

SPEAKER RECOGNITION: STATO DELL'ARTE IN ITALIA VALUTAZIONE DEI CORPORA, DEI METODI E DELLE PROFESSIONALITÀ COINVOLTE

Luciano Romito, Vincenzo Galatà

Laboratorio di Fonetica – Dip.to di Linguistica – Università della Calabria
luciano.romito@unical.it, vgalata@libero.it

1 SOMMARIO

Sabrina Antognoli¹ in uno studio del 1996 scrive che in una sentenza ad una cosca mafiosa su 52 pagine sono presenti ben 52 richiami ad intercettazioni sia telefoniche che ambientali e solo 15 richiami a testimonianze dirette. Questo indica che l'accusa prima, e la sentenza in seguito, fondano gran parte delle proprie deduzioni e conclusioni proprio sulle intercettazioni coatte di voci. L'intercettazione è oggi diventata una tra le più importanti e usate tecniche investigative ma di contro non esiste una figura professionale pronta, preparata e riconosciuta per assolvere questo compito. Non esiste neppure un percorso formativo universitario o professionale che possa creare competenza; inoltre è assolutamente assente qualunque attenzione da parte dello Stato o delle Istituzioni al problema. Questo 'vuoto' è spesso facile preda di pseudo-periti e pseudo-esperti con pseudo-metodi e per chiunque appartenga al fantasioso popolo di poeti, sognatori e navigatori.

Il presente lavoro si inserisce in un ampio progetto di ricerca del Laboratorio di Fonetica (Università della Calabria) sullo Speaker Recognition (SR d'ora in poi) in ambito forense. Tale progetto, denominato Primula Rossa e privo di alcun finanziamento, comprende una sezione riguardante l'anagrafe e una sezione riguardante la valutazione dei metodi attraverso un corpus costruito ad hoc.

Nello specifico il progetto prevede una:

1. anagrafe dei corpora per lo SR presenti in italiano con analisi e specifiche di ogni singolo corpus;
2. costruzione di un corpus acquisito con caratteristiche e apparati normalmente utilizzati nelle intercettazioni al fine di valutare i diversi metodi utilizzati oggi in Italia;
3. anagrafe delle ricerche e dei lavori pubblicati negli ultimi cinque anni in Italia sull'argomento, nonché dei convegni e dei progetti finanziati;
4. anagrafe dei fondi investiti per la formazione di personale dipendente delle forze dell'ordine e per corsi regionali o di perfezionamento istituiti dalle singole Università;
5. anagrafe delle professionalità coinvolte, ovvero di tutti coloro che utilizzano la fonetica forense e le tecniche di SR in ambito forense;
6. anagrafe delle competenze richieste dai Tribunali e dalle Procure per poter utilizzare ed adoperare lo SR;
7. anagrafe di tutti i metodi utilizzati presso le Procure e le Aule di Tribunale e non solo quelli conosciuti in ambito scientifico o perché divulgati attraverso pubblicazioni scientifiche;
8. valutazione dei metodi.

In questo lavoro ci occuperemo pertanto dei punti 1-7 che costituiranno la base di partenza e la naturale continuazione del lavoro qui intrapreso per lo sviluppo del punto 8.

2. ALCUNE PREMESSE STORICO-INTRODUTTIVE

In Italia il problema della comparazione tra voci nell'ambito delle indagini forensi, è stato e viene ancora affrontato con due differenti metodi di indagine: uno che potremmo definire -linguistico e l'altro acustico-strumentale.

I primi lavori di SR effettuati furono sicuramente di tipo linguistico, inteso qui nel senso più ampio del termine, dalle analisi dialettologiche a quelle sociolinguistiche fino alle analisi glottologiche². Con la nascita della dialettometria si iniziò ad utilizzare la statistica anche in dialettologia ed in linguistica³.

Per quel che attiene il metodo acustico-strumentale bisogna fare un salto indietro nel tempo e fare riferimento alla nascita del *sonografo* o *spettrografo*⁴. Attraverso tale attrezzatura, e soprattutto con l'uso dei sonogrammi, fu

¹ Antognoli S. (1996), *Il procedimento penale come <<evento linguistico>>*. *Analisi sociolinguistica di un processo di mafia in Toscana*, Università di Firenze, Facoltà di Scienze Politiche, Tesi di Laurea quadriennale in "Partiti politici e gruppi di pressione", op. cit. in Bellucci P. (2006), *A onor del vero. Fondamenti di linguistica giudiziaria*, UTET Libreria, Torino, pag. 59.

² Si fa in tal caso riferimento a Trumper J.B. (1979), *Sociolinguistica giudiziaria. Preliminari di metodi e applicazioni*, CLESP Editrice, Padova.

³ Quanto appena affermato è ovviamente il contrario di quanto scrive a pag. 14 il dott. D. Zavattaro (A.A. 2004-2005), <<Fino agli anni '60 l'approccio alla comparazione della voce era di tipo esclusivamente linguistico, gli esperti erano fonetisti ed il livello di scientificità modesto>>, dove si lascia intendere che le variabili considerate fossero esclusivamente di tipo qualitativo e non quantitativo.

sviluppato un metodo per il riconoscimento del parlatore. Tale metodo fece ingresso in un aula di tribunale con Kersta nel 1962. Lo stesso Kersta pubblicò i risultati di una ricerca basata sull'identificazione del parlatore e conìò, forse troppo ottimisticamente, il termine "voice-print" (impronta vocale) sulla scia di quello già esistente di "finger-print" (impronta digitale). Ben presto però ci si rese conto che tale strumento, e il metodo ad esso legato, utilissimi per effettuare studi sulla co-articolazione vocalica, erano inadatti per una comparazione fonica⁵. Dal loro primo utilizzo e a causa del termine con cui venivano identificati molte furono le accuse mosse come dimostrano tra l'altro gli studi condotti⁶ e, di non minore importanza, il rapporto di un'equipe di scienziati del *Technical Committee on Speech Communication* dell'*Acoustical Society of America* a seguito della controversia scaturita da innumerevoli perizie giudiziarie:

"Today's consensus suggests that speaker identification by voice pattern is subject to error at a high, and yet undetermined, rate. Courts determinations may also depend on the apparent validity of exhibits brought in evidence. Spectrographic evidence may often display features that are overwhelmingly influenced by the words spoken rather than by the speaker's identity. Judge and jury may therefore be misled in understanding the evidence and in assessing an expert's testimony."⁷

Nonostante ciò comunque, come vedremo in seguito, in Italia l'uso del solo confronto dei sonogrammi come metodo di SR è ancora abbastanza diffuso⁸.

Ben presto fu chiaro che l'unica strada da intraprendere era quella della condivisione di competenze e conoscenze per uno scopo ed un fine comune. Iniziarono così intense e fruttuose collaborazioni tra linguisti, fisici e ingegneri⁹.

Eventi quali l'incessante richiesta di comparazioni foniche e il sempre minor tempo concesso agli esperti, interruppero tale percorso. Ciò che all'inizio rappresentava uno stimolante campo di ricerca e di verifica, diventò ben presto un servizio da offrire alla Magistratura. Quest'ultima, forse per una non completa comprensione del problema o forse per una necessità dettata dal momento storico¹⁰, preferì sempre più incarichi non collegiali e risposte più celeri e meno vaghe statisticamente. In un certo qual modo quindi l'apporto scientifico che l'esperto poteva dare come aiuto alla Magistratura diventò ben presto una relazione tecnica che oggi spesso influenza e condiziona la stessa sentenza emessa dal Tribunale.

Crediamo in sostanza che questa sia la prima causa dell'allontanamento delle Università e degli Enti di ricerca dal mondo delle comparazioni foniche in ambito forense (ne sono un esempio i dati più avanti riportati relativi ai corsi tenuti, alle pubblicazioni, ai convegni e a quanto altro possa ricondurre alla sfera accademica). In questo vacuum presero il sopravvento nuove figure con formazione professionale che, come vedremo, risulta assai diversificata.

Da una forte spinta iniziale fatta di idee tutte da sviluppare e di interessanti teorie linguistiche, si è giunti ad una moltitudine di metodi differenti, alcuni con base scientifica, altri secondo noi discutibili e assolutamente senza alcun fondamento scientifico ma comunque tutti normalmente utilizzati nei Tribunali italiani.

Sono stati fondati laboratori, cooperative e società¹¹ che effettuano su richiesta misure a tal proposito. In alcuni casi, il perito è infatti solo un "intermediario" tra la magistratura e i laboratori¹² che accetta l'incarico e affida ad una

⁴ Lo *spettrografo* (o *sound-spectrograph*) nacque prevalentemente come apparecchiatura da destinare ad uso militare durante la seconda guerra mondiale e fece la sua prima apparizione in una pubblicazione scientifica nel 1945 grazie a Ralph K. Potter un ricercatore dei laboratori Bell.

⁵ Cfr. Rose P. (2002), *Forensic Speaker Identification*, Taylor & Francis, London: <<the studies that have been produced over the years have run the gambit in type, parameter and result. A quick review of the available published data would leave one with the impression that the spectrographic method of voice identification was only somewhat more accurate than flipping a coin>>.

⁶ Cfr. Young e Cambell (1967), Stevens et al. (1968), Hazen (1972a e 1972b), Tosi et al. (1972), Bolt et al. (1974).

⁷ Cfr. Bolt R., Cooper F.S., David E.E. Jr., Denes P.B., Pickett J.M., Stevens K.N. (1970) 'Speaker identification by speech spectrograms: A scientists' view of its reliability for legal purposes.' *Journal of the Acoustical Society of America*, 4 (2): 597-612, pag. 602.

⁸ Quanto detto risulta ancor più vero consultando un concorso per ufficiali, pubblicato nella G.U. della Repubblica Italiana 4° serie Speciale n. 99 del 29/12/1995 e a cui si farà riferimento più avanti, dove al comma 5 *Metodi di identificazione vocale*, ritroviamo nei programmi concetto di "impronta vocale", spettrografo elettroacustico e sonogramma.

⁹ Ne è un esempio il primo convegno patrocinato dall'AIA tenutosi a Padova il 14 e 15 settembre 1978, che ha portato ad una raccolta degli atti della tavola rotonda in un volume a cura di F. Ferrero dal titolo *L'Identificazione della persona per mezzo della voce*, Edizioni Scientifiche Associate, Roma, a cui presero parte fisici, linguisti, ingegneri.

¹⁰ Parliamo degli anni della tensione il cui inizio ufficiale viene fatto coincidere con gli otto attentati avvenuti il 9 agosto 1969 o degli anni del terrorismo o di piombo quando il 25 gennaio 1971 fece la sua prima comparsa un commando di brigatisti con la stella rossa a cinque punte.

¹¹ Tra questi i riferimenti più comuni sono la sezione "Voce" della Fondazione Ugo Bordoni e l'Istituto Elettrotecnico Nazionale "Galileo Ferraris" di Torino.

compagnia o un laboratorio privato il segnale da misurare consegnando infine il risultato con una breve relazione più burocratica che metodologica al Magistrato.

Gli esperti di linguistica forense, in Italia, dovrebbero uniformare la propria disciplina alla criminalistica adottando metodi e obiettivi che siano assolutamente ripetibili e verificabili, come d'altronde dovrebbe essere ogni scienza che tale voglia essere definita. Solo in questi ultimissimi anni nascono sporadici esempi di quanto detto, si veda ad esempio SMART¹³ e IDEM¹⁴ per SR o ANFAF¹⁵ per le trascrizioni giudiziarie.

Ma tornando allo SR, sotto il profilo scientifico riteniamo che l'unica strada oggi percorribile sia quella del linguista¹⁶ con competenze di signal processing dove l'approccio degli ingegneri deve servire per la costruzione di strumenti informatici quali software, hardware, filtri e algoritmi, l'apporto dei matematici per la costruzione di strumenti statistici¹⁷.

Ciò non toglie tuttavia che in Italia ci siano molti ingegneri o fisici (in particolare il pensiero va al compianto Franco Ferrero che più di ogni altro ha dato una spinta alle analisi linguistico-fonetiche in Italia) che hanno il grande merito di aver acquisito competenze linguistiche e fonetiche e quindi a pieno titolo svolgono analisi di signal processing e contestualmente di linguistica. Siamo tuttavia in accordo con Furui (2005) che la comparazione fonica non può essere ridotta ad un semplice confronto di numeri completamente scollegati dal segnale linguistico:

<<significant advances in speech and speaker recognition are not likely to come solely from research in statistical pattern recognition and signal processing. Although these areas of investigation are important, the significant advances will come from studies in acoustic-phonetics, speech perception, linguistics, and psychoacoustics. Future systems need to have an efficient way of representing, storing and retrieving "knowledge" required for natural conversation.>>

3. ANAGRAFE DEI CORPORA PER LO SPEAKER RECOGNITION PRESENTI IN ITALIANO

In Campbell Jr. J. & Reynolds D. (1999) e in Melin H. (1999), per l'italiano figura un unico e solo corpus per la valutazione dei sistemi di riconoscimento del parlatore, ma a distanza di più di cinque anni la situazione non è molto diversa. Da numerose ed approfondite ricerche effettuate¹⁸, allo stato attuale tra i corpora in lingua italiana creati appositamente per scopi inerenti il riconoscimento/identificazione del parlatore, o eventualmente utilizzabili per scopi di SR in ambito forense troviamo sicuramente FOCUS e ViPS. A questi possono, o meglio, vanno necessariamente aggiunti i loro predecessori SIVA e CALÌ. Ci preme, in questa sede, mettere in evidenza le differenze che caratterizzano i corpora individuati riportando per ciascuno di essi alcune tra le informazioni più importanti che ci hanno spinto a creare un nuovo corpus.

SIVA (Speaker Identification and Verification Archives)¹⁹ è un database raccolto e creato dalla FUB dal 1994 al 1996 per la verifica del parlatore su linea telefonica raccolto attualmente su singolo DVD. Tutti i files sono campionati a 8kHz, 8 bit con compressione mu-law. Il DVD è diviso in 4 directory contenenti:

1. MU (Male Users; 18 parlatori, 20 sessioni)
2. MI (Male Impostors; 411 parlatori, 1 sessione)
3. FU (Female Users; 18 parlatori, 20 sessioni)
4. FI (Female Impostors; 411 parlatori, 1 sessione)

Ogni sessione prevede la produzione di 28 parole, un dialogo unidirezionale e la lettura di un breve brano foneticamente bilanciato con copertura nazionale. La FUB non garantisce tuttavia che il database sia "privo di errori"²⁰.

Il corpus CALI' (CALler Identification on mobile and fixed network)²¹ è un database creato dalla FUB e dall'ISCTI nel 1999 per la verifica del parlatore su linea telefonica e GSM e attualmente contenuto in 10 Cd-Rom. Il database

¹² Si veda più avanti il § 6.5 6.6 *La figura dell'intermediario.*

¹³ Statistical Methods Applied to the Recognition of the Talker, quest'anno al terzo step, è un progetto di ricerca finanziato dalla Commissione Europea nell'ambito del programma OISIN cui partecipano diversi paesi europei, nel quale viene valutata la possibilità di riconoscere le voci di parlatori con approccio statistico.

¹⁴ IDentification Method è un ambiente software creato dalla Fondazione Ugo Bordoni per l'analisi del segnale vocale ai fini dell'identificazione del parlatore.

¹⁵ Associazione Nazionale di Fonetica Acustica e Forense.

¹⁶ Come suggerito tra l'altro da Fraser, H. (2003), Issues in Transcription: Factors affecting the reliability of transcripts as evidence in legal cases. *Forensic Linguistics*, 10(2), pp. 203-226, per le questioni legate alle trascrizioni giudiziarie.

¹⁷ Segnali molto simili alla forma d'onda sono presenti in cardiologia, in geologia, in neurologia e non per questo ingegneri e matematici sono considerati bravi cardiologi o geologi.

¹⁸ Senza escludere i siti di organizzazioni non-profit che gestiscono e valutano le risorse linguistiche ovvero ELRA (European Language Resources Association) ed ELDA (Evaluations and Language resources Distribution Agency).

¹⁹ Falcone M., Gallo A., (1996), The SIVA speech database for speaker Verification: description and evaluation, *Proceedings of ICSLP'96*, Philadelphia, pp. 1902-1906.

²⁰ Cfr. <http://www.fub.it/voice/corpora/SIVA/>

contiene 20 chiamate per parlatore (10 su linea GSM e 10 su linea PSTN) per 100 parlatori maschili e 100 parlatori femminili. All'interno del corpus ciascun parlatore pronuncia 18 item relativi a sequenze di cifre, brevi frasi e spelling di nomi. La distribuzione geografica del corpus copre il territorio nazionale.

I corpora FOCUS (FOrensic CorpUS)²² e ViPS (Voci [italiane] Presumibilmente Simili)²³ possono essere considerati in un certo senso come uno solo. ViPS è un database creato dalla FUB e dall'ISCTI nel 2003 per la verifica del parlatore con limitato numero di parlatori ma con diverse qualità di segnale²⁴ contenuto su singolo DVD. Come già desumibile dal nome assegnato al corpus la caratteristica che lo contraddistingue è quella di contenere una serie di parlatori aventi caratteristiche comuni quali ad esempio estrazione dialettale, età ecc. La similitudine delle voci è tuttavia fornita a discrezione dell'operatore, mentre la copertura geografica del campione è limitata a soli 12 parlatori di Roma e provincia aventi una media di 40 anni. Come si accennava ViPS nasce sulla base del corpus denominato FOCUS del 1999 includendolo ed arricchendolo. FOCUS comprende infatti anch'esso 12 parlatori tutti di sesso maschile. Il legame tra i due corpora sta nel fatto che 4 parlatori presenti in FOCUS sono presenti con nuove registrazioni in ViPS che, a sua volta, include anche gli 8 del primo cui se ne aggiungono altri 8 in ViPS. Il materiale registrato prevede in entrambi i casi ripetizioni di testo letto, brani di dialogo elicitato attraverso intervista, parlato spontaneo disponibile in quattro versioni: ortofonico, PSTN, GSM ed Elaborazione software.

A questi potremmo aggiungere una serie di altri corpora che tuttavia difficilmente possono rivelarsi utili per gli obiettivi che con questo lavoro ci siamo prefissati. A titolo di cronaca abbiamo quindi:

- AVIP (Archivio delle Varietà di Italiano Parlato), un corpus che comprende, per la parte che in questa sede potrebbe interessare, testi parlati semi-spontanei raccolti attraverso la tecnica del Map Task.²⁵ Il corpus consta di 3,5 ore circa di parlato e contiene materiali di italiano regionale di Toscana (pisano), Campania (napoletano) e Puglia (barese) integralmente trascritti ortograficamente. Una parte del materiale trascritto è stato analizzato ed annotato (a livello fonetico, fonologico della varietà, fonologico standard, intonativo) con un segmentatore semiautomatico di analisi fonetica (Segwin).
- CLIPS (Corpora e Lessici di Italiano Parlato e Scritto), un vasto corpus contenente parlato e scritto. CLIPS è destinato prevalentemente alle applicazioni in ambito tecnologico. Il corpus CLIPS dovrebbe contenere, al termine del progetto, 100 ore di italiano parlato di varie tipologie (parlato dialogico, parlato letto, parlato televisivo, conversazioni telefoniche e corpora speciali), stratificato dal punto di vista diafasico e diatopico. I materiali del corpus sono stati raccolti in 15 città (Torino, Bergamo, Milano, Venezia, Parma, Genova, Firenze, Perugia, Roma, Napoli, Bari, Lecce, Catanzaro, Palermo e Cagliari) ritenuti rappresentativi sia da un punto di vista di varietà di italiano sia da un punto di vista demografico al fine di fornire uno spaccato quanto più possibile realistico dell'eterogenea realtà linguistica dell'italiano.

Infine, come più sopra si accennava, passando in rassegna i siti dell'ELRA (European Language Resources Association) e dell'ELDA (Evaluations and Language resources Distribution Agency) abbiamo individuato, sempre per la lingua italiana, i seguenti corpora che di seguito verranno menzionati:

- SPK, un database di italiano parlato di cifre isolate e cifre connesse realizzato presso l'Istituto per la Ricerca Scientifica e Tecnologica (ITC/IRST) di Trento. SPK nasce per scopi di verifica e riconoscimento del parlatore in applicazioni commerciali. Il materiale è stato raccolto da 100 parlatori (30 femminili e 70 maschili, da 23 a 50 anni) la cui provenienza è prevalentemente dal nord-est Italia. Sono presenti 5 sessioni di registrazione per parlatore. Per ogni sessione sono state acquisite quattro ripetizioni di cifre rigorosamente effettuate in camera silente.
- FIXED0IT - DB1 è un corpus che contiene frasi foneticamente bilanciate e frasi application oriented. Il Fixed Network Speech SpeechDat(M) Corpus (italiano) versione 1.0 è stato registrato all'interno e per scopi del progetto SpeechDat(M) (LRE-63314), finanziato dalla Commissione Europea. Le registrazioni effettuate attraverso interfaccia ISDN (8 kHz, 8 bits, A-law coding) contengono il parlato di circa 1000 parlatori (ca. 500 maschi e 500 donne) e sono state disegnate per supportare la creazione di servizi telematici a controllo vocale. Ciascun parlatore ha prodotto almeno 39 item, tra cui cifre isolate e connesse, numeri naturali, importi monetari, spelling di parole, frasi contenenti ora e data, *yes/no questions*, nomi di città, parole comuni di applicazioni, frasi foneticamente bilanciate. La maggior parte degli item è letta ed alcuni sono pronunciati spontaneamente. Oltre a ciò vengono fornite informazioni legate al parlatore quali età e provenienza regionale. Di questo corpus è presente un sub-set che contiene soltanto le frasi foneticamente bilanciate ovvero il DB2 "Phonetically rich sentences sub-set".
- Italian SpeechDat (II) MDB-250 comprende 375 parlatori registrati attraverso linea telefonica mobile. L'MDB-250 è organizzato in 6 CD in formato ISO 9660 i cui campioni sonori sono archiviati in sequenze di 8-bit 8 kHz A-law.

²¹ Falcone M., Ridolfi P. (1999), CALÌ: un corpus vocale per l'autenticazione del parlatore su rete telefonica fissa e mobile, *X Giornate di Studio del GFS*, Napoli.

²² Falcone M., Barone A., (2000), FOCUS: un corpus vocale di voci simili per lo studio della identificazione del parlatore in ambito forense, *28° Convegno AIA*, Trani.

²³ Falcone M., Barone A., (2003), ViPS: Voci (italiane) Presumibilmente Simili, *XIV Giornate di Studio del GFS*, Viterbo.

²⁴ Si tratta in questo caso dello stesso segnale ritrasmesso con una bocca artificiale su reali linee telefoniche.

²⁵ L'altra parte contiene testi parlati elicitati mediante immagini in bambini normoudenti e ipoacusici.

Ciascuna frase pronunciata è archiviata in un file ed ogni file è accompagnato da un file etichetta ASCII SAM contenente le informazioni descrittive più rilevanti. Ciascun parlatore ha pronunciato diversi item che vanno da cifre isolate e connesse, numeri telefonici, codici PIN ecc. a frasi e parole foneticamente bilanciate. Il corpus consta di 3 speaker sotto i 16 anni, 147 tra 16 and 30, 149 tra 31 e 45, 56 tra 46 e 60, 48 oltre 60.

- Italian SpeechDat(II) FDB-3000 contiene più di 3000 parlatori (1494 maschi, 1546 femmine) registrati attraverso rete telefonica mobile. Ha le stesse caratteristiche dell'MDB-250 ma con una differente distribuzione di parlatori: 133 speaker sotto i 16 anni, 757 tra 16 and 30, 862 tra 31 e 45, 626 tra 46 e 60, 482 oltre 60.

- Italian Speech Corpus 1 (Appen) contiene registrazioni di 202 parlanti nativi italiani (112 maschili, 90 femminili) in ufficio e posto pubblico chiuso su 4 canali, in ambiente con rumore di fondo da basso a medio. Ogni campione è archiviato a 16-bit 22.05 kHz PCM WAV. Ogni parlatore ha letto item consistenti di 100 parole di comando e 100 frasi foneticamente bilanciate. La distribuzione in termini di età è: 22 speaker tra 18 e 19, 141 tra 20 e 30, 34 tra 31 e 45, e 5 tra 45 e 60 e vengono fornite informazioni sul luogo di nascita del parlatore. Il database è stato creato per lo più per applicazioni di speech recognition ed applicazioni a controllo vocale.

Al termine di questa tediosa seppur significativa e necessaria carrellata di corpora, emergono una serie di considerazioni che, in vista di una valutazione dei diversi metodi di riconoscimento del parlatore utilizzati oggi in Italia in ambito forense, ci hanno spinto a creare un nuovo corpus.

Il riconoscimento del parlatore in ambito forense è spesso caratterizzato da situazioni estreme, sia in termini di qualità sia in termini di quantità del materiale da analizzare. Con il termine "estremo" ci riferiamo qui al fatto che il materiale a disposizione di chi deve eseguire un'operazione di riconoscimento del parlatore è quasi sempre troppo scarso in termini di durata della voce da analizzare, troppo scarso in termini di qualità della registrazione (sia essa legata a rumori di varia natura, o sia essa legata al canale di trasmissione o al formato di registrazione compresso che oramai dilaga).

IDEAZIONE E CREAZIONE DEL CORPUS PRIMULA

Il corpus PRIMULA è un corpus ristretto di voci calabresi ideato e creato presso il Laboratorio di Fonetica dell'Università della Calabria per la valutazione delle metodologie e dei sistemi di riconoscimento del parlatore con particolare attenzione all'ambito forense scaturite dalla nostra inchiesta di cui ai prossimi §§. Si tratta per il momento di un progetto pilota che potrà portare, se i risultati derivanti dal suo utilizzo saranno confortanti, ad un progetto di più ampio respiro che abbia valenza a livello nazionale negli interessi dell'intera comunità.

Ciò che sostanzialmente differenzia Primula, come già in parte accennato, rispetto ai corpora individuati e descritti nel precedente §, è la modalità con cui lo stesso è stato creato. Durante la fase di ideazione del corpus e successivamente durante la fase di acquisizione dello stesso si è ritenuto di dover simulare una situazione reale al fine di avere, a prodotto finito, situazioni simili o quantomeno assai vicine a quelle che si presentano di norma nella maggior parte dei casi forensi.

Proprio in virtù di ciò una parte delle registrazioni è stata effettuata con attrezzature normalmente utilizzate per le intercettazioni. Ciò è stato possibile grazie alla collaborazione del Commissariato della Polizia di Stato di Siderno e al supporto tecnico/logistico della ditta ACTIVE Service Unit srl di Lodi che ci ha cortesemente noleggiato la microspia e che ha provveduto alla sua installazione sull'autovettura personale di uno degli autori del presente lavoro. È stato così possibile registrare lo stesso materiale prodotto sia attraverso la microspia installata sull'autovettura sia attraverso un cellulare collegato con un telefono fisso presso il nostro laboratorio dove la registrazione veniva acquisita su un registratore DAT. Il materiale registrato e così derivato ha portato quindi ad avere una intercettazione ambientale (in automobile) e una registrazione telefonica (tra utenza cellulare e utenza di rete fissa).

Non a caso quando viene richiesta o affidata una perizia o una consulenza di comparazione fonica, nella maggior parte dei casi viene fornita al perito o consulente una o più intercettazioni ambientali con il quesito di verificare l'appartenenza o meno di una o più voci ad un indagato. Nei casi in cui viene a mancare la voce del sospettato si procede quindi a prelievo di saggio fonico che può essere effettuato o tramite linea telefonica²⁶ o semplicemente dal vivo tramite un semplice microfono collegato ad un registratore.

All'interno del corpus abbiamo pertanto la voce di 5 interlocutori maschili di simile statura, peso e classe di età. I tipi di registrazione sono Ortofónico (in camera silente), Ambientale (in auto) e Telefonico. Per ogni tipo abbiamo registrazioni di lettura di tre frasi foneticamente bilanciate ripetute da 10 a 50 volte, lettura di 10 frasi singole e diverse sessioni di parlato spontaneo sia in dialetto calabrese che italiano regionale. Per studiare e verificare l'influenza del canale abbiamo effettuato come già detto la stessa identica registrazione sia in modalità ambientale (intercettazione) in auto che attraverso il telefono cellulare. Per studiare e valutare l'influenza del rumore e l'intensità del locutore in presenza di rumore abbiamo la stessa registrazione in strada, ad una fermata di autobus, in un aula universitaria molto rumorosa e in automobile con finestrino aperto. Una visione più dettagliata e schematizzata della struttura del corpus si può evincere dal grafico di seguito presentato.

²⁶ In tal caso nel corpus è prevista per ciascun parlatore una conversazione telefonica in macchina registrata da rete fissa su registratore DAT.

Tutte le registrazioni sono state acquisite, nonostante la presenza dei diversi canali di registrazione, in formato *.wav con una frequenza di campionamento di 44100 Hz, 24-bit in modalità monoaurale, mentre le registrazioni frutto dell'intercettazione risultano essere in un formato proprietario compresso di tipo CCITT A-Law a 8Khz 8-bit Mono.

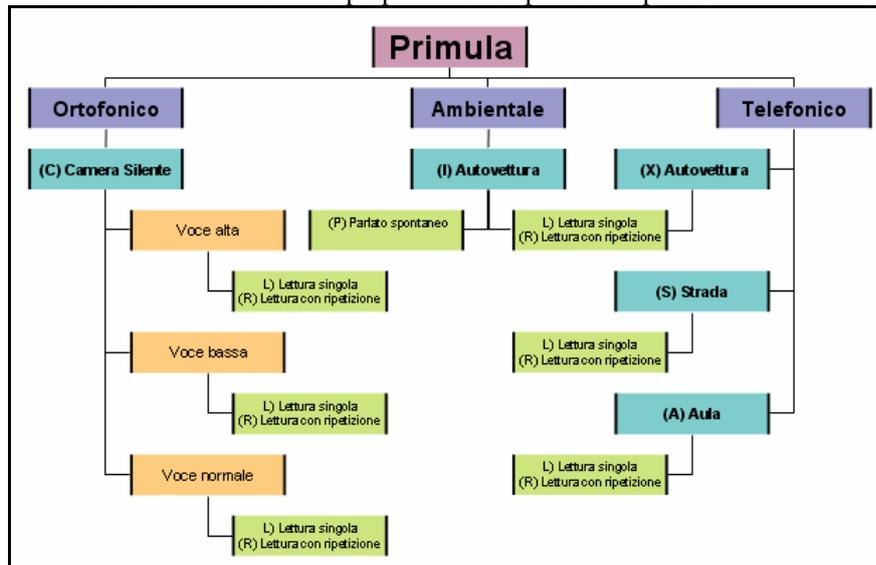


Figura 1: Strutturazione e organizzazione del corpus Primula.

4. ANAGRAFE DELLE RICERCHE E DEI LAVORI PUBBLICATI NEGLI ULTIMI CINQUE ANNI IN ITALIA SULL'ARGOMENTO, NONCHÉ DEI CONVEGNI E DEI PROGETTI FINANZIATI²⁷

Per il reperimento delle informazioni di cui al presente paragrafo ci siamo prevalentemente serviti della rete effettuando ricerche sia tra le banche dati disponibili presso il nostro ateneo sia navigando tra i numerosi siti sulle scienze forensi che oramai dilagano in internet grazie anche agli in-put spesso fantascientifici di alcune serie televisive che alimentano questo mito.

I LAVORI PUBBLICATI IN ITALIA²⁸ NEGLI ULTIMI CINQUE ANNI

Una delle domande del nostro questionario che tratteremo al § 6. chiedeva agli intervistati se, e in che misura (pubblicazioni, ricerche ecc.), contribuissero all'avanzare delle conoscenze sulla questione del riconoscimento del parlatore in ambito forense. Tranne qualche sporadico caso²⁹ la quasi totalità dei periti, effettua essenzialmente solo perizie senza intervenire in alcun modo nella ricerca di nuovi parametri.

La situazione relativa ai lavori pubblicati negli ultimi cinque anni è significativa della situazione dello SR in Italia.

Di seguito viene presentata la lista in ordine decrescente:

- Jona Lasinio G., Rossi C., Bove T. (2004), The speaker recognition problem, *Proceedings XLII Scientific Meeting of the Italian Statistical Society (SIS)* (Sessioni plenarie, Sessioni specializzate), Coop.Libreria Editrice Università di Padova, Padova, 429-440.
- Romito L., Galatà V. (2004), Towards a protocol in speaker recognition analysis, *Forensic Science International*, 1st Workshop of Mediterranean Academy of Forensic Sciences, Vol 146S pp.S107-S111.
- Alexander A., Botti F., Dessimoz D., Drygajlo A. (2004), A The effect of mismatched recording conditions on human and automatic speaker recognition in forensic applications, *Forensic Science International*, 1st Workshop of Mediterranean Academy of Forensic Sciences, Vol 146, pp S95-S99.
- Botti F., Alexander A., Drygajlo A. (2004), On compensation of mismatched recording conditions in the Bayesian approach for forensic automatic speaker recognition, *Forensic Science International*, 1st Workshop of Mediterranean Academy of Forensic Sciences, Vol 146, pp S101-S106.
- Mori, L. & Paoloni, A. (2004), Sulla sociolinguistica forense: la costituzione di corpora vocali per l'analisi della velocità di articolazione. In *Costituzione, gestione e restauro di corpora vocali* (De Dominicis, A./Mori, L./Stefani, M.), Roma: Esagrafica, pp. 75-80.

²⁷ Sicuramente qualche cosa ci sarà sfuggita, riponiamo quindi la nostra fiducia e speranza nel lettore attento che possedendo informazioni diverse sarà così cortese da segnalarcele.

²⁸ Con il termine pubblicati in Italia ci riferiamo in questo caso a pubblicazioni frutto di convegni o workshop o altro che hanno avuto luogo in Italia.

²⁹ Che nel tempo va ben oltre gli ultimi cinque anni sui quali stiamo indagando.

- Bove T., Giua P.E., Forte A., Rossi C. (2002), Un metodo statistico per il riconoscimento del parlatore basato sull'analisi delle formanti, *Statistica*, anno LXII, n° 3.
- PAOLONI, A. (2002), La voce come elemento di identificazione della persona. In: A. De Dominicis (a cura di), *La voce come bene culturale*, Roma, Carocci: 125-139.

Come è facile notare, su sette lavori pubblicati, ben cinque affrontano il problema prettamente dal punto di vista statistico, mentre solo due trattano problemi linguistici. Tale dato non si riflette affatto poi nella prassi quotidiana dove la parte statistica spesso non viene assolutamente considerata (cfr. § 7.2)

I CONVEGNI SULL'ARGOMENTO

La cosa che colpisce maggiormente è che, nonostante i giornali ogni giorno facciano riferimento a casi giudiziari frutto di intercettazioni³⁰, non vi siano convegni dedicati esaminando il problema in questione.

Andando a ritroso qualcosa del genere può essere rintracciato nel Convegno del 1° MAFS³¹ Workshop tenutosi ad Isola Capo Rizzuto l'11-12 giugno 2004 e organizzato dallo Studio Indagini Mediche E Forensi (SIMEF) di Reggio Calabria: in tale occasione fu appositamente istituita una sezione nel programma del convegno che potesse ospitare un nostro intervento a quello di altri due ricercatori svizzeri. Risulterà in tal caso divertente osservare il logo dell'associazione all'interno del quale manca qualsiasi riferimento all'elemento voce.

L'argomento ha suscitato interesse durante la conferenza TAL2006, *Uomini e macchine: un colloquio possibile* dove all'interno della sessione "Sessione III – Intelligence" ha trovato spazio un intervento del Dirigente del Servizio della Polizia Scientifica Tommaso Bove, *Metodologie concernenti il riconoscimento del parlatore applicate alle attività preventive e giudiziarie*.

Altro incontro degno di menzione va individuato nella Conferenza TIPI (Tecnologie Informatiche nella Promozione della lingua Italiana) tenutasi presso il Ministero delle Comunicazioni dal 25-26 giugno 2002 all'interno del quale è stato presentato un intervento di Serafino Liberati (Generale dei Carabinieri) dal titolo *Dall'impronta digitale all'impronta vocale: lavori in corso*.

Oltre a ciò va ricordato che ogni anno, all'interno del convegno annuale dell' AISV³² una sezione viene dedicata al riconoscimento del parlatore, ma i partecipanti sono sempre gli stessi e cioè due o tre persone al massimo.

I PROGETTI FINANZIATI

Negli ultimi cinque anni l'unico progetto italiano finanziato di cui si ha conoscenza è SMART (Statistical Methods Applied to the Recognition of the Talker) di cui sono stati già finanziati negli anni scorsi due step, mentre è attualmente in corso il terzo. Si tratta di un progetto di ricerca finanziato dalla Commissione Europea nell'ambito del programma OISIN e coordinato dal Ministero dell'Interno - Dipartimento della Pubblica Sicurezza - Direzione Centrale della Polizia Criminale - Servizio Polizia Scientifica a cui partecipano diversi paesi europei e 20 partecipanti: 11 membri della Polizia Scientifica italiana, 1 di quella spagnola, 1 di quella francese, 1 di quella belga, 2 ricercatori del C.N.R e 4 Docenti Universitari italiani³³.

5. ANAGRAFE DEI FONDI INVESTITI PER LA FORMAZIONE DI PERSONALE DIPENDENTE DELLE FORZE DELL'ORDINE E PER CORSI REGIONALI O DI PERFEZIONAMENTO ISTITUITI DALLE SINGOLE UNIVERSITÀ

Partendo da un concorso per ufficiali, pubblicato nella G.U. della Repubblica Italiana 4° serie Speciale n. 99 del 29/12/1995 pag. 22, con particolare riferimento al comma 5 *Metodi di identificazione vocale*, viene da chiedersi quale corso di studi o di formazione possa offrire le seguenti conoscenze/competenze³⁴:

- concetto di "impronta vocale"
- riconoscimento del parlatore
- analisi dialettologica e sociolinguistica
- analisi strumentale dei parametri spettrali e temporali del segnale vocale

³⁰ Cfr. Lionetti F., *La voce intercettata: 2006 anno record*, Tesi di Laurea, relatore prof. Romito L., a.a. 2006-2007.

³¹ Mediterranean Academy of Forensic Sciences, orientata più alla criminalistica.

³² Associazione Italiana di Scienze della Voce.

³³ Che sono: Carla Rossi (Docente presso il Dipartimento di Matematica - Università degli Studi di Roma "Tor Vergata"), Giovanna Jona Lasinio (Docente presso il Dipartimento di Statistica, Probabilità e Statistiche Applicate - Università degli Studi di Roma "La Sapienza"), Angelo M. Luciano (Docente presso il Dipartimento di Elettronica - Università degli Studi "Federico II" di Napoli), Luciano Romito (Docente presso il Dipartimento di Linguistica - Università degli Studi della Calabria).

³⁴ Nella sezione riguardante i programmi del concorso si fa esclusivo riferimento a laureati in fisica o ingegneria elettronica evidenziando quindi l'assoluta mancanza di considerazione per figure quali quella del linguista o del fonetista nonostante, come riportato, vi sono discipline quali dialettologia e sociolinguistica, concetti di "impronta vocale" e di acquisizione del "saggio fonico".

- spettrografo elettroacustico e sonogramma
- operazioni di confronto vocale automatico e semiautomatico
- software esistente e dedicato al riconoscimento
- limiti del metodo strumentale
- archiviazione e classificazione delle voci
- miglioramento qualitativo del segnale vocale (filtraggio)
- acquisizione del “saggio fonico”.

In Italia vengono attivati innumerevoli master in criminalistica o intelligence³⁵, ma al momento attuale **nessuno** di questi prevede tra le proprie discipline di insegnamento la voce, o la linguistica in genere. È stato rintracciato sul web solo un corso di “Riconoscimento Vocale” organizzato da un Istituto privato di Palermo.

Divisi per città quindi abbiamo i seguenti master:

- Roma - Università degli Studi di Roma "La Sapienza":

- Master Internazionale in Scienze Criminologico Forensi,
- Master in Teorie e Metodi dell'Investigazione Criminale- Facoltà di Scienze della Comunicazione,
- Master in Criminologia e Scienze Strategiche,
- Master in Scienze Forensi (Criminologia-Investigazione-Security-Intelligence),

- Roma - Libera Università degli Studi "S. Pio V":

- Master in Scienze Criminologiche, Investigative e della Difesa,

- Varese - Università Carlo Cattaneo – LIUC:

- Master in Criminologia Forense - MA.CRI.F.,

- Messina - Università degli studi di Messina:

- Master in Criminologia,

- Varese - Università dell'Insubria:

- Master in Criminologia,

- Siena - Università di Siena:

- Master in Criminologia,

- Campobasso - Università degli Studi del Molise:

- Master in Criminologia: Teorie della Devianza, Strategie della Prevenzione e Tecniche Investigative,

- Parma - Università degli Studi di Parma:

- Master universitario di II livello in scienze forensi.

Anche per quanto riguarda i Corsi di Laurea, almeno dalle informazioni in nostro possesso, nessun insegnamento del settore scientifico disciplinare di L-Lin/01 Linguistica e Glottologia è previsto.

- L'Aquila:

- Corso di Laurea in Scienze dell'Investigazione,

- Perugia - Università di Perugia:

- Corso di laurea in Scienze per l'investigazione e la sicurezza - Facoltà di Scienze della Formazione,

- Milano - Università Cattolica di Milano:

- Laurea specialistica con percorso "criminalità e sicurezza", Facoltà di Sociologia.

Come è facile notare, non esiste allo stato attuale nessun percorso universitario o istituzionale che preveda la formazione di una figura professionale che si occupi di tutte le problematiche legate alla voce o che possa fornire ai candidati di concorsi quali quello sopra menzionato le competenze richieste nel bando.

6. MODALITÀ DI ACQUISIZIONE DEI DATI: IL QUESTIONARIO E LE INTERVISTE

Per il reperimento delle informazioni necessarie allo sviluppo dei punti 5-7 delle fasi di attuazione del presente progetto, ovvero:

- anagrafe delle professionalità coinvolte; di tutti coloro che utilizzano la fonetica forense e le tecniche di SR in ambito forense;

- anagrafe delle competenze richieste dai Tribunali e dalle Procure per poter utilizzare ed adoperare lo SR;

- anagrafe di tutti i metodi utilizzati presso le Procure e le Aule di Tribunale e non solo quelli conosciuti in ambito scientifico o perché divulgati attraverso pubblicazioni scientifiche;

è stato somministrato a tutti i periti/consulenti fonici, o conosciuti come tali, un questionario telefonico di 66 domande suddivise in tre tipologie: una dedicata all'individuazione delle professionalità coinvolte e operanti nel campo delle comparazioni foniche in ambito forense, una dedicata alla metodologia utilizzata e al modus operandi dell'intervistato, e un'altra di tipo informativo-generale.

Vista l'assenza di un Albo professionale specifico (cfr §§ 6.2 e 6.3), per il reperimento dei nominativi da contattare la nostra inchiesta è iniziata a livello locale con una lista di nomi di cui avevamo conoscenza diretta e in corso d'opera,

³⁵ Non si fa riferimento qui ai Master specifici su discipline cliniche, psicologiche o economiche.

dietro nostra esplicita richiesta, è stato chiesto agli stessi intervistati un elenco di nominativi di altri periti di loro conoscenza. Tale elenco ovviamente non era obbligatorio ma affidato alla discrezione dell'intervistato.³⁶

Le interviste sono state effettuate per telefono³⁷ previa presentazione e illustrazione degli scopi del progetto, del dichiarato consenso da parte dell'intervistato e successivo appuntamento. Le interviste hanno avuto una durata variabile, a seconda della disponibilità dei nostri interlocutori, da un minimo di 15 min. ad un massimo di 30 min³⁸. A ciclo di interviste concluso³⁹, durato all'incirca 6 mesi e completamente auto-finanziato, è stata raggiunta, con nostra grande sorpresa, quota 94 nominativi distribuiti su quasi tutto il territorio nazionale⁴⁰. Tuttavia dei 94 nominativi non tutti si sono resi disponibili declinando qualsiasi forma di collaborazione⁴¹, mentre un'altra parte ci è risultata non reperibile per una serie di motivi che qui non verranno menzionati.

Va inoltre precisato che agli intervistati è stata garantita la massima riservatezza e che nel trattamento dei dati si sarebbe fatto l'impossibile per garantire l'assoluto anonimato in merito alle risposte fornite. La disparità e la varietà dei casi ha spesso reso complicato il trattamento statistico delle risposte ma nonostante ciò abbiamo sempre preferito garantire la riservatezza dell'interessato.

Il campione risulta pertanto essere composto per il 52% da soggetti intervistati, 19% da soggetti che si sono rifiutati di collaborare e un 29% da soggetti non reperibili come si può tra l'altro facilmente evincere dal grafico di seguito proposto. Nel prosieguo dell'indagine tutte le analisi e tutte le valutazioni verranno pertanto effettuate ovviamente su coloro che si sono resi disponibili a "chiacchierare" con noi senza però dimenticare che proprio quel 19% di rifiuti e probabilmente il 29% dei non reperibili potrebbero in alcuni casi falsare in grande misura i nostri dati come di volta in volta metteremo in evidenza.

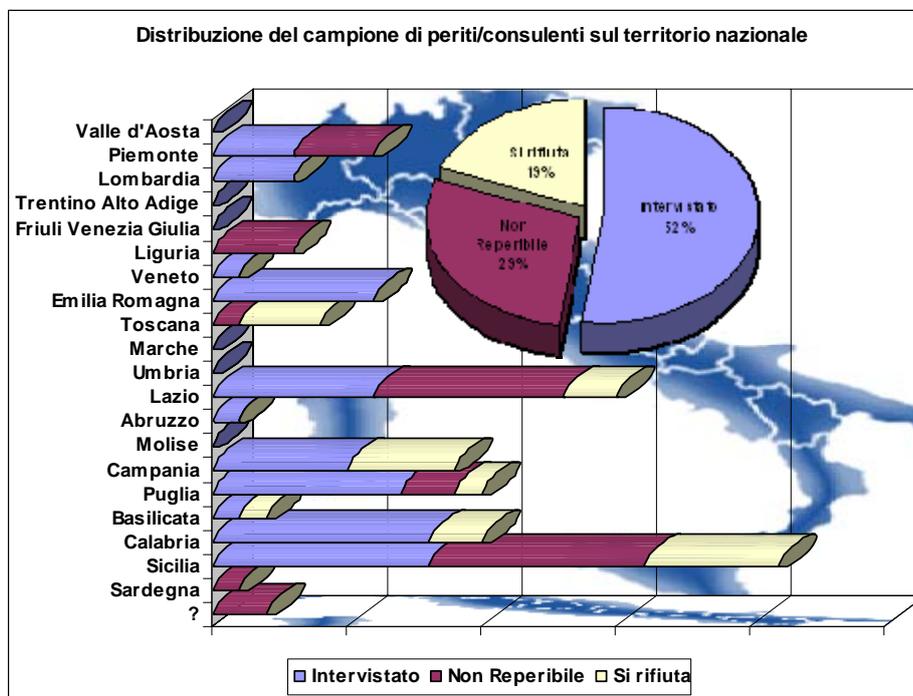


Figura 2: Il campione di indagine.

6.1 ANAGRAFE DELLE PROFESSIONALITÀ COINVOLTE

Da una prima analisi dei dati emerge che il 57% degli intervistati è in possesso di una laurea, ben il 33% è in possesso di un diploma di scuola secondaria superiore, il 4% ha un titolo conseguito al conservatorio e il 6% si divide invece tra

³⁶ Facciamo notare che il numero di periti complessivo raggiunto potrebbe tuttavia non essere completo in quanto ad ogni nuova intervista la lista dei nominativi veniva costantemente aggiornata ed ampliata.

³⁷ La scelta del canale telefonico è stata voluta e qualsiasi altra forma di somministrazione del questionario è stata da parte nostra rifiutata al fine di avere informazioni il più possibile veritiere e spontanee.

³⁸ Non nascondiamo che in alcuni casi c'è stata talmente tanta interazione che l'intervista è durata oltre 1 ora.

³⁹ Per ovvi limiti di tempo non ci è stato possibile protrarre oltre l'inchiesta.

⁴⁰ Il fatto che in alcune regioni manchino nella presente inchiesta figure che si occupino di SR è data dal fatto che, per quanto il censimento sia stato meticoloso, non sempre si è riusciti ad avere sufficienti informazioni al reperimento dei nominativi per quelle regioni.

⁴¹ In tutti i casi su richiesta degli interessati abbiamo anche provveduto ad inviare via mail una nostra lettera di presentazione con una descrizione delle finalità dell'intervista.

diploma di scuola secondaria inferiore (scuola media) e altro. Inoltre vorremmo sottolineare che, dal punto di vista delle professionalità coinvolte, quel 48% dei non intervistati per rifiuto o perché non reperibili falsa, a nostro avviso, enormemente il risultato. Riteniamo infatti che dietro al rifiuto si possano nascondere proprio quelle persone senza alcun titolo e con una professionalità dubbia. Non si capirebbe altrimenti la scelta adottata.

Dai dati in nostro possesso, comunque, l'Italia, risulta essere divisa in due: poco più del 50% è in possesso di competenze attestate (laurea ed altro) mentre il restante 50% risulta avere piuttosto competenze *acquisite sul campo*.

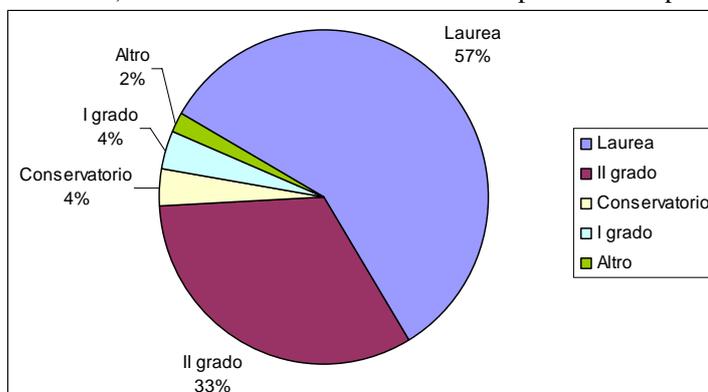


Figura 3: Titoli di studio.

Analizzando nel dettaglio i dati, scopriamo che per entrambi i gradi di istruzione, laurea e scuola secondaria superiore (non trattiamo ovviamente qui quel 6% di Scuola Media ed altro, perché lo riteniamo assolutamente inutile), prevalgono determinati curricula di studio come ben si evince dai grafici di seguito riportati. Nel primo caso prevale di gran lunga la competenza tecnica, ovvero l'ingegnere con una percentuale che arriva addirittura al 60%, o il perito industriale, nel caso del diploma di scuola superiore che raggiunge il 50%.

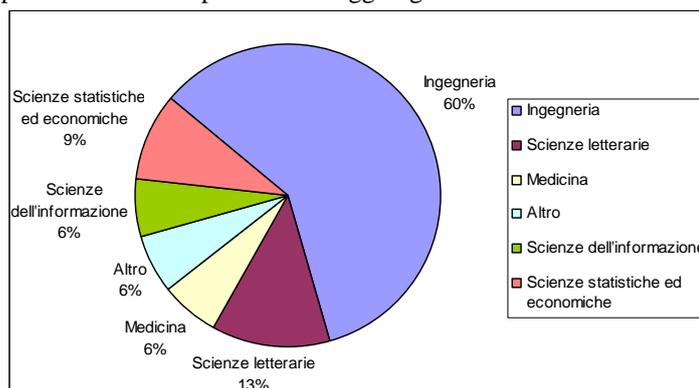


Figura 4: Titoli di studio: Laurea.

Senza voler essere provocatori, è comunque alquanto strano che proprio la figura che dell'elemento "voce" ne fa il proprio oggetto di studio, cioè il linguista, sia scarsamente rappresentata (13% nel raggruppamento denominato Scienze Letterarie).

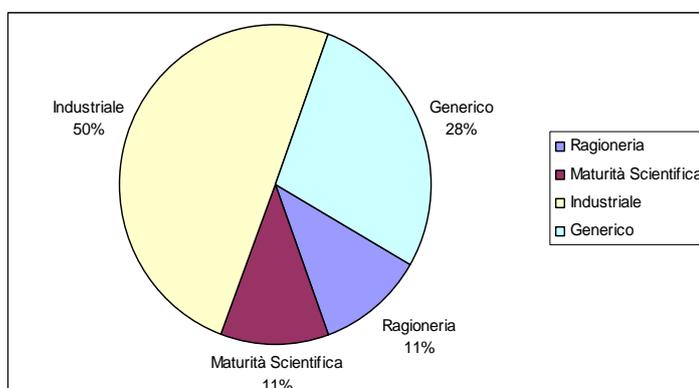


Figura 5: Titoli di studio: Diploma di Scuola Superiore

Un altro dato che lascia presupporre come spesso le conoscenze siano state acquisite sul campo viene fornito dalla distribuzione dell'età degli intervistati. La figura del perito fonico esiste ormai da tanti anni, non è un caso quindi, che oltre il 50% degli intervistati abbia dai 50 anni in su. Il perito più giovane ha solo 29 anni ed è diplomato, come anche il più anziano che ne ha 72.

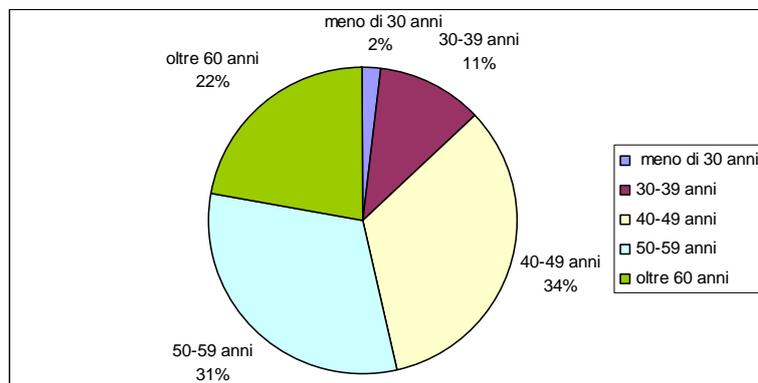


Figura 6: Età degli intervistati.

Se analizziamo però la correlazione tra l'età e il titolo di studio, si nota che l'età media di un laureato è 53 anni mentre quella di un diplomato è 49. Anche se non vi è una grande differenza, osservando la tabella seguente si evince che oggi sempre più spesso ci si avvicina alla *professione* del perito fonico con il solo titolo di studio del Diploma di Scuola Superiore.

Laureati	Età (in anni)	Diplomati ⁴²
0%	<30	2%
4%	da 30 a 39	7%
21%	da 40 a 49	9%
17%	da 50 a 59	13%
15%	>60	5%

Tabella 1: Correlazione tra età e titolo di studio.

6.2. ANAGRAFE DELLE COMPETENZE RICHIESTE DALLA MAGISTRATURA

Per quanto riguarda le competenze richieste dalla magistratura, basti considerare che la prassi comunemente adottata, in genere per la nomina di un perito, è legata *al passaparola* o al *rapporto di fiducia*⁴³ (quindi non al titolo posseduto o all'età). A tal proposito risulterà forse interessante sapere che il 67% degli intervistati ha dichiarato di essere stato convocato direttamente dalla magistratura per **il suo primo incarico** (63% di laureati e 36% di diplomati), mentre solo il 33% ha dichiarato di aver **prima** appreso "le competenze necessarie" presso lo studio di un perito fonico.

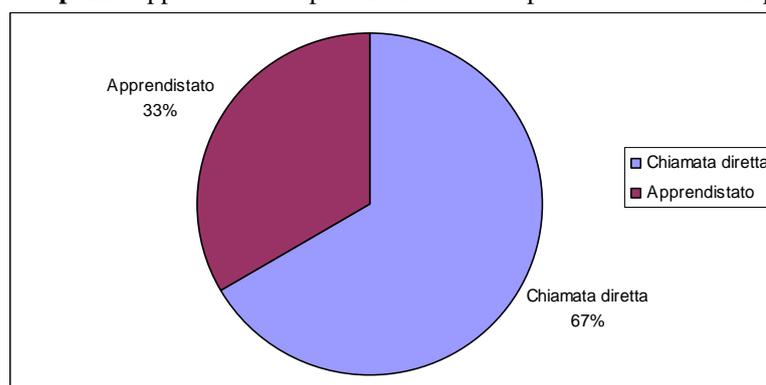


Figura 7: Modalità di avviamento all'attività di perito.

⁴² Viene qui conteggiato tra i diplomi anche il Conservatorio.

⁴³ Come ci è stato risposto dalla quasi totalità degli intervistati (solo 2 persone dichiarano di aver iniziato la loro attività grazie all'iscrizione presso l'Albo Generale del Tribunale).

È interessante anche in questo caso la correlazione con il titolo di studio. Ci si aspetterebbe infatti che l'apprendistato riguardi solo gli intervistati diplomati, invece il dato si divide quasi al 50%⁴⁴.

	Apprendistato	<u>Chiamata diretta</u>
Laureati	15%	42%
Diplomati	18%	21%
Altro ⁴⁵	0%	4%

Tabella 2: Correlazione tra titolo di studio e modalità di avviamento all'attività di perito.

Da questi dati, l'unica cosa ipotizzabile è che la Magistratura non ha un'idea chiara su quale competenza ricercare per nominare un perito o un consulente ma basa la propria scelta sul piano della fiducia e della conoscenza personale.

A questo punto non possiamo non affrontare la questione dell'incarico e di tutte le problematiche ad esso connesse.

6.3. L'AFFIDAMENTO DELL'INCARICO PERITALE

Un problema che in Italia è molto sentito, riguarda l'incarico peritale e l'esistenza di un Albo Professionale.

Bisogna precisare che in Italia esistono due tipi di *albo*: uno custodito presso il Tribunale Civile, chiamato *Albo dei consulenti tecnici*, e il secondo custodito presso il Tribunale Penale, chiamato *Albo dei periti*.

L'iscrizione all'Albo prevede una «speciale competenza nella materia»⁴⁶. Tale *competenza* deve essere dimostrata attraverso la presentazione di **titoli e documenti** che la attestino. Ad oggi questo è l'unico criterio previsto dalla normativa volto ad accertare la competenza del candidato. Un Comitato presiede alla scelta degli iscritti e provvede alla revisione periodica dell'Albo, al fine di cancellare eventuali iscritti per i quali sia venuto meno uno dei requisiti necessari o sia sorto un impedimento per l'esercizio di perito⁴⁷.

A tutto ciò però devono essere aggiunte due informazioni fondamentali: 1) l'elenco delle competenze previste all'interno degli Albi non prevede la figura di **perito fonico o linguistico** o simile, e 2) che l'art. 67 disp. att. c.p.p. prevede la possibilità di nomina di un perito non iscritto all'Albo. Tale scelta non causa la nullità dell'atto, ma implica solo l'obbligo per l'autorità procedente di motivare tale decisione. Ne consegue che il ricorso all'Albo ai fini della nomina non è obbligatorio. Quindi, se, come già osservato, l'iscrizione all'Albo non garantisce automaticamente delle competenze, la non iscrizione, a nostro parere, la garantisce ancor meno.

La prima considerazione che deve essere effettuata è: se per iscriversi all'Albo è necessario avere titoli che attestino le proprie competenze, per una persona non iscritta (sia perché l'albo non esiste, sia per scelta personale) decade sia giuridicamente che moralmente la necessità di avere titoli che attestino tale competenza? In Italia ciò che si evince, almeno dai dati certi in nostro possesso⁴⁸, dei periti nominati dal Tribunale il 4% dichiara di avere un titolo di studio di I grado, il 33% un titolo di studio di II grado e il 4% un titolo di studio conseguito al conservatorio. Quindi un 40% di persone che oggi in Italia svolge perizie foniche con titoli al più di Scuola Media Superiore⁴⁹.

Dai dati forniti in Commissione giustizia alla Camera dal sottosegretario Luigi Li Gotti nel 2005⁵⁰, sono stati spesi ben 307.346.676,72 euro in intercettazioni di comunicazioni sia ambientali che telefoniche, ben una volta e mezzo rispetto a quanto servirebbe ai rettori per far funzionare al meglio tutte le Università. A questo dato deve essere aggiunta la mole di liquidazioni di consulenze, filraggi, trascrizioni ecc.⁵¹. Nonostante questo enorme capitolo di spesa però, non esiste, allo stato attuale, né un protocollo certificato sulle operazioni peritali da fare, né corsi professionalizzanti, né un Albo professionale (aggiungiamo inoltre che ben il 67% delle perizie e delle consulenze vengono assegnate dal Tribunale e dalla Procura, non dalla Difesa).

6.5. LA FIGURA DEL PERITO VS. LA FIGURA DEL CONSULENTE TECNICO

Avendo affrontato nel precedente § la questione dell'incarico peritale e la questione dell'albo ci pare opportuno operare qui un distinguo tra due figure del sistema giudiziario italiano, ovvero quella del perito e quella del consulente tecnico.

⁴⁴ Come informazione aggiuntiva, vorremmo sottolineare che gli intervistati con titolo di studio di scuola media inferiore, sono stati chiamati e nominati direttamente dalla Magistratura.

⁴⁵ Compreso il titolo di scuola secondaria inferiore (scuola media).

⁴⁶ (art. 69 norme di attuazione c.p.p.).

⁴⁷ (art. 68 norme di attuazione c.p.p.).

⁴⁸ Si rammenta che manca qui all'appello un 19% di rifiuto all'intervista, il che lascerebbe presupporre quanto stiamo dicendo, e un 29% sul quale purtroppo possiamo dire poco.

⁴⁹ Per un dettaglio sui tipi di diploma più attestati si veda il precedente § 6.1.

⁵⁰ Cfr. www.giustizia.it.

⁵¹ Dal campione intervistato è emerso che mediamente vengono effettuate qualcosa come ca. 600 comparazioni foniche l'anno.

La nomina di un perito è motivata dalla necessità del giudice di avvalersi di competenze tecniche specializzate che lui stesso non possiede in prima persona ma che gli sono indispensabili per accertare i fatti e formulare un giudizio, come regolamentato dall'art. 220.1 c.p.p.

La perizia assume perciò un carattere di *eccezionalità*, dal momento che il perito viene nominato esclusivamente quando la sua presenza è necessaria per colmare una lacuna di **competenze tecniche** da parte dell'autorità, che deve pronunciarsi su questioni talvolta **estremamente specialistiche**⁵².

L'apprezzamento della necessità o meno di un'indagine peritale è oggetto di valutazione discrezionale da parte del giudice, che "preferibilmente" ma non necessariamente, sceglie fra i nominativi iscritti nell'apposito Albo (art. 221 c.p.p.).

Tale necessità deve essere considerata anche nell'ottica delle parti, per le quali la perizia è una garanzia di giudizio tecnico qualificato, nonché di decisione corretta e ponderata da parte dell'autorità. Ciò non significa però che il perito sia un "giudice tecnico", poiché la perizia, una volta acquisita agli atti del procedimento, non ha potere vincolante ai fini della decisione dell'autorità, che la considera alla stregua degli altri elementi probatori che vengono valutati, utilizzati oppure non condivisi. In tal senso la perizia è un "**mezzo di prova neutro**", indipendente dalle parti, le quali dispongono del potere di nomina solo riguardo alla figura del consulente tecnico. Esse hanno, infatti, facoltà di nominare un *consulente tecnico "di parte"* che assista alle operazioni del perito, ma che non è un perito "di parte".

Per ovvi motivi d'imparzialità e di conflitto di interessi, l'art. 222 c.p.p. vieta che l'incarico di perito venga attribuito a chi sia già stato nominato consulente tecnico nell'ambito dello stesso procedimento o di un procedimento connesso.

Il conferimento dell'incarico, che avviene tramite ordinanza del giudice (art. 224 c.p.p.), consente al perito di intraprendere gli accertamenti necessari a rispondere per iscritto ai quesiti oggetto della perizia.

Va ricordato ancora che la prestazione dell'ufficio di perito è obbligatoria (art. 221.3 c.p.p) e il rifiuto dell'incarico (art. 366 c.p.), così come la falsa perizia sono perseguibili a norma di legge (art. 373 c.p.)⁵³.

La procedura penale opera una differenza fondamentale fra perizia e consulenza tecnica. La prima, come precedentemente riportato è disposta dal giudice, la seconda dalle parti.

Dall'art. 225 c.p.p. emerge chiaramente la caratteristica "di parte" del consulente tecnico (d'ora in poi CT), che può essere nominato dal Pubblico Ministero per integrare le indagini peritali o dalle parti private, a **proprie spese**, allo scopo di integrare l'attività della difesa. Il giudizio dei consulenti di parte può contrapporsi a quello formulato sulla medesima questione da periti e consulenti del Pubblico Ministero.

Il consulente tecnico è vincolato al rispetto di determinati doveri professionali **nel solo interesse della parte che lo ha nominato** (a questo riguardo il 68% del nostro campione ha risposto di mantenere sempre lo stesso atteggiamento sia quando si tratta di incarico di perizia e sia quando si tratta di incarico di parte). I reati previsti (artt. 380 e 381 c.p.) sono la collusione con la parte avversaria, la prestazione contemporanea del patrocinio o della consulenza a favore di parti contrarie, l'assunzione di patrocinio o consulenza della parte avversaria dopo aver assistito una parte e senza il consenso di questa ultima.

I consulenti tecnici partecipano allo svolgimento della perizia sin dalla prima fase di formulazione dei quesiti da parte dell'autorità, e anche alle operazioni peritali, con facoltà di proporre al perito indagini, così come di presentare riserve e osservazioni, che devono entrare a far parte della relazione peritale. Si potrebbe dire pertanto che il consulente tecnico esercita in alcuni casi una vera e propria funzione di "controllo" sull'attività del perito nominato dall'autorità, come regolamentato all'art. 230 c.p.p..

Contrariamente a quanto accade per i periti, non vi sono vincoli normativi riguardo ai tempi di nomina dei consulenti tecnici, che possono essere nominati anche una volta conclusa la perizia. In tal caso i consulenti hanno facoltà di esaminare la relazione peritale, nonché di richiedere al giudice di esaminare la persona, il luogo o l'oggetto della

⁵² Qui inevitabilmente torna in mente quanto precedentemente affermato: frutto di esperienza personale o frutto di meriti acquisiti e titoli attestanti?

⁵³ Anche questo rientra tra i motivi che probabilmente allontanano gli Accademici e i ricercatori dall'interessarsi della questione. Questi ultimi, per spostarsi, sono costretti ad anticipare tutte le spese (di solito percependo l'onorario con minimo un anno di ritardo) e a svolgere un lavoro con molta celerità. Per questa ragione il PM o il Giudice tendono spesso a preferire degli esperti "locali" e quindi facilmente rintracciabili, quasi dei "dipendenti" che conoscono la vita del Tribunale (solo il 33% degli intervistati dichiara di accettare incarichi in tutto il territorio nazionale). Una volta conferito l'incarico, l'esperto della disciplina, dovrà infatti conoscere in qualche modo il Tribunale in quanto sarà suo compito, previa nullità, avvisare le parti (con raccomandata A/R e non con fax), richiedere copia dell'incarico o copia della trascrizione di udienza relativa all'incarico per prendere visione del quesito posto e delle autorizzazioni ottenute/concesse (ovviamente precedentemente richieste), in seguito avere conoscenza di mille sigle (R.I., p.p., r.g.n.r., DDA, ecc.) che condurranno l'esperto presso l'Ufficio Reperti della Procura competente (non necessariamente sempre nello stesso Tribunale), dove dovrà poi ottenere autorizzazione ad estrarre copia o a prelevare i reperti (a volte su supporti in uso solo ed esclusivamente alle forze dell'ordine cui segue un'altra serie di procedure e di autorizzazioni ecc.). Insomma, tutta questa fase di "lavoro" non riconosciuta e liquidata come semplice vacanza deve necessariamente rientrare tra le competenze dell'esperto.

perizia, al fine di verificare, nell'interesse della parte che li ha nominati, la correttezza e l'affidabilità delle stesse indagini peritali.

È necessario, anche in questo caso, fare alcune considerazioni:

1. il ricorso agli albi non è assolutamente obbligatorio e le autorità possono nominare chiunque senza alcun controllo. La garanzia è data dal fatto che ogni parte ("a proprie spese") può nominare il proprio consulente tecnico⁵⁴;
2. non è inusuale trovare un CT molto esperto e un Perito assolutamente inadatto e incompetente, quindi la funzione di controllo diventa o materialmente impossibile o molto vincolante (a questo proposito ben l'81% del nostro campione ha dichiarato di aver avuto come controparte una perizia inaccettabile da un punto di vista scientifico, e ben il 63% è stato nominato per ri-eseguire una consulenza o una perizia precedentemente eseguita da altri periti);
3. così come nell'insegnamento anche nelle perizie oggi è diventata importante la presentazione in Aula. La competenza è sempre da tenere in alta considerazione ma alcuni periti hanno difficoltà ad esprimere adeguatamente questa professionalità all'interno di un processo. È necessaria una preparazione adeguata per gli 'attacchi' delle parti. Spesso il perito in Aula deve 'difendere' la propria posizione nei confronti della parte avversa. È così che vengono richieste nullità di perizia, non per il metodo utilizzato ma perché non è stata comunicata per mezzo di raccomandata l'inizio delle operazioni peritali, o perché mancano a verbale alcuni dati riguardanti il giuramento di collaboratori se questi prendono parte al giudizio o all'interpretazione dei dati, o ancora se il numero delle pagine della relazione non corrisponde con quanto dichiarato nel verbale di deposito della relazione ecc. Spesso una delle parti cerca la contraddizione nel perito, si pensi alle perizie di trascrizione che a volte sono la parte preliminare di una comparazione. Tali trascrizioni possono riguardare anche ore e ore di intercettazione e quindi migliaia di pagine di relazione, è assolutamente immaginabile e impensabile che il perito ricordi esattamente ogni riga di trascrizione e ogni fatto inerente quella precisa parte di registrazione.

6.6 LA FIGURA DELL'INTERMEDIARIO

Altro fenomeno che ci ha molto sconcertato, contrariamente a quanto fino ad ora affermato dalla normativa vigente, è la presenza assai diffusa e attestata di una figura che potremmo definire "intermediario". Con questo termine ci riferiamo ad un 19% degli intervistati che ha palesemente dichiarato di avvalersi di altre figure più specializzate per l'espletamento di funzioni peritali che richiedono particolari competenze.

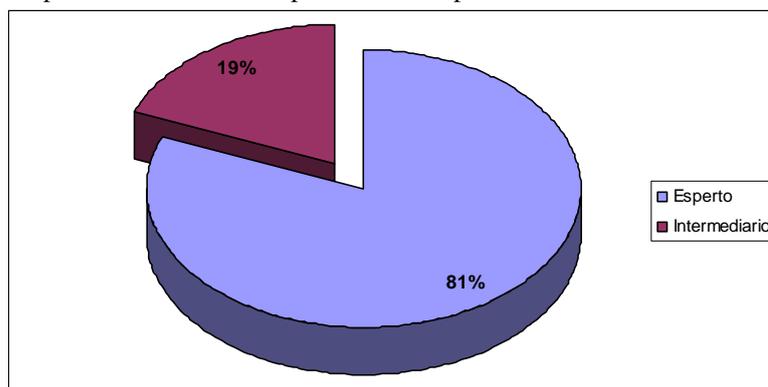


Figura 8: La figura de "l'intermediario" e la figura dell'esperto

Questo stato di incertezza conferma a pieno la situazione di aleatorietà in cui questo campo di indagine viene, forse volutamente, lasciato. Da ciò ne deriva che solo l'81% dei periti è in grado di svolgere il compito assegnatogli dal giudice o dal magistrato. Da ciò alle considerazioni precedentemente riportate si aggiungono le seguenti:

- su che base avviene la nomina del perito da parte del giudice o del magistrato?
- che titoli, qualifiche o competenze deve avere una persona per poter espletare un incarico di comparazione della voce in ambito forense?

Per quel che attiene la prima considerazione la risposta è già stata in parte data nei precedenti §§.

Riguardo la seconda considerazione invece, le risposte fornite al nostro questionario, sono alquanto diversificate e sintomatiche di un stato di assoluta incertezza che vige nel campo delle indagini foniche. Per molti intervistati è necessario avere una comprovata esperienza nel campo⁵⁵, per altri sono necessarie conoscenze di natura tecnica, di acustica, di fisiologia dell'udito ecc., altri ancora legano le competenze al tipo di metodo utilizzato. Sintetizzando, ciascuno degli intervistati difende e valorizza la propria categoria e le proprie competenze. Un dato certo è che allo

⁵⁴ Non serve citare qui alcuni casi giudiziari italiani recenti dove una difesa con grandi disponibilità economiche ha letteralmente fatto "razzia" di esperti nominando un pool composto da nomi illustri e riconosciuti a livello nazionale e costringendo il tribunale e il Pubblico Ministero alla nomina del perito tra le persone *rimaste* disponibili.

⁵⁵ Nel senso di aver già svolto in precedenza altri incarichi di questo tipo.

stato attuale (cfr. §§ 4-5) non esiste nulla di concreto che porti alla definizione di una vera e propria figura di “perito fonico”.

7. ANAGRAFE DELLE METODOLOGIE UTILIZZATE⁵⁶

I metodi rilevati attraverso il nostro questionario sono, come si potrà immaginare visto le competenze interessate, tanti e molto differenti tra di loro.

Da un punto di vista prettamente scientifico l’approccio può essere diverso interessando una sezione che potremmo definire linguistica ed una sezione parametrica. La prima con dati e variabili qualitative e la seconda con dati e variabili quantitative.

La sezione LINGUISTICA include caratteristiche soggettive come il timbro, la cadenza, l’intonazione, la costruzione sintattica e morfologica, l’analisi dialettologica, l’inventario fonologico, la presenza di alcune realizzazioni fonetiche ecc. La sezione definita PARAMETRICA include invece la frequenza fondamentale, le frequenze formantiche, la transizione formantica, la durata di alcuni segmenti ecc.

Alcuni metodi operano una analisi **discreta** altri invece una analisi **continua**. In alcuni l’operatore (per esempio nelle prove di ascolto) è **naive** in altri invece è un **tecnico**. Alcuni metodi utilizzano una statistica molto elaborata e forniscono risultati numerici altri invece non utilizzano assolutamente alcuna statistica e fornisco opinioni e considerazioni personali.

Come è facile notare è veramente difficile presentare tutti i metodi presenti in Italia quindi abbiamo deciso di creare delle categorie, delle famiglie di metodi.

Così facendo possiamo innanzitutto individuare i metodi automatici⁵⁷ e quelli semiautomatici; in secondo luogo i metodi soggettivi 24% che possono essere divisi in uditivi⁵⁸ per un 13% e di confronto dei sonogrammi 11%; infine, ultima categoria, i metodi parametrici che arrivano addirittura ad un 65%. Bisogna però considerare che in questa categoria sono state incluse tutte quelle figure che abbiamo definito più sopra come “intermediari” e che nell’affidarsi ad esterni si affidano a persone che dichiaratamente fanno uso di metodi parametrici. Inoltre, nella categoria “parametrico” abbiamo incluso coloro che fanno utilizzo di sole medie di formanti (4%) e di ellissi formantiche (11%) ecc. tutti i metodi che in qualche modo utilizzano delle misure e quindi dei parametri.

Esiste poi un 11% di metodi **assolutamente incerti e non classificabili**. In questa sede abbiamo deciso di non commentare assolutamente tali metodi, ci preme solo sottolineare che questo dato dovrebbe preoccupare molto la Magistratura Italiana.

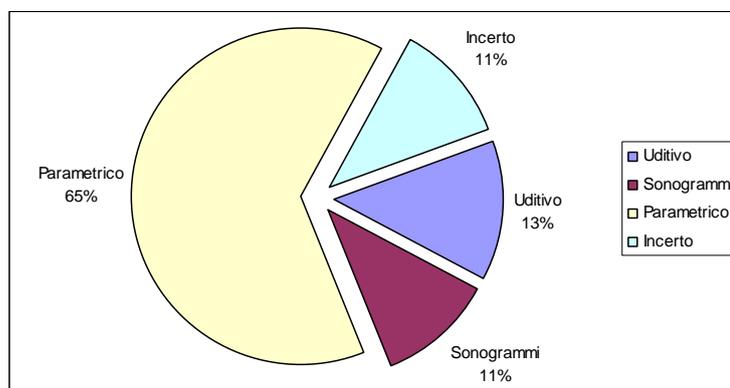


Figura 9: Metodi divisi in classi.

Prima di analizzare nello specifico, i singoli metodi, un dato che ci sembra interessante in contrapposizione a quello appena presentato, è quello relativo alla domanda fatta agli intervistati circa il *metodo che secondo loro non dovrebbe mai essere utilizzato a scopo di SR* in ambito forense. I dati di seguito riportati nel grafico sono abbastanza significativi della conoscenza dei limiti legati ai metodi stessi data la presenza di un abbondante 58% che non ha saputo o voluto dare una risposta in tal senso, della assoluta incertezza riguardo la disciplina, della variegata competenza che spazia dalla linguistica alla fisica alla medicina e all’economia. Quindi come per la domanda relativa a *quali competenze*

⁵⁶ Ricordiamo che la valutazione dei metodi attraverso il corpus Primula sarà oggetto del nostro prossimo lavoro.

⁵⁷ Che dalle nostre interviste si sono del tutto rivelati assenti al contrario di quanto invece ci saremmo aspettati avendo riscontrato tra i periti laureati che oltre il 60% proviene da percorsi scientifici ingegneristici, informatici o di altro genere (cfr. § 6.1).

⁵⁸ Non si tratta qui in nessun caso di prove di ascolto costruite con metodo statistico ma solo ascolti simultanei di due registrazioni fidandosi quindi esclusivamente del proprio orecchio e della propria esperienza. Ovviamente il risultato di tale indagine è meramente soggettiva e il Magistrato o il Tribunale dovrà fidarsi ciecamente della competenza del proprio Consulente o Perito.

dovrebbe avere secondo Lei un perito fonico?, anche in questo caso ogni intervistato difende il proprio metodo senza però esprimere giudizi netti sull'uso di metodi differenti o alternativi.

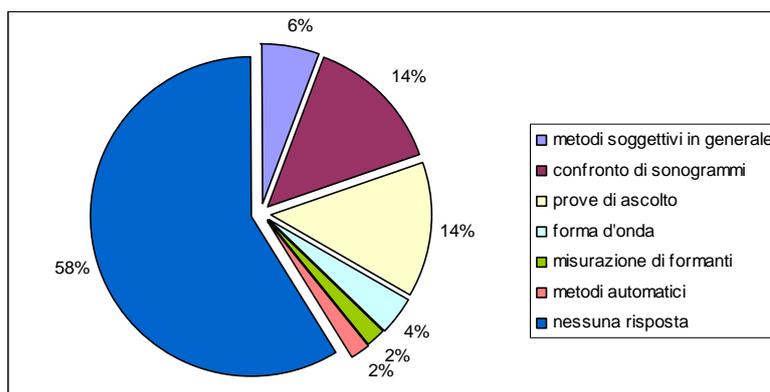


Figura 10: I metodi che secondo gli intervistati non vanno utilizzati a scopo di SR

Osservando il grafico si evince, tuttavia, come i metodi soggettivi⁵⁹ siano i più criticati in quanto per definizione considerati non oggettivi. L'intervistato in questo caso utilizza oggettivo come sinonimo di 'frutto di misure'. È singolare che molti tra i metodi così definiti *parametrici* sono molto meno precisi e molto più soggettivi di una corretta e organizzata prova uditiva.

È interessante notare il senso di incertezza anche in quel 62% degli intervistati che ha dichiarato che allo stato attuale in Italia non esiste una base di consenso comune su come, e secondo quali criteri, una perizia di comparazione fonica debba essere effettuata.

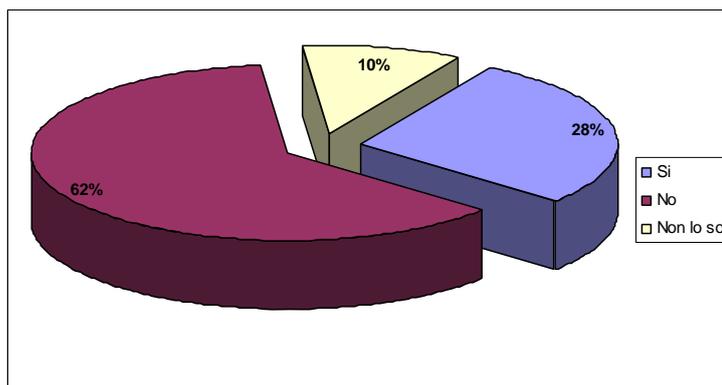


Figura 11: Esistenza di una base di consenso comune su come effettuare una perizia di comparazione fonica

Ovvio che questa è solo una situazione italiana, senza citare, infatti, i lavori riguardanti un protocollo nelle perizie di comparazione degli autori di questo lavoro, sottolineiamo che esiste un decalogo redatto dalla IAFPA (International Association for Forensic Phonetics and Acoustics) e approvato durante l'Annual General Meeting dell'associazione tenutosi ad Helsinki nel 2004⁶⁰.

IL MATERIALE UTILIZZATO

Altra questione complicata e alquanto confusa si riscontra andando ad esaminare il tipo di materiale, ovvero il *cosa* viene misurato o valutato in un'espressione verbale o un campione di voce sottoposto ad accertamento.

In Italia la maggior parte degli intervistati utilizza, nei metodi parametrici, solo le vocali (60% sia vocali toniche che atone e 40% solo vocali toniche)⁶¹.

Il tempo minimo⁶² che viene richiesto come materiale oscilla da non meno di 10 s ad un valore preferibilmente oltre i 30 s. È inutile sottolineare qui la risposta più ricorrente del tipo "è ovvio che più materiale si ha più la comparazione sarà attendibile".

⁵⁹ Naturalmente utilizzati come unico metodo per lo SR.

⁶⁰ Cfr. decalogo riportato in Appendice.

⁶¹ All'interno di questo dato vi è comunque chi nella non distinzione tra toniche e atone opera una distinzione sul piano della migliore realizzazione fonetica.

⁶² In termini di durata effettiva del campione di voce in esame.

IL TIPO DI COMPARAZIONE EFFETTUATA ED IL TRATTAMENTO STATISTICO DEI DATI

Un dato molto importante nel caso dello SR è il tipo di comparazione. Abbiamo chiesto al nostro campione se nel metodo di SR la comparazione viene effettuata tra due voci o se invece viene utilizzata una comunità di confronto. Come si evince dal grafico il 58% effettua una comparazione tra solo due voci cercando quindi la similitudine tra caratteristiche acustiche. Questo ovviamente è un dato molto importante perché non tiene conto della variabilità intra-parlatore ed inter-parlatore e non tiene conto neppure della grande variabilità della voce umana. Anche due automobili dello stesso colore, della stessa marca e dello stesso modello sono simili, anzi identiche, ma non sono affatto la stessa automobile⁶³.

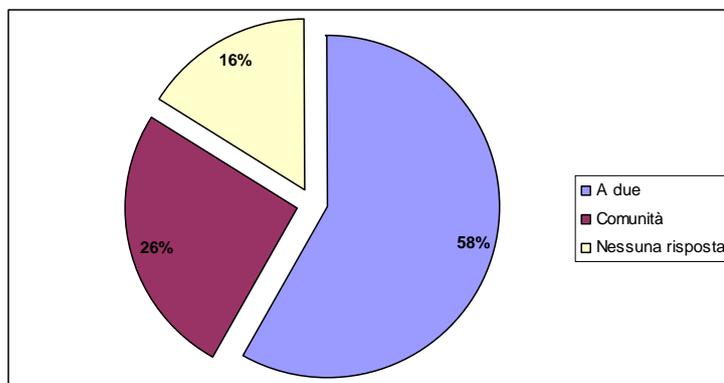


Figura 12: Tipo di comparazione effettuata

L'analisi statistica che viene effettuata⁶⁴ è strettamente legata ai dati estrapolati; quindi si nota per il confronto a due: un 6% utilizza il confronto delle medie (a volte per singola formante di ogni vocale con percentuali differenti per ogni confronto, a volte medie di medie formantiche ecc. ma comunque senza analisi statistica); a questo deve essere aggiunto un 38% che non utilizza nessuna statistica; invece un 22% utilizza analisi di Varianza, T-student per piccoli campioni, distanze di Hotelling, di Mahalanobis ecc. e solo il 26% utilizza un campione di confronto grazie ad un pacchetto informatico chiamato Spread e commercializzato dalla FUB. Vi è poi un 2% che fa uso di un pacchetto statistico sviluppato col progetto SMART che più sopra abbiamo menzionato.

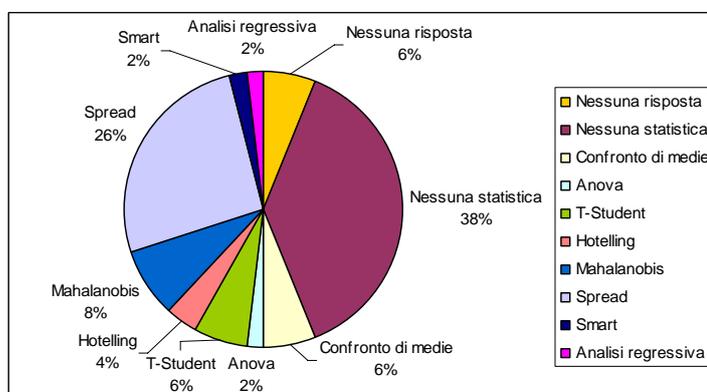


Figura 13: Il tipo di statistica utilizzata nel trattamento dei dati estrapolati.

LA RISPOSTA AL QUESITO DELL'INCARICO AFFIDATO

Questa è forse la parte più importante dello SR e della relazione peritale, in quanto è quella parte non tecnica, notoriamente letta in aula durante il contraddittorio. Al quesito posto dal Magistrato: *le due voci appartengono allo stesso locutore?*, i periti intervistati rispondono con un si/no secco il 6%; con una percentuale di similitudine il 36%; con un grado di compatibilità il 9%; con un giudizio lungo una scala di similitudine che va dall'improbabile al molto probabile il 21%; con un rapporto di verosimiglianza e quindi un calcolo dell'errore di Falsa Identificazione il 13%.

⁶³ Ringrazio il mio amico Andrea per questo esempio a parer mio molto esemplificativo.

⁶⁴ Per una discussione approfondita sui dati e sui parametri si veda Romito L. (2000), *Manuale di Fonetica Articolatoria, Acustica e Forense*, Centro Editoriale e Librario - Università degli Studi della Calabria, Cosenza.

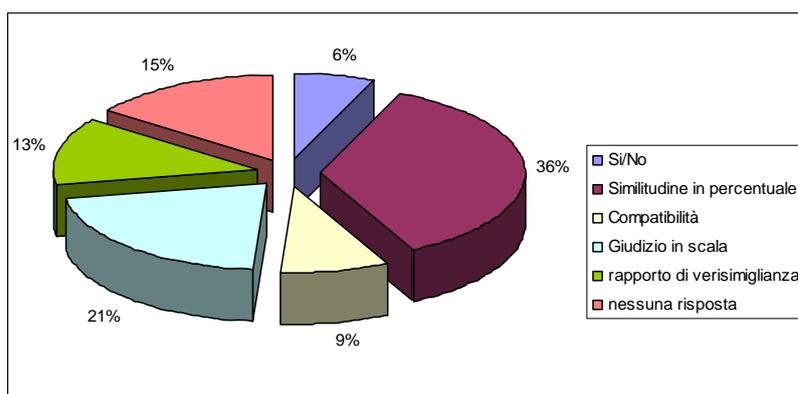


Figura 14: Il tipo di risposta al quesito del Magistrato

8. CONCLUSIONI

Quali conclusioni trarre da questo lavoro? Inutile dire che molto di quanto riscontrato non è altro che una conferma di quanto ci aspettavamo. In Italia si intercetta più che in ogni altro stato europeo. Le intercettazioni riguardano essenzialmente la voce (rispetto al video) e sono di tipo ambientale e telefonico, sia da utenza fissa che da utenza mobile. Le Istituzioni italiane investono enormi risorse economiche per la prima fase delle intercettazioni ma sembrano assolutamente non curarsi delle fasi successive.

Il trattamento sembra identico a quello utilizzato in Italia per la formazione di nuovi ricercatori e studiosi. Si investono molte risorse economiche in percorsi universitari, borse di studio, assegni di ricerca e dottorati senza però in seguito creare un contesto o un *humus* tale da far fruttare il proprio investimento. Altri stati trovano sul mercato delle persone già pronte da inserire immediatamente nel percorso di ricerca o in quello produttivo, risparmiando su tutta la fase della formazione (fuga dei cervelli).

Nelle intercettazioni avviene più o meno la stessa cosa. A fronte di un'enorme spesa per l'intercettazione di segnale sonoro, a fronte di una grande domanda (si consideri che solo per le comparazioni foniche, considerate la punta di diamante delle perizie foniche, e per solo la metà del campione, le perizie effettuate in Italia sono almeno 600 l'anno) non si registra un'adeguata risposta da parte delle Istituzioni. Come si può notare sono stati pubblicati sette articoli in 5 anni, è stato riconosciuto un solo finanziamento dalla Comunità Europea (SMART), nessun corso istituzionale è stato attivato né come corso di laurea, né come master di primo o di secondo di livello. In questa situazione di assoluta vaghezza e indecisione, senza un Albo Professionale che in qualche modo funga da controllo o da Osservatorio Scientifico, era ovvio aspettarci i risultati ottenuti.

Il perito italiano proviene da un percorso di studi veramente variegato. Non esiste un limite rispetto ai titoli, quindi abbiamo periti con la licenza di scuola media e periti con Master e Ph.D in Italia e all'estero. I titoli variano anch'essi all'interno delle categorie segnalate. Tra la categoria Scuola Superiore abbiamo Maturità classica o Scientifica, Diploma di Ragioneria, di Perito Tecnico o Industriale. Anche per il perito laureato la situazione non cambia, abbiamo infatti periti con laurea in Economia, in Statistica, in Medicina, in Lettere, in Ingegneria, in Giurisprudenza o con il diploma di Conservatorio Musicale.

L'età varia da 29 a 72 anni. La diffusione geografica è anch'essa molto *variegata*, abbiamo periti locali che lavorano solo nella propria città e al massimo nella propria provincia ed esperti che invece accettano incarichi su tutto il territorio nazionale. In questo campione assolutamente non 'normale' dal punto di vista statistico, è ovvio aspettarci una estrema variabilità nei metodi utilizzati. Sottolineando qui quel 11% che utilizza metodi assolutamente senza alcun supporto scientifico e che hanno lo stesso valore di una moneta lanciata in aria, i metodi utilizzati variano un po' in tutto. Sono differenti i dati di partenza, viene considerato tutto il segnale, o tutte le vocali, o solo le vocali toniche, oppure vengono misurate tutte le formanti con l'intensità di ogni picco, la larghezza di banda di ogni formante e la durata di ogni singolo elemento oppure solo le prime due formanti per la costruzione di un quadrilatero vocalico dove porre le aree di dispersione vocalica in due dimensioni. I metodi oscillano dal metodo soggettivo di ascolto di due voci, al metodo completamente automatico. Nella comparazione si passa da una risposta frutto di complicate analisi statistiche e di calcolo dell'errore fino a risposte senza alcuna statistica. Eppure ben il 38% degli intervistati (ma ricordiamo i periti che si sono rifiutati e quelli non reperibili, i cui dati in casi come questi risultano essere veramente importanti) dichiara che l'attività di perito fonico è la loro attività unica e principale. Così come è generale tutto ciò che riguarda questo campo e questo settore, in alcuni casi è generale anche la figura del perito stesso. All'interno di questa percentuale (38%) esiste una figura di Perito Generale, cioè un perito che fornisce alla magistratura una relazione tecnica su molti e differenti campi come la grafia, la trascrizione, la comparazione di voci, di volti, l'infortunistica stradale, la comparazione di bossoli ecc., e una figura invece ancora più Generale e meno esperta (19% del nostro campione) che funge da tramite tra la magistratura e le strutture di ricerca e i laboratori di misura. Insomma: *la giustizia è uguale per tutti ma dipende dal metodo.*

Avremmo voluto presentare i dati riguardanti tutti i periti italiani, quindi anche quel 19% che ha rifiutato e quel 29% che si è reso in qualche modo irreperibile.

Questo lavoro non ha assolutamente lo scopo di giudicare o valutare le singole competenze o titoli di studio, ma mira solamente a presentare la grande variabilità di metodi e di conoscenze che oggi vengono utilizzate in Italia, a mettere in luce tutto il vasto mondo delle intercettazioni e in questo caso delle comparazioni foniche in ambito forense.

Riteniamo che a garanzia di tutte le parti coinvolte un processo, sia necessario addivenire ad un protocollo metodologico comune, qualunque esso sia. Che abbia un accordo comune ed una base di consenso larga e scientifica. Riteniamo che la risposta al quesito debba avere una formulazione per quanto più possibile precisa e simile in tutti i casi di comparazione fonica siano essi presentati a Milano o a Palermo e questo in modo da non doverci trovare di fronte a percentuali diverse, metodi diversi e parametri differenti nonostante si disponga dello stesso identico materiale. La fonetica forense deve diventare una scienza da inserire nella criminalistica e quindi come tale deve dotarsi di tutte quelle che sono le caratteristiche essenziali di una scienza: accordo sui dati da estrapolare, accordo sulle metodologie di estrapolazione delle misure, accordo sulla statistica da utilizzare ed infine accordo sulla formulazione da offrire o fornire al magistrato competente. Ovviamente in questo difficile processo la Magistratura e le Istituzioni giocano un ruolo essenziale e fondamentale.

9. APPENDICE

Riportiamo qui il *Code of Practice* dell'IAFPA approvato dall'AGM tenutosi ad Helsinki nel 2004 disponibile sul sito <http://www.iafpa.net/code.htm>.

1. Members should act in all circumstances with integrity, fairness and impartiality.
2. Recognising the varied array of casework subsumed under the interests of IAFPA (eg speaker identification/elimination, speaker profiling, voice line-ups, transcription, authentication, signal enhancement, sound propagation at crime scenes), Members should maintain awareness of the limits of their knowledge and competencies when agreeing to carry out work.
3. Members should not enter into any arrangements in which remuneration is dependent on the outcome of the case.
4. Members should make clear, both in their reports and in giving evidence in court, the limitations of forensic phonetic and acoustic analysis.
5. In reporting on cases where an opinion or conclusion is required, Members should make clear their level of certainty and give an indication of where their conclusion lies in relation to the range of judgements they are prepared to give.
6. (a) Members should exercise particular caution if carrying out forensic analysis of any kind on recordings containing speech in languages of which they are not native speakers.
(b) In carrying out forensic speaker identification/elimination work, Members should exercise particular caution if the samples for comparison are in different languages.
(c) Members should exercise particular caution if carrying out authenticity or integrity examinations of recordings that are not claimed to be original.
7. Members undertaking forensic phonetic and acoustic analyses or operations of all kinds should state in their reports the methods they have followed and provide details of the equipment and computer programs used.
8. Members, in making their analysis, should take due account of the methods available at the time and of their appropriateness to the samples under examination.
9. Members should not attempt to do psychological profiles or assessments of the sincerity of speakers.
10. Members' reports should not include or exclude any material which has been suggested by others (in particular by those instructing them) unless that Member has formed an independent view.

10. RINGRAZIAMENTI

Un vivissimo ringraziamento va naturalmente a tutti coloro che con l'intervista hanno reso questo lavoro possibile e senza i quali non saremmo mai riusciti a fare il punto sulla situazione italiana in merito al problema dello Speaker Recognition in ambito forense.

Ringraziamo anche l'Ing. Andrea Paoloni per i suggerimenti che ci ha fornito.

11. BIBLIOGRAFIA

Bolt R., Cooper F.S., David E.E. Jr., Denes P.B., Pickett J.M., Stevens K.N. (1970) 'Speaker identification by speech spectrograms: A scientists' view of its reliability for legal purposes.' *Journal of the Acoustical Society of America*, 4 (2): 597-612.

Campbell Jr. J., Reynolds D. (1999), Corpora for the evaluation of speaker recognition systems, *ICASSP'99*, Phoenix, USA, March 15-19, pp. 829-832.

Falcone M., Barone A. (2003), ViPS: Voci (italiane) Presumibilmente Simili, *XIV Giornate di Studio del GFS*, Viterbo.

- Falcone M., Barone A., (2000), FOCUS: un corpus vocale di voci simili per lo studio della identificazione del parlatore in ambito forense, 28° *Convegno AIA*, Trani.
- Falcone M., Gallo A., (1996), The SIVA speech database for speaker Verification: description and evaluation, *Proceedings of ICSLP'96*, Philadelphia, pp. 1902-1906.
- Falcone M., Ridolfi P. (1999), CALÌ: un corpus vocale per l'autenticazione del parlatore su rete telefonica fissa e mobile, *X Giornate di Studio del GFS*, Napoli.
- Fraser, H. (2003), Issues in Transcription: Factors affecting the reliability of transcripts as evidence in legal cases, *Forensic Linguistics*, 10(2), pp. 203-226.
- Furui S. (2005), 50 years of progress in speech and speaker recognition, *Proc. SPECOM 2005*, Patras, Greece, pp.1-9.
- Lionetti F., *La voce intercettata: 2006 anno record*, Tesi di Laurea, relatore prof. Romito L., a.a. 2006-2007.
- Melin H. (1999), Databases for Speaker Recognition: Activities in COST250 Working Group 2, *Proc. COST250 Workshop on Speaker Recognition in Telephony*, Rome.
- Nakamura M., Furui S. and Iwano K. (2006), Acoustic and linguistic characterization of spontaneous speech, *Proc. Speech Recognition and Intrinsic Variation Workshop (SRIV2006)*, Toulouse, France.
- Nakamura M., Furui S. and Iwano K. (2007), Acoustic and linguistic characterization of spontaneous speech, *Proc. Symposium on Large-Scale Knowledge Resources (LKR2007)*, Tokyo, Japan, pp.163-168.
- Nakamura M., Iwano K. and Furui S. (2007), The effect of spectral space reduction in spontaneous speech on recognition performances, *Proc. ICASSP2007*, Honolulu, Hawaii, USA, SPE-L7.5.
- Romito L. (2000), *Manuale di Fonetica Articolatoria, Acustica e Forense*, Centro Editoriale e Librario - Università degli Studi della Calabria, Cosenza.
- Rose P. (2002), *Forensic Speaker Identification*, Taylor & Francis, London.
- Trumper J.B. (1979), *Sociolinguistica giudiziaria. Preliminari di metodi e applicazioni*, CLESP Editrice, Padova.
- Zavattaro D., (2004-2005), *Articulation rate e l'identificazione del parlatore a scopo forense*, Tesi di Dottorato in Scienze Forensi, Università Tor Vergata, Roma.