

PASSATO PRESENTE E FUTURO NELLE ANALISI DI SPEAKER RECOGNITION

Luciano Romito
Università degli Studi della Calabria

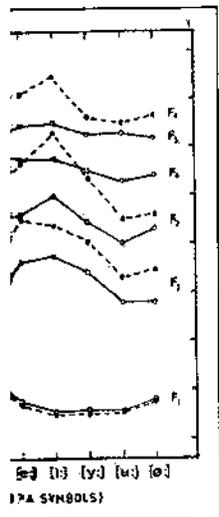
1. INTRODUZIONE

Nel voler tracciare una storia delle comparazioni foniche, delle indagini peritali sulla voce, del riconoscimento del parlatore o della ricerca della sua area linguistico-dialettale di provenienza in ambito forense, il nome di Franco Ferrero e la sua figura di fisico prestatore alla linguistica sono, e devono necessariamente essere, il nostro punto di partenza. Fu lui infatti il curatore degli Atti dei primi due convegni dedicati al problema delle indagini peritali e la voce. Il primo dal titolo "L'identificazione della persona per mezzo della voce" risale al 1979 e il secondo fu l'incontro di Venezia del 1981, con una sezione del quarto simposio della Federazione delle Società di Acustica Europee dal titolo Speech and Speaker recognition.

Osservando i lavori appena citati, si rileva, in maniera inequivocabile, come il fulcro della discussione in atto in quegli anni, fosse la ricerca di una metodologia, di un trait-d'union tra discipline molto differenti tra loro. Si nota come vi fosse il tentativo di far convergere metodiche acustiche, analisi linguistiche e dialettologiche a sfondo umanistico e teorie musicali. Ancora oggi, dopo venti anni, i Laboratori di Fonetica trovano spazio in alcune Università all'interno di Facoltà di Lettere e Filosofia, in altre all'interno di Dipartimenti di Informatica o di Fisica, in altre ancora si ritagliano uno spazio interdipartimentale o a volte anche interfacoltà. E ancora oggi la figura dell'esperto fonico in ambito forense è ricoperta a volte da un ingegnere, a volte da un fisico o da un medico, o ancora da un linguista, un fonetista o un matematico.

2. GLI ANNI '70-80

Negli anni '70 e '80 le perizie nei sequestri Lancia e Getty, nel caso Moro, o ancora nella strage di Peteano e gli esperti come Iba, Ferrero, Trumper ed in seguito Paoloni, Federico e Piazza hanno evidenziato come si fosse alla ricerca di convogliare le conoscenze dei diversi settori utilizzando la fisica acustica, la teoria musicale, l'indagine dialettologica, la fisiologia dell'apparato buccale, la sociolinguistica, la fonetica articolatoria, la fonetica acustica, la psicolinguistica ecc. Ogni singola perizia è una vera e propria ricerca scientifica che affronta il problema del riconoscimento da differenti, e a volte opposti, punti di vista; si cercano le caratteristiche acustiche generali del segnale riconducibili al parlatore, caratteristiche che esulano dal segnale linguistico, come le frequenze formantiche o il valore della frequenza fondamentale, lo spettro medio totale



fornire al cantante
cale, le sfumature
restazione possano
la specificità della
one sonora che da

ttavia l'ascolto per
to diversi aspetti:
ello dell'orchestra
na certa intensità

(in corrispondenza

gio con l'esigenza
e esigenze sono in
sce affrontare una

di 69, 1989
i, *Acustica*, vol.58,

ic Press, 1982

ecc, o al contrario si ricercano particolari produzioni di singoli segmenti strettamente legati alla produzione linguistica evitando ogni sorta di generalizzazione. Vengono studiati processi di sonorizzazione, di malocclusione, di affricativizzazione ecc; vengono considerati singole *particolarità* presenti sia nella voce anonima sia in quella nota da comparare.

Le ricerche linguistico-dialettologiche (chiamate più spesso glottologiche) sono basate essenzialmente sull'ascolto e sulla capacità ed esperienza nonché conoscenza dei dialetti italiani dell'esperto linguista o dialettologo. Il risultato di tali analisi è l'identificazione di un'un'area di indagine, di una zona di provenienza dell'anonimo. Le analisi acustiche, invece, si basano sulla misura di frequenze formantiche estrapolate da vocali presenti nel segnale anonimo comparate con le stesse presenti nel segnale noto. Tali analisi vengono effettuate utilizzando diversi metodi per la misura (FFT, LPC, Analisi Cepstrale, Estrazione dei Picchi, lettura dei sonogrammi, nonché estrazione di Formogrammi o altro), diversi strumenti (spettrografi, sonografi, software dedicati, ecc.), diverse strategie per l'estrapolazione (spettri medi, steady state, mid point, zone stazionarie, rapporto tra picchi e valli ecc.), e soprattutto utilizzando diverse procedure statistiche (Anova, χ^2 , T-student, Mahalanobis, Hotelling ecc.)¹ normalmente utilizzate per comparazione di dati numerici, dialettologici, di studio degli elementi rappresentativi di una comunità, di distribuzioni di variabili socio-linguistiche ecc. Sempre negli stessi anni vengono effettuate analisi basate sull'andamento melodico. Viene riportato su un pentagramma l'andamento della curva intonativa sia della voce anonima sia di quella nota. Ogni perizia, in questi anni, è una ricerca, un motivo per mettere in pratica metodologie e un tentativo di sperimentare nuovi scambi interdisciplinari.

3. DAGLI ANNI '80 AD OGGI

Dagli anni '80 ad oggi la situazione è molto cambiata. Il telefono negli anni '70-80 viene usato per rivendicazioni politiche di attentati, per contrattazione durante i sequestri; l'interlocutore anonimo di solito parla molto e le Perizie richieste dalla Magistratura sono poche. Oggi il telefono viene usato per gli stessi motivi anche se l'interlocutore anonimo parla molto meno e spesso legge dei comunicati; inoltre l'uso si è affargato alle estorsioni, alle ricattazioni, alla concussione e corruzione, alle ingiurie, alla vendita ed all'acquisto di sostanze stupefacenti; sono cambiati anche i canali, infatti si passa dalla telefonia fissa ai telefoni personali, ai cellulari TACS, GSM ecc. Conseguentemente la richiesta di Perizie, consulenze ed accertamenti tecnici è molto aumentata.

Per quanto riguarda i metodi, sulla scorta di altre categorie di periti (balistici, chimici, ecc.), è invalsa la voglia di racchiudere in un'unica figura tutte le competenze prima così

¹ Fu per questo motivo, vista le diverse competenze richieste che in quegli anni molti incarichi peritali furono collegiali.

ben differenziate; inoltre la grande richiesta di accertamenti da parte della Magistratura, porta ad un notevole accorciamento nei tempi di espletamento delle indagini peritali².

In questi anni l'uso di differenti metodi non riflette più il tentativo sicuramente ammirevole di una ricerca metodologica, di una applicazione e traduzione in numeri e percentuali di teorie fonetiche e fonologiche, ma spesso è solo il riflesso di una grande confusione, confusione che a volte, maliziosamente, viene usata per discreditarla o invalidare dei risultati già raggiunti. I nuovi metodi riflettono l'uso di una o dell'altra parte in causa in un processo, sono poco o molto garantisti e spesso si arriva in aula con tre o quattro perizie tutte con risultati differenti raggiunti con metodi altrettanto differenti.

4. IL PROSSIMO FUTURO

È per questo motivo che da più parti ormai si avverte la necessità di un protocollo scientifico nelle analisi di speaker recognition. Tale necessità non è avvertita solo in ambito scientifico, ma risponde ad una richiesta di tutela e di garanzia dalle parti in causa. Spesso, come appena detto, accade di arrivare ai più alti gradi di giudizio con consulenze che giungono a conclusioni diverse attraverso diverse metodologie, nonostante sia noto che un esperimento può essere ritenuto scientifico se, e solo se, risponde ad un protocollo metodologico conosciuto e se può essere ripetuto, ottenendo gli stessi risultati in altri laboratori, e da parte di altri operatori.

A tal pro, quindi, la discussione dovrebbe iniziare dallo stabilire se la perizia fonica è la ricerca, all'interno del segnale, di una particolarità che possa essere riconosciuta e comparata con una voce nota o anonima, o se invece non possa essere l'applicazione di una serie di procedure note e programmate.

Nel primo caso ogni perizia assumerebbe il ruolo di una ricerca più che di una consulenza tecnica, nel secondo caso invece, si potrebbe tentare la programmazione di un protocollo.

Personalmente ritengo che alcuni passi nell'espletamento di una perizia fonica possano essere schematizzati. I risultati sicuramente sarebbero più confrontabili se si accettasse un solo metodo. È ormai indubbio che le metodologie percettive, sia uditive sia visive (confronto dei sonogrammi) per quanto importanti e sicuramente utili sono troppo soggettive e legate quindi all'esperto che effettua la perizia. Il risultato può essere eccellente, o assolutamente infondato ma comunque quel che è peggio è che non è confrontabile, non è possibile ripetere l'esperimento e raggiungere gli stessi risultati. Accettando quindi l'ipotesi dell'indagine oggettiva, e dell'uso delle frequenze formantiche delle vocali e del valore della frequenza fondamentale nella comparazione, si potrebbe differenziare una perizia fonica in alcuni passi fondamentali. Un primo passo dovrebbe

² In questo periodo è anche cambiato molto il rapporto Magistrato-Esperto. Il primo ha richieste precise, tempi sempre più brevi e compensi sempre più risicati (si pensi che la parcella di un perito oggi ammonta a 3,46 Euro lorde l'ora, e spesso dopo la conclusione della perizia il perito viene liquidato anche con due anni di distanza), il secondo invece non è più uno studioso della disciplina, ma spesso è un perito di professione che vive di Consulenze e Perizie.

essere basato sul controllo dei reperti; affermando in maniera abbastanza decisa che non sempre è possibile effettuare delle perizie foniche, ma che un segnale per poter essere analizzato deve soddisfare alcune importanti caratteristiche acustiche. Un secondo passo dovrebbe riguardare l'estrapolazione dei dati, l'individuazione delle porzioni da analizzare; anche in questo caso bisognerebbe prevedere dei margini di *garanzia* al di sotto dei quali qualsiasi indagine fonica avrebbe risultati scientificamente inaccettabili. Un terzo passo, più tecnico, dovrebbe riguardare la misura; in questa fase si dovrebbe scegliere il metodo più testato che garantirebbe sicuramente l'uniformità delle indagini peritali. Il quarto ed ultimo passo potrebbe infine essere la comparazione statistica sul campione di dati estrapolato dai segnali oggetto di indagine; anche in questo caso dovrebbero essere definiti dei limiti e delle procedure standard; i limiti e le procedure dovrebbero essere dettate e stabilite da esperti nel settore come matematici e informatici in accordo con fonetisti e linguisti.

4.1 La Prima Fase

Questa è sicuramente la fase più importante di tutta l'operazione in quanto migliorare una cattiva registrazione non è sforzo da poco e spesso, dopo aver effettuato complesse analisi del segnale e aver utilizzato sofisticate apparecchiature per il filtraggio, i risultati sono pessimi.

Una voce o una discussione, può essere registrata dal vivo (intercettazione ambientale) o attraverso il telefono (intercettazione telefonica).

La registrazione tecnicamente può essere magnetica o digitale. Nel primo caso i registratori normalmente utilizzati dalle Procure, sono gli RT-2000 e gli RT-4000. Queste apparecchiature, dotate di un dispositivo per l'attivazione automatica della registrazione, presentano l'inconveniente di *scrivere* sul nastro, nella banda di frequenza di 2130 Hz, tutte le informazioni relative alla telefonata che si sta intercettando come l'ora, il giorno e l'anno, il numero di telefono selezionato, ecc. La registrazione digitale attraverso l'RT-6000 invece rimuove sia l'inconveniente del taglio di banda, sia tutti gli altri inconvenienti legati alla registrazione magnetica.

Nel caso della registrazione magnetica oltre al taglio nella banda di frequenza appena descritto, le caratteristiche della registrazione telefonica dipendono dal registratore stesso e dal canale utilizzato. Dal registratore dipende la velocità di registrazione utilizzata, che normalmente viene fissata a 2.38 cm/sec. e che notoriamente è riconosciuta come una caratteristica dei registratori di uso domestico³.

Rispetto al canale invece, questo introduce nel segnale sia distorsioni che alterazioni oltre ad un rumore costante di linea e ad una larghezza di banda che raggiunge nel migliore dei casi il limite massimo di 3 kHz. Le distorsioni e le alterazioni vengono introdotte da quelle variabili che potremmo in qualche modo definire *difficilmente*

³ Tale scelta, in Procura, viene effettuata per economizzare sul numero delle bobine, e per poter registrare più conversazioni possibili sullo stesso nastro. È per lo stesso motivo che spesso vengono utilizzate tutte e quattro le piste presenti e a volte anche il lato B della bobina. Tutto ciò ovviamente non fa che peggiorare di molto la qualità del segnale registrato.

controllabili e cioè quelle che intervengono lungo il canale creato tra il telefono ed il registratore. Nel dettaglio: lo stato del microfono del telefono o del supporto (o "cornetta"), il cavo telefonico, la centrale meccanica/numerica della Telecom collegata all'utenza di nostro interesse, il traffico viaggiante sullo stesso cavo, i ponti radio utilizzati ecc. Ciò fa sì che spesso il materiale sonoro consegnato nelle mani dei periti risulta essere una pessima registrazione. Va aggiunto inoltre, che spesso al perito viene consegnato un segnale registrato cinque, sei o anche dieci anni prima e che, com'è facile immaginare, tali registrazioni, non sono certo state conservate in luoghi asciutti e adeguatamente protetti da agenti atmosferici, da campi elettromagnetici e ciò peggiora ulteriormente le cose.

Nel caso delle intercettazioni ambientali, queste possono essere effettuate attivando il microfono del telefono, e quindi vale quanto detto precedentemente sulla qualità della registrazione con l'aggiunta che, mentre nell'intercettazione telefonica i microfoni di entrambi i parlatori sono vicini alla bocca, e quindi almeno sotto il profilo dell'intensità il segnale risulta essere più che soddisfacente, nel caso delle intercettazioni ambientali attraverso il microfono del telefono, i parlatori hanno posizioni differenti, possono muoversi in un ambiente e possono essere di numero superiore a due.

Negli altri casi, le intercettazioni ambientali si effettuano attraverso microspie o microfoni unidirezionali. Questi, possono essere allocati direttamente su una persona che funge da informatore o in un ambiente che si vuole controllare. Nel caso dell'ambiente, importantissima è la scelta del luogo dove posizionare la microspia. Questa se deve trasmettere segnali per un numero di ore superiore a dieci o quindici, deve essere alimentata direttamente con corrente (nessuna batteria, infatti, riesce ad associare alle dimensioni ridotte una carica ed una tensione costante per un numero maggiore di ore). A tal pro quindi il luogo deve trovarsi vicino ad una presa di corrente o ad un apparecchio alimentato a corrente, e ciò aggiunge rumore al segnale registrato.

Quanto detto avvalorata la necessità, prima ancora di iniziare qualsiasi operazione di trascrizione o di comparazione che abbia un fondamento scientifico, di controllare che la registrazione soddisfi alcune caratteristiche tecniche essenziali.

Tali caratteristiche sono in primo luogo il rapporto fra l'energia del segnale (S) e quella del rumore di fondo (R) ciò sta ad indicare che il rumore di fondo non deve essere così *forte* da impedire l'analisi o la trascrizione di una registrazione (tale valore dovrebbe essere fissato ed utilizzato come limite). Durante una telefonata, normalmente, il parlante mantiene sempre la stessa distanza dal microfono che è allocato nella cornetta e questo fa sì, che un'intercettazione telefonica produca di norma registrazioni caratterizzate da un buon livello d'intensità e con un rapporto segnale rumore soddisfacente. Purtroppo la stessa cosa non può essere detta per le intercettazioni ambientali, dove il parlatore non mantenendo sempre la stessa distanza dal microfono causa delle registrazioni con livelli d'intensità estremamente variabili che oscillano dalla distorsione allo zero.

Un protocollo ed un equippe di esperti, oltre al limite da stabilire, dovrebbe prevedere delle indicazioni precise riguardo le situazioni, per esempio può capitare di avere delle registrazioni effettuate in presenza di una forte musica in sottofondo con un valore di S/R

(rapporto tra il segnale ed il rumore) molto elevato nonostante il materiale sia assolutamente inutilizzabile.

La seconda importante caratteristica è la larghezza di banda, cioè quell'intervallo di frequenza che viene utilizzato per la registrazione. Anche in questo caso un protocollo dovrebbe fissare dei limiti. L'ultima caratteristica, per alcuni versi la più importante ma anche la più difficile da trattare, è la durata. Tale parametro ovviamente non deve essere considerato nelle operazioni di trascrizione ma solo nelle analisi linguistico-dialettologiche⁴ e nelle comparazioni a fini di identificazione del parlatore. È ovvio che per poter avere un numero sufficiente di dati in modo che, sia la comparazione che l'analisi linguistica abbia un senso, il materiale registrato deve avere una certa durata, d'altronde porre un limite così come per il rapporto S/N o per la larghezza di banda non è facile visto che tale parametro è strettamente legato alla velocità d'eloquio del parlatore. In 5 secondi una persona con una consistente velocità di eloquio può produrre materiale sufficiente per una qualsiasi analisi, al contrario un'altra, nello stesso intervallo può produrne talmente poco da inficiare qualsiasi indagine.

4.2 La Seconda Fase

In questa seconda fase, avendo valutato il segnale 'analizzabile', si passa alla scelta delle vocali da misurare, infatti parametri utilizzati a fini identificativi sono il valore della frequenza fondamentale⁵ e i valori delle prime tre frequenze formantiche⁶ delle vocali. La scelta delle misure formantiche, per quanto possa apparire semplice, pone invece una serie di problemi. Prendiamo avvio affrontando la questione fondamentale, ossia quella dell'individuazione delle vocali.

Individuare una vocale implica:

- 1) determinare l'esatto inventario fonologico del sistema linguistico da analizzare;
- 2) definire lo status di ogni vocale all'interno del sistema: quali tra tutte le vocali presenti possono e devono essere misurate; stabilire una dipendenza tra il sistema di vocali toniche ed il sotto sistema di vocali atone; considerare teoricamente prima, e verificare praticamente attraverso le misurazioni, i processi quali la riduzione, la centralizzazione e in generale gli effetti di coarticolazione;

La mancanza di una chiara visione circa lo status fonetico-fonologico di una vocale, all'interno del proprio sistema di riferimento, può portare a fraintendimenti estremamente

⁴ Già dalla nascita delle perizie glottologiche, quando il segnale a disposizione era considerevolmente lungo, le analisi dialettologiche miravano solo ad individuare la zona di provenienze dell'anonimo e non il riconoscimento dello stesso, oggi, invece, spesso nelle consulenze si legge che il riconoscimento è avvenuto anche grazie ad una indagine dialettologica effettuata su una telefonata della durata spesso non superiore a 20 o 30 secondi.

⁵ La frequenza fondamentale (F_0) rappresenta la frequenza più bassa presente nello spettro vocale (cfr § 2.1). È compresa tra 80 e 200 Hz per le voci maschili e tra 100 e 300 Hz per quelle femminili ed è legata alle caratteristiche anatomico-fisiologiche del parlatore ed al suo atteggiamento fonatorio, dipende quindi dal peso, dall'altezza nonché dalla lunghezza, elasticità e spessore delle pliche vocaliche. È per tale motivo che viene ritenuta molto importante nell'identificazione del parlatore.

⁶ Le frequenze formantiche sono delle concentrazioni di energia caratteristiche dei suoni soprattutto vocalici, e definiscono la qualità di ogni singola vocale.

il materiale sia quell'intervallo di caso un protocollo più importante ma le non deve essere analisi linguistico-tore. È ovvio che comparazione che una certa durata, rza di banda non è io del parlatore. In ridurre materiale so intervallo può

si passa alla scelta ono il valore della e⁶ delle vocali. La e invece una serie ale, ossia quella

analizzare;

tra tutte le vocali a tra il sistema di icamente prima, e i la riduzione, la

ico di una vocale, enti estremamente

mente lungo, le analisi oscimento dello stesso, prazie ad una indagine

o vocale (cfr § 2.1). È gata alle caratteristiche so, dall'altezza nonché muta molto importante

soprattutto vocalici, e

gravi su una linea che va dal considerare uguali unità che non lo sono, fino al considerare come indipendenti ricorrenze della stessa unità strutturale. Perfino nei casi di più apparentemente facile interpretazione, dal punto di vista dell'inventario del sistema linguistico, come nel caso dell'italiano parlato, la soluzione non è sempre univoca. Ad esempio la prassi in campo peritale consiste nell'analisi dettagliata di un iposistema vocalico considerato aprioristicamente pentavocalico del tipo /i, e, a, o, u/ considerando uguali unità quali /e, ε/ ed /o, o/. Come già altre volte osservato esistono, molti tipi di italiano parlato, caratterizzati da un sistema fonologico con sette vocali /i, e, ε, a, o, o, u/. Ipotizzare un sistema pentavocalico in alcune zone dell'Italia sia settentrionale che meridionale, porta a considerare simili realizzazioni quali ['pesca] *frutto* e ['pesca] voce del verbo *pescare*, ignorando che la vocale [e] e la vocale [ε] sono diverse e quindi, caratterizzate da valori formantici differenti (/e/ F1= 360, F2=2040; /ε/. F1= 560, F2=1840⁷). Ovviamente quanto detto vale anche per la coppia di vocali /o/ ed /o/ per esempio di ['botte] *recipiente per il vino* e ['botte] *percasse*. Se ciò, pone dei seri problemi nell'inventario del sistema linguistico dell'italiano, ancora più attenzione deve essere rivolta a quei sistemi, come per esempio i dialetti, dove l'inventario fonologico può anche essere molto diverso da quello dell'italiano. Alcuni esempi potrebbero essere i dialetti meridionali dove non solo esistono opposizioni del tipo /e ε/ ma anche /i i⁸/.

Non considerando queste importanti distinzioni, le misurazioni relative a due unità differenti vengono considerate come appartenenti alla stessa unità, aumentando inequivocabilmente l'errore e spesso inficiando l'analisi stessa.

Un problema aggiuntivo è l'esecuzione di atti linguistici mistilingui, cioè la produzione di frasi miste in italiano ed in dialetto⁹, da parte dello stesso parlante.

Quando ci si imbatte in produzioni linguistiche mistilingui, il problema è sia teorico, perché nasce dalla legittimità di considerare sullo stesso piano sistemi appartenenti a due codici diversi, che pratico, perché deve essere valutato il rapporto genetico e/o strutturale tra i vari codici in gioco e i rapporti di affinità tipologiche. In breve, solo dopo un'attenta analisi dell'inventario fonologico dei due codici ma soprattutto della distribuzione fonetica delle aree di esistenza dei segmenti vocalici¹⁰, il linguista può decidere di utilizzare le misurazioni di vocali estrapolate da brani dialettali e quelle estrapolate da brani italiani come vocali di un unico sistema linguistico.

⁷ Questi valori sono stati estrapolati da Giannini e Pettorino 1992 pag.161 tav III.

⁸ cfr. Trumper, Romito 1991, Romito et al.1995 e §1.1.5.3.

⁹ Oggi tra l'altro gli atti linguistici mistilingui possono anche essere albanese-italiano, slavo-italiano, Wolof-Serer o altro e italiano per la nota forte immigrazione.

¹⁰ Cfr. Trumper, Romito, Maddalon 1991, dove si affronta, sia fonologicamente che foneticamente, il problema della comparazione tra sistemi vocalici appartenenti a lingue diverse o a dialetti differenti che presentano però lo stesso numero di vocali. Gli esperimenti vengono effettuati su sistemi eptavocalici come il Toscano, Napoletano, Veneto e Calabrese settentrionale (area di Lausberg) e pentavocalici come italiano parlato a Catanzaro, a Cosenza ecc. le tabelle ed i grafici presenti a pp.61-68 sono molto esplicativi.

Una volta individuato l'iposistema vocalico della comunità cui appartiene il parlante, il passo successivo consisterà nella segmentazione delle unità di tale iposistema. Le unità che normalmente vengono considerate, sono le vocali nel continuum fonico. A questo punto nasce il problema sulla scelta delle vocali da considerare. Scegliere di analizzare le sole vocali toniche o quelle atone, o considerarle come facenti parte di un unico sistema è un problema che necessita approfondimenti fonetici per essere risolto.

La letteratura in merito differenzia le vocali toniche da quelle atone e soprattutto differenzia lo spazio occupato all'interno del quadrilatero vocalico dal sistema delle vocali toniche rispetto a quello occupato dal *sottoinsieme* delle vocali atone. Queste ultime, infatti, subiscono maggiormente tutti i processi coarticolatori come la riduzione, la centralizzazione, la neutralizzazione fino ad arrivare alla completa cancellazione. Il target di una vocale viene raggiunto solo durante la produzione di un segmento che abbia uno status ben preciso nella catena fonica cioè una vocale tonica. Solo tali segmenti caratterizzati da un accento dinamico raggiungeranno un target articolatorio. Le vocali atone al contrario, spesso vengono considerate degli elementi di passaggio tra consonante e consonante. Esse quindi subiscono limitazioni sia dal punto di vista articolatorio, che energetico e di durata¹¹. Tali influenze ovviamente si riflettono sulla produzione acustica facendo rilevare valori formantici diversi (a volte anche molto diversi), da quelli rilevati nella misurazione di vocali toniche (cfr. Romito 1997).

Per quanto riguarda la produzione di vocali atone in posizione finale, il problema è diverso, infatti, non solo si registra una riduzione con neutralizzazione dell'opposizione /e/ ed /o/ ma il sistema in alcuni casi può ridurre il numero degli elementi come avviene per esempio in alcuni dialetti meridionali

Concludiamo questa fase molto importante, dell'identificazione dei segmenti da analizzare, affermando che *prima* di iniziare qualsiasi analisi sul materiale sonoro *bisogna* analizzare la varietà linguistica in questione per individuare l'inventario fonologico (non possiamo in alcun modo presupporre aprioristicamente l'esistenza di un unico inventario fonologico caratterizzato da 5 o 7 vocali) e le regole prosodiche che sottostanno a tale codice e, solo in seguito, proseguire con le analisi e le misurazioni delle vocali identificate.

4.3 La Terza Fase

La terza fase riguarda l'analisi in senso esterno, cioè l'identificazione della porzione steady-state¹² all'interno della vocale segmentata e la seguente estrapolazione dei dati.

Il processo di misurazione dei valori formantici delle vocali segmentate, interessa almeno due aspetti fondamentali: l'operatore e la strumentazione utilizzata. Nel caso della strumentazione non di alta qualità, una volta specificate le caratteristiche degli strumenti

¹¹ Dai dati presenti in Romito 1997 la percentuale di allungamento di una vocale tonica rispetto ad una vocale atona nel dialetto cosentino oscilla dal 57.4 al 126.2% in una lista di parole. Sicuramente tale dato aumenta se si considera il parlato spontaneo.

¹² Si presuppone che durante un parlato spontaneo l'apparato vocale resti congelato in una precisa impostazione, per almeno 20-25 ms, durante la produzione di una vocale tonica. Tale porzione, caratterizzata da una stabilità di valori formantici, viene definita *steady state*.

utilizzati, può essere calcolato il margine di errore che si riflette nel risultato, tale errore sarà comunque costante in tutte le misure effettuate. Infatti una approssimazione nelle misure formantiche di un certo peso sarà presente su tutte le misure in ugual modo. Nel caso in cui l'operatore, invece, fosse inesperto o con nessuna competenza linguistica, la validità del risultato verrebbe sicuramente messa in crisi. Non vi è, infatti, nessuna possibilità di valutare, a posteriori, l'errore presente nelle misure. Quindi nonostante il metodo venga ritenuto oggettivo parametrico, la fase dell'identificazione della vocale e della sua porzione steady-state da misurare e la fase dell'estrapolazione dei dati sono tuttora legate esclusivamente all'esperienza, alla capacità ed alla competenza dell'operatore.

A questo proposito quindi, è necessario che il perito o l'esperto si avvalga di collaboratori o operatori con competenze linguistiche, che espliciti chiaramente quale analisi sia stata adoperata nella sua indagine e le caratteristiche tecniche della strumentazione utilizzata in modo da permettere a chiunque di poter riprodurre l'esperimento *giungendo agli stessi risultati*.

4.4 La Quarta Fase

Questa fase comprende l'Analisi Statistica e la comparazione. Tale analisi dovrebbe dare come risultato non tanto delle percentuali di similitudine, ma delle risposte complesse che in qualche modo includano la probabilità di errore del sistema e del protocollo, legato alla specificità della voce analizzata, alla inadeguatezza dei parametri misurati ecc. In conclusione dovrebbe dare tutte le informazioni necessarie al Magistrato (incluso la probabilità di falso riconoscimento e di falsa reiezione) consentendogli quindi di giungere ad una conclusione sull'appartenenza o meno di due voci allo stesso parlatore.

5. CONCLUSIONI

Concludendo, crediamo che si debba procedere al più presto alla definizione di un protocollo metodologico per l'espletamento delle indagini foniche sulla voce. Tale protocollo dovrebbe essere scientificamente rigoroso e redatto da persone esperte e competenti nelle diverse discipline operanti nelle indagini peritali. Nel dettaglio il protocollo dovrebbe:

1. prevedere l'accertamento della qualità del materiale sonoro. Fissare dei limiti che stabiliscano quando tale materiale è accettabile sotto il profilo scientifico e quando invece, considerata la possibilità di elevato errore che ne conseguirebbe, non è utilizzabile per effettuare qualunque analisi. Alcune delle caratteristiche che un segnale dovrebbe soddisfare potrebbero essere il rapporto segnale/rumore, la larghezza di banda, la durata del segnale da analizzare e il numero dei parametri da estrapolare. Tale punto del protocollo potrebbe agevolmente essere affrontato e stilato da un fisico acustico o ad un ingegnere esperto in trattamento del segnale;

2. stabilire esattamente come priorità l'individuazione dell'inventario fonologico della produzione linguistica da esaminare (sistema penta, eptavocalico ecc.) prima di effettuare qualunque altra operazione peritale;
3. stabilire esattamente quali segmenti analizzare: vocali toniche o vocali atone;
4. stabilire esattamente come trattare gli enunciati mistilingui: effettuando comparazioni differenziate dialetto-dialetto ed italiano-italiano ecc.
5. stabilire esattamente l'analisi statistica da effettuare.

Una volta fissato il protocollo metodologico per il riconoscimento del parlatore, il passo successivo potrebbe essere quello di istituire un Albo dei periti fonici.

Riguardo il metodo, il riconoscimento del parlatore riveste un ruolo importante nelle ricerche fonetiche. Gli esperimenti spaziano dall'individuazione di nuovi parametri al miglioramento delle tecniche di estrapolazione di quelli già esistenti. I nuovi parametri identificati sono: la misura e l'identificazione del VOT (Voice onset Time) sia positivo che negativo; la durata della fase di tenuta delle occlusive in un parlato spontaneo; l'analisi della gestione temporale di singole unità, il rapporto tra pause e articolazione: speech rate, o la velocità di produzione di singole sillabe: articulation rate ecc; il numero e la dimensione spaziale delle pause (sia piene che vuote) durante una conversazione spontanea; e infine l'analisi della struttura formantica dei segmenti che utilizzano risonatori fissi come le fosse nasali (segmenti nasali in contesti tonici ecc.).

È indiscutibile che la necessità di un incremento della ricerca per rendere sempre più attendibile il risultato, per abbassare la soglia dell'errore o, ancora, per ricercare nuovi parametri che diano risposte più certe ai quesiti posti dalla Magistratura. È però parimenti indiscutibile il fatto che ogni ricerca ha bisogno di un periodo di sperimentazione in laboratorio e certo non nelle aule dei tribunali. Solo dopo questo periodo di sperimentazione effettuata da diversi ricercatori in diversi laboratori, una pubblicazione scientifica può dare valore e rendere 'utilizzabile' una nuova procedura. Una garanzia a tutto può essere solo la individuazione di un protocollo metodologico e la preparazione di figure professionali scientificamente preparate, esaminate ed inserite in un Albo professionale dei periti fonici alle quali la Magistratura può fare riferimento ed affidamento per il proprio lavoro.