

L'identificazione delle persone nell'ambito giudiziario è compito del tribunale. Ne consegue che l'esperto non deve esprimersi in termini di identificazione positiva o negativa, ma in termini di rapporto di verisimiglianza. Purtroppo l'attualità ci porta ad esprimere preoccupazione per le modalità utilizzate da alcuni esperti nello svolgimento del compito loro demandato, sia per quanto riguarda la scelta del metodo di identificazione, sia per la presentazione delle conclusioni raggiunte. Sarebbe di grande utilità fare chiarezza su come la consulenza di riconoscimento del parlante debba essere svolta e come i risultati debbano essere presentati.

Andrea PAOLONI, laureato in Ingegneria Elettronica, dal 1974 lavora presso la Fondazione Ugo Bordoni dove attualmente ricopre la carica di Responsabile dei Progetti sul "Trattamento Automatico della Lingua". Ha collaborato con l'Autorità Giudiziaria in circa cinquecento consulenze sull'analisi dei nastri, sulla trascrizione di segnali rumorosi e sul riconoscimento del parlante; ha collaborato e collabora con il RaCIS nello studio delle tecnologie di fonetica forense, è tra gli ideatori del sistema di riconoscimento del parlante denominato IDEM.

SUL RICONOSCIMENTO DEL PARLANTE IN AMBITO FORENSE

di Andrea PAOLONI

Premessa

Le prove scientifiche rivestono un ruolo sempre più importante nel processo penale. Ne consegue il rischio che il termine "scientifico" possa fare pensare che il risultato ottenuto attraverso l'espletamento della prova scientifica, abbia la caratteristica della verità. Le prove scientifiche sono soggette, come tutte le altre tipologie di prova, ad errore.

Per alcune di esse tuttavia esiste un metodo codificato e ampiamente validato (uno *standard*), seguendo il quale l'errore rimane confinato in un ambito ristretto e soprattutto noto. Prendiamo ad esempio l'impronta digitale: il confronto è soggetto a precise norme da rispettare affinché il confronto sia valido con errore definito. La Cassazione, con la sentenza n. 10567 del 13 novembre 1985 (Sezione 2, Pres. Savina, Rel. Della Penna), ha stabilito che l'identità dei due impronte è ritenuta prova solo "attraverso l'esistenza di almeno 16-17 punti caratteristici uguali per forma e posizione" e che "siano state trovate almeno 16 corrispondenze tra le impronte a confronto". Le corrispondenze vengono evidenziate nelle due impronte nell'ambito della consulenza come si può vedere nella figura 1.



Figura 1 – evidenza delle 16 corrispondenze che devono essere trovate tra due impronte affinché l'identità sia ritenuta una prova.

La probabilità che esista un'impronta con le stesse caratteristiche appartenente a un altro individuo è da ritenersi pressoché trascurabile. Naturalmente anche per questa prova esistono problemi legati alla possibilità di realizzare false impronte tramite calco e alla inutilizzabilità dell'impronta in quanto corrotta o disponibile in frammento troppo piccolo.

Un'altra tecnologia, quella del DNA, è altrettanto definita nelle modalità di analisi (misura delle polimerasi) e nelle modalità di identificazione basate sulla statistica Bayesiana volta a definire un rapporto di verisimiglianza (Likelihood Ratio, LR).

L'approccio Bayesiano, riassumibile nella nota formula:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

riscritta in termini di rapporto di scommessa (odds):

$$O(A/B) = LR \times O(A)$$

ci dice che la probabilità a posteriori (dopo l'esperimento) è data dalla probabilità a priori (prima dell'esperimento) moltiplicata per il rapporto di verisimiglianza (LR). Ora il problema non è solamente il calcolo di LR, ma anche la stima della probabilità a priori. In alcuni casi, come per il DNA, dove LR è molto elevato, la probabilità a priori ha un peso limitato, nel senso che basta presupporre che una qualsiasi persona (compreso il sospettato) sia il possibile possessore del DNA (1/intera popolazione mondiale) per giungere a probabilità di identificazione rilevanti.

Purtroppo vi sono molte prove scientifiche la cui standardizzazione non è compiuta e questo probabilmente è dovuto alla non unicità del metodo identificativo utilizzato o dalla sua non ancora risolta soggettività. Questo è, ad esempio, il caso della psichiatria forense, ma anche del confronto delle manoscritte e del confronto delle voci nell'identificazione del parlante.

Per queste due ultime biometrie bisogna sottolineare che, a differenza del DNA e dell'impronta digitale, che sono caratteristiche fisiologiche impossibili o quasi da modificare da parte della persona sospettata, la manoscrittura e la voce possono essere alterate dal sospettato, trattandosi di biometrie comportamentali, le quali prevedono l'apprendimento e consentono un controllo da parte del soggetto.

Il superamento della variabilità dei campioni, sia per la mano scritta sia per il confronto vocale, avviene utilizzando la statistica che consente di confrontare la similarità dei due campioni con quella della popolazione di riferimento. Così operando il risultato presenta sia incertezze nella valutazione della similarità, sia incertezze nella definizione della popolazione di riferimento.

A questo punto riteniamo opportuno ricordare che l'identificazione delle persone nell'ambito giudiziario è compito del tribunale. Ne consegue che l'esperto non deve esprimersi in termini di identificazione positiva o negativa, a seconda del risultato dell'esperimento da lui effettuato, ma in termini di rapporto di verisimiglianza, dato che andrà a incrementare o decrementare l'ipotesi identificativa formulata dal giudice. Questo il modo con cui l'esperto si deve esprimere secondo quanto stabilito dall'ENFSI (*European Network of Forensic Science Institutes*). Come abbiamo già anticipato, il rapporto di verisimiglianza (LR) è il rapporto tra l'operatore che stima la similarità dei campioni a confronto, ovvero quanto sia simile il campione del noto rispetto a quello dell'anonimo - al numeratore - e di un operatore che stima la tipicità, ovvero quanto il campione all'esame sia comune nella popolazione di riferimento - al denominatore. Supponendo che i valori di similarità siano dello stesso ordine di grandezza, il rapporto dipende da quanto il campione sia comune all'interno della popolazione.

Facciamo un esempio usando un parametro antropometrico, l'altezza: il fatto che i campioni a confronto abbiano altezza di 2 m è certamente molto più significativo di quanto lo sarebbe se l'altezza fosse di 1,70 m grazie alla diversa "tipicità". Purtroppo non tutti gli esperti si esprimono in termini di rapporto di verisimiglianza, riteniamo che alcuni di essi nemmeno ne conoscano l'esistenza. Qualche esperto si limita a fornire il numeratore del rapporto, ossia la somiglianza dei campioni del noto con quello dell'anonimo, a volte spacciandola per probabilità di identificazione. Sul tema preparazione degli esperti chiamati a svolgere il ruolo di periti o consulenti ricordiamo il lavoro del prof. Trumper [Trumper et alii 1996] dove erano riportate alcune affermazioni come la seguente: "*sufficiente probabilità è determinata dalla eterogeneità dei fonemi*" (pp 130/94 Tribunale di Cosenza), affermazioni che fanno capire come la preparazione di qualche consulente sia, a dir poco, inadeguata.

Il quadro rappresentato ci porta ad esprimere preoccupazione per le modalità utilizzate da alcuni esperti nello svolgimento del compito loro demandato, sia per quanto riguarda la scelta del metodo di identificazione, sia per la presentazione delle conclusioni raggiunte. **Sarebbe di grande utilità fare chiarezza su come la consulenza di riconoscimento del parlante debba essere svolta e come i risultati debbano essere presentati.** Purtroppo la ragione per cui questa chiarezza si presenta di difficile attuazione e non è stata ancora tentata da nessun organo preposto alla valutazione delle consulenze scientifiche la standardizzazione delle procedure,

va ricercata, a nostro avviso, nell'assenza di un consenso unanime su come debba essere svolta una perizia fonica.

Sul tema ci si è limitati (IAFPA 2007) a mettere al bando il metodo sonografico, così come proposto dal professor Tosi (nel 1979), ma rimangono in gioco diversi metodi: il metodo soggettivo di ascolto, il metodo fonico acustico, i metodi basati sulla misura delle formanti (IDEM 1992, SMART 2000) e i metodi automatici. Alcuni esperti aggiungono ai precedenti metodi il metodo linguistico, volto a stabilire se due voci appartengono ad una stessa area linguistica comune, ovvero abbiano lo stesso idioletto. Osserviamo inoltre che alcuni dei metodi citati non si prestano al calcolo della rapporto di verisimiglianza LR e che qualunque sia il metodo che l'esperto ha ritenuto di utilizzare, la mancanza di un protocollo *standard* per ciascun metodo non consente di capire come l'esperto ha operato in quanto ciascun metodo viene applicato con modalità molto diverse nei diversi laboratori.

Sarebbe pertanto necessario procedere alla stesura di una linea guida, condivisa a livello internazionale, che specifichi nel maggior dettaglio possibile le diverse modalità dell'esecuzione di una perizia fonica al fine di rendere le consulenze foniche relative al riconoscimento del parlante più facilmente valutabili da parte del magistrato.

Sui metodi

I più diffusi metodi di identificazione del parlante proposti dalla letteratura scientifica possono essere ricondotti ai quattro seguenti:

1. i metodi d'ascolto o uditivi;
2. i metodi basati sul confronto dei sonogrammi;
3. i metodi basati sull'analisi dei parametri acustico-fonetici;
4. i metodi automatici basati su codifiche del segnale vocale.

I metodi di ascolto o uditivi possono essere ulteriormente suddivisi [Nolan 1997] in un primo criterio basato su un reiterato ascolto, da parte dell'esperto, dei campioni di voce in esame al fine di individuare eventuali elementi di natura linguistica, fonatoria o acustica comuni alle due voci. Queste ultime possono essere ascoltate sia in sequenza sia alternativamente a discrezione dell'operatore che, sulla base degli elementi recepiti, esprimerà un giudizio sulla attribuzione o meno ad uno stesso parlante delle voci ascoltate.

Un secondo criterio è quello basato sul confronto delle voci effettuato da una squadra di ascoltatori, anche non esperti. Il materiale fonico in questo caso è costituito da un insieme di voci comprendenti la voce da identificare, le voci sospette ed eventualmente alcune voci estranee (popolazione di riferimento), prelevate da parlanti aventi caratteristiche fonatorie simili a quelle delle voci in esame; si formano così dei veri propri *test* vocali costituiti da coppie di frasi ottenute raggruppando due a due, in tutte le possibili combinazioni, i campioni di voce da valuta. Ciascun operatore dopo l'ascolto di ogni coppia dovrà esprimere un giudizio di similarità o meno delle voci a confronto. L'elaborazione statistica dei giudizi espressi dagli ascoltatori consente di giungere a conclusioni di tipo sostanzialmente qualitativo [Anil 2005]. E' tuttavia possibile, utilizzando questo secondo metodo, stimare un rapporto di verisimiglianza: LR sarà il rapporto tra la similitudine media attribuita alla coppia voce anonima-voce del sospettato, e la similitudine media delle coppie di voci diverse.

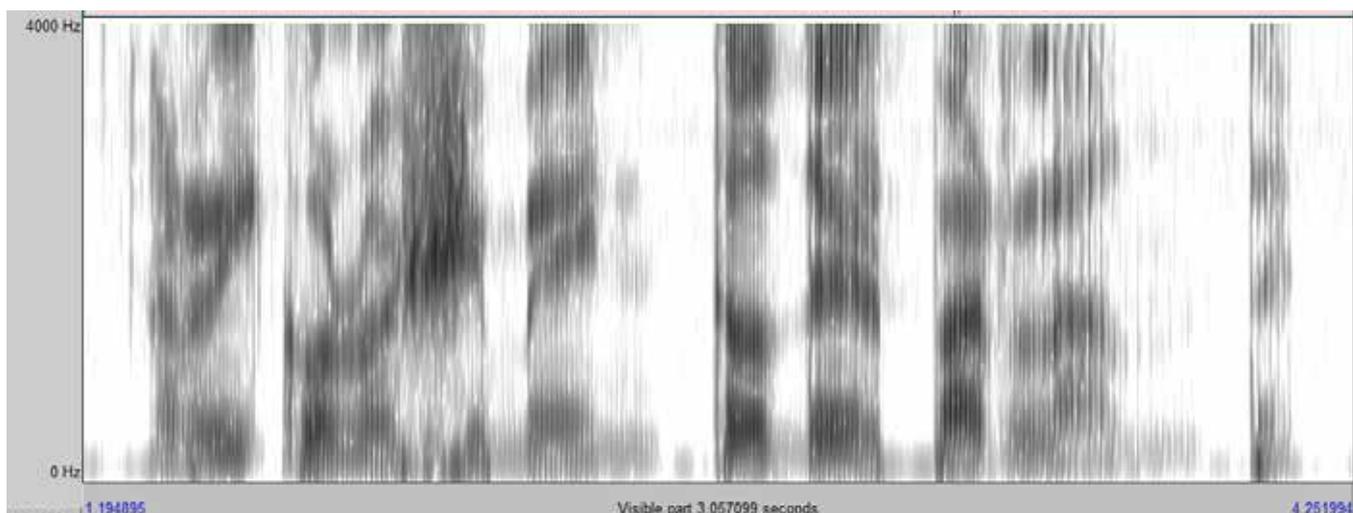


Figura 2 – tecnica del confronto tra sonogrammi. Il tempo è riportato sull'asse delle ascisse, la frequenza su quello delle ordinate. Le diverse gradazioni di grigio forniscono la misura del livello energetico del segnale.

Un secondo metodo è quello che fa uso di una tecnica un tempo diffusa di identificazione parlante, **tecnica che prevede il confronto tra sonogrammi** ovvero tra quei particolari diagrammi tridimensionali (vedi figura 2) che forniscono una rappresentazione grafica dell'andamento temporale della frequenza e dell'intensità del segnale vocale. Il tempo è riportato sull'asse delle ascisse, la frequenza su quello delle ordinate mentre le diverse gradazioni di grigio od un'opportuna scala di colori forniscono la misura del livello energetico il segnale.

Questi sonogrammi contengono sia informazioni di carattere linguistico, comuni alle emissioni verbali di tutti i parlanti una data lingua, sia informazioni sulle caratteristiche acustiche e fonologiche individuali di ciascun parlante.

L'analisi spettrografica ha avuto un notevole sviluppo d'impiego in ambito giudiziario a opera soprattutto di Kersta [Kersta 1962] e Tosi [Tosi 1979]. Noi non illustreremo ulteriormente questo metodo in quanto screditato a livello internazionale (vedi fig. 3) [IAFAPA 2007].

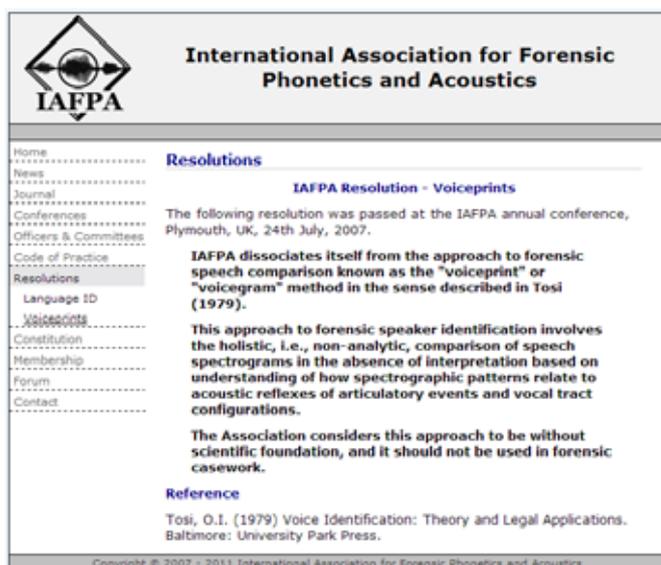


Figura 3 – risoluzione dell'IAFAPA del 24 luglio 2007 sul metodo di confronto della voce conosciuto come "voiceprint" o "voicegram".

Un terzo metodo è quello basato sull'analisi acustico fonetica

[Falcone 1995], [Paoloni 1998]: si tratta di estrarre dai campioni di voce a confronto particolari caratteristiche acustiche, quali la durata delle occlusive, le formanti delle vocali, la velocità di articolazione, la frequenza fondamentale, che si ritengono maggiormente rappresentative della voce l'individuo e meno influenzate dal canale di trasmissione. Questi parametri assunti come elementi di caratterizzazione della voce dovrebbero presentare al contempo piccole variabilità nello stesso individuo (variabilità intra parlante) e variabilità maggiore tra individui diversi (variabilità inter parlanti) [Wolf 1972], [Federico 1987].

Il metodo può essere descritto in tre diverse fasi operative: una prima fase riguarda la scelta dei campioni di voce da analizzare. Questa operazione è assai delicata perché occorre selezionare, tra il materiale fonico disponibile, parole o frasi che oltre a essere effettivamente rappresentative della popolazione di riferimento, debbono anche possedere caratteristiche qualitative (rapporto segnale rumore, larghezza di banda e durata) tali da consentire l'estrazione corretta delle misure strumentali.

La seconda fase riguarda la misura dai suddetti campioni dei parametri spettrali e temporali atti a caratterizzare la voce. La terza fase infine, la fase di decisione, pone a confronto le misure effettuate al fine di calcolare il rapporto di verisimiglianza LR.

Il quarto metodo infine è quello cosiddetto automatico

[Drygajlo 2007] che fa uso di parametri di tipo spettrale originati da una qualche codifica del segnale audio come i coefficienti LPC e i coefficienti cepstrali calcolati ad intervalli regolari utilizzando finestre di ampiezza opportuna. I parametri di tipo spettrale sono di facile e veloce stima, perché la loro misura può essere completamente automatizzata. Dopo aver calcolato i parametri di tipo spettrale si procede ad elaborare un modello statistico di detti parametri e successivamente a mettere a confronto i modelli statistici dei diversi campioni (vedi figura 4).

È evidente che tanto più i modelli statistici sono simili tra loro, tanto più è probabile che i campioni provengano dalla stessa persona. Principale limite di questo metodo è che risente in maniera significativa dell'influenza del canale di trasmissione, inteso in senso lato, ovvero sia del rumore dell'ambiente di emissione sia del tipo di codifica alla quale il segnale stato già sottoposto.

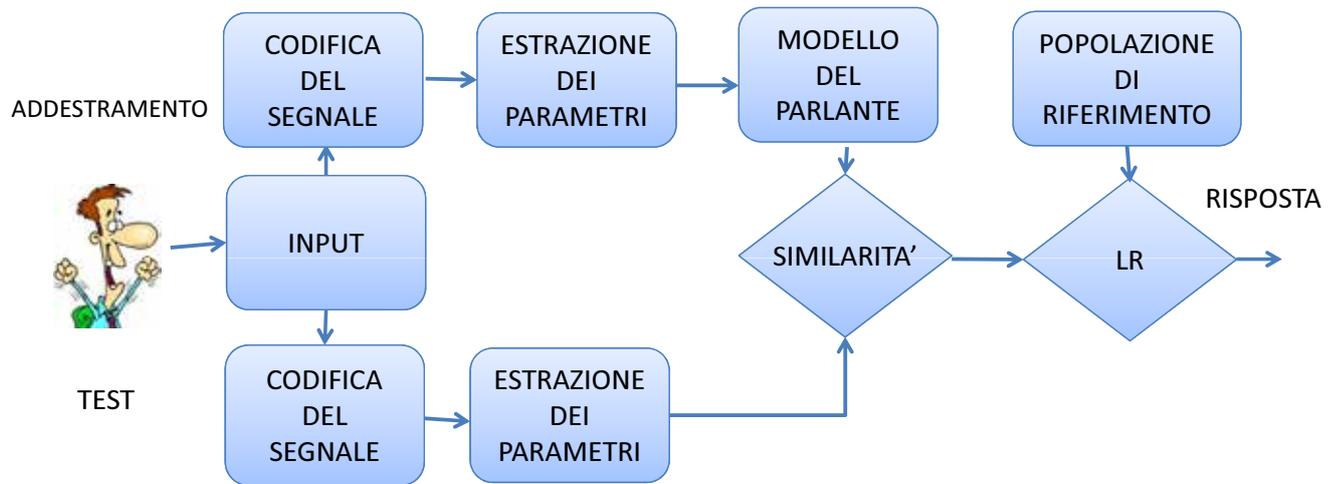


Figura 4 – schema a blocchi dei metodi automatici.

Problemi di decisione

Relativamente alla decisione sull'identità del parlante anonimo la letteratura internazionale suggerisce uno schema di tipo *bayesiano*: l'esperto, dopo aver effettuato i suoi calcoli, dovrà fornire un moltiplicatore, il rapporto di verosimiglianza (LR), utilizzando il quale la Corte aumenterà o diminuirà la probabilità di identificazione che ha ritenuto di assegnare all'imputato prima della prova vocale. Il rapporto di verosimiglianza avrà al numeratore la misura della "similarità" tra la caratteristica dell'imputato e quella della traccia ed al denominatore la "tipicità" ovvero quanto la caratteristica presa in esame sia rara, al limite unica, all'interno della popolazione di riferimento.

La scelta del campione della voce anonima, dopo averne accertato l'idoneità al confronto, influenza il risultato ma più ancora lo influenza la scelta della popolazione di riferimento. Sorge il problema, in parte analogo a quello relativo alla scelta delle voci nella composizione di un *line up*: bisogna decidere se la popolazione di riferimento debba essere un campione dell'intera popolazione dei parlanti oppure un campione di parlanti aventi caratteristiche simili a quelle della voce anonima: ad esempio stesso ambiente sociale, stesso sesso, stesso livello culturale, stessa parlata regionale, ecc.

Concludendo

Al fine di fornire un contributo efficace, verificabile e confrontabile, la consulenza fonica dovrebbe utilizzare uno o più dei metodi precedentemente descritti ciascuno dei quali dovrebbe essere formalizzato in modo da consentire a tutte le parti in causa la sua valutazione e la ripetitività dell'esperimento. Per quanto riguarda la prova soggettiva di ascolto questo significa che debba essere utilizzato solo il secondo metodo. Una squadra di ascolto fornisce un risultato ripetibile, all'interno delle variazioni statistiche, rispetto a un'altra squadra di ascolto. Nel caso di metodi che utilizzino misure acustiche, come ad esempio la frequenza fondamentale e le frequenze formanti, la consulenza dovrà indicare il *file* utilizzato, dove è stata effettuata ciascuna misura, nonché quale *software* sia stato utilizzato per l'analisi; nel caso invece di utilizzo del sistema automatico, dovranno essere resi disponibili i campioni utilizzati nel confronto e la popolazione di riferimento, la cui origine dovrà

essere descritta in dettaglio. Per quanto attiene la risposta al quesito, l'esperto dovrà fornire la risposta in termini di rapporto di verosimiglianza (LR) riportato in forma numerica e in forma qualitativa utilizzando la scala suggerita ad esempio da Rose [Rose 2005].

Ci auguriamo che l'evolversi della tecnologia consenta di scegliere una metodica definita e di definire lo *standard* da utilizzare nelle consulenze di identificazione del parlante. ©

BIBLIOGRAFIA

- Anil A., Dessimoz D., Botti F., Drygajlo A., 2005 "Aural and Automatic Forensic Speaker Recognition in Mismatched Conditions", The International Journal of Speech, Language and the Law, vol. 12, Dec., pp. 214-234
- Drygajlo A., 2007 "Forensic Automatic Speaker Recognition", IEEE Signal Processing Magazine, 24 (2): 132-135 (2007).
- Falcone M, De Sario N., 1994 "A PC speaker Identification System for Forensic Use: IDEM" ESCA Workshop on Automatic Speaker Recognition, Identification, and Verification, Martigny, Switzerland.
- Kersta L.J., Voiceprint Identification, Nature, vol. 196, pp. 1253-1257, 29 Dicembre 1962;
- Koenig B. E., 1993, Selected Topics in Forensic Voice Identification, Crime Laboratory Digest, vol. 20, n. 4, pp. 78-81;
- Nolan F., 1997, Speaker recognition and forensic phonetics. In: W. Hardcastle and J. Laver (eds), A Handbook of Phonetic Science. Oxford: Blackwell, pp. 744-767. Koenig B., 1993, "Selected Topics in Forensic Voice Identification", Crime Laboratory Digest, vol. 20, n. 4, 1993 pp.78-81
- Paoloni A.1997, Il riconoscimento del parlante, Detective&Crime Magazine / Criminalistica –Le indagini fonetiche;
- Paoloni A., Falcone M., Federico A., 1998, The Parametric Approach in Forensic Speaker Recognition, Proceedings of the COST 250 Workshop on Speaker Recognition by man and machine: directions for forensic applications, Ankara, Turkey, ed. by Demirekler M., Saranlı A., Altınçay H., Paoloni A., pp.45-51
- Paoloni A.2003, Note sul riconoscimento del parlante nelle applicazioni forensi con particolare riferimento al metodo parametrico IDEM, Riv. Italiana di Acustica, Vol. 27 n. 3-4;
- Rose P., Forensic Speaker Identification, Taylor & Francis, 2002.
- Tosi Oscar, 1979, Voice Identification. Theory and Legal Applications, University Park Press, Baltimore
- Trumper J., Maddalon M., Romito L., 1996, Atteggiamento della Magistratura nei confronti delle perizie foniche in Collana degli atti dell'AIA pp 35 – 46
- Wolf J.J. 1972, "Efficient acoustic parameters for speaker recognition" J.A.S.A., Vol. 51, N° 6 , pp.2044-2056. ♦