

MARCO FARINELLA, MARCO CARNAROGLIO, FABIO CIAN

Una nuova idea di “impronta vocale” come strumento identificativo e riabilitativo

A new concept of “voiceprint” as identification and rehabilitation tool

The concept of voiceprint always arouses many perplexities: it has been criticized several times in the past and it has therefore been archived. What if we made a mistake? In this study we tried to make a completely innovative use of it, observing it from a new perspective. We analysed the voices of more than a thousand subjects. We divided the spectrum of the voice into four categories, extrapolating refined biological information. In particular we analysed the adduction of the vocal cords, the use of the oral cavity, the mucosal wave and the physiological aspects of the apparatus. This new idea of voiceprint has proved to be an unexpectedly effective rehabilitation tool. We do not rule out that it may be reconsidered for use in forensic field in the future.

Keywords: voiceprint, FFT, laryngeal stress, identity, rehabilitation.

1. Introduzione

La voce dell'uomo è un fenomeno fortemente correlato alla sua identità. Le sensazioni provate, l'educazione ricevuta, le esperienze passate, la psicologia, le convinzioni raggiunte, la cultura acquisita, le condizioni fisiche, le intenzioni, le credenze, le emozioni e molto altro sono elementi che indubbiamente caratterizzano il parlato. Tuttavia, analizzando unicamente il processo fisico, possono essere sussunti gli elementi significativi che intervengono nel processo fonatorio. Essi sono: il cervello che determina il programma motorio, il corpo fisico che produce il fenomeno vibratorio e la voce, ovvero l'evento sonoro finale. I tre ambiti sono stati ampiamente indagati dal punto di vista scientifico da svariate discipline che, da un lato, hanno permesso di comprenderne i meccanismi di base e, dall'altro, hanno messo in luce la complessità del fenomeno e la necessità di nuove forme di approccio concettuale e strumentale.

Sappiamo che il connettoma, cioè l'organizzazione strutturale e funzionale di neuroni e sinapsi fra aree corticali e subcorticali (White, Southgate, Thomson & Brenner, 1986; Varshney, Chen, Paniagua, Hall & Chklovskii, 2011) è un meccanismo assai articolato che, oltre ad occuparsi della gestione pratica dei movimenti, racchiude anche intricati aspetti psicologici ed emotivi. L'effetto di questi ultimi elementi sul sistema nervoso è noto da lungo tempo perché rappresenta il tema principale nell'Arte che, da sempre, ne sfrutta l'impatto emozionale pur avendo solo una conoscenza superficiale della fisiologia che ne sta alla base.

La foniatra e la logopedia, invece, hanno approfondito scientificamente la vocalizzazione umana, puntando costantemente la propria attenzione sui meccanismi patologici. Ciò ha permesso di comprendere e correggere molte delle disfunzioni legate alla voce ma ha portato anche a trascurare parzialmente gli aspetti eufonici ed eutonici della prestazione. Di conseguenza, gli studi sulla funzionalità laringea risultano molto rari, datati e deficitari. Inoltre i soggetti-cavia coinvolti sono quasi sempre poco numerosi e sottoposti ad osservazioni di durata limitata (Stolze, 1995).

Infine, la fisica acustica si è occupata di analizzare minuziosamente il fenomeno sonoro, scomponendo la voce in parametri sempre più dettagliati e precisi (Kelly, Forth, Kent, Gerlach & Alexander, 2019). Questo tipo di pensiero ha certamente aperto nuovi orizzonti sulla concezione del segnale audio, però ha altresì reso faticoso definire con certezza l'identità del parlante, le sue emozioni ed ancor più le sue intenzioni e sensazioni, parametri che richiedono chiavi di lettura interdisciplinari di notevole difficoltà. Tutto questo impone che si avvicendino ricerche ininterrotte e che si aprano continui dibattiti sui risultati raggiunti, al fine di arrivare ad accertare uno standard di analisi che possa essere preciso, veritiero e condivisibile (Hamidi, Satori, Laaidi & Satori, 2020).

La decodifica della voce è comunque materia antica. 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. In molte epoche, vicine e lontane, le parole sono state usate come armi investigative, spesso strumentalizzandole in interpretazioni apofeniche, come è successo nei periodi più bui della recente storia dell'uomo e sotto le dittature. È però inammissibile, in sede giudiziaria, che l'opinione soggettiva abbia rilevanza ai fini processuali. Ecco perché, con l'avvento della tecnologia, è diventato indispensabile studiare metodi efficaci per identificare la falsità, appagando così la necessità di una valutazione *super partes* adatta ad un'aula di tribunale (IlPost.it, 2015).

Inizialmente l'attenzione si è rivolta verso gli effetti secondari legati alla fisiologia. Tralasciando i primi maldestri tentativi di Angelo Mosso e Cesare Lombroso, possiamo datare il poligrafo moderno (più volgarmente chiamato *macchina della verità*) al 1913, quando è nato per mano di William Moulton Marston, venendo poi successivamente perfezionato da Larson. Purtroppo la sua efficacia non è mai stata dimostrata pienamente (De Cataldo Neuburger, 2007; Germanà Tascona, 2005) e, per questo motivo, è rifiutato come prova dall'ordinamento giuridico moderno, sebbene sia tollerato da alcune giurisdizioni statunitensi come prova indiziaria, quando concordata fra le parti.

L'interesse degli addetti ai lavori si è quindi inevitabilmente spostato verso il suono e, in particolare, sull'*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 (Saleem, Subhan, Naseer, Bais & Imtiaz, 2020). Numerosi studi hanno indagato tutti questi parametri utilizzando la stima della frequenza, i modelli di misture gaussiane (GMM) e quel-

li 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 tecniche utilizzate (*cohort model*, modelli ambientali, algoritmi di riduzione del rumore, etc...), permangono molti problemi (Zetterholm, 2003). 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, French, 2019). In ambito forense ciò rende indispensabile l'affiancamento di esperti con abilità d'ascolto specifiche (Romito, Galatà, 2008) affinché l'esito dell'analisi possa ritenersi accettabile in termini probabilistici (Grimaldi, d'Apolito, Gili Fivela & Sigona, 2014). L'impronta vocale è stata quindi abbandonata come indice di riconoscimento in favore di un calcolo matematico che possa orientare una giuria riguardo all'identità di un parlante, lasciando alla disamina umana qualsiasi giudizio in merito alla veridicità del contenuto a cui si rifà la conversazione (Tonini, 2005). Questo contributo quindi, utilizzando un vecchio strumento, propone una diversa interpretazione del fenomeno sonoro durante la verbalizzazione, nell'attesa che la scienza individui metodologie d'integrazione fra le letture strumentali e la percezione umana che possano affiancare, in maniera affidabile, l'indispensabile apporto del tecnico forense.

2. *Scopo del lavoro*

Lo scopo del presente lavoro è approfondire la correlazione fra l'intera gamma di frequenze emesse con la voce umana e la fisiologia che l'ha prodotta, per individuare una configurazione atta ad identificare e migliorare gli aspetti funzionali ed emotivi peculiari di una persona. Ciò permette di dare nuova vita al superato concetto di impronta vocale che, opportunamente mappata (alla stregua del genoma umano) potrebbe svelare, con precisione sorprendente, peculiarità fisiche ed emotive di un soggetto.

3. *Prima fase*

3.1 Premessa

Partendo dal presupposto che l'uomo riproduce con la voce soltanto ciò che riesce ad udire (Tomatis, 1987) e lo fa attraverso un procedimento esclusivamente biologico, l'idea è di trovare una stretta correlazione fra l'intero spettro sonoro e la fisiologia che lo ha prodotto, individuando così una personale impronta acustica.

Se il tratto vocale maschile fosse un cilindro ricurvo, di sezione costante, la sua risonanza sarebbe caratterizzata perlopiù da modi dispari a partire da 500 Hz. Nel caso di voci femminili o di bambini la fondamentale risulterebbe un po' più acuta ma seguente la stessa logica. Nella realtà si aggiungono l'oscillazione delle corde vo-

cali (diverse da individuo a individuo) che caratterizza la nota fondamentale emessa e la deformazione volontaria del tratto che sposta le formanti vocaliche dando leggibilità al linguaggio. Il processo si completa con l'onda della mucosa che produce armonici sopracuti in funzione della dinamica e della tensione cordale, nonché del contributo delle cavità del cranio che possono filtrare, rinforzare o assorbire una vasta gamma di frequenze in base al loro utilizzo.

Di conseguenza, considerando la regione anatomica, il corpo vibrante, le modalità di risonanza, la pressione aerobica e quella muscolare è dunque possibile individuare precise correlazioni fra essi e lo spettro di Fourier, il quale denuncia così non solo la presenza frequenziale del fenomeno sonoro ma anche le condizioni psicobiologiche del corpo che le ha prodotte (Rohmert, 1995).

3.2 Raccolta dati

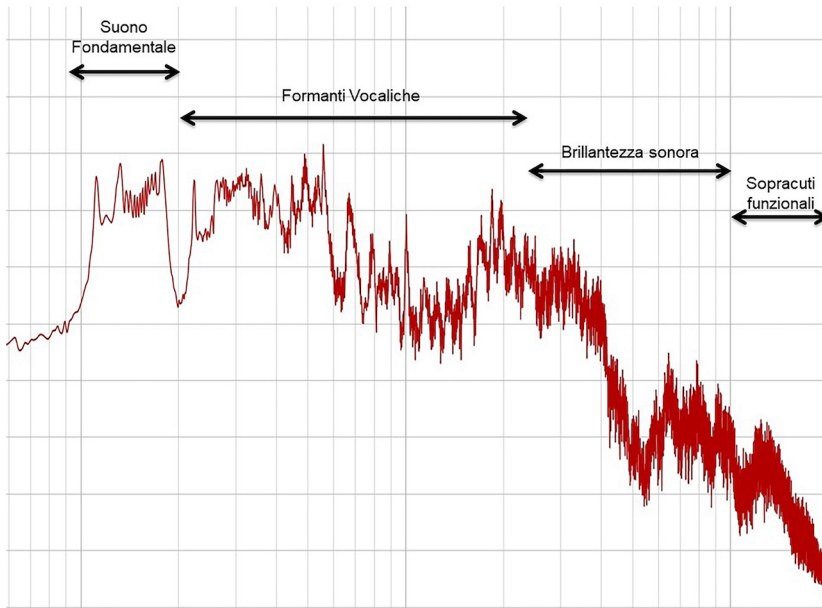
Fra il 2003 e il 2020, attraverso un'osservazione funzionale, si è indagato l'atteggiamento laringeo (seguendo i principi espressi da Rohmert, 1981) di un campione composto da 1061 persone (33% uomini e 67% donne) nell'atto della fonazione. Le età medie sono state di 47 anni (fra i 7 e i 91 anni) per i maschi e di 48 anni (fra i 12 e gli 85 anni) per le femmine.

Le condizioni di ascolto delle singole voci sono state omogenee nel complesso ed identiche per gruppi di 10-12 persone circa. La strumentazione utilizzata per le analisi è comprensibilmente variata nel corso dei 17 anni di ricerca, adeguandosi al progredire della tecnologia ma tutte le attrezzature impiegate (hardware e software) sono sempre state di fascia professionale e manovrate da operatori qualificati.

Inoltre, avendo a disposizione i soggetti (e non semplicemente registrazioni di conversazioni telefoniche) le principali analisi effettuate riguardano: la determinazione della *frequenza fondamentale* (f_0) e dell'ampiezza di emissione vocale ed il relativo andamento nel tempo (curva di intonazione e di intensità); l'identificazione delle *perturbazioni del periodo fondamentale* (*Jitter*) e dell'*ampiezza* (*Shimmer*), lo studio del *bilancio energetico spettrale* (rapporto fra componente *periodica* ed *aperiodica* nel segnale: *Harmonic to Noise Ratio*), il calcolo della presenza di *diplofonia* (semplice o multipla) e degli *arresti momentanei* dell'emissione (*Vocal Breaks*), nonché la *fonetografia* per l'identificazione dell'*estensione e tessitura* vocale.

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 di elaborazione dei dati l'intera gamma delle frequenze udibili dall'uomo è stata suddivisa nelle seguenti aree (identificative 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) (Rohmert, Rehders & Rohmert, 1990).

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



3.3 Analisi ed elaborazione dei dati

I risultati più interessanti sono derivati dalla tradizionale analisi prosodico-intonativa applicata però a tutte le aree, monitorando quindi la derivata di f_0 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. Utilizzando un passaggio d’aria durante il meccanismo di rilassamento della *chiusura glottale* (Greene, Mathieson, 2001) è stata misurata la risonanza laringea a riposo di ogni partecipante. Questa indagine ha avuto esiti con risultati sostanzialmente sovrapponibili indipendentemente dal numero di misurazioni eseguite per singolo soggetto. A questo riferimento teorico si è attribuito un valore arbitrario di *affaticamento glottico* pari allo 0%. Comparando statisticamente le trasformate di Fourier (FFT) reali dell’intera frase concordata (“le mie aiuole”) con il modello ideale si sono evidenziate le discordanze dovute alle costrizioni laringee tipiche di ciascun soggetto considerato, attribuendo ad ognuno 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).

3.4 Risultati (prima fase)

L'idea di dare una nuova vita al concetto di impronta vocale si è dimostrata sensata. Lavorando con voci analizzate a pieno spettro si è potuto attingere ad una ricchezza frequenziale impensabile per le conversazioni telefoniche. Ciò ha permesso di identificare delle correlazioni sonore con la biologia dei parlanti che normalmente vengono trascurate in ambito forense e considerate solo in campo medico in relazione alle patologie.

Paragonando quindi l'intero spettro sonoro delle voci con la 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. Questo, 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.

4. *Seconda fase*

4.1 Premessa

Per valutare la qualità dei dati raccolti e la bontà delle conclusioni, si è deciso di eseguire un secondo test sul modello di riferimento appena individuato. Si è provato quindi ad invertire l'ottica d'indagine e, 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*. Le persone impiegate nell'esperimento sono state per il 30% uomini e 70% donne con una età media di 45 anni in entrambi i sessi. Il reclutamento dei volontari è avvenuto volutamente in maniera casuale all'interno delle seguenti categorie professionali: insegnanti, artisti, casalinghe, disoccupati, professionisti, manager e dirigenti, impiegati, imprenditori, operatori sanitari, medici, psicologi, rieducatori, doppiatori, operatori dello spettacolo, sportivi, personal trainer, formatori. Circa il 50% di essi necessitava di rieducazione, il restante 50% partiva da una condizione di salute. Tutti i partecipanti erano accomunati da uno spiccato interesse per la voce (parlata o cantata).

4.2 Raccolta ed elaborazione dei dati

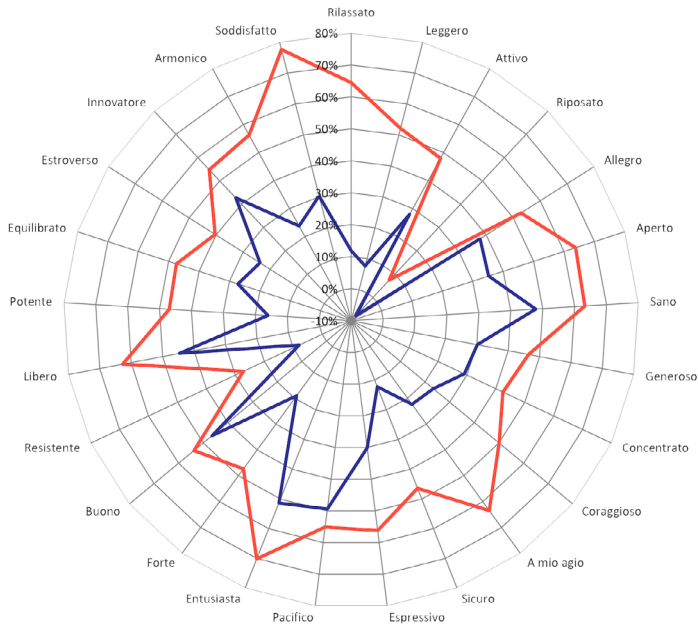
I dati sono stati raccolti da gruppi eterogenei di 10-12 persone convocate per singola seduta. Prima di iniziare, ogni partecipante ha compilato un questionario riguardante il proprio stato emotivo e sensoriale. Per la valutazione è stato usato un test con venticinque coppie di aggettivi contrapposti (ispirato a Costa, McCrae, 1992), disposti in maniera casuale. Il modulo per l'indagine è stato preparato da una psicologa professionista esterna alla struttura, nel rispetto

di tutti i criteri necessari e convenzionalmente accettati per la corretta stima dei risultati. Il gruppo è stato poi sottoposto ad un *training* uditivo intensivo (Rohmert, 1990). Ad ogni aderente è stato chiesto di vocalizzare singolarmente per 45 minuti durante i quali il proprio ascolto è stato sollecitato dai parametri funzionali identificati nel *pattern sonoro* (Rohmert, 1991). L'idea era di elicitarne sensorialmente le voci verso un *punto zero fisiologico* di affaticamento e sordarne la retroazione, funzionale ed emotiva, impattante sulla percezione del sé, sull'identità e sul conseguente modello comportamentale. La percezione delle *formanti*, unita alle stimolazioni binaurali, monaurali e sincroniche hanno progressivamente codificato il suono parlato e cantato dei soggetti cavia. Sfruttando il concetto di *sintonizzazione cerebrale* decodificato all'Università di Cambridge (Leong, Byrne, Clackson, Georgieva, Lam & Wass, 2017) si è assistito ad una modificazione plastica della struttura cerebrale degli aderenti con l'attivazione di onde Alpha (8-13 Hz), Theta (4-7 Hz) e Delta (<3,5 Hz) rilevabili da elettroencefalogramma (EEG). Al termine l'intero gruppo è stato nuovamente invitato a compilare, in forma anonima, il test riguardante la percezione soggettiva, emotiva e sensoriale dell'esperienza vissuta.

4.3 Risultati (seconda fase)

Il risultato certamente più sorprendente si è raggiunto 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*



5. Conclusioni

Oggi la comunicazione rappresenta oltre il 50% dei problemi della nostra società al punto che è considerata anche il fattore critico delle strategie d'impresa delle aziende (Passerini, Tomatis, 2007). Il motivo è che ormai ci troviamo ad abitare un mondo caratterizzato da un eccessivo rumore di fondo, un frastuono d'informazioni in cui tutti gridano la propria verità e in cui è quasi impossibile distinguere la qualità del messaggio. "Paradossalmente, l'ipercomunicazione ha creato un blocco della comunicazione stessa. Più si grida, meno si comunica e meno ci si capisce" (Passerini, Tomatis, 2007: 17). Nella società di oggi, oltre alle sordità fisiche e psicologiche (sia nei confronti dell'ambiente circostante, che di se stessi), siamo arrivati a quantificare i danni prodotti persino da un inquinamento semiotico, caratterizzato da nuove malattie sociali, come il "Tautismo: un autismo tautologico" (Passerini, Tomatis, 2007: 18), una sorta quindi di impossibilità di comprendersi. È facile immaginare cosa abbia creato questa difficoltà di comunicazione nelle aule di tribunale (Grevi, 2003).

Arrivati ormai alla babele si è cercato in tutti i modi di tornare a capirsi, riscoprendo la semantica delle parole, mutuando termini da altre lingue e persino coniando nuovi vocaboli. Nulla però è servito a riportare una piena comprensione del prossimo perché è impossibile arrivare ad interagire efficacemente con gli altri se prima non si giunge a conoscere appieno la propria capacità percettiva. Solo così un dialogo può diventare un rapporto profondo fatto di parole essenziali. Infatti,

all'interno di una buona comunicazione è importante tenere conto del fattore soggettivo di chi emette e di chi riceve l'informazione. "Nei processi comunicativi esiste una sorta di autocomunicazione" (Passerini, Tomatis, 2007: 30) che ha a che fare più con il nostro "percepito" che con la realtà. Si potrebbe dire quindi che le persone, più che vere informazioni, si scambiano continuamente impressioni. Questo aspetto è ciò che maggiormente mina la comprensione della voce umana, diventando anche un fattore invalidante nel caso di analisi strumentali (Zetterholm, 2003). Migliorare consapevolmente l'ascolto è, allora, l'unica possibilità che un tecnico forense ha per svelare il vero senso della comunicazione.

Ne consegue che l'attenzione degli studiosi si è concentrata sulle caratteristiche del suono emesso, tentando di individuare le affinità che intercorrono fra la voce umana e la rosa di emozioni che guida gli atteggiamenti. Da subito la materia ha rivelato tutta la sua complessità rendendo particolarmente arduo il compito (Federico, Paoloni, 2007). Si è tentato quindi di riferirsi ad elementi che potessero essere più facilmente identificabili e che potessero diventare indici per il riconoscimento delle emozioni in corso. Tuttavia, la percezione umana è assai complessa. L'uomo infatti può incamerare fino a 60 milioni d'informazioni al secondo, anche se le persone mediamente ne traggono, dall'ambiente circostante, appena 2-3 mila ed il loro cervello cosciente riesce a gestirne contemporaneamente soltanto 5-9 (Miller, 1956). Da questi numeri si evince quanto qualunque azione istintiva possa essere più efficace rispetto a qualsiasi atto tecnico volontario. Il nostro principio base, dunque, è quello di attingere alla ricchezza delle risorse istintive dove si trova il 95% dei dati utili ad innescare reazioni automatiche. Dei 60 milioni di informazioni al secondo che l'uomo riceve il 54% riguarda il corpo fisico e solo il 46% viaggia attraverso il suono. Di conseguenza il linguaggio corporeo viene spesso preso come principale indicatore nelle analisi comportamentali ed emotive. Tuttavia, in questo 54%, ad eccezione di una piccola parte di indicatori (che riguarda il colore della pelle, la quantità di sudorazione, etc...), la maggior parte delle informazioni è facilmente travisabile se non addirittura manipolabile, *in primis* la postura, come insegnano le tecniche di comunicazione aziendale. Il residuo 46% delle informazioni invece, spesso trascurato e decisamente sottovalutato, viene trasportato dal suono. Qui, soltanto il 7%, che si riferisce al valore semantico delle parole, è di fatto manipolabile e possibile fonte di menzogna, mentre il restante 39% riguarda suoni non verbali, istintivi, che difficilmente si prestano alla falsità. In questo ambito vengono esplicitati anche i bisogni più intimi ed i pensieri più nascosti. Si tratta quindi di una preziosa rappresentazione sensoriale, dinamica e sinergica dell'organizzazione cerebrale dell'interlocutore, al punto che si potrebbe parafrasare il vecchio detto "dimmi come parli e ti dirò chi sei!".

Imparare a leggere l'impronta vocale, quindi, risulta una risorsa preziosissima, considerato che essa trasporta oltre 26 milioni di informazioni al secondo, di natura purissima! Di fronte a numeri del genere è impossibile, per chiunque debba analizzare una voce, non soffermarsi a riflettere sull'importanza del saper dirigere adeguatamente la propria attenzione uditiva sulle frequenze che la caratterizzano

e sui relativi parametri maggiormente ricchi di significato. Il timbro vocale è una precisa e scrupolosa carta d'identità che mostra al nostro interlocutore una realtà che va oltre le parole pronunciate. Considerando poi che il suono, a differenza dell'impronta digitale, oltre ad identificare il soggetto ne definisce dettagliatamente anche ogni sua sensazione, atteggiamento e attitudine, comprendere i segnali vocali diventa uno strumento più che mai necessario in ambito forense.

Durante una deposizione tutto ciò che avviene si fa principalmente con la voce. Occorre tener presente che scambiare informazioni durante un colloquio non significa necessariamente comunicare. La vera comunicazione, infatti, fa leva sulle intenzioni inconscie prima ancora che diventino azioni e lo fa attraverso un profondo processo di comprensione, mettendo al centro dell'attenzione i propri interlocutori. In questo scenario saper ascoltare e, di conseguenza interagire, risulta quindi fondamentale per accertare la verità. Il tecnico forense non dovrebbe giudicare il contenuto della dichiarazione, compito per il quale viene deputata un'apposita giuria. Per colui che è designato ad analizzare i parametri sonori è indispensabile tornare a dare voce a tutti i livelli di comunicazione, svelando il piano biologico alla stregua di quello culturale. Del resto, se l'era dell'*ipercomunicazione* oggi ci impone di riscoprire la genuinità del messaggio, perché non fare il passaggio successivo abbracciando la neuro-comunicazione anche in tribunale?

La voce, nel ventaglio delle sue ventimila frequenze e grazie agli oltre ventisei milioni di informazioni empatiche al secondo che esse portano con sé, è certamente ascrivibile all'identità individuale della persona. È quindi necessario imparare a leggerla in tutto il suo contenuto. 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 un'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 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 stimare 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 (Rohmert, 1989). 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.

Riferimenti bibliografici

- COSTA, P.T., McCRAE, R.R. (1992). *NEO PI-R professional manual*. Odessa: Psychological Assessment Resources.
- DE CATALDO NEUBURGER, L. (2007). *La prova scientifica nel processo penale*. Padova: Cedam.
- FEDERICO, A., PAOLONI, A. (2007). *Riconoscimento del parlante*. In *Media Duemila*, 250, 47-55.
- GERMANÀ TASCONA, N. (2005). Sub art. 188. In CORSO, P. (a cura di). *Commento al codice di procedura penale*. Piacenza: La tribuna.
- GOLD, E., FRENCH, P. (2019). International practices in forensic speaker comparisons: second survey. In *International Journal of Speech Language and the Law*, 26(1), 1-20. <https://doi.org/10.1558/ijll.38028>
- GREENE, M., MATHIESON, L. (2001). *The Voice and its Disorders*. New York City: Wiley.
- GREVI, V. (2003). *Prove*. In CONSO, G., GREVI, V. (a cura di). *Compendio di procedura penale*. Padova: Cedam.
- GRIMALDI, M., D' APOLITO, S., GILI FIVELA, B. & SIGONA, F. (2014). Illusione e scienza nella fonetica forense: una sintesi. In *Mondo digitale*, 13(53), 1-9.
- HAMIDI, M., SATORI, H., LAAIDI, N. & SATORI, K. (2020). Conception of Speaker Recognition Methods: A Review. In *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*. IEEE, 1-6.
- IlPost.it*. Come ingannare la macchina della verità. (2015). <https://www.ilpost.it/2015/04/12/macchina-della-verita-funziona/> Accessed 06.04.21.
- KELLY, F., FORTH, O., KENT, S., GERLACH, L. & ALEXANDER, A. (2019). Deep neural network based forensic automatic speaker recognition in VOCALISE using x-vectors. In *Audio Engineering Society Conference: 2019 AES International Conference on Audio Forensics*. Porto, Portogallo, 18-20 giugno 2019.
- LEONG, V., BYRNE, E., CLACKSON, K., GEORGIEVA, S., LAM, S. & WASS, S. (2017). Speaker gaze increases information coupling between infant and adult brains. In *Proceedings of the National Academy of Sciences of the United States of America*, 114(50), 13290-13295. <https://doi.org/10.1073/pnas.1702493114>
- LINDH, J. (2004). Handling the “Voiceprint” Issue. In *Proceedings FONETIK 2004*, Stockholm, Sweden, 26-28 May 2004, 72-75.
- MILLER, G.A. (1956). *The magical number seven, plus or minus two: Some limits on our capacity for processing information*. In *Psychological Review*, 63, 81-97.
- PASSERINI, W., TOMATIS, A. (2007). *Management dell'ascolto. Tutto ciò che le persone e le organizzazioni devono sapere prima di comunicare e per comunicare meglio*. Milano: FrancoAngeli editore.

- ROHMERT, G. (1995). *Il cantante in cammino verso il suono. Leggi e processi di autoregolazione nella voce del cantante*. Treviso: Diastema libri.
- ROHMERT, W. (1981). Physische Beanspruchung durch muskuläre Belastungen. In SCHMIDTKE, H. (1981), *Lehrbuch der Ergonomie*, München: Hanser.
- ROHMERT, W. (1989). *Grundzüge des funktionalen Stimmtrainings*. Köln: Schmidt.
- ROHMERT, W. (Ed.) (1990). *Beiträge zum 1. Kolloquium Praktische Musikphysiologie*. Köln: Schmidt.
- ROHMERT, W. (Ed.) (1991). *Beiträge zum 2. Kolloquium Praktische Musikphysiologie*. Köln: Schmidt.
- ROHMERT, W., REHDEERS, H. & ROHMERT, G. (1990). Alexandertechnik. Auswirkungen der Körperaufrichtung auf Klangspektrum der menschlichen in Stimme. In ROHMERT, G. (1990), *Beiträge zum 1. Kolloquium Praktische Musikphysiologie*. Köln: Schmidt.
- ROMITO, L., GALATÀ, V. (2008). Speaker Recognition in Italy: Evaluation of Methods used in Forensic cases. In *Language Design*, Special issue, 1, 229-240.
- SALEEM, S., SUBHAN, F., NASEER, N., BAIS, A. & IMTIAZ, A. (2020). Forensic speaker recognition: A new method based on extracting accent and language information from short utterances. In *Forensic Science International: Digital Investigation*, 34, 300982. <https://doi.org/10.1016/j.fsidi.2020.300982>
- STOLZE, H. (1995). Prefazione. In ROHMERT, G. (1995), *Il cantante in cammino verso il suono. Leggi e processi di autoregolazione nella voce del cantante*. Treviso: Diastema libri.
- TOMATIS, A. (1987). *L'Oreille et la voix*. Robert Laffont. [Trad. it., TOMATIS A. (1993), *L'orecchio e la voce*. Milano: Baldini e Castoldi].
- TONINI, P. (2005). *Manuale di procedura penale*. Milano: Giuffrè.
- VARSHNEY, L.R., CHEN, B.L., PANIAGUA, E., HALL, D.H. & CHKLOVSKII, D.B. (2011). Structural properties of the *Caenorhabditis elegans* neuronal network. In *PLoS computational biology*, 7(2), e1001066. <https://doi.org/10.1371/journal.pcbi.1001066>
- WHITE, J.G., SOUTHGATE, E., THOMSON, J.N. & BRENNER, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. In *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, 314(1165), 1-340. <https://doi.org/10.1098/rstb.1986.0056>
- ZETTERHOLM, E. (2003). Voice Imitation. A Phonetic Study of Perceptual Illusions and Acoustic Success. PhD Dissertation, Lund University.