

UNA NUOVA IDEA DI “IMPRONTA VOCALE”

COME STRUMENTO IDENTIFICATIVO E RIABILITATIVO

Autori: Marco Farinella, Marco Carnaroglio, Fabio Cian
Istituto Mod.A.I.® di Torino – segreteria@istitutomodai.it

INTRODUZIONE

Da sempre l'uomo usa il proprio istinto empatico per decifrare i segnali verbali della comunicazione, nel tentativo di riconoscere i propri simili, comprenderne le intenzioni o svelarne le menzogne. L'opinione soggettiva, però, in sede giudiziaria è irrilevante ai fini processuali. Con l'avvento della tecnologia è diventato quindi indispensabile studiare l'*impronta vocale*. Timbro, intonazione, posizione degli accenti, velocità di elocuzione, durata ed epentesi vocalica e consonantica, variazione intra e interlinguistica sono diventati i nuovi campi di ricerca. Numerosi studi hanno indagato tutti questi parametri utilizzando la *stima della frequenza*, i modelli di *misture gaussiane* (GMM) e quelli di *Markov nascosti* (HMM), gli algoritmi di *pattern matching*, le *reti neurali*, le *matrici di rappresentazione*, la *quantizzazione vettoriale*, e gli *alberi di decisione*. Tuttavia, nonostante le varie teorie formulate e le metodiche utilizzate (*cohort model*, *modelli ambientali*, algoritmi di riduzione del rumore, etc...), permangono molti problemi. Le condizioni di ripresa, i sottofondi ambientali, i comportamenti, le inflessioni linguistiche, gli umori, lo stato di salute e l'età del soggetto possono infatti inficiare sia la fase di *raccolta*, sia quella di *verifica* degli elementi, rendendo alquanto controversa la qualità dei risultati ottenuta elettronicamente (Gold e French 2019). In ambito forense ciò rende indispensabile l'affiancamento di esperti con allenamento all'ascolto e conoscenze tecniche specifiche (Romito, Galatà, 2008) affinché l'esito dell'analisi possa ritenersi accettabile in termini probabilistici (Grimaldi, d'Apolito, Gili Fivela, Sigona 2014). Questo contributo propone una diversa e personale interpretazione del fenomeno sonoro durante la verbalizzazione, nell'attesa che la scienza individui nuove metodologie d'integrazione fra le letture strumentali e la percezione umana che possano affiancare, in maniera affidabile, l'indispensabile contributo del tecnico forense.

SCOPO DEL LAVORO

Lo scopo del presente lavoro è approfondire la correlazione fra l'*impronta vocale* e la *fisiologia* che l'ha prodotta, per individuare una *configurazione frequenziale* atta ad identificare e migliorare gli aspetti *funzionali* ed *emotivi* peculiari di una persona.

MATERIALI E METODI

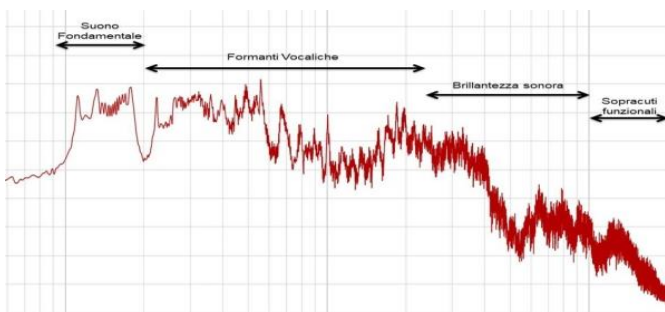


Figura 1: In rosso lo spettro di Fourier di una voce maschile che pronuncia la frase “le mie aiuole”. In nero la suddivisione delle aree frequenziali per l'analisi vocale.

Fra il 2003 e il 2020, attraverso un'osservazione funzionale, si è indagato l'atteggiamento laringeo di un campione, composto da 1061 persone (33% uomini e 67% donne), nell'atto della fonazione. Le età medie sono state di 47 anni (fra i 7 e i 91 anni) per i maschi e di 48 anni (fra i 12 e gli 85 anni) per le femmine. Le caratteristiche acustiche delle voci sono state raffrontate con la condizione di salute *fisica e psicologica*, con le *abitudini sociali* e i *comportamenti professionali* degli individui esaminati. Per praticità d'analisi l'intera gamma delle *frequenze udibili* dall'uomo è stata suddivisa nei seguenti gruppi (identificativi della regione anatomica in cui si sviluppano): *suono fondamentale* (prodotto dalle pliche vocali), *formanti vocaliche* (proprie della cavità orale), *brillantezza sonora* (influenzata dall'onda della mucosa), *sopracuti funzionali* (frutto della fisiologia degli apparati). Si è

proceduto con la tradizionale analisi *prosodico-intonativa* applicandola però a tutti i gruppi, monitorando quindi la derivata di F0 e dell'intera gamma di armonici nelle bande di maggiore rilevanza. Le indagini sono servite per enucleare un *pattern sonoro* ideale, di perfetto funzionamento laringeo, da attribuire ad un ipotetico adulto, in perfetta salute, scevro da traumi fisici e psicologici, in condizioni emotive di equilibrata serenità. A questo riferimento *teorico* si è attribuito un valore arbitrario, di *affaticamento glottico*, pari allo 0%. Successivamente, utilizzando un passaggio d'aria durante il meccanismo di rilassamento della *chiusura glottale* (Greene e Mathieson, 2001) è stata misurata la risonanza laringea normale delle laringi a riposo. Comparando statisticamente le *trasformate di Fourier* (FFT) reali mediate dell'intera frase con il *modello ideale* si sono evidenziate le discordanze dovute alle costrizioni laringee tipiche del soggetto considerato ed attribuito a ciascuno un valore di *affaticamento vocale* espresso in percentuale. Tale punteggio rappresentava la condizione fonatoria quotidiana dell'individuo, nel periodo osservato, generando anche una nuova concezione di “impronta vocale” più vicina a fattori biologici che ad elementi linguistici (Lindh, 2004). Infine, per valutare la qualità dei dati raccolti e la bontà delle conclusioni, si è deciso di testare anche il modello di riferimento, invertendo l'ottica d'indagine. Fra il 2005 e il 2011, 171 volontari sono stati sottoposti ad un'unica stimolazione percettiva che inducesse la loro voce ad assomigliare al *pattern sonoro ideale*. L'idea era di elicitarne sensorialmente le voci verso un *punto zero fisiologico di affaticamento* e sondarne la retroazione, funzionale ed emotiva, impattante sulla percezione del sé, sull'identità e sul conseguente modello comportamentale. Per la valutazione è stato usato un test con venticinque coppie di aggettivi contrapposti (ispirato a Costa e McCrae, 1992), disposti in maniera casuale, che gli esaminandi compilavano prima e dopo la pratica di vocalizzazione.

RISULTATI

Paragonando l'intero spettro sonoro delle voci con il FFT ideale si è determinato il grado di funzionalità delle prestazioni vocali senza ricorrere ai tradizionali esami clinici che, per la loro complessità, sono eseguibili soltanto in apposite strutture sanitarie con personale qualificato. Ciò, di fatto, ha snellito notevolmente la procedura. Inoltre si è evidenziato come la percentuale di discostamento dal *modello di riferimento* tenda a mantenersi costante per ogni esaminato a parità di condizioni psicofisiche, caratterizzando quindi l'oratore, pur non essendo ancora riusciti a parametrizzare la lettura strumentale per il riconoscimento univoco del parlante senza l'intervento umano. Tuttavia, il risultato certamente più sorprendente è stato ottenuto utilizzando il *pattern vocale ideale* come stimolazione per il miglioramento della prestazione fonatoria. Infatti, l'esito dell'esperimento ha messo in luce la sua grande capacità rieducativa (priva di effetti collaterali) e la forte retroazione sull'identità dei volontari stimolati. L'affaticamento glottico si è notevolmente ridotto ed i test psicologici hanno conseguentemente evidenziato una variazione drastica del quadro emotivo, con due importanti traguardi: l'eliminazione dei vocaboli con un'accezione negativa nella descrizione del proprio stato psicofisico (ad es. teso, pesante, stanco, triste...) ed un incremento della propriocezione dei soggetti verso gli aggettivi positivi (attivo, riposato, allegro, concentrato...) con percentuali di miglioramento che vanno dal 6% fino al 53%, con una media globale di poco inferiore al 30% ottenuta in un'unica seduta di stimolazione.

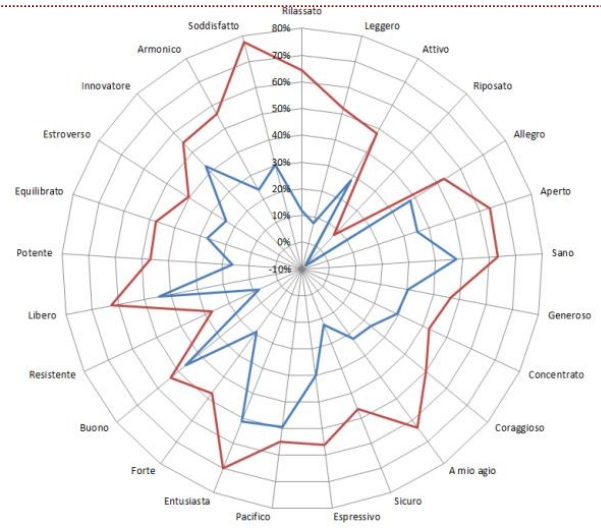


Figura 2: In blu la condizione emotiva dei soggetti prima della stimolazione, in rosso il cambiamento propriocettivo al termine della seduta

CONCLUSIONI

La voce, nel ventaglio delle sue ventimila frequenze e grazie alle oltre ventisei milioni d'informazioni empatiche al secondo che esse portano con sé, è certamente ascrivibile all'identità individuale della persona. La muscolatura del tratto vocale reagisce istintivamente alle emozioni come primaria difesa dei polmoni. Questo fa sì che la risonanza di ogni persona sia diversa e assimilabile ad una impronta digitale. Il vissuto di un individuo resta registrato nella sua biologia, caratterizzandone la personalità emotiva, comportamentale e quindi anche quella sonora globale. Per esaminare la condizione *psicofisica* di un individuo si ritiene pertanto necessario considerare l'intera configurazione frequenziale, valutando il differenziale della sua *impronta vocale mediata di una intera asserzione* (intesa nell'accezione del presente studio) da una *trasformata di Fourier ideale*. Perfezionando questa tecnica, non si esclude una futura applicazione in ambito forense, non tanto per validare le intercettazioni (normalmente troppo disturbate da rumori parassiti che interferiscono con l'esame audio), quanto per valutare il grado di *stress laringeo* durante una dichiarazione giurata. L'incremento di tensione muscolare nel pronunciare una frase (rilevato paragonando l'*analisi della voce* durante una deposizione con la propria *impronta laringea*) potrebbe risultare assai prezioso, come parametro oggettivo, per valutare la veridicità di una dichiarazione. Tuttavia, in attesa di creare una procedura automatizzata e sicura per l'osservazione tecnica ipotizzata, questa nuova concezione di *impronta vocale ideale* può già essere utilizzata con successo come strumento riabilitativo, sia dal punto di vista fisiologico, sia psicologico. Infatti, quando è impiegata come *modello acustico*, all'interno di una *pedagogia sensorialmente orientativa*, essa è in grado di *rieducare* velocemente una funzione glottica deficitaria e di retroagire fortemente sull'umore, con effetti duraturi di miglioramento sulla percezione del sé. La voce così riabilitata aumenta il suo grado di funzionalità fisiologica e l'identità delle persone trattate risulta più definita, equilibrata e risoluta.

BIBLIOGRAFIA

- Costa P.T., McCrae R.R., in *Psychological Assessment Resources* (1992)
- Federico A., Paoloni A., in *Media Duemila 250* (2007)
- Gold E., French P., in *International Journal of Speech Language and the Law* (2019)
- Greene M. e Mathieson L., "The Voice and its Disorders" (2001)
- Grimaldi M. et al., in *Mondo digitale* (2014)
- Hamidi M. et al., in *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)* (2020)
- Kelly F. et al., in *Audio Engineering Society Conference: 2019 AES International Conference on Audio Forensics* (2019)
- Lindh J., "Handling the Voiceprint Issue" (2004)
- Romito L., Galatà V., in *Language Design*, vol. Special issue I, (2008)
- Saleem S. et al., in *Forensic Science International: Digital Investigation* (2020)
- Zetterholm E., "Voice Imitation. A Phonetic Study of Perceptual Illusions and Acoustic Success", PhD thesis, Lund University. (2003)