno dei quali conserva il 88% dell'informazione relativa al sistema in esame.

3.5 Analisi delle componenti indipendenti (ICA)

L'analisi delle componenti indipendenti (ICA) è un metodo per trovare le componenti indipendenti in dati multidimensionali descritti da una statistica multi-variata. E' una metodologia che è stata introdotta negli anni ottanta [Hérault J. et al., 1984] ed è stata sviluppata per risolvere il problema del cocktail party. L'ICA effettua il riconoscimento e la separazione di sorgenti incognite statisticamente indipendenti, supponendo che i segnali registrati siano una mistura lineare di tali sorgenti [Hyvärinen et al., 1997 e . Hyvärinen et al., 2001]. Utilizzando il concetto di non gaussianità e una statistica del quarto ordine si effettuano le separazioni (vedi Appendice A).

L'ICA ha fornito risultati interessanti quando applicata ai sistemi dinamici, quali gli eventi sismici che accompagnano le eruzioni vulcaniche, i fasci di particelle che mostrano moti collettivi negli acceleratori, il suono emesso dagli strumenti musicali, fornendo informazioni sulla complessità dei sistemi dinamici in esame e riuscendo a descrivere in modo efficace il loro comportamento [De Lauro et al., 2005].

L'analisi delle componenti indipendenti è stata applicata ai segnali acustici ed ai segnali di vibrazione meccanica acquisiti lungo il tratto trachea/laringe mentre diversi operatori emettevano

suoni vocalici. I segnali di input, le misture, sono stati filtrati tra 70 Hz e 2200 Hz e sono stati riportati per ciascuna vocale (Fig.13, Fig.15, Fig.17, Fig.19, Fig.21). I risultati ottenuti dall'analisi dei segnali acustici e dei segnali laser sono relativi ad una donna di 37 anni che emette suoni delle cinque vocali italiane a, e, i, o, u.

L'analisi ICA sui segnali acustici, su quattro misture di input associate a due a due alla stessa misura in tempi diversi lungo il tratto trachea-laringe, estrae per ciascuna vocale due componenti indipendenti. Le componenti indipendenti estratte per la vocale a sono 820 Hz e 1030 Hz (Fig.14), per la vocale e sono 622 Hz e 1882 Hz (Fig. 16), per la vocale i sono 215 Hz e 423 Hz (Fig.18), per la vocale o sono 625 Hz e 1040 Hz (Fig.20) e per la vocale u sono 448 Hz e 897 Hz (Fig.22).

Si può concludere che l'ICA riesce a separare le prime due formanti (F1, F2) per ciascuna vocale ad eccezione della i che presenta valori delle formanti successive alla F1 a frequenze superiori a 2200 Hz e per la quale vengono estratte la prima formante F1 e una frequenza F0, cioè una frequenza vicino al valore della frequenza di pitch (f_0) ma leggermente più alta, dovuta al restringimento che il flusso d'aria proveniente dai polmoni incontra alla fine della trachea per la presenza delle pieghe vocali (De lauro et al., 2007).

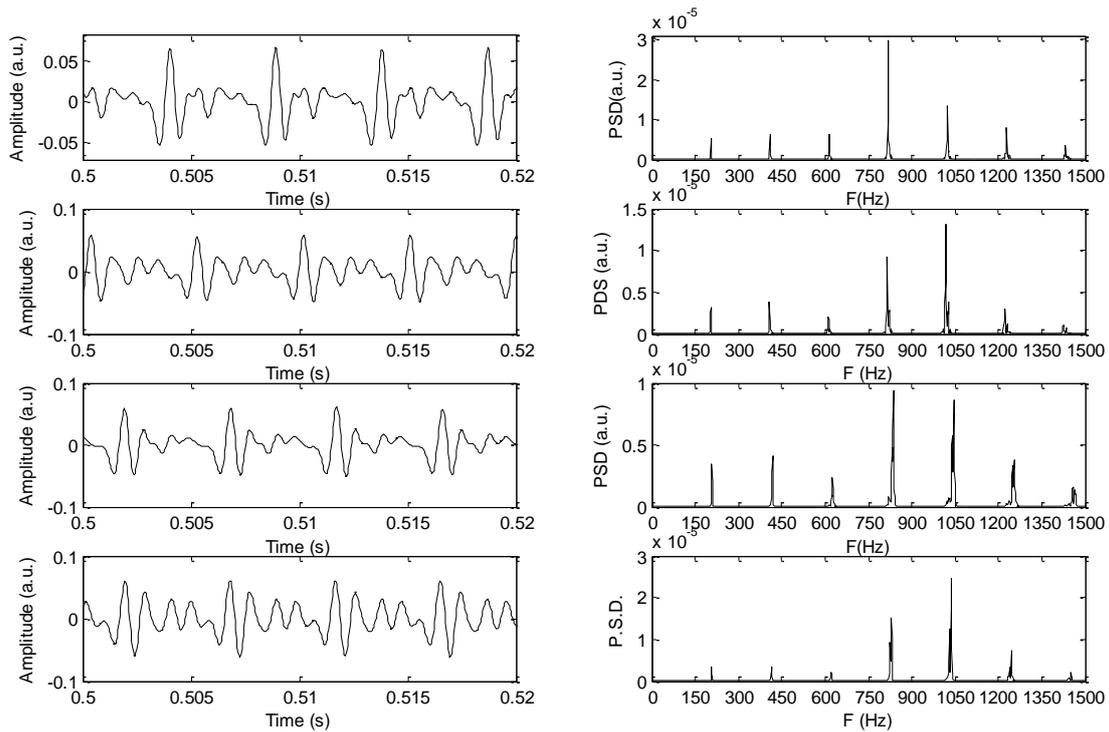


Fig.13: Misture segnali acustici vocale a.

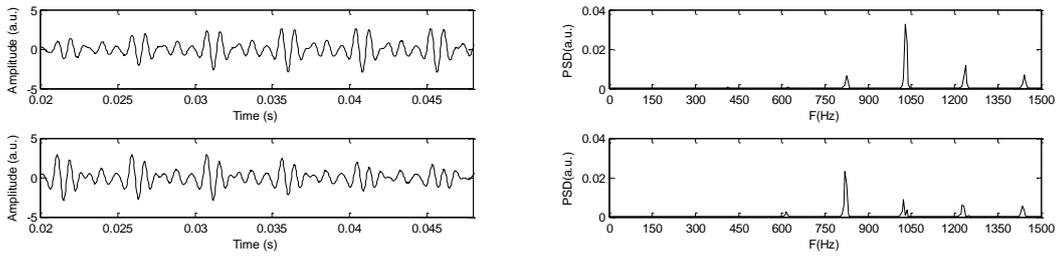


Fig.14: Analisi componenti indipendenti segnali acustici vocale a.

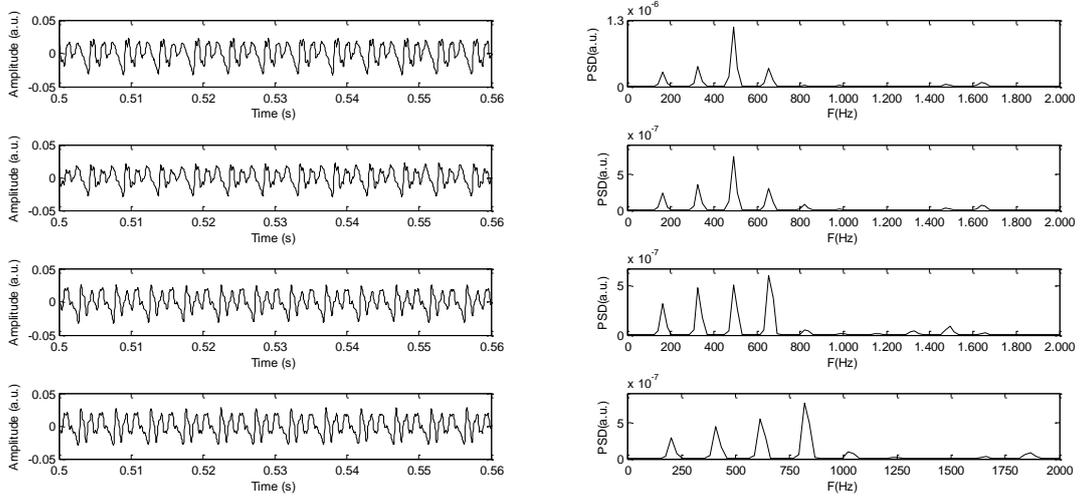


Fig.15: Misure segnali acustici vocale e.

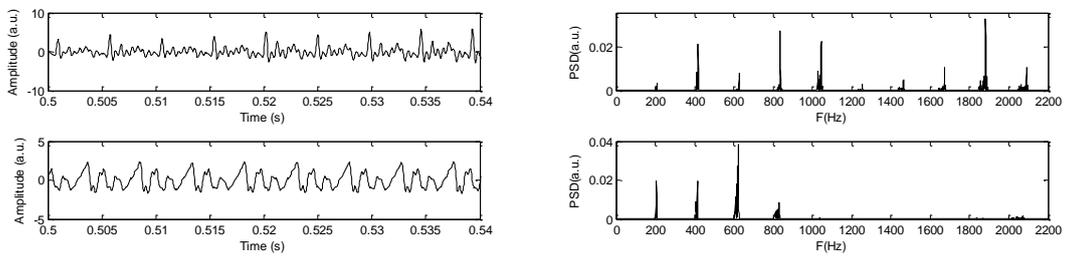


Fig.16: Analisi componenti indipendenti segnali acustici vocale e.

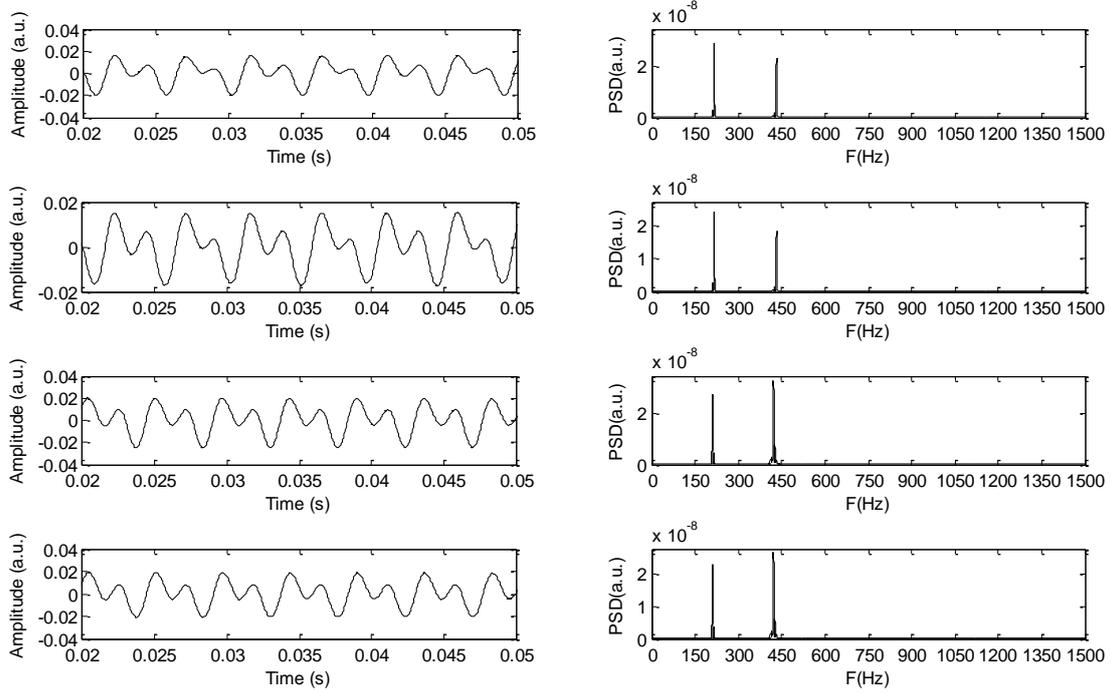


Fig.17: Misture segnali acustici vocale i.

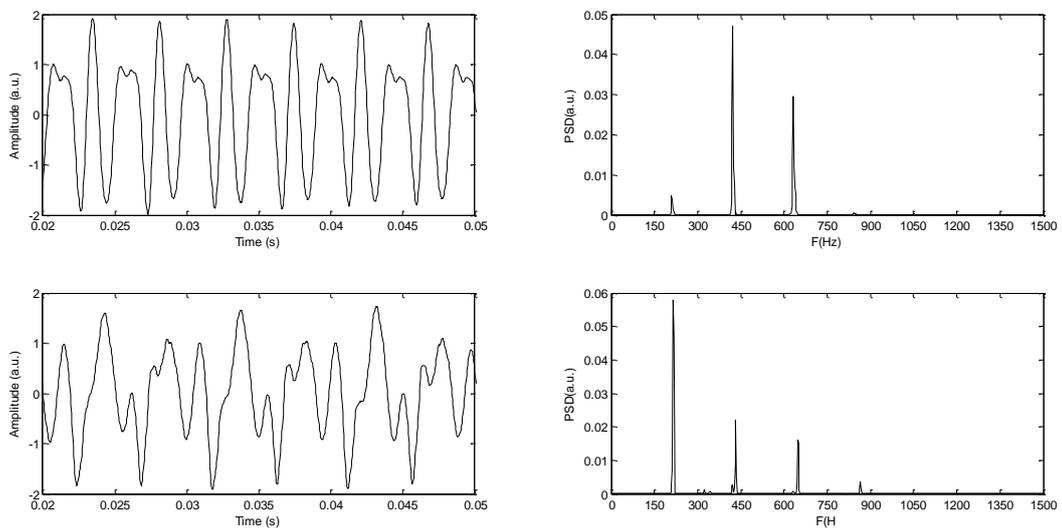


Fig.18:Analisi componenti indipendenti segnali acustici vocale i.

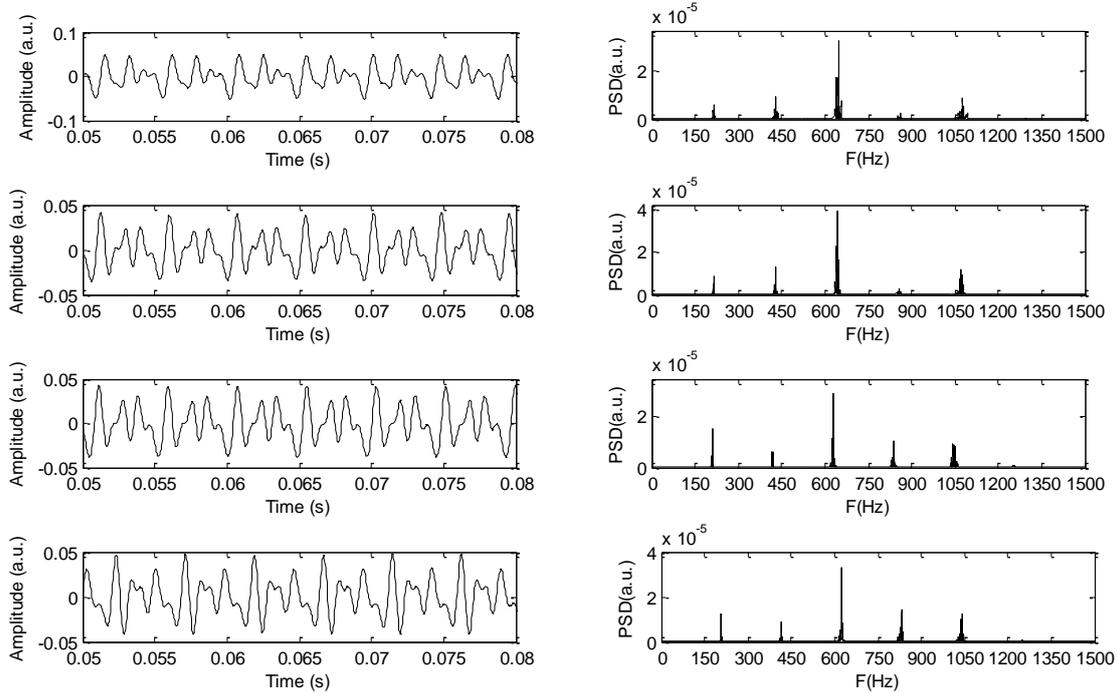


Fig.19: Mixture vocale segnali acustici vocale o.

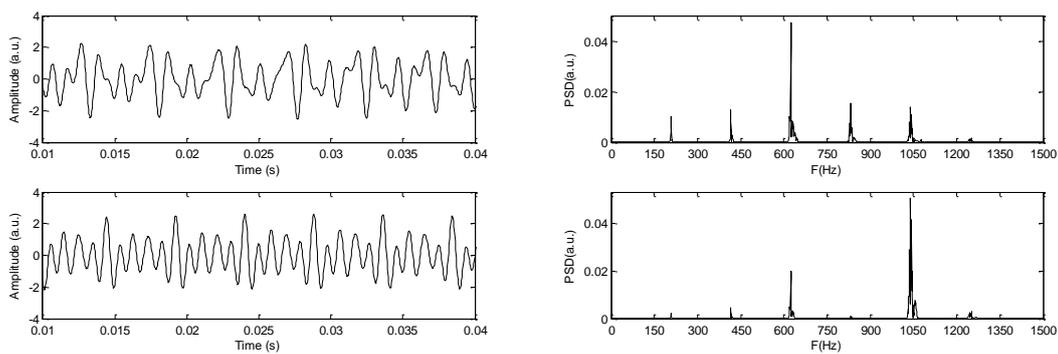


Fig.20: Analisi componenti indipendenti segnali acustici vocale o.

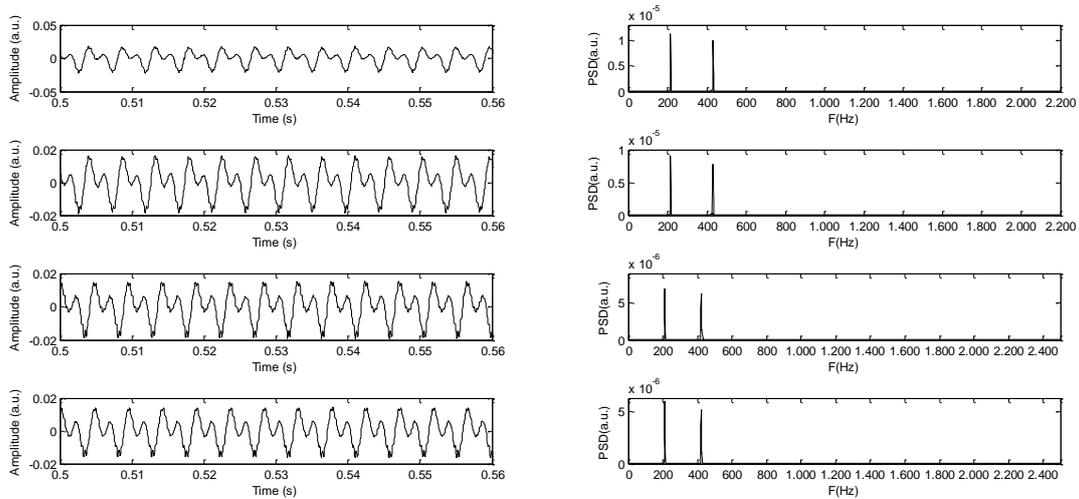


Fig.21: Misure segnali acustici vocale u.

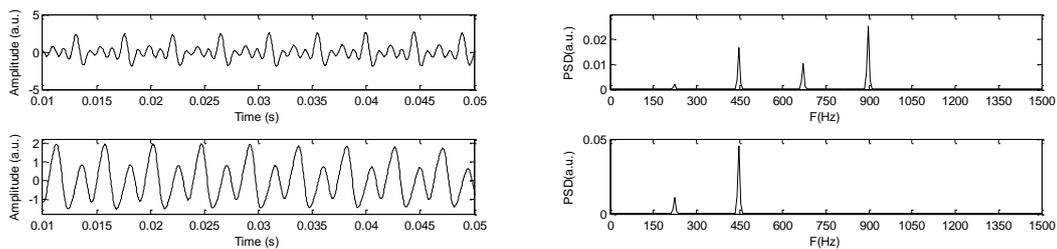


Fig.22: Analisi componenti indipendenti segnali acustici vocale u.

Vengono riportati i risultati ottenuti dall'analisi ICA dei segnali laser, con quattro miscele di input. Le miscele sono, a due a due, pezzi registrati in tempi diversi di uno stesso segnale associate ad una unica misura di vibrazione meccanica effettuata in due punti della trachea con i vibrometri laser. Con questa tecnica per ciascuna vocale vengono estratte due componenti indipendenti. La prima per tutte le vocali ha un valore compreso nella finestra [204-220]Hz e corrisponde proprio alla frequenza di pitch, mentre la seconda componente indipendente ha un valore compreso nella finestra [406-440]Hz. La prima e seconda componente estratta corrispondono al primo e secondo modo di Landau della "canna" trachea. Il primo modo associato alla frequenza di pitch compare nello spettro con la massima ampiezza, ad indicare che nella trachea si origina un segnale indipendente con frequenza caratteristica corrispondente alla frequenza fondamentale. Anche l'analisi ICA conferma l'ipotesi che nella produzione dei suoni la trachea non svolge un ruolo passivo, ma è la sua oscillazione sorgente del suono stesso.

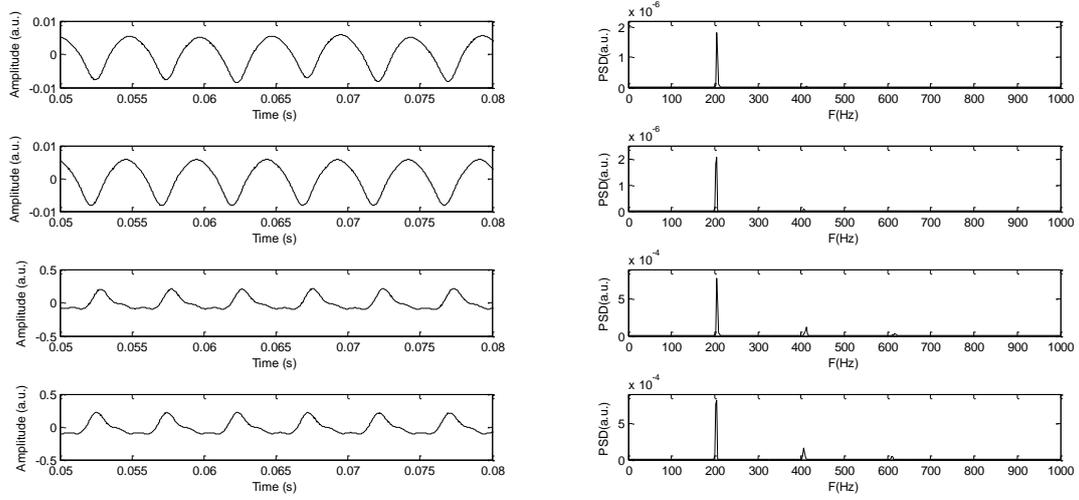


Fig.23: Mixture segnale laser vocale a.

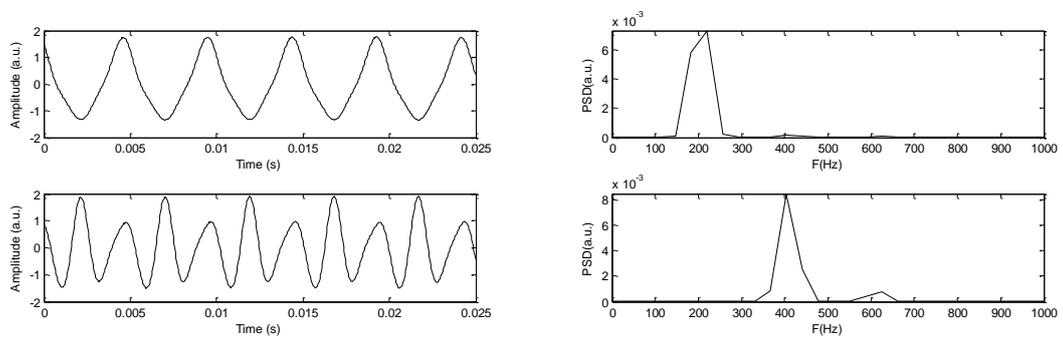


Fig.24: Analisi componenti indipendenti sul segnale laser vocale a.

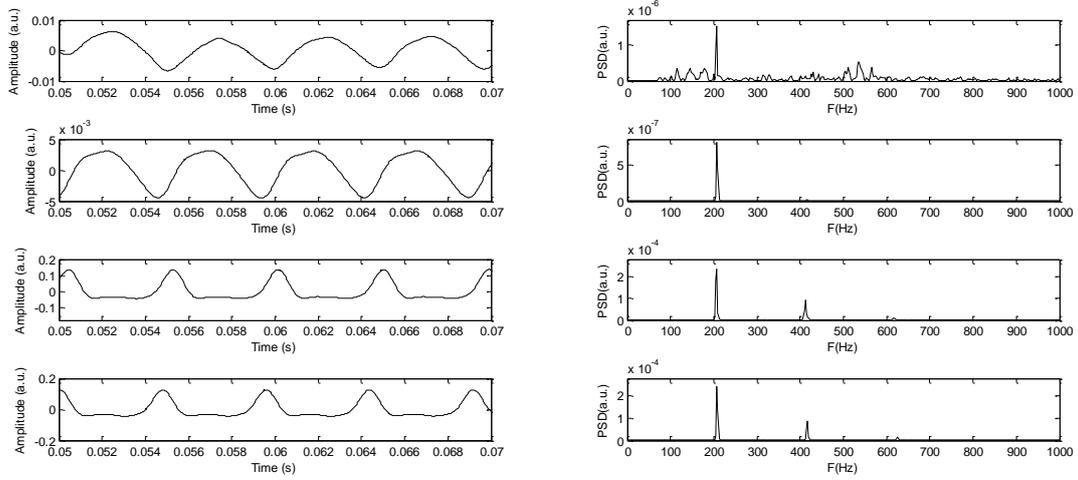


Fig.25: Mixture segnale laser vocale e.

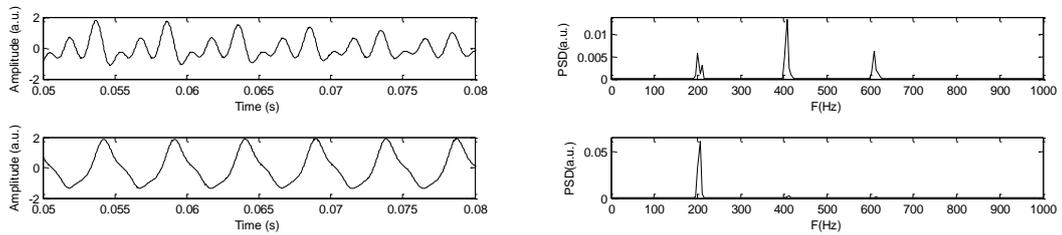


Fig.26: Analisi componenti indipendenti sul segnale laser vocale e.

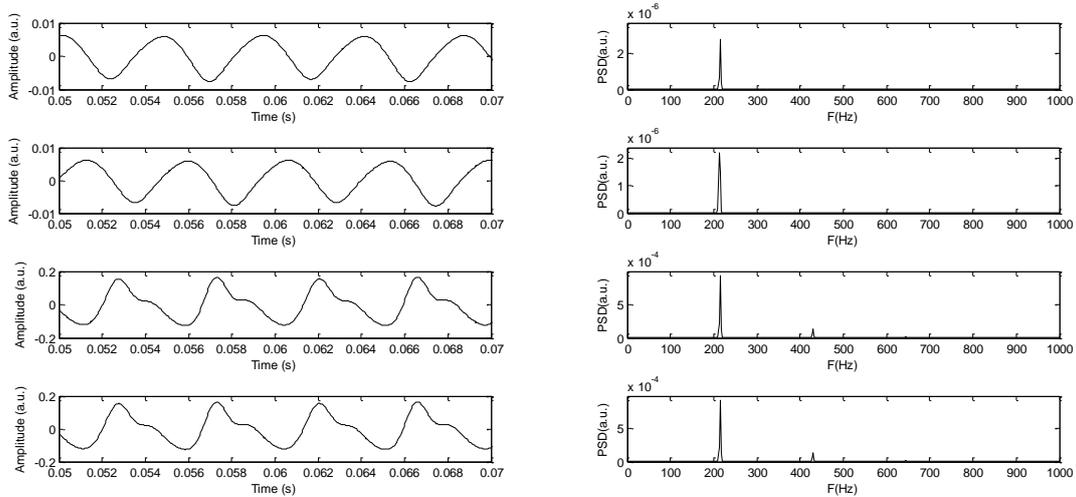


Fig.27: Misture segnale laser vocale i.

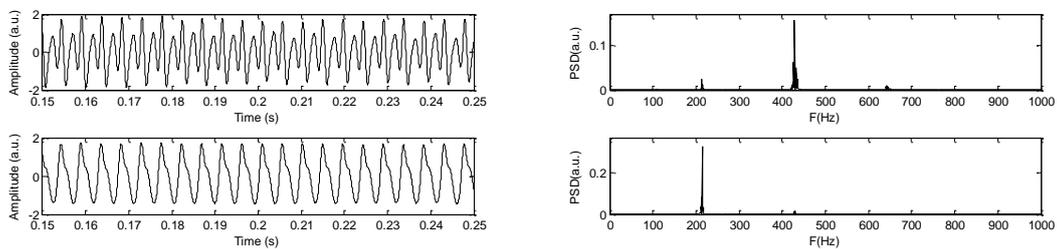


Fig.28: Analisi componenti indipendenti sul segnale laser associato alla vocale i.

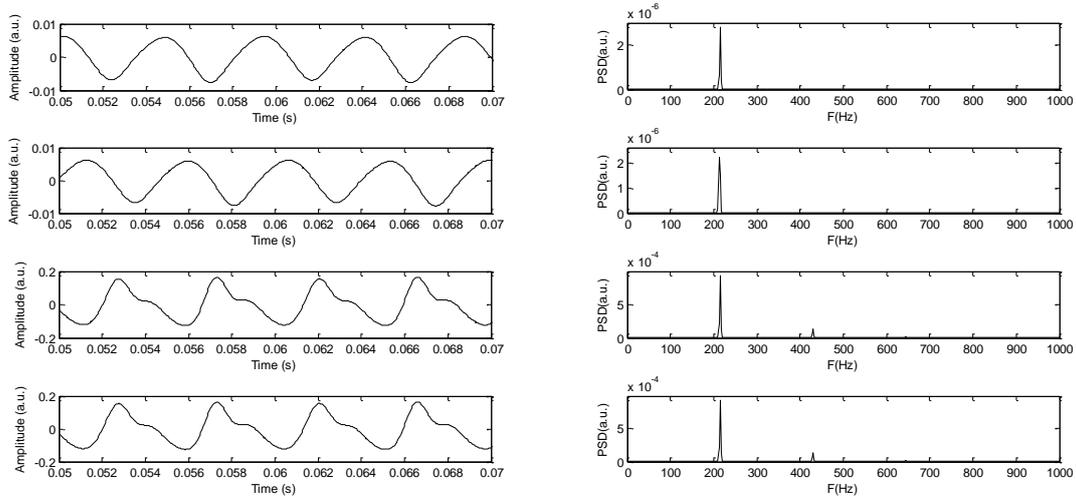


Fig.29: Mixture segnale laser vocale o.

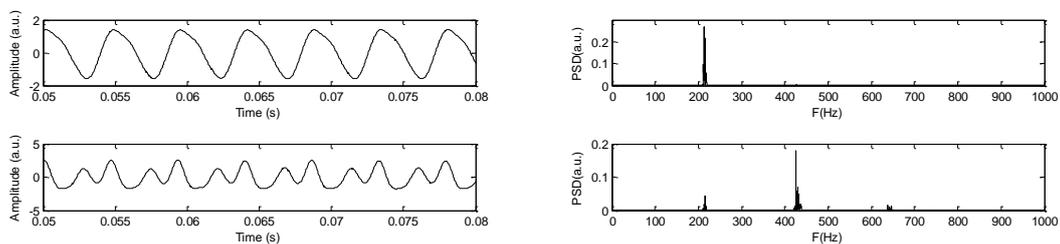


Fig.30: Analisi componenti indipendenti sul segnale laser associato alla vocale o.

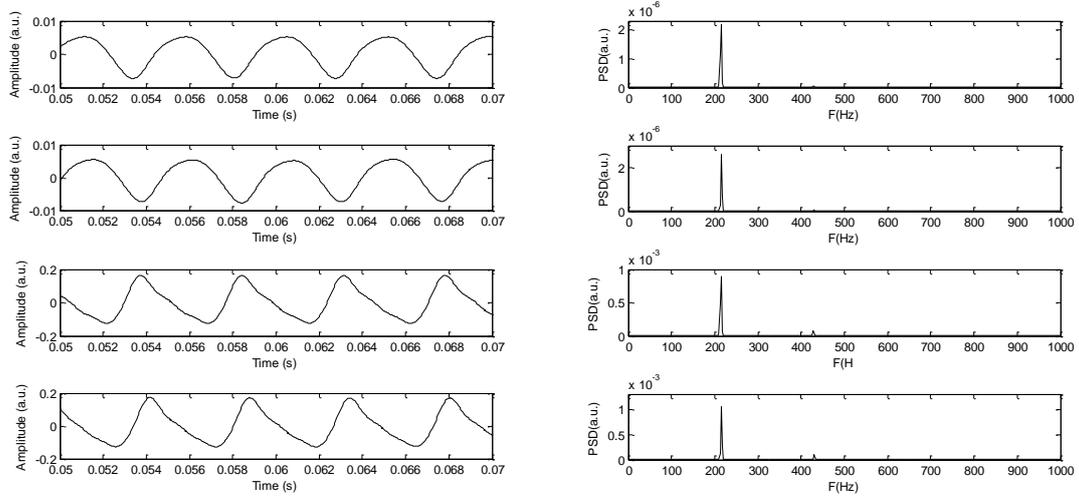


Fig.31: Mixture segnale laser vocale u.

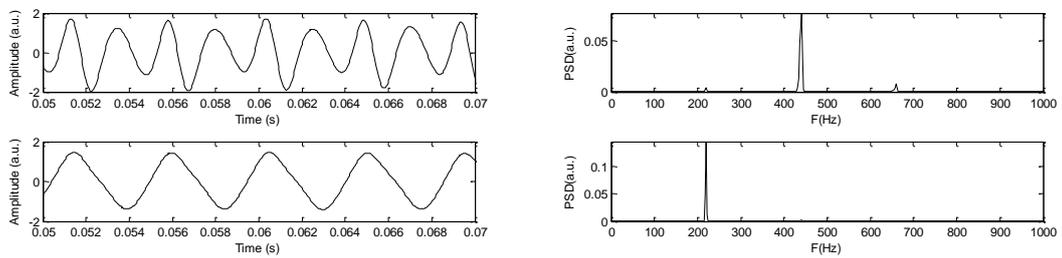


Fig.32: Analisi componenti indipendenti sul segnale laser associato alla vocale u.

In conclusione il campo di vibrazione meccanica e acustico risultano essere composti da segnali indipendenti nel tempo. L'attivazione dei modi indipendenti (modi di Landau) è non lineare ma la loro sovrapposizione è lineare.

3.6 Modello fenomenologico del vocal tract

L'analisi effettuata sull'emissione dei suoni vocalici suggerisce di schematizzare il tratto vocale come una serie di tubi acustici di lunghezza e sezione differente collegati in serie. La parte del vocal tract inferiore viene schematizzata con una canna d'organo che presenta inferiormente l'apertura attraverso la quale il flusso d'aria proveniente dai polmoni giunge nella trachea e superiormente essa presenta un restringimento che fisiologicamente è rappresentato dalle pieghe vocali. Il flusso d'aria proveniente dai polmoni mette in oscillazione la "canna" trachea che, in accordo con i risultati sperimentali, oscilla alla frequenza fondamentale di pitch (f_0). Il flusso d'aria giunto alla sommità della trachea trova un restringimento del corpo della canna dovuto alla presenza delle corde vocali, in tal punto quindi si genera un'altra instabilità idrodinamica. Le corde vocali entrano in interazione con il flusso d'aria proveniente dal basso ed incominciano ad oscillare in modo solidale con la trachea. Sperimentalmente osserviamo al di sopra delle pieghe vocali, che il restringimento fa aumentare la pressione a cui è sottoposto il flusso d'aria e in tal modo il segnale si arricchisce di modi superiori alla frequenza fondamentale. Si osservano: uno shift delle frequenze verso valori più alti ed un aumento delle ampiezze delle frequenze superiori [De lauro et al., 2007]. Il segnale acustico, dovuto alla presenza della parte superiore del vocal tract, costituita dalla cavità orale e dalla cavità nasale, amplifica alcune frequenze e si arricchisce di frequenze più alte, con caratteristiche spettrali simili a quelle mostrate da un sassofono.

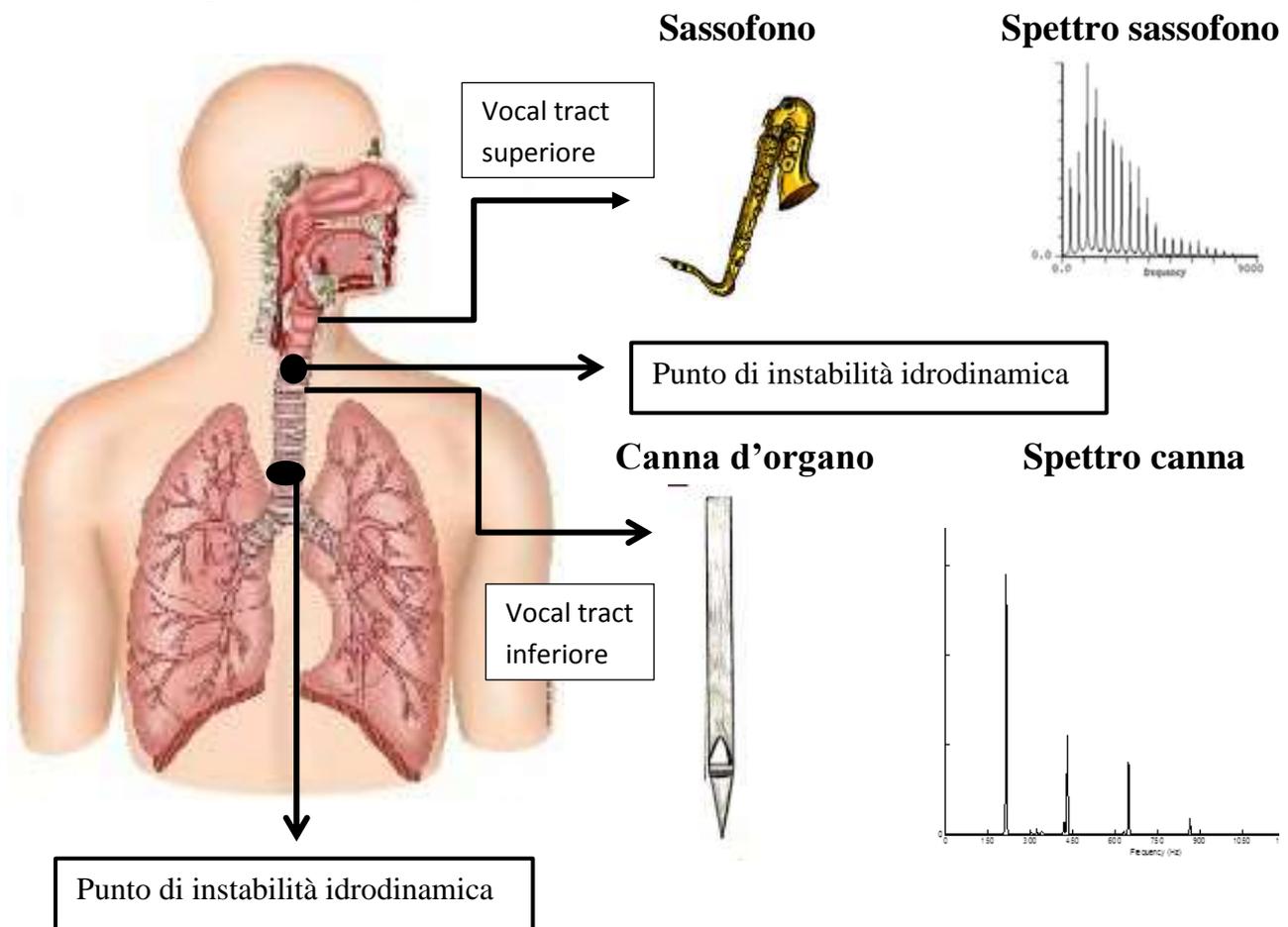


Fig.33: Modello fenomenologico del vocal tract.

In figura (Fig. 33) si riporta in dettaglio il modello fenomenologico del vocal tract, che nasce dall'analisi sperimentale effettuata in questo lavoro di ricerca. Il tratto vocale superiore viene schematizzato come un sassofono, perché dall'osservazione dello spettro dei segnali acustici registrati si nota una certa similarità con lo spettro di questo strumento musicale evidente in Fig.33. Procedendo verso il basso, al livello della glottide, si giunge ad un primo punto di instabilità idrodinamica in cui il flusso d'aria proveniente dai polmoni trova un restringimento a causa delle corde vocali e si determinano i fenomeni suddetti. Il vocal tract inferiore è stato schematizzato, sempre tenendo conto dello spettro dei segnali laser registrati e delle analisi fatte, come una canna che suona attivando al massimo due modi di Landau. Infine, un secondo punto di instabilità a livello della giunzione trachea-polmoni si osserva ed è dovuto al flusso d'aria generato dai polmoni che mette in oscillazione la trachea.

Capitolo 4

Le oscillazioni meccaniche del cuore

In questo capitolo vengono studiate ed analizzate le oscillazioni meccaniche associate ad un altro sistema biologico che è il cuore, misurando attraverso due vibrometri laser doppler le oscillazioni che esso induce sulla cassa toracica.

Il segnale associato all'oscillazione meccanica "indotta" nel torace dall'attività cardiaca, è certamente "a priori" molto complicato, ma se le diagnosi nel passato sono state fatte con la auscultazione vuol dire che esso presenta degli elementi di riconoscibilità ben precisi. L'attività di ricerca in corso mira a fornire un esame clinico alternativo all'elettrocardiogramma che permetta di dare informazioni su situazioni di normalità o di malattia del cuore.

L'elettrocardiogramma (ECG) fornisce informazioni dello stato di "salute" del cuore analizzando la sua attività elettrica, invece nell'attività di ricerca proposta il cuore viene investigato attraverso le sue oscillazioni meccaniche.

4.1 L'apparato cardiocircolatorio

Il sistema cardiocircolatorio è costituito dal cuore e da un complesso sistema di vasi sanguigni entro i quali scorre il sangue: le arterie, che partono dal cuore e dividendosi in rami di diametro sempre più piccolo, giungono sino ai capillari dove avvengono gli scambi tra sangue e tessuti.

Questo apparato rende possibile il trasporto in tutto il corpo delle sostanze di nutrimento e dell'ossigeno, per questo motivo esso è strettamente legato all'apparato respiratorio e ai suoi organi principali, i polmoni, grazie ai quali può rifornirsi di ossigeno ed eliminare le sostanze di rifiuto, principalmente rappresentate dall'anidride carbonica.

4.1.1 Il sangue e il cuore

Il sangue è composto da una parte liquida (plasma), dai globuli rossi (che hanno il compito di trasportare l'ossigeno), dai globuli bianchi (che difendono l'organismo dai batteri e da altri elementi estranei pericolosi alla salute) e dalle piastrine (importanti per la coagulazione del sangue).

Il sangue è un efficiente mezzo di trasporto in grado di portare in tutto il corpo varie sostanze. Le sue funzioni principali sono:

- trasportare l'ossigeno dai polmoni alle cellule;
- trasportare l'anidride carbonica ai polmoni perché la espellano;
- trasportare le sostanze nutritive;
- trasportare le sostanze di rifiuto (come l'urea), dai tessuti ai reni per farle espellere;
- trasportare acqua e sali minerali ovunque siano necessari;
- mantenere costante la temperatura del corpo;
- trasportare gli ormoni che regolano varie funzioni dell'organismo;
- difendere il corpo dai microbi, grazie all'azione dei globuli bianchi;

- rendere possibile la coagulazione del sangue in caso di ferite o escoriazioni del nostro corpo, grazie all'azione delle piastrine.

L'organismo ha bisogno di una pompa abbastanza potente per mantenere il sangue in circolazione in tutto il corpo. Questo compito è svolto dal cuore, un organo cavo che si trova al centro della gabbia toracica e che è composto da tessuto muscolare striato e involontario, il miocardio.

Il cuore si divide in quattro cavità, due atri e due ventricoli (Fig.1). L'atrio di destra riceve il sangue venoso, proveniente dai tessuti, e lo immette nel ventricolo destro che lo pompa nei polmoni.

L'atrio sinistro, invece, accoglie il sangue proveniente dai polmoni, lo immette nel ventricolo sinistro che lo pompa nelle arterie per raggiungere i tessuti. Le valvole, tricuspide e mitralica, tra gli atri e i ventricoli, assicurano che il sangue scorra in un'unica direzione.

Quando le cavità atriali e ventricolari ricevono il sangue, il muscolo cardiaco si rilassa per far spazio al liquido, questa fase è chiamata diastole. Successivamente, il cuore si contrae, determinando l'espulsione del sangue dai ventricoli verso i vasi sanguigni, nella fase chiamata sistole. Il numero di battiti al minuto, che in genere a riposo corrisponde a circa 70-80, costituisce la frequenza cardiaca. La gittata cardiaca invece è la quantità di sangue che viene messa in circolo dal cuore in un minuto.

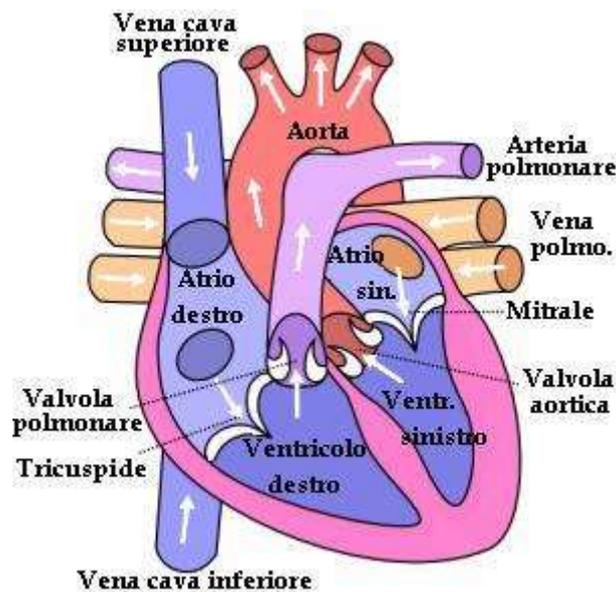


Fig.1:Il cuore e le sue cavità.

Nell'apparato circolatorio si distinguono una piccola circolazione e una grande circolazione (Fig.2). La prima inizia nel ventricolo destro e porta ai polmoni il sangue venoso ricco di anidride carbonica, attraverso l'arteria polmonare. Nei polmoni avviene lo scambio tra anidride carbonica ed ossigeno e il sangue arterioso ricco di ossigeno torna all'atrio sinistro del cuore, attraverso le vene polmonari. La grande circolazione, invece, inizia nel ventricolo sinistro dove il sangue arterioso, ossigenato, passa nell'arteria aorta e da qui raggiunge le varie parti del corpo fino ai capillari dove avvengono gli scambi gassosi e di sostanze tra sangue e tessuti. Poi, attraverso le vene cave

superiore e inferiore, il sangue, ricco di anidride carbonica, torna nell'atrio destro del cuore. [tratto dal sito: www.iocsanmarcello.gov.it/.../wp.../Lapparato-cardiocircolatorio.pdf]

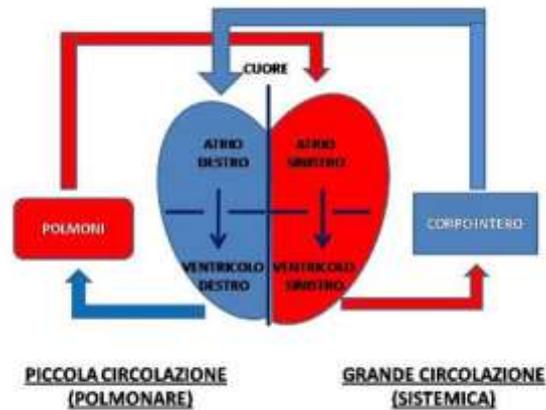


Fig.2: Piccola e grande circolazione.

4.1.2 L'attività meccanica del cuore

L'attività meccanica del cuore è stimolata da una attività elettrica che si origina all'interno dell'organo e si propaga in modo ritmico a tutto il muscolo cardiaco.

L'attività elettrica del cuore può essere rilevata dalla superficie del corpo sotto forma di varie onde che corrispondono a differenti momenti dell'attività cardiaca (contrazione e rilassamento) e prendono il nome di elettrocardiogramma (Fig.3).

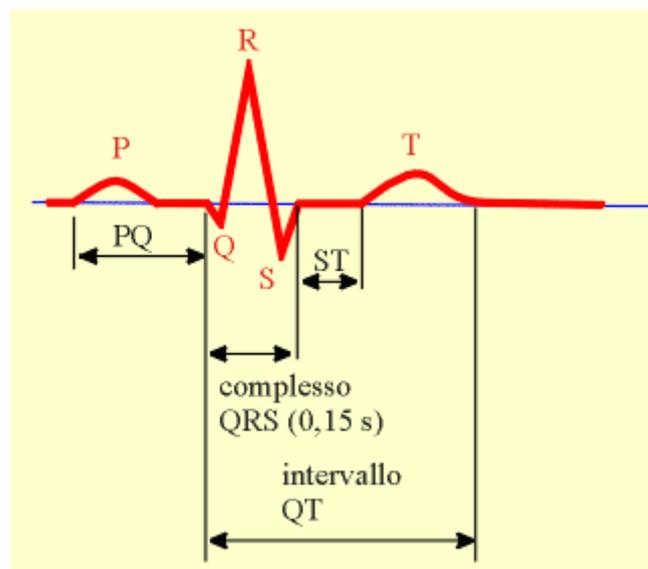


Fig.3: Un esempio di segnale elettrocardiografico.

L'elettrocardiogramma è un esame non invasivo che permette di evidenziare patologie sospettate durante l'esame clinico oppure decorse in maniera silente e quindi riscontrate casualmente.

L'interpretazione di questo grafico consente al cardiologo di rilevare situazioni anormali o di malattia: ritmi diversi da quello normale, sovraccarico di lavoro del muscolo cardiaco, condizioni di inadeguata ossigenazione (ischemia) o infarto del miocardio [tratto dal sito: www.centrolottainfarto.org/saperne...cuore/59-lelettrocardiografia.html].

4.2 Il ciclo cardiaco

Per ciclo cardiaco si intende l'insieme degli eventi che caratterizzano l'attività meccanica del cuore. Il ciclo cardiaco consta di due fasi, sistole e diastole, ciascuna delle quali si divide in sottofasi [Ganong et al., 2011]. Per convenzione si considera il ciclo cardiaco dal punto di vista dei ventricoli, importanti perché sono quelli che immettono sangue in circolo. In condizioni di riposo ogni ciclo cardiaco dura 0,8 secondi; sotto sforzo la durata del ciclo si riduce quindi aumenta la frequenza cardiaca. Dei 0,8 secondi del ciclo cardiaco 0,5 secondi sono occupati dalla diastole ventricolare (compresi gli ultimi 0,11 s di presistole o sistole atriale) e 0,3 secondi sono occupati dalla sistole ventricolare.

Per parlare del ciclo cardiaco occorre considerare le variazioni di pressione che si registrano nel ventricolo, nell'atrio e nel vaso arterioso ad esso connessi.

Contemporaneamente possono anche essere valutati le variazioni di riempimento delle camere cardiache e i suoni del cuore che si associano a determinati eventi. I suoni cardiaci sono quattro, di cui solo due udibili con fonendoscopio.

E' possibile considerare un ciclo cardiaco destro e sinistro che si differenziano per i valori della pressione, essendo le pressioni in gioco molto minori a destra, perché le resistenze dei vasi polmonari sono molto più basse. In genere, tuttavia, si considera il ciclo cardiaco sinistro.

Il ciclo cardiaco inizia per convenzione dalla sistole atriale, che è l'ultima delle fasi di riempimento (diastole) del ventricolo. La contrazione delle pareti, causa una vibrazione, il quarto tono, che non è udibile normalmente.

La sistole atriale genera un leggero aumento della pressione nell'atrio che determina il completo svuotamento dell'atrio stesso (il sangue passa per gradiente di pressione al ventricolo). Il volume di sangue che a seguito della contrazione passa al ventricolo non contribuisce in maniera significativa al riempimento dello stesso, che è avvenuta per la maggior parte precedentemente. Normalmente però questo volume di sangue ha la funzione di far aumentare la pressione nel ventricolo tanto da stabilire un gradiente di pressione opposto al precedente (la pressione diventa maggiore nel ventricolo che nell'atrio), così il sangue tende a refluire dal ventricolo verso l'atrio, spingendo i lembi della valvola atrioventricolare alla chiusura. La chiusura della valvola atrioventricolare determina l'insorgenza del I° tono, causato dall'oscillazione della pareti delle camere cardiache, udibile nel quinto spazio intercostale, a livello della linea margino-sternale destra o sinistra per le due valvole. Il ventricolo a questo punto diventa una cavità chiusa.

Nel frattempo, il cuore subisce una contrazione dal basso verso l'alto dovuta alla propagazione dell'onda di depolarizzazione. L'onda di depolarizzazione dei ventricoli è rilevata dall'ECG come complesso QRS. Si ha così una prima fase di sistole ventricolare isovolumetrica, perché essendo il ventricolo una cavità chiusa e piena di liquido incompressibile, il volume non può cambiare. Le

fibre in contrazione sviluppano perciò solo tensione e non lavoro; la tensione sviluppata consente alla pressione nel ventricolo di superare la pressione nell'aorta, generando l'apertura della valvola semilunare. A causa della tensione che si sviluppa nel ventricolo la valvola atrioventricolare chiusa viene spinta con i suoi lembi verso la cavità atriale, facendo aumentare leggermente la pressione anche nell'atrio che si trova in diastole.

La pressione nell'aorta nel momento di apertura della valvola semilunare è quella minima raggiunta ed è detta diastolica o anche postcarico, ovvero carico che il ventricolo deve vincere per l'eiezione del sangue. Il volume di sangue presente nel ventricolo al momento della sistole è detto volume telediastolico ed è di circa 120/130 ml. Il volume massimo di sangue contenuto nel ventricolo alla fine della diastole o telediastolico è il precarico.

Una volta aperta la valvola semilunare, quindi, il sangue può fuoriuscire dal ventricolo secondo gradiente di pressione, permettendo una variazione del volume del ventricolo. Si ha perciò una fase detta sistole isotonica in cui il ventricolo continua a contrarsi con accorciamento delle fibre e diminuzione del volume.

La sistole isotonica produce lavoro, che è quello per spostare un certo volume di sangue verso l'arteria.

Nella fase di sistole isotonica si ha l'efflusso del sangue che si compie in due modi, con accelerazione e con decelerazione. Nella prima fase il sangue esce dal ventricolo a velocità crescente (questo non implica che la maggior parte del sangue venga espulso in questa fase) e preme sulle pareti dell'arteria facendole dilatare. In questo momento la pressione nell'aorta è massima (pressione sistolica); dopo di che il sangue tenta di refluire nel ventricolo per gradiente pressorio, perché la pressione nel ventricolo è scesa. Per fare questo il sangue rallenta la sua corsa (efflusso con decelerazione), si arresta e riparte in direzione opposta verso il ventricolo dove determina la chiusura della valvola semilunare. Circa 70 ml di sangue sono usciti dal ventricolo (gittata sistolica = volume di sangue espulso da un ventricolo con una sistole), mentre 50-60 ml restano nel ventricolo e costituiscono il volume telesistolico (alla fine della sistole). Questa quantità è una importante riserva di sangue che il ventricolo può usare sotto sforzo per aumentare la gittata sistolica.

La chiusura della valvola semilunare genera il 2° tono cardiaco, che è più breve e più acuto del precedente in quanto generato dalle vibrazioni delle pareti del vaso. E' udibile ponendo il fonendoscopio nel secondo spazio intercostale presso la linea margino-sternale destra per l'aorta, sinistra per la polmonare.

Come già accennato, la differenza di durata/frequenza tra primo e secondo tono dipende dal meccanismo di propagazione del suono stesso; infatti il 2° tono viene propagato attraverso la parete delle arterie che è molto più sottile della parete del miocardio attraversata dal primo, che giunge perciò più ovattato.

Il reflusso del sangue che provoca la chiusura della valvola semilunare genera un piccolo calo di pressione, favorito anche dalla dilatazione dell'arteria. La chiusura della valvola, accompagnata dal ritorno causato dalla elasticità delle pareti del vaso, determina tuttavia un nuovo piccolo aumento di pressione, dando luogo all'onda pressoria detta dicrota. Il ventricolo, a questo punto, torna ad essere una cavità chiusa, le fibre si rilassano in maniera isometrica (non si ha variazione della lunghezza delle fibre). Questo rilassamento determina il crollo della pressione nel ventricolo, tanto che essa diviene minore di quella dell'atrio dove nel frattempo si è accumulato il sangue, che per il gradiente di pressione preme sulla valvola atrioventricolare causandone l'apertura. Il sangue passa, quindi,

per gradiente di pressione dall'atrio al ventricolo, causando la distensione delle fibre di quest'ultimo (che è proporzionale alla quantità di sangue che entra e quindi al ritorno venoso).

Il riempimento del ventricolo avviene in 2 fasi: si ha un iniziale riempimento rapido, quando il gradiente è molto forte, che provvede a immettere circa i 2/3 del volume telediastolico. Il riempimento rapido è responsabile del 3° tono cardiaco, che è dovuto ai vortici creati dal sangue che entra nel ventricolo ad alta velocità.

Poi vi è un riempimento lento che immette circa tutto il terzo restante; infine si ha la sistole atriale, che, come ricordato, a riposo non ha funzioni rilevanti al fine del riempimento ventricolare.

In caso di aumento della frequenza cardiaca la durata del ciclo cardiaco si riduce e con essa la fase di riempimento lento che addirittura può scomparire; in queste condizioni la sistole atriale assume importanza ai fini del riempimento, per immettere nel ventricolo l'ultimo terzo di sangue. Se la frequenza cardiaca supera i 180 battiti al minuto anche la fase di riempimento rapido viene intaccata e la sistole atriale diventa ancora più importante, anche se non riesce a garantire comunque l'introduzione di un normale volume telediastolico nel ventricolo e la gittata sistolica si riduce (un aumento della frequenza cardiaca si ripercuote sulla diastole più che sulla sistole, perché la sistole non può durare meno di un certo periodo).

La frequenza cardiaca può aumentare fino a 180 battiti/min in presenza di tachicardia ventricolare, un'aritmia caratterizzata da un notevole aumento della frequenza. In questo caso la sistole ventricolare risulta sfasata rispetto alla sistole atriale, che può avvenire quando la valvola atrioventricolare è chiusa e il ventricolo non si riempie, oppure quando è aperta e allora è efficace ed essenziale per il mantenimento della gittata sistolica. Nel secondo caso quindi la sistole atriale potrà contribuire al riempimento ventricolare mentre nel primo caso no. Si osserverà che nel secondo caso la gittata sistolica del ventricolo sarà piuttosto elevata mentre nel primo caso sarà molto bassa. Se la frequenza dei ventricoli è molto elevata si potranno avere dei battiti inefficaci, in quanto nel ventricolo la pressione non riuscirà a salire al punto tale da superare la pressione diastolica aortica.

Durante la sistole del ventricolo, la pressione nell'atrio continua a salire per accumulo di sangue in arrivo dalle vene. La pressione nell'atrio, comunque, sarà molto più bassa che nel ventricolo.

Durante la fase di diastole atriale, inoltre, la pressione nell'atrio diventa anche negativa, facilitando in questo modo il ritorno venoso al cuore. Se la pressione nell'atrio diventa negativa aumenta infatti il gradiente che richiama verso di esso il sangue e l'atrio funge da pompa aspirante.

Il valore di pressione nell'atrio, inoltre, varia in funzione del ciclo respiratorio. Infatti, la pressione negativa nell'atrio è secondaria alla depressione intrapleurica, dovuta alle opposte tendenze elastiche del polmone (che tende a collassare) e della gabbia toracica (che tende a espandersi) tenuti adesi dalle due pleure (parietale e viscerale) accollate con liquido intrapleurico. Prima dell'inspirazione la pressione intrapleurica ha un valore di -2 mmHg; durante l'inspirazione questo valore scende a -6 mmHg perché il polmone e la gabbia toracica si espandono e la forza di retrazione elastica del polmone aumenta. Le variazioni della pressione intrapleurica si ripercuotono anche sulle vene cave e gli atri dove la pressione durante l'inspirazione si riduce (durante l'inspirazione aumenta quindi il ritorno venoso al cuore). Durante l'espirazione il polmone e la gabbia toracica tornano alla loro posizione di equilibrio e la forza di retrazione elastica del polmone si riduce, la pressione intrapleurica si riporta a -2 mmHg con conseguenze anche sulla pressione degli atri (il ritorno venoso al cuore si riduce).

Il ciclo cardiaco è brevemente illustrato nella figura 4 [Philippe Primo Caimmi, Elena Grossini - Dispense di Fisiologia- ARACNE editrice S.r.l., 2009].

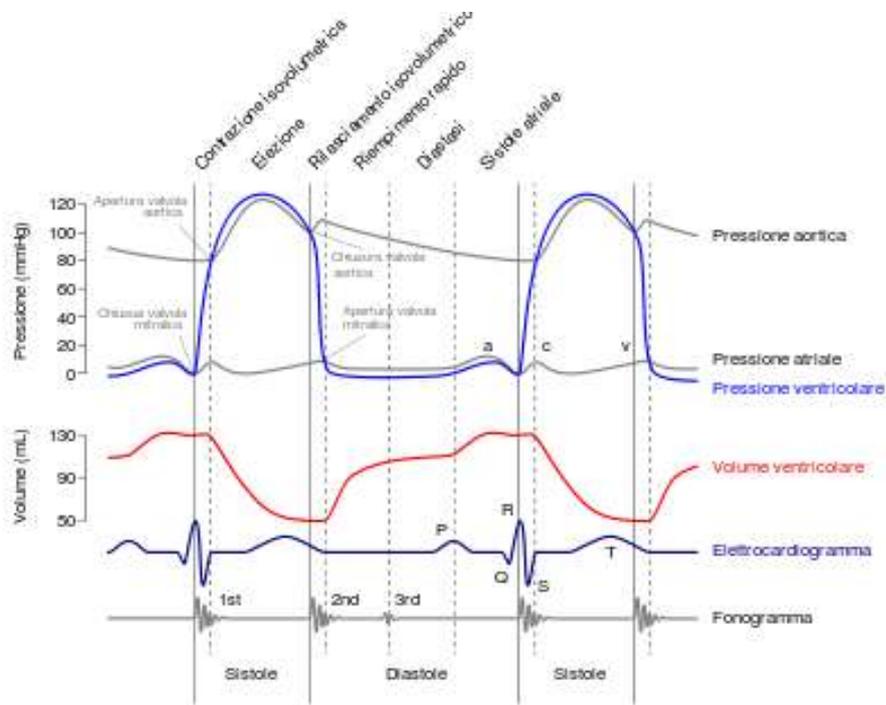


Fig.4: Eventi meccanici (pressioni e flussi) nel ventricolo sinistro, nell'atrio sinistro e nell'aorta durante il ciclo cardiaco. I toni cardiaci.

4.3 Data set

Gli esperimenti sono stati condotti su uomini e donne di età diversa. Le oscillazioni meccaniche del cuore sono state misurate con due vibrometri laser Polytec: PDV-100 e OFV 5000, con banda di frequenza di acquisizione fino a 20 KHz , con frequenza di campionamento pari a 25600 Hz, con $\pm 4V$ di fondo scala e con sensibilità configurata a 5 (mm/sec)/V. I punti di misura sono quelli tipicamente utilizzati per l'auscultazione come visualizzato in Fig.6.

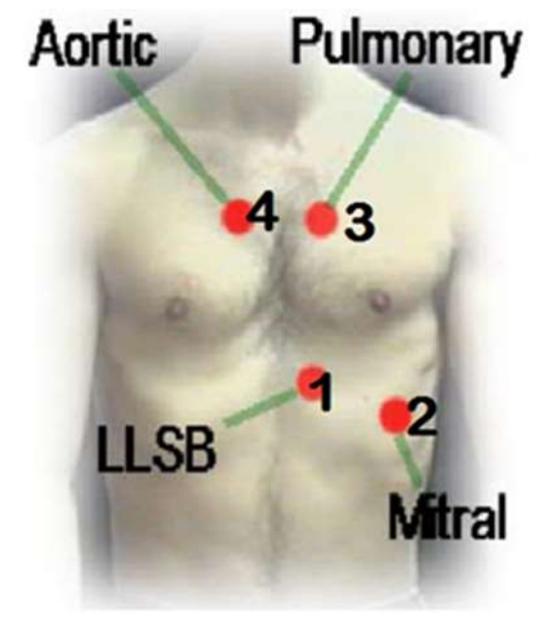


Fig.6: Punti di misura.

E' stata misurata la velocità di oscillazione di una lamina posta sul torace dell'operatore, che permette di riflettere al meglio il fascio laser che incide sul torace.

4.3.1 Analisi spettrale dei segnali

In figura (Fig.7) sono rappresentati i segnali registrati nei quattro punti di interesse, partendo dal punto 1 in alto a destra al punto 4 in basso a destra con i rispettivi spettri. E' evidente l'oscillazione ad 1 sec che è stata messa in evidenza filtrando i segnali con un filtro passa basso a 10 Hz.

Il segnale presenta degli spike evidenti che sono dovuti allo spekle noise perché la superficie su cui incide il fascio laser non è liscia.

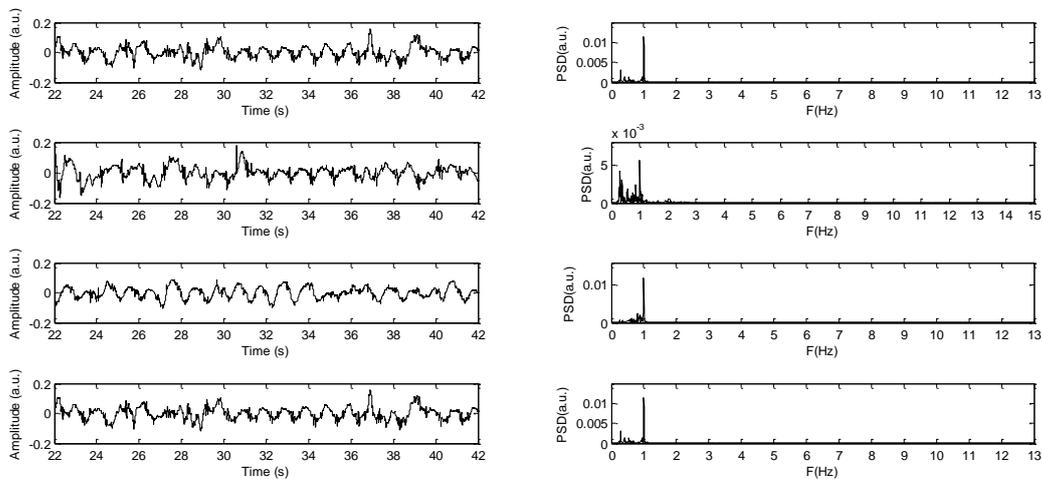


Fig.7: Segnali di vibrazione meccanica registrati nei quattro punti di interesse con i relativi spettri filtrati con un filtro passa basso a 10Hz.

I segnali originari, successivamente sono stati filtrati con un butterworth di ordine 3 passa banda con frequenza tra 3 e 30 Hz, in modo tale da guardare solo alle frequenze cardiache caratteristiche. Vengono riportati in figura i segnali registrati filtrati e la densità di potenza spettrale dei segnali nei quattro punti di misura. (Fig.8)

L'analisi spettrale mette in evidenza picchi caratteristici tra gli 8-18Hz. In questa analisi il ricordo della componente a 1s si osserva nell'intertempo medio caratteristico del battito. L'analisi spettrale mostra che l'ampiezza della componente ad 1Hz ha un'ampiezza maggiore rispetto alle ampiezze delle altre frequenze cardiache caratteristiche.

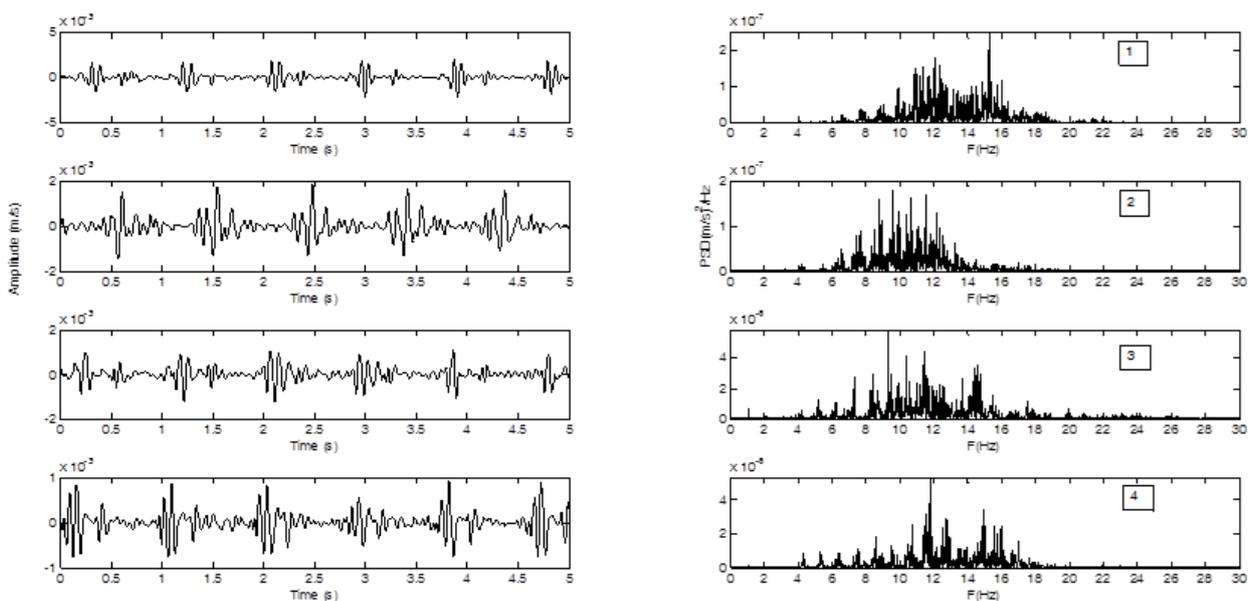


Fig.8: Segnali filtrati e rispettivi spettri di potenza nei quattro punti di registrazione.

4.4 Tecniche di analisi applicate al segnale cardiaco

Lo strumento privilegiato per l'analisi dei segnali, sin dalla sua introduzione, è l'analisi spettrale di Fourier, la quale fornisce una rappresentazione del segnale nel dominio delle frequenze. Per applicare ragionevolmente questa tecnica, come sappiamo, il sistema in esame deve essere lineare e i segnali analizzati devono essere strettamente periodici o stazionari e il segnale deve essere scomponibile come somma di sinusoidi. Raramente segnali che descrivono fenomeni naturali soddisfano queste condizioni, perciò negli anni si sono introdotte tecniche alternative per l'analisi dei fenomeni naturali ai quali corrispondono segnali non stazionari e non lineari [Rangaraj, 2002].

4.4.1 Spettrogramma

Questo è il metodo più elementare, consiste nel fare un'analisi di Fourier in una finestra limitata temporalmente e scorrere lungo tutto l'asse del tempo. Essendo però basato sull'analisi di Fourier, devono essere rispettate le condizioni di stazionarietà e linearità entro la finestra temporale. La scelta delle dimensioni della finestra diventa quindi determinante.

E' riportato lo spettrogramma (Fig. 9, Fig. 10) relativo al segnale registrato e filtrato nel punto di misura 2, in cui è evidente la presenza di due frequenze cardiache caratteristiche a 12 e 15 Hz molto probabilmente associate al moto longitudinale e trasversale del cuore. La finestra temporale scelta è stata pari 0,8 s.

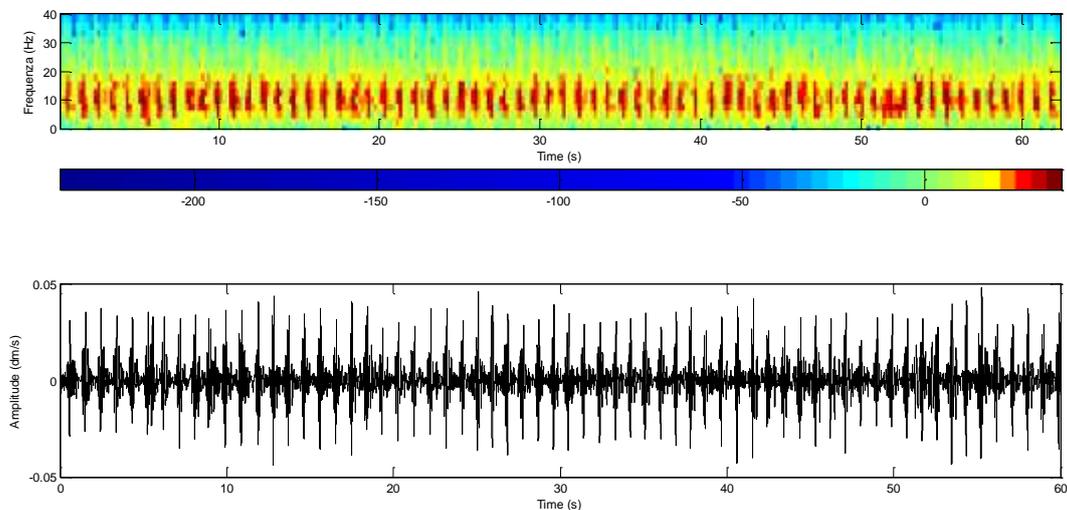


Fig. 9: Segnale registrato e filtrato e relativo spettrogramma nel punto di misura 2.

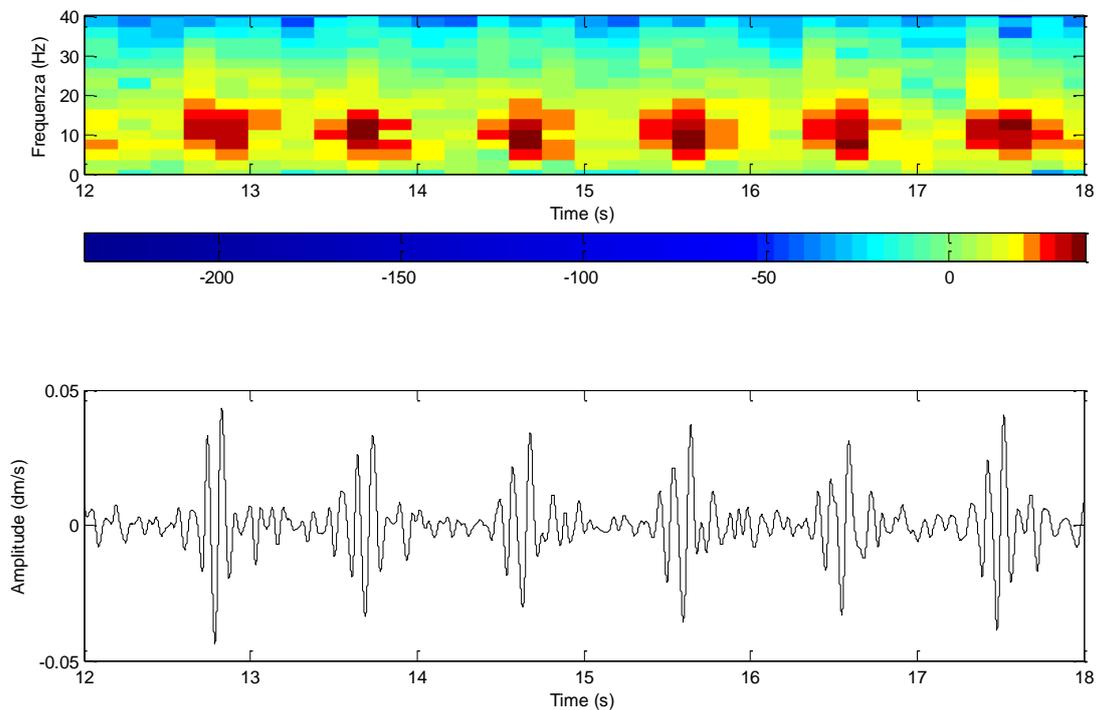


Fig.10: Zoom del segnale registrato e filtrato e relativo spettrogramma nel punto di misura 2.

4.4.2 Analisi delle componenti indipendenti (ICA)

L'analisi delle componenti indipendenti è stata applicata ai segnali registrati e filtrati nei quattro punti di misura indicati in figura 6. Vengono riportati i risultati relativi al segnale filtrato e registrato nel punto di misura 3 suddiviso in 50 tracce di circa 1s relative ad un uomo di 67 anni.

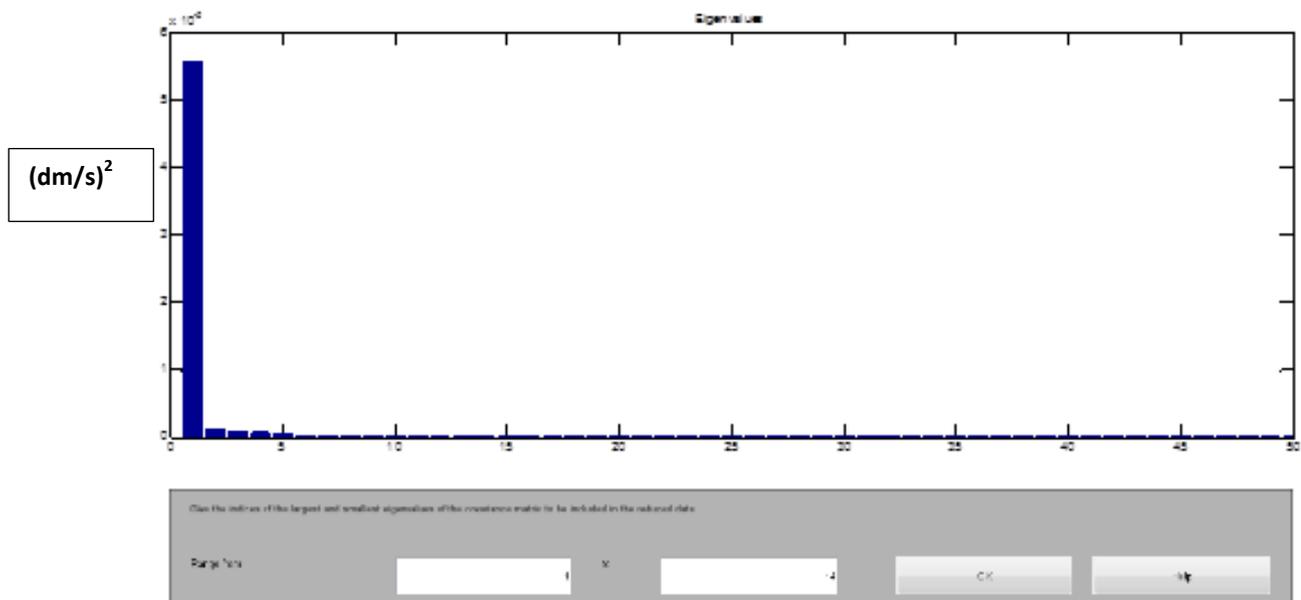


Fig.11: Analisi PCA effettuata sul segnale registrato e filtrato nel punto di misura 3.

L'analisi tramite PCA mette in evidenza che il primo autovalore contiene il 92% dell'informazione del sistema in esame e considerando i primi 5 autovalori si conserva il 99%, per questo motivo sono state estratte cinque componenti tramite l'analisi ICA di cui solo due fisicamente significative.

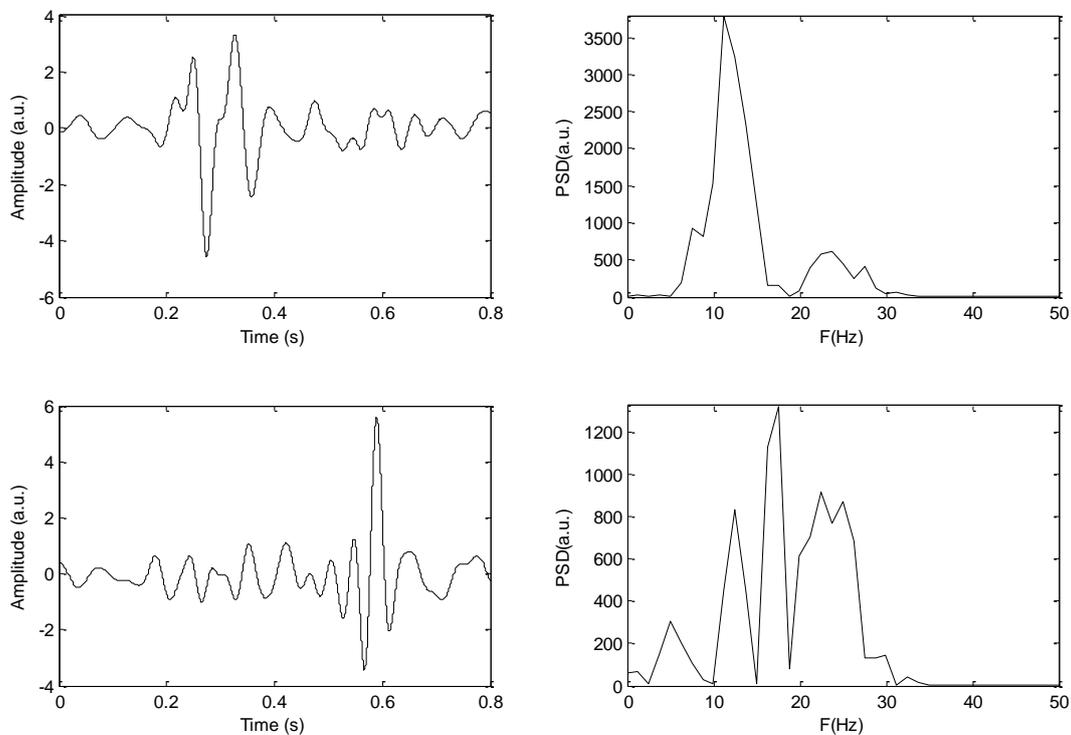


Fig.12: Analisi ICA: L'analisi ICA mostra le due componenti indipendenti a 11Hz e 18 Hz.

Tramite ICA sono stati separati due modi indipendenti piccati a 11Hz e 18Hz probabilmente associati al moto longitudinale e trasversale che rappresentano il movimento meccanico complessivo del cuore. Per poter confermare questa ipotesi si è avviata una collaborazione con il reparto di cardiologia dell'ospedale di Salerno "OO.RR. San Giovanni di Dio Ruggi d'Aragona" e si è pensato di effettuare misure contemporanee con tre vibrometri laser disposti ortogonalmente tra loro in modo da studiare i piani di oscillazione associati alle onde indipendenti.

4.4.3 La trasformata di Hilbert

Il segnale di oscillazione meccanica, associato alla vibrazione del torace dovuta all'attività cardiaca, registrato nei quattro punti di misura di figura 6 è stato suddiviso in tracce di 0.8 secondi che sono state allineate rispetto al primo massimo. In Fig. 13 è mostrato l'allineamento delle 54 tracce del segnale registrato nel punto uno. E' facile osservare un buona correlazione tra le varie tracce, con coefficiente di correlazione pari a 0.9.

Per le 54 tracce selezionate è stata determinata la trasformata di Hilbert. La trasformata di Hilbert è una particolare rappresentazione del segnale che non realizza un cambiamento del dominio di definizione, quindi a partire da una funzione del tempo la trasformata di Hilbert è ancora una funzione del tempo. In dettaglio la trasformata di Hilbert è descritta nella Appendice B.

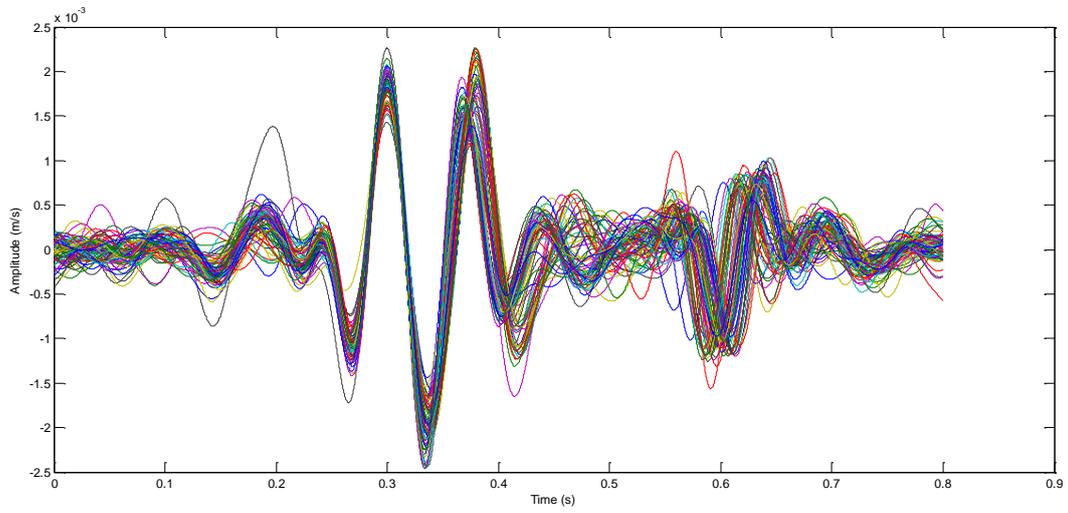


Fig. 13: Allineamento di 54 tracce del segnale registrato nel punto.

Nella figura 14 sono riportate ancora una volta le 54 tracce sulle quali sono sovrapposte le cinquantaquattro trasformate di Hilbert relative ad ogni singola traccia.

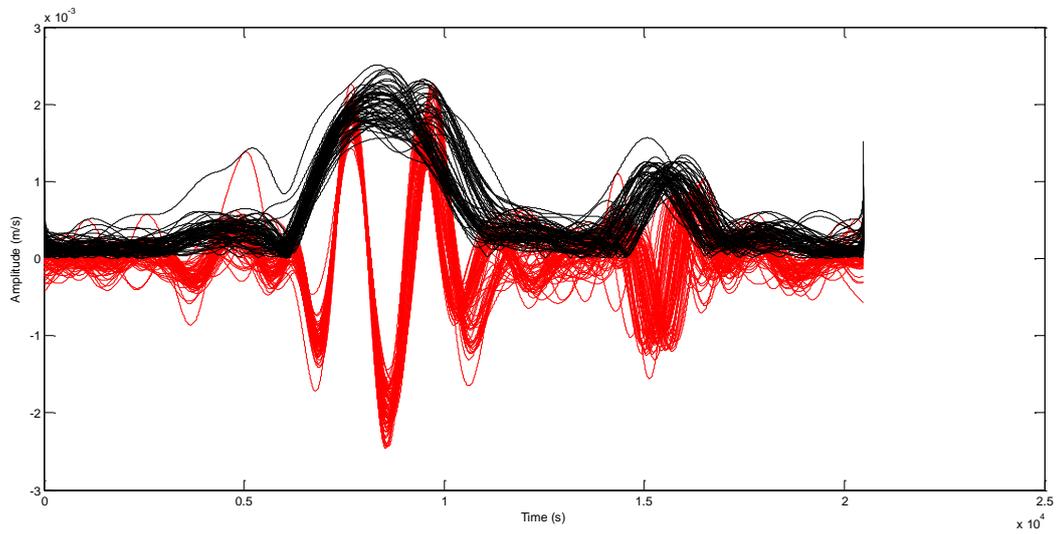


Fig.14: Le 54 tracce selezionate e le relative trasformate di Hilbert.

La figura 15 mostra la traccia media tra le tracce selezionate e la trasformata di Hilbert media tra le trasformate delle singole tracce.

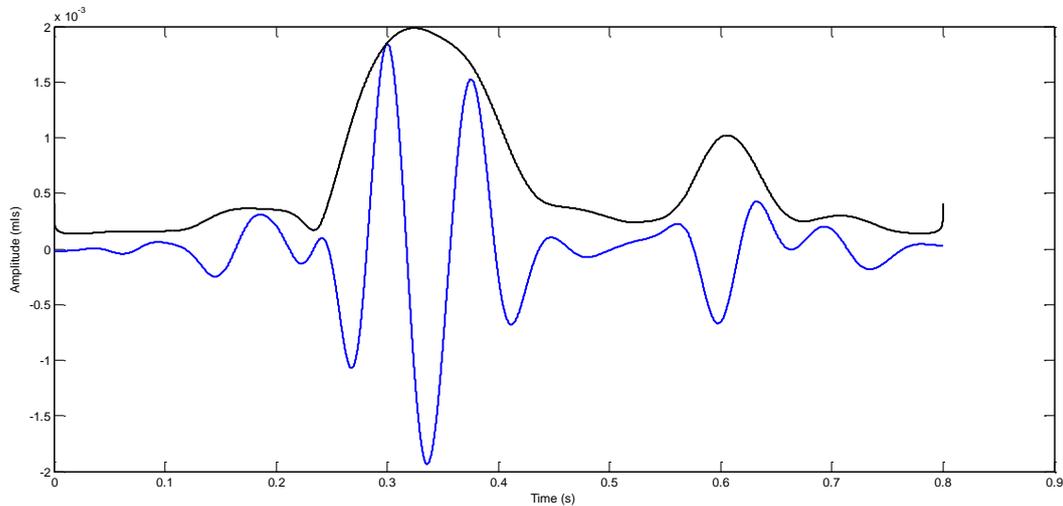


Fig.15: Traccia media e trasformata di Hilbert media.

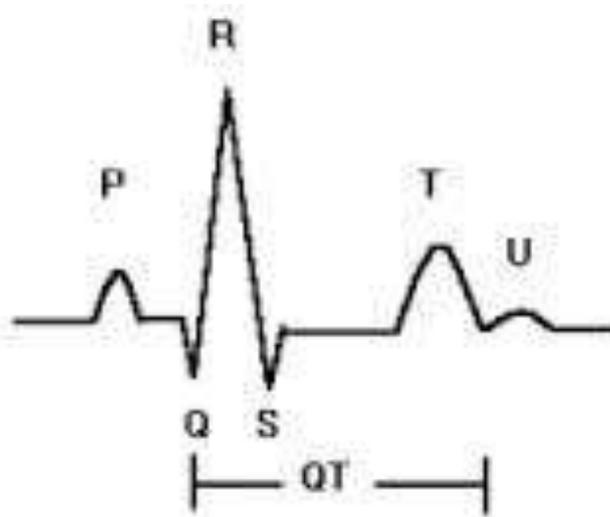


Fig. 16: Segnale elettrocardiografico.

Dalla semplice confronto visivo tra il segnale ottenuto tramite trasformata di Hilbert e il classico segnale ECG sono evidenti caratteristiche comuni nel loro andamento temporale.

Possiamo concludere che l'ECG, essendo ben riprodotto in forma d'onda dalla trasformata di Hilbert, rappresenta l'involuppo del segnale meccanico, fornendo informazioni globali del cuore. Viceversa il segnale meccanico riproduce la struttura in dettaglio del sistema e quindi conserva la

ricchezza di tutta la fenomenologia. Ulteriori sviluppi della tecnica possono risultare molto utili nella prevenzione e nel riscontro di eventuali anomalie.

Conclusioni

In questo lavoro di ricerca sono state studiate le complesse fenomenologie associate a sistemi biologici attraverso sistemi fisici vibranti non lineari per poter comprendere al meglio le loro caratteristiche e per produrre un modello teorico. I sistemi biologici studiati sono l'apparato fonatorio e il cuore. Si tratta di sistemi che generano oscillazioni globalmente organizzate da una complessa interazione tra un sistema fluido e un sistema solido. Entrambe le fenomenologie sono associate a moti collettivi che permettono di descrivere il singolo sistema in esame con pochi gradi di libertà.

Trattandosi di sistemi non lineari, lo studio fenomenologico dei segnali temporali associati comporta l'analisi attraverso metodologie diverse dall'analisi Fourier. La trasformata di Fourier diventa una importante pre-analisi, ma le analisi effettive sono state effettuate guardando i segnali nel dominio del tempo. Tecniche nate con la teoria dell'informazione come l'ICA e metodologie scaturite dalla teoria del caos e dei sistemi dinamici sono state applicate ai segnali esaminati.

Nel caso dell'apparato fonatorio sono stati studiati in dettaglio i campi di vibrazione acustica e meccanica di vari operatori, diversi per età e sesso, mentre emettono suoni vocalici. Entrambi i campi contengono informazioni sia sul sistema fluido che su quello solido coinvolti nella generazione delle autoscillazioni, per cui mettere in evidenza le differenze e le similitudini tra i campi analizzati rappresenta un passo essenziale per modellare il sistema.

Il campo di vibrazione meccanica misurato con tecniche ottiche non invasive (vibrometri laser) lungo il tratto trachea/laringe ha evidenziato un'oscillazione meccanica indipendente della trachea.

Le caratteristiche spettrali del campo di vibrazione meccanica, dalle analisi effettuate, è simile a quello di una canna d'organo aperta. Si tratta di un sistema dinamico non lineare di bassa dimensionalità in cui vengono attivati solo due modi di Landau. Il segnale laser associato alle oscillazioni del tratto tracheale risulta essere un segnale di base che oscilla alla frequenza di pitch fo. Questo risultato fa ipotizzare che l'emissione dei suoni non sia attribuibile alla sola differenza di pressione attraverso la glottide come suggerisce la teoria classica, ma che la trachea stessa è sorgente del suono vocalico. Quest'idea è stata confermata dallo studio e dal confronto del transiente d'attacco dei segnali registrati relativi al campo acustico e meccanico, la trachea inizia ad oscillare in modo indipendente e prima dei segnali acustici fino a raggiungere il regime stazionario quando tutto l'apparato vocale globalmente autoscilla. Il segnale tracheale si è valutato partire prima del segnale a livello della glottide confermando ancora una volta l'ipotesi fatta su di esso.

Il campo di vibrazione acustica è stato investigato tramite un sistema di registrazione ed una catena microfonica. I risultati sperimentali ottenuti dall'analisi dei segnali acustici, che contengono gli effetti dovuti alla presenza del vocal tract superiore, costituito dalla cavità orale e dalla cavità nasale, mette in evidenza che alcune frequenze vengono amplificate e che il segnale si arricchisce di frequenze più alte, con caratteristiche spettrali simili a quelle mostrate da un sassofono. Nel vocal tract si possono individuare due punti di instabilità idrodinamica: il primo punto è quello in cui

l'aria proveniente dai polmoni è immessa nella trachea e il secondo è il punto in cui il flusso d'aria giunge alla fine della trachea dove incontra un restringimento dovuto alla presenza delle corde vocali.

Questi risultati sperimentali hanno permesso di giungere ad un modello teorico del vocal tract, quando sono emessi suoni vocalici. Il tratto vocale superiore viene schematizzato come un sassofono, perché dall'osservazione dello spettro dei segnali acustici registrati si nota una certa similarità con lo spettro di questo strumento musicale. Procedendo verso il basso, al livello della glottide, si giunge ad un punto di instabilità idrodinamica in cui il flusso d'aria proveniente dai polmoni trova un restringimento a causa delle corde vocali che incominciano ad oscillare solidalmente con la trachea. Al di sopra delle pieghe vocali, la pressione del flusso d'aria è aumentata a causa del restringimento e si osserva: uno shift delle frequenze verso valori più alti ed un aumento delle ampiezze delle frequenze superiori (De lauro et al., 2007) . Il vocal tract inferiore è stato schematizzato come una canna che suona attivando al massimo due modi di Landau. Infine, un secondo punto di instabilità a livello della giunzione trachea polmoni si osserva ed è dovuto al flusso d'aria generato dai polmoni che mette in oscillazione la trachea.

L'obiettivo successivo dell'attività di ricerca è quello di tradurre i risultati sperimentali con un modello fisico-matematico del vocal tract attraverso il quale si possa sintetizzare in modo naturale il suono vocalico emesso da un generico operatore.

Le oscillazioni meccaniche associate ad un altro sistema biologico: il cuore, sono state studiate con due vibrometri laser doppler registrando le oscillazioni che esso induce sulla cassa toracica. L'attività cardiaca viene generalmente monitorata attraverso l'elettrocardiogramma (ECG) che fornisce informazioni sul funzionamento del cuore analizzando la sua attività elettrica, invece nell'attività di ricerca proposta il cuore viene investigato attraverso le sue oscillazioni meccaniche.

L'analisi effettuata sul segnale cardiaco ha evidenziato che si tratta ancora di un sistema non lineare a bassa dimensionalità caratterizzato da due oscillazioni indipendenti a 11 Hz ed a 18 Hz. Queste frequenze cardiache caratteristiche sono probabilmente associate al moto longitudinale e trasversale che rappresentano il movimento meccanico complessivo del cuore. Una collaborazione con il reparto di cardiologia dell'ospedale di Salerno è stata avviata per confermare quest'ipotesi.

Infine, applicando la trasformata di Hilbert ai segnali che descrivono l'attività meccanica cardiaca si è osservato un segnale che conserva caratteristiche molto simili a quelle del segnale elettrocardiografico. L'ECG, essendo ben riprodotto in forma d'onda dalla trasformata di Hilbert, rappresenta l'involuppo del segnale meccanico e ci fornisce informazioni globali sul cuore. Il segnale meccanico riproduce la struttura cardiaca in dettaglio conservando la ricchezza di tutta la fenomenologia. Ulteriori sviluppi della tecnica possono risultare molto utili nella prevenzione e nel riscontro di eventuali anomalie cardiache.

L'obiettivo successivo della ricerca in tale ambito suggerisce di poter sviluppare tecniche di confronto tra il segnale elettrocardiografico e quello di oscillazione meccanica cardiaca per ottenere un metodo alternativo all'analisi del funzionamento del cuore che potrebbe essere di supporto e/o arricchire la conoscenza delle sue caratteristiche di funzionamento.

APPENDICI

Appendice A: Analisi delle Componenti Indipendenti (ICA)

L'analisi delle componenti indipendenti è una tecnica computazionale e statistica che trasforma un set di segnali, rappresentati da una statistica multidimensionale, in una base di vettori statisticamente indipendenti, riducendo al minimo la perdita di contenuto informativo.

L'ICA nasce nell'ambito della blind source separation, dove il fine è di individuare un numero di segnali statisticamente indipendenti conoscendo le loro misture.

Il tipico esempio che si fa in merito è costituito dal “problema del cocktail party”. Immaginiamo un certo numero di persone che parlano contemporaneamente in una stanza e che vi siano sistemati dei microfoni, ognuno dei quali è capace di ricevere una diversa sovrapposizione delle voci emesse dalle persone. Lo scopo dell'ICA è di ricostruire, attraverso l'uso delle sole misture acquisite dai microfoni, i singoli segnali vocali indipendenti emessi dalle persone.

Il problema è così formalizzato: si considerano alcuni segnali emessi da differenti sorgenti fisiche e captati da un certo numero di sensori che, posti in differenti posizioni, registrano una mistura di segnali delle sorgenti con pesi differenti. In tale modello si assume che il mixing sia istantaneo, in modo tale da trascurare qualsiasi ritardo temporale dovuto alla propagazione del suono nel mezzo dalla sorgente al sensore. Si può scrivere, pertanto:

$$x = As + n \quad [A.1]$$

dove x è il vettore osservato m -dimensionale, s è un vettore casuale n -dimensionale le cui componenti sono assunte mutuamente indipendenti, A è una matrice $m \times n$ da stimare e n è il rumore additivo, che spesso viene ommesso perché incluso nella somma come segnale sorgente. Si assume, inoltre, che il numero di misture m sia maggiore o, al meno, uguale al numero di sorgenti n e che i segnali di una sola sorgente abbiano distribuzione gaussiana, poiché risulta impossibile separare due o più sorgenti gaussiane [14].

In tale procedura, si aggiorna una matrice B $m \times n$ così che il vettore

$$y = Bx \quad [A.2]$$

risulta essere una stima $y \approx s$ dei segnali delle sorgenti indipendenti originali.

Esistono numerosi algoritmi per l'implementazione dell'ICA. In seguito si farà uso dell'algoritmo a punto fisso chiamato FastICA che cerca una direzione, ad esempio un vettore unitario w tale che la proiezione $w^T x$ massimizza l'indipendenza della singola sorgente stimata y , misurata qui dall'approssimazione della negentropy:

$$JG(w) = [E\{G(w^T x)\} - E\{G(v)\}]^2 \quad [A.3]$$

dove w è un vettore (peso) m -dimensionale, x rappresenta la mistura dei segnali, v è una variabile casuale gaussiana standardizzata e $E\{(w^T x)^2\} = 1$. Massimizzando JG , si ottiene una componente indipendente. Si devono, in ogni caso, precedentemente, “sbiancare” i dati: si deve, cioè, trasformare linearmente la variabile x in una variabile a media zero $v = Qx$ in modo tale che $E\{(v v^T)\} = I$. Tale pre-processing dei dati può essere sempre eseguito, ad esempio, dall'Analisi delle Componenti Principali [12].

Per ricercare un vettore riga w , si applica il seguente algoritmo a punto fisso:

$$\mathbf{w}^* = E[\mathbf{v}g(\mathbf{w}_i^T \mathbf{v})] - E[g'(\mathbf{w}_i^T \mathbf{v})]\mathbf{w}_i$$

$$\mathbf{w}_i = \mathbf{w}_i^* / \|\mathbf{w}_i^*\| \quad [\text{A.4}]$$

con $g(\cdot)$ è un'opportuna nonlinearity, in ogni caso $g(y) = \tanh(y)$ e $g'(y)$ è la sua derivata rispetto a y . Tale algoritmo a unico punto determina, in tal modo, soltanto una delle componenti indipendenti. Per stimare le altre, lo si deve iterare tramite vari punti con vettori peso $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$. Per evitare che i vettori convergano allo stesso massimo, si devono decorrelare i vari risultati $\mathbf{w}_1^T \mathbf{x}, \mathbf{w}_2^T \mathbf{x}, \dots, \mathbf{w}_n^T \mathbf{x}$ dopo ogni iterazione.

Se si assume che i dati siano stati sbiancati, si ha

$$\mathbf{W} = \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-0.5} \quad [\text{A.5}]$$

in cui \mathbf{W} è la matrice dei vettori $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ e la radice quadrata inversa è ottenuta dalla decomposizione agli autovalori

$$(\mathbf{W}^T \mathbf{W})^{-0.5} = \mathbf{E} \mathbf{D}^{-0.5} \mathbf{E}^T$$

dove \mathbf{E} è la matrice degli autovettori e \mathbf{D} è quella diagonale degli autovalori.

Appendice B: Trasformata di Hilbert

La Trasformata di Hilbert è una particolare rappresentazione che, contrariamente ad altre trasformate (Fourier, Laplace, Z, ...) non realizza un cambiamento del dominio di definizione. In altre parole, a partire da una funzione del tempo $s(t)$, la trasformata di Hilbert $\tilde{s}(t)$ è ancora una funzione del tempo. Essa è definita:

$$\tilde{s}(t) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \left[\frac{s(\tau)}{t-\tau} d\tau \right] \quad [B.1]$$

Si ha la convoluzione tra $s(t)$ e $1/t-\tau$ e P.V. indica l'integrale va inteso come valore principale di Cauchy.

L'operazione di trasformazione di Hilbert è concettualmente ed operativamente semplice: $\tilde{s}(t)$ si ottiene come uscita da un filtro, detto di Hilbert, caratterizzato dalla seguente funzione di trasferimento:

$$H_H(f) = \begin{cases} -j & \text{per } f > 0 \\ 0 & \text{per } f = 0 \\ j & \text{per } f < 0 \end{cases} \quad [B.2]$$

in termini di parte reale e parte immaginaria, si ha:

$$|H_H(f)| = \begin{cases} 1 & f \neq 0 \\ 0 & f = 0 \end{cases} \quad [B.3]$$

$$\arg\{H_H(f)\} = \begin{cases} -\pi/2 & \text{per } f > 0 \\ 0 & \text{per } f = 0 \\ \pi/2 & \text{per } f < 0 \end{cases} \quad [B.4]$$

in termini di modulo e fase.

Nel dominio della frequenza si ha dunque:

$$\tilde{S}(f) = H_H(f)S(f) \quad [B.5]$$

dove $S(f)$ è la trasformata di Fourier di $s(t)$ mentre $\tilde{S}(f)$ è la trasformata di Fourier di $\tilde{s}(t)$.

La trasformata di Hilbert $\tilde{s}(t)$ può dunque essere ottenuta dapprima utilizzando la [B.5] e quindi antitrasformando, secondo Fourier, il risultato del prodotto, oppure direttamente nel dominio del tempo, come integrale di convoluzione tra il segnale $s(t)$ e la risposta impulsiva $h_H(t)$ del filtro di Hilbert [B.1]. Quest'ultima si ricava facilmente e vale:

$$h_H(t) = 1/\pi t \quad [B.6]$$

In considerazione del fatto che $H_H(f)$ è una funzione puramente immaginaria e dispari, era lecito attendersi che $h_H(t)$ fosse puramente reale e dispari. Tutto ciò dalle proprietà della trasformata di Fourier.

L'operazione di trasformazione inversa di Hilbert, che consente di riottenere il segnale $s(t)$ a partire da $\tilde{s}(t)$, richiede, in realtà, una nuova trasformata di Hilbert. E' immediato, infatti, verificare che

$$\tilde{\tilde{s}}(f) \approx H_H(f) \tilde{S}(f) = H_H(f) H_H(f) S(f) = -S(f) \quad [B.7]$$

Se ne conclude che la trasformata di Hilbert della trasformata di Hilbert restituisce, a meno del segno, il segnale originale. A rigore, quindi, l'operazione di antitrasformazione è identica a quella di trasformazione con l'introduzione di un cambiamento di segno nel risultato. Va a questo punto precisato che quanto sopra vale per tutti i valori di $f \neq 0$. In $f = 0$, infatti, in ragione delle [B.2], [B.3] e [B.4], il valore di $S(0)$, se diverso da zero, viene annullato, e non potrà essere più recuperato (in particolare dall'operazione di antitrasformazione [B.7]). Questa puntualizzazione definisce la condizione che deve essere verificata affinché un segnale $s(t)$ sia "Hilbert trasformabile": è necessario che la sua trasformata di Fourier sia nulla in $f = 0$. Visto che:

$$S(0) = \int_{-\infty}^{+\infty} s(t) dt$$

è proporzionale al valor medio del segnale $s(t)$, si può concludere che la classe dei segnali per cui è applicabile la trasformata di Hilbert è quella dei segnali a valor medio nullo.

In realtà, la definizione fornita per il filtro di Hilbert è ideale: le funzioni di trasferimento [B.2], [B.3] e [B.4], infatti, presentano una transizione brusca per $f = 0$. Un filtro reale (e quindi realizzabile) potrà soltanto approssimare tale definizione nell'intorno dell'origine per cui, onde evitare la comparsa di distorsione nel segnale ricostruito, si dovrà imporre che il segnale $s(t)$ non solo presenti valor medio nullo ma, in aggiunta, non abbia componenti armoniche significative per un opportuno intervallo di frequenze nell'intorno dell'origine.

BIBLIOGRAFIA

1. Airy G.B. (1830). "On certain condition under which a perpetual motion is possible", T. Cam. Philos. Soc. 3, 369–372.
2. Alipour-Haghighi F. and I. R. Titze. Simulation of particle trajectories of vocal fold tissue during phonation. In I. R. Titze and R. C. Scherer, editors, *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control*, pages 183-190. Denver Center for the Performing Arts, Denver, 1983.
3. Avanzini F. Computational issues in physically-based sound models. PhD thesis, Università degli Studi di Padova, Dipartimento di Elettronica e Informatica, 2002.
4. Avanzini F., Maratea S., and Drioli C. Physiological control of low-dimensional glottal models with applications to voice source parameter matching. *Acta Acustica united with Acustica*, 92(5):731-740, September/October 2006.
5. Avanzini F. Low-dimensional modeling of the vocal folds: analysis of large-amplitude oscillations and application to a modified one-mass model. *Speech Communication*, 2007.
6. Berry D. A. and Titze I. R. Normal modes in a continuum model of vocal fold tissues. *Journal of the Acoustical Society of America*, 100:3345-3354, November 1996.
7. Cataldo E., Leta F. R., Lucero J. C. and Nicolato L. (2006). "Synthesis of voiced sounds using low-dimensional models of the vocal cords and time-varying subglottal pressure", *Mech. Res. Commun.*, 33, 250.
8. Childers and Wong C.F. Measuring and modeling vocal source-tract interaction. *IEEE Transactions on Biomedical Engineering*, 41(7):663-671, July 1994.
9. De Lauro E., De Martino S. , Falanga M., Ciaramella A. and Tagliaferri R. (2005) Complexity of time series associated to dynamical systems inferred from independent component analysis, *Phys. Rev. E* 72, 046712.
10. De Lauro E., De Martino S. , Esposito E., Falanga M., and Tomasini E. P, "Analogical model for mechanical vibrations in flue organ pipes inferred by independent component analysis", *J. Acous. Soc. Am.* 122(4), 2413-24 (2007)..
11. Falanga M. (1999). *Oscillatori autosostenuti e dinamica eruttiva di Stromboli*. Tesi di Laurea in Fisica, Università degli Studi di Salerno.

12. Falanga M. (2003). Costruzione di modelli fisici a partire da dati osservazionali. Tesi di Dottorato di Ricerca in Fisica (XV ciclo- I ciclo nuova serie), Università degli Studi di Salerno.
13. Fant G. (1960). *Acoustic Theory of Speech Production*, The Hague, Netherlands: Mouton, 2nd edition (1970).
14. Flanagan J. L., *Speech Analysis, Synthesis and Perception*. New York: Academic Press, 1965
15. Ferrero F. E., Magno-Caldognetto E., Vaggies K. and Lavagnoli C. (1978). "Some acoustic Characteristics of Italian vowels", *J. Italian Linguistics Phonology and Phonetics* 3,87–96.
16. Ferrero F.E., Magno Caldognetto E. and Cosi P. (1996). "Sui piani formantici acustici e uditivi delle vocali di uomo, donna, bambino", in A. Peretti, P. Simonetti (acd), *Atti del XXIV Convegno Nazionale dell'Associazione Italiana di Acustica*, Trento, pp. 169–178.
17. Fletcher N.H. and Rossing T.D. (2000), *The Physics of Musical Instruments*, Eds. Springer-Verlag, New York. 7 Fraser A.M., Swinney H.L. (1986). "Independent coordinates for strange attractors from mutual information", *Phys. Rev. A* 33, 1134-1139.
18. Ganong ,Barrett ,Barman ,Boitano , Brooks - *Fisiologia medica di Ganong*- Editore: Piccin Edizione: XII (2011) Volume: Unico.
19. Hyvärinen A., Oja J., A fast fixed-point algorithm for independent component analysis, *Neural Computation*, 9,7,1483-1492, 1997.
20. Hyvärinen A., Karhunen J., Oja J., *Independent Component Analysis*, Wile & Son, Inc, 2001.
21. Ishizaka K. and Flanagan J. L. (1972). "Synthesis of Voiced Sounds From a Two-Mass Model of the vocal cords", *Bell. System. Technical. J.* 51, 1233–1268.
22. Héroult J. and Ans B., *Circuits neuronaux à synapses modifiables: décodage de messages composites per apprentissage non supervise*. C.- R. de l'Accadémie des Sciences, 299 (III-13), 525-528, (1984)
23. Ishizaka K. and Matsudaira M. Fluid mechanical considerations of vocal fold vibration. In *Monogr. 8. Speech Commun. Res. Lab., Santa Barbara*, 1972.
24. Iwarsson J., Thomasson M. and Sundberg J. (1998). "Effects of lung volume on the glottal voice source" *J. Voice*, 12, 424–433).
25. Kent R. D. (1993). "Vocal tract acoustics", *J. Voice* 7, 97–117.
26. Larsen O. N. and Goller F. (2002). "Direct observation of syringeal muscle function in songbirds and a parrot", *J. Exp. Biol.*, 205(Pt 1) 25-35.
27. Maratea S. *Regole per il controllo di modelli glottali a bassa dimensionalità*. Master's thesis, Università degli Studi di Padova, Dipartimento di Elettronica e Informatica, 2003-2004.

28. Magnani S.- Vivere di voce: l'arte della manutenzione della voce per chi parla, recita e canta- Franco Angeli srl 2010
29. Marchetto E.- Inversione di un modello fisico dell'apparato fonatorio mediante programmazione dinamica e reti – Università degli studi di Padova-2007.
30. Maurel A. , Ern P., Zielinska B.J.A. and Wesfreid J. E. (1996). "Experimental study of self-sustained oscillations in a confined jet", *Phys. Rev. E*, 54(4), 3643–3651.
31. Primo Caimmi P., Grossini E. - Dispense di Fisiologia- ARACNE editrice S.r.l., 2009.
32. Sondhi M. M. and Schroeter J. A hybrid time-frequency domain articulatory speech synthesizer. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-35(7):955-966, July 1987.
33. Story Brad H., Titze Ingo R. Voice simulation with a body-cover model of the vocal folds. *Journal of the Acoustical Society of America*, 97(2):1249-1260, February 1995.
34. Story Brad H., Titze Ingo R. and Hoffman Eric A. "Vocal tract area functions from magnetic resonance imaging", *J. Acoust. Soc. Am.* 100, 537 (1996).
35. Teager H. M. and Teager S.M. (1983). A Phenomenological Model for Vowel Production in the Vocal Tract, in R.G. Daniloff (ed.) *Speech Sciences: Recent Advances*, College Hill Press, San Diego, CA.
36. Teager H. M. and Teager S.M. (1990). A Phenomenological Model for Vowel Production in the Vocal Tract, in *Speech Production and Speech Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic, London) pp.241-262.
37. Titze I. R. (1973). "The human vocal cords: A Mathematical Model part.I", *Phonetica* 28, 129–170.
38. Titze I. R. (1974). "The human vocal cords: A Mathematical Model part.II" *Phonetica*, 29, 1–21.
39. Titze I. R. and Story B.. Acoustic interactions of the voice source with the lower vocal tract. *Journal of the Acoustical Society of America*, 101(4), April 1997.
40. Titze I. R. and Story B. H. Rules for controlling low-dimensional vocal fold models with muscle activation. *Journal of the Acoustical Society of America*, 112(3): 1064-1076, September 2002.
41. Van den Berg J. (1958). "Myoelastic-aerodynamic theory of voice production", *J. Speech Hear. Res.*, 1, 227–244.
42. Verge M. P., Fabre B., Mahu W. E. A., Hirschberg A., van Hassel R.R., Wijnands A. P. J., de Vries J. J. and Hogendoorn C. J. (1994). "Jet formation and jet velocity fluctuations in a flue organ pipe", *J. Acoust. Soc. Am.*, 95, 1119-32.
43. Wegel R.L. (1930). "Theory of vibration of the larynx", *J. Acoust. Soc. Am.* 1, 1–21.

44. Willis R. (1833). "On the theory of larynx", T. Cam. Philos. Soc. 4, 323–352.
45. Wong D., Ito M. R., and Cox N. B. Observation of perturbations in a lumped element model of the vocal folds with application to some pathological cases. Journal of the Acoustical Society of America, 89(1):383394, January 1991.
46. Rangaraj M. Rangayyan , "Biomedical signal analysis"- University of Calgary, Canada – IEE Press Series on Biomedical Engineering 2002.

Siti internet:

www.iocsanmarcello.gov.it/.../wp.../Lapparato-cardiocircolatorio.pdf

www.centrolottainfarto.org/saperne...cuore/59-lelettrocardiografia.html

PRODUZIONE SCIENTIFICA

1. G. Buccheri, E. De Lauro, S. De Martino, M. Falanga, Experimental investigation of trachea oscillation and its role in the pitch production, sottomesso a Journal of Sound and Vibration.
2. G. Buccheri, E. De Lauro, S. De Martino, M. Esposito M. Falanga C. Fontanella- Identification of soil redistribution by using ^{137}Cs for characterizing landslides-prone areas: a case study in Sarno-Quindici, Italy – on line: Febbraio 2014 – DOI:10.1007/s12665-014-3120-1- Environmental Earth Sciences.
3. G. Buccheri, E. De Lauro, S. De Martino, M. Falanga, M. Palo, From the standard Strombolian activity to the volcanic crises at Stromboli (Italy): an overview of the seismic signatures, pubblicato in: "Complex Monitoring of Volcanic Activity: Methods and Results", V. Zobin (ed). Nova Science Publishers Inc., New York, 2012, ISBN: 978-1-62417-985-3.
4. G. Buccheri, E. De Lauro, M. Falanga, S. Petrosino, Source characterization of lp events and background seismic noise at Campi Flegrei during 2006 swarm. In: D. Slejko et al.. Geodinamica. p. 200-206, Gruppo Nazionale della Terra Solida, ISBN: 9788890210112, Potenza, 20-22 Novembre 2012