

# La competenza linguistica nelle trascrizioni forensi: l'intelligibilità, l'oggettività e il rapporto segnale/rumore

Luciano Romito  
Professore Associato di Fonetica e Fonologia  
(Università della Calabria – Laboratorio di Fonetica)

<<quando parlo dovete ascoltare soprattutto le mie pause. Dicono più delle mie parole>><sup>1</sup>

Un atto comunicativo è un passaggio di informazione da un emittente ad uno o più riceventi. Una tale definizione non rende giustizia al ruolo fortemente attivo che viene svolto dal ricevente. La stessa, potrebbe, quindi, essere riproposta come: la percezione è l'interpretazione da parte di un ricevente di una informazione prodotta da un emittente.

È necessario dividere il processo percettivo in due fasi ben distinte: la prima riguarda l'estrapolazione dell'informazione acustica del segnale percepito, ciò porta all'identificazione dei segmenti e al raggruppamento di questi in unità astratte superiori come sillabe, parole e frasi; nella seconda fase, forse la più complessa, detta di accesso lessicale, le parole vengono riconosciute e quindi avviene il processo di comprensione/interpretazione del (intero) messaggio ricevuto<sup>2</sup>.

In questa sede ci concentreremo su quest'ultima fase e tenteremo di dimostrare il ruolo fortemente attivo dell'ascoltatore nel processo di percezione<sup>3</sup>.

Tenteremo inoltre di dimostrare come sia ancora più attivo (nel senso interpretativo/immaginario) il ruolo di una persona esterna all'atto comunicativo, che redige in tempi a volte diversi e a volte molto lontani dal momento in cui l'atto è avvenuto, una trascrizione letterale su carta di un atto registrato (o intercettato) nato per essere orale.

Gli aspetti segmentali della comunicazione vengono studiati da molti anni ormai; scienze quali la linguistica, la sociolinguistica, la fonetica, la fonologia ma anche la semantica, la pragmatica<sup>4</sup> e tante altre, non fanno altro che basare le proprie indagini e le proprie ipotesi sull'atto linguistico. Al contrario tutto ciò che fa parte della percezione del linguaggio non dal punto di vista meccanico o funzionale è oggetto di studio solo da poco tempo da discipline come la psicolinguistica, la psicoacustica<sup>5</sup>, o acustica psicologica e la fonetica percettiva.

## L'atto Comunicativo

La psicolinguistica in particolare ha approfondito tre aree differenti: il riconoscimento delle parole, la comprensione delle frasi e la comprensione del testo o del discorso. La prova che esiste una comprensione a livelli separati e soprattutto una comprensione a livello superiore è data dal fatto che alla fine di un discorso, l'ascoltatore non ricorda l'esatta sequenza delle parole, ma un *succo*, un concentrato, una comprensione di fatto superiore alla comprensione delle singole parole.

---

<sup>1</sup> Peter Høeg, *I quasi adatti*, Mondadori 2003:9.

<sup>2</sup> Cfr. Wundt 1989 oltre un secolo fa distingueva nella percezione “i fatti esterni” legati alla produzione ed alla percezione dei singoli suoni linguistici dai “fatti interni” che riguardano i fenomeni del pensiero.

<sup>3</sup> In accordo con tutti i sostenitori delle teorie attive come Liberman, Lindbloom ecc, riteniamo che <<percevoir le discours, n'est pas chose passive et automatique. Celui qui aperçoit assume une fonction sélective en répondant à certains aspects de la situation globale et non à d'autres il répond aux stimuli selon une organisation qu'il leur impose. Et il remplace la stimulation absente au contradictoire d'une manière compatible avec ses besoins et son expérience passée>> cfr. Miller G.A in *Langage et communication*, Parigi 1956:III.

<sup>4</sup> La scienza che studia il contesto si chiama pragmatica e uno degli autori che ha contribuito al progredire di detta scienza è sicuramente Hymes.

<sup>5</sup> L'acustica psicologica o psicoacustica è un settore della psicologia che studia il suono inteso come rappresentazione psichica di una realtà esterna, esistente quindi solo nella nostra mente che lo attribuisce a oggetti del mondo esterno, localizzandolo in determinate posizioni dello spazio. Mentre l'acustica fisica studia il modo in cui gli oggetti vibrano trasmettendo impulsi meccanici all'orecchio e l'acustica fisiologica si interessa a come l'orecchio trasforma questi impulsi meccanici in impulsi nervosi, l'acustica psicologica studia come l'effetto di onde elastiche sull'organo del senso dell'udito viene rappresentato mediante il suono, inteso come sensazione psichica.

Una comunicazione orale avviene sempre in un preciso *contesto*, in una determinata *situazione*, tra specifiche *persone*, utilizzando un preciso *codice* e, oltre ad una cooperazione che si instaura tra emittente e ricevente<sup>6</sup>, avvalendosi sempre dell'apporto di tutta una serie di conoscenze condivise.

Le variabili che condizionano un atto comunicativo sono (cfr. Hymes 1967): il numero dei partecipanti ad una discussione<sup>7</sup>, la forma, il contenuto e i risultati che si intendono raggiungere. La situazione è intesa come luogo, come scena e stato psicologico sia del parlante che del ricevente; il canale è la scelta da parte dell'emittente (una lettera scritta, una comunicazione orale in presenza o una telefonata). Il codice scelto tra tutti quelli conosciuti dall'emittente e dal ricevente come l'italiano o l'inglese o diafasicamente la lingua standard, il dialetto regionale, il gergo, un codice nascosto o cifrato.

Una discussione orale coinvolge molto di più di quanto possa coinvolgere una comunicazione scritta. L'emittente è portato sempre a verificare l'andamento della discussione e il destinatario può favorirne il successo utilizzando segnali etichettati come back-channels (ad esempio 'mmh' o 'si ti seguo', 'si, si', o ancora 'esatto'<sup>8</sup>). Tali segnali di conferma stimolano l'emittente a proseguire la conversazione, in mancanza di questi l'emittente può tornare indietro e riformulare il proprio pensiero in modo da renderlo più comprensibile (usando per esempio parafrasi), può rendere più accettabile il proprio discorso, può eliminare parti ritenute troppo ostiche o a volte per raggiungere lo scopo che va al di là del significato dell'atto linguistico, può addirittura cambiare completamente versione dicendo esattamente il contrario di quanto aveva detto precedentemente (si pensi per esempio alla conversazione tra due amanti, o tra un impiegato ed il proprio datore di lavoro insomma in tutte quelle situazioni di subalternità). È da tenere a mente per il prosieguo di quanto affermeremo in seguito, che il contesto al fine di una buona interpretazione di un atto linguistico è una delle variabili più importanti.

La situazione riguarda sia il parlante che il ricevente. In una comunicazione il parlante è il centro, tutto ruota intorno a lui, lo spazio è organizzato in funzione della sua posizione, il tempo in base al *tempo* del parlante (cfr. deissi spaziali e temporali). La lingua deve essere intesa come un *processo*, un mezzo di interazione all'interno di una *situazione* e di un *contesto*. Il mezzo attraverso il quale il parlante definisce il suo stato sociale, le sue sensazioni, i suoi sentimenti, tutte cose che si situano al di sopra della lingua<sup>9</sup> e che l'ascoltatore sa ben decifrare, il linguaggio infatti ha una duplice funzione comunicativa, quella interna o intrinseca alle parole stesse e quella esterna o estrinseca alle parole, quella che si riferisce ai sentimenti, alle impressioni e alle aspettative del parlante.

Una ulteriore variabile è costituita dalle conoscenze condivise. Ogni atto comunicativo, ogni interazione orale fa riferimento a delle conoscenze condivise. Queste conoscenze sono condivise (o ritenute tali dal parlante) da tutti i partecipanti alla conversazione<sup>10</sup>. Al bar è facile ascoltare frasi del tipo: "Hai visto che alla fine ha deciso Moratti su Adriano?". Di una frase così prodotta può essere percepito il segnale acustico (e quindi questo può anche essere trascritto), ma nessuna interpretazione può essere possibile se non si hanno conoscenze del mondo calcistico e degli attuali avvenimenti sportivi, cioè se tra parlante ed ascoltatore non si condividono le stesse conoscenze. È inutile sottolineare che tali conoscenze non vengono veicolate sul piano segmentale della comunicazione linguistica.

Quanto appena riportato evidenzia, se ce ne fosse bisogno, come una conversazione sia il frutto di una collaborazione voluta, di una ricerca e di un lavoro che coinvolge sia l'emittente sia il destinatario, il loro intervento tra cooperazione, condivisione e avvicinarsi di turni si fondono senza più differenziarsi tra loro e

---

<sup>6</sup> cfr. Firth 1957 "la conversazione è un rituale, molto di più di quanto si pensi. [...] non siete liberi di dire quello che volete (il contesto condiziona). Siamo nati individui, ma per soddisfare le nostre esigenze dobbiamo diventare persone sociali".

<sup>7</sup> Un emittente ed un ricevente come un dialogo o una conversazione telefonica; un emittente e un pubblico come una conversazione, una lezione o una conferenza.

<sup>8</sup> Oltre che con atti linguistici tali segnali possono anche essere inviati con il movimento del capo, o con la chiusura degli occhi o in tantissimi altri modi che dipendono dalla cultura di appartenenza dell'emittente e del destinatario e quindi dal contesto in cui si sta svolgendo l'interazione linguistica.

<sup>9</sup> Secondo quanto scrive Lyons 1977, addirittura le informazioni gestuali possono trasmettere informazioni che vanno esattamente nella direzione opposta a quelle verbali. Un esempio potrebbe essere produrre un atto linguistico tipo 'è buona! Vero?' ed intanto con il gesto della testa o della mano, lasciar intendere esattamente il contrario.

<sup>10</sup> Ciò ovviamente non vale per l'operatore esterno alla conversazione, che non partecipa attivamente, e complica il suo compito di trascrittore.

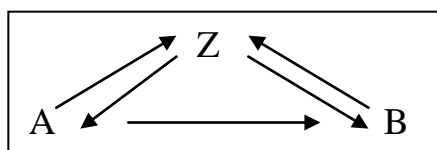
fornendo un unico prodotto che è la conversazione. Tutto questo non si evince da una analisi del *solo* piano segmentale.

## La percezione

Un atto comunicativo avviene secondo uno schema di tipo:

$$A \rightarrow B$$

Dove la freccia indica il passaggio di informazione dall'emittente A al ricevente B. In linea con quanto già riportato e per evidenziare il ruolo attivo del ricevente, lo schema potrebbe essere integrato dall'accesso ad una conoscenza comune Z:



Così l'emittente A, facendo riferimento ad una conoscenza Z comune, progetta ed in seguito produce un atto linguistico. Il ricevente B recepisce il segnale acustico prodotto da A e facendo riferimento alla stessa conoscenza Z interpreta e comprende il messaggio. L'ascoltatore B, nel promuovere questa operazione, può contare su due importanti fattori: 1) le caratteristiche acustiche del messaggio, l'identificazione dei singoli segmenti sonori<sup>11</sup>, la propria conoscenza del codice, il raggruppare i singoli suoni in unità astratte maggiori come le sillabe, le parole e le frasi; ma, nel caso in cui le informazioni acustiche risultassero insufficienti, allora può fare riferimento 2) alla conoscenza Z, alla lingua, al sistema<sup>12</sup>.

Un segnale prodotto come nell'esempio del Bar e cioè: "Hai visto che alla fine ha deciso Moratti su Adriano?" è un segnale system-oriented, cioè fortemente basato sia nella sua produzione che nella sua percezione e interpretazione, sulla conoscenza del sistema, della lingua e della situazione calcistica. Un segnale così prodotto lascia molto spazio alla libera ricostruzione e interpretazione del segnale vocale. Un segnale invece orientato e progettato per essere solo ed esclusivamente acustico viene definito output-oriented ed è l'unico segnale privo di interpretazione. La produzione linguistica, così come la percezione, oscilla sempre tra un livello di massima ricostruzione e quindi di massima interpretazione, ed un livello senza ricostruzione. La normale produzione linguistica è frutto di un compromesso tra il segnale acustico, le conoscenze pregresse e le aspettative.

L'ascoltatore B accede al sistema e alla conoscenza Z per correggere distorsioni acustiche nel segnale prodotto da A, per sopperire a parziali mancanze dovute alla cattiva percezione, al rumore o a qualcosa che si sovrappone. L'ascoltatore interpreta e *indovina* statisticamente la porzione di segnale mancante o mal percepita e mal identificata. È l'intero processo che porta all'interpretazione e alla comprensione dell'atto acustico.

L'ascoltatore attua il processo appena descritto spesso anche in condizioni favorevoli solo per anticipare A, per rendere più fluida e veloce la percezione e quindi la conversazione. Spesso la percezione diventa un susseguirsi di previsioni e di conferme. Non viene esaminata tutta la sequenza acustica del segnale di A, ma solo la porzione ritenuta più importante. Non appena B ritiene di avere sufficienti motivi (prove) per ritenere di aver compreso, si innesca il processo di previsione che verrà in seguito confermato dall'intero significato dell'atto linguistico. L'ascolto è lineare ma l'interpretazione è globale. L'ascoltatore B *registra* grandi parti di segnale acustico lo *riversa* su una memoria di lavoro e solo in seguito decide da dove far iniziare l'analisi acustico-interpretativa.

La probabilità di occorrenza è co-determinata nell'atto di percezione dell'ascoltatore.

Un esempio anche se banale potrebbe essere quello di leggere velocemente il brano riportato di seguito, ci accorgeremo di quanta vasta sia la ricostruzione del segnale e la correzione anche involontaria:

<sup>11</sup> L'isolamento e il riconoscimento delle parole è il primo passo della comprensione, di seguito si passa all'attribuzione di una struttura sintattica agli elementi identificati.

<sup>12</sup> In questo caso si fa riferimento alla teoria H-H di Lindbloom, lungi da noi l'idea che quanto affermato finora non sia stato già abbondantemente scritto e trattato.

Da uno studio della Uvinrestia di Cmabirgde, non ha ipomrtanza in che ordine le lettere siano in una parola, l'unica cosa importante è che la prima e l'ultima lettera siano al posto giusto. Il resto può essere tolto o aggiunto e si può leggere senza problemi. Questo perché la mente umana non legge ogni lettera ma la parola come un'insieme. Diveramente, no?

In un esperimento, condotto presso il Laboratorio di Fonetica dell'Università della Calabria, è stato fatto ascoltare in camera silente, ad un campione di 60 ascoltatori, una serie di frasi 'sporcate con del rumore'. Le ipotesi di partenza dell'esperimento sono: 1) in mancanza di informazione acustica necessaria all'identificazione di una porzione del segnale, il cervello nella sua fase di percezione/comprendimento ricostruisce il segnale mancante sulla base delle proprie conoscenze sintattiche, morfologiche, semantiche e ovviamente conoscendo un contesto adeguato; 2) la percezione non è un atto lineare ma globale, il cervello può ricostruire una porzione di segnale mancante riconoscendo una porzione di segnale temporalmente successiva<sup>13</sup>.

Per quanto riguarda la prima ipotesi, una frase del tipo "a colazione di solito mangio un XXXXXX" prevede nella fase della ricostruzione, la sostituzione della parola coperta da rumore e acusticamente non riconoscibile, non con tutto il vocabolario, ma con una sua precisa parte, cioè tutto ciò che è commestibile; quindi 'tavolo' non potrà mai essere una opzione valutata in quanto sarebbe scorretta semanticamente una frase del tipo "a colazione di solito mangio un tavolo". Inoltre tra tutto ciò che è commestibile, il cervello riconosciuto e interpretato l'articolo "un", prenderà in considerazione solo ed esclusivamente ciò che è commestibile e che è di genere maschile. Quindi 'pera' così come 'tavolo' non rientrerà tra le opzioni della possibile ricostruzione perché è sintatticamente scorretta una frase del tipo "a colazione di solito mangio un pera". La ricerca delle opzioni possibili prosegue restringendo sempre di più il campo, basando la ricerca sulla lunghezza del segnale non percepito, sulla cadenza dell'accento, su qualche indice fonetico ascoltato come una /t/ o altro per giungere alla fine con una buona approssimazione che la parola da inserire deve essere una parola scelta tra tutto ciò che è commestibile, che è di genere maschile, che è all'incirca trisillabico, che è una parola piana e che contiene una o più /t/.

In questo caso ovviamente aiuta molto la percezione e l'interpretazione anche il contesto. Diverso è il caso invece in cui l'informazione è veicolata solo ed esclusivamente dal canale acustico come l'esempio: "ieri ho incontrato XXXX che passeggiava in via XXXXXXXX". In questo caso non ho alcuna informazione rilevabile né dal contesto né dalla struttura sintattica o semantica.

Riguardo la seconda ipotesi, l'esperimento è basato su una frase altamente prevedibile, come: "il pomeriggio d'estate mi piace schiacciare un pisolino" dove la parola *schiacciare* è stata inizialmente solo coperta da rumore, in seguito sostituita da una parola assonante, di pari durata e con struttura accentuale simile come *snocciolare*. Ben il 98,2% del campione ha trascritto "il pomeriggio d'estate mi piace schiacciare un pisolino"<sup>14</sup> sia nel primo che nel secondo caso e quindi non ciò che ha sentito, ma ciò che in seguito ad una previsione ad un rimando alla conoscenza della lingua, ha ritenuto la frase più probabile. Un *sonnellino* può solo essere *schiacciato*. Qualunque parola temporalmente simile e con uno schema accentuale simile venga prodotta prima di *pisolino* verrà interpretata come *schiacciare*. Inoltre la parola 'schiacciare' è stata interpretata riconoscendo una parola temporalmente successiva. Questo vuol dire che la percezione è stata inizialmente ad un primo livello "il pomeriggio d'estate mi piace XXXXXXXXX un pisolino" in seguito dopo avere effettuato una analisi di tutto il segnale, grazie alla parola 'pisolino', è stato possibile ricostruire una porzione precedente del segnale.

### **Cosa disturba la percezione?**

Il segnale vocale in fase percettiva è caratterizzato da fattori come l'intelligibilità<sup>15</sup>, la naturalezza, la gradevolezza e l'accettabilità.

La percezione può essere disturbata in ognuna di queste caratteristiche ovviamente con differenti conseguenze. I problemi fondamentali riguardo la percezione vengono da fattori interni come la non corretta

<sup>13</sup> Per una analisi dettagliata e per una lettura approfondita dei dati si veda Romito L. 2004 I Convegno AISV.

<sup>14</sup> Sono condizioni artificiali quelle in cui l'ascoltatore non possiede altra informazione se non quella contenuta nella forma acustica dei suoni.

<sup>15</sup> L'intelligibilità del segnale è caratteristica percettiva che definisce la più o meno corretta comprensibilità del segnale trasmesso>>

stima da parte dell'emittente in riferimento alla conoscenza Z condivisa, o da fattori esterni come l'intensità del segnale o il rumore che si sovrappone al segnale stesso.

Volendo quindi schematizzare quanto fin qui riportato, l'ascoltatore durante un processo di percezione muove la sua attenzione come l'ago di una bilancia tra il livello acustico delle informazioni segmentali e il livello del sistema (cioè della ricostruzione e dell'esperienza). Il segnale mancante viene ricostruito in base ad alcuni indici che potremmo definire di *predicibilità, acustico e della conoscenza*.

Ma a cosa è legata la percezione della singola parola? Sicuramente alla frequenza d'uso, una parola più è frequente più viene riconosciuta anche in situazioni di rumore, e in base al contesto, una parola inserita in un contesto appropriato è più facilmente riconoscibile<sup>16</sup>. L'intelligibilità è strettamente legata alla sua ridondanza.

In base alle conoscenze acquisite, la percezione può avvenire con maggiore o minore successo<sup>17</sup>. Già Bruce nel 1956:176 scriveva <<sembra che la frequenza d'occorrenza di una parola influenzi il tempo necessario al suo accesso nel lessico. In generale le parole più frequenti vengono trovate più rapidamente e vengono identificate più facilmente in condizioni precarie d'ascolto [...]>><sup>18</sup>. Alcuni esperimenti hanno provato che oltre alla frequenza delle parole riveste una notevole importanza la grandezza del vocabolario. L'intelligibilità del discorso in una situazione di ascolto rumorosa aumenta moltissimo se il discorso è costruito con un vocabolario molto ristretto, a patto che l'ascoltatore sia a conoscenza della estrema riduzione del vocabolario. È come se la scelta tra le diverse opportunità di interpretare una parola sia molto ridotta e quindi sia più facile interpretare e trascrivere correttamente un intero discorso. In alcuni esperimenti condotti su liste di parole *sporcate* da rumore è stato evidenziato che appena l'ascoltatore si accorge che il campo semantico della lista è ristretto e se il campo appena identificato è alla sua portata, la comprensione e la capacità di *indovinare* tutte le parole ascoltate aumenta in maniera esponenziale<sup>19</sup>. Ciò non avviene ovviamente in tutti gli ascoltatori nello stesso istante, ma appena nell'ascoltatore-trascrittore nasce o si forma la consapevolezza della restrizione contestuale e se la restrizione è alla sua portata, il numero delle parole riconosciute nonostante il rumore aumenta immediatamente.

Quando il contesto è difficile e vario invece, per raggiungere lo stesso risultato bisogna abbassare moltissimo il livello del rumore, cioè bisogna far veicolare la maggior parte dell'informazione sul piano segmentale ed acustico<sup>20</sup>.

Più aumenta il rumore e più aumenta la necessità di interpretazione e quindi l'accesso al sistema, e l'allontanamento dall'identificazione di indici acustici. La stessa cosa accade se (come in ogni normale conversazione) non vi sia nessuna restrizione contestuale. La scelta diventa vastissima e l'unico appiglio possibile è, oltre al lato acustico, il contesto.

### ***Gli indici acustici e le illusioni percettive***

Da alcuni esperimenti effettuati (cfr science 1981:212) risulta che le consonanti vengono percepite meglio rispetto alle vocali. Le consonanti occlusive, ad esempio, vengono percepite con una bassa percentuale di errore. Anche quando la produzione acustica si allontana dal target (dalla norma) un ascoltatore avrà pochi problemi ad identificare il singolo fono, anzi spesso non percepisce la devianza. Con le vocali la situazione è completamente diversa: l'ascoltatore rileva senza difficoltà piccole diversità nel segmento acustico, ma le sue identificazioni degli stimoli devianti sono meno sicure e fortemente influenzate dal contesto<sup>21</sup>.

---

<sup>16</sup> Sottostà a quanto detto la superiorità della parola. Secondo gli studi di Wheler 1970 e Reicher 1969 alcune parole come alcuni suoni hanno una preferenza ad essere identificate per prima e in maniera più facile. Esistono due modelli quello ad attivazione (cfr. Morton 1969-1970) e quello a ricerca (cfr. Forster 1976).

<sup>17</sup> È ovvio che il trascrittore dovrebbe fare poco uso di questo parametro.

<sup>18</sup> Cfr Savin 1963. È ovvio, comunque, che le parole più frequenti sono quelle che con maggiore probabilità si trovano nella conoscenza di tutti i parlanti del codice in questione.

<sup>19</sup> Per esempio avendo capito il campo semantico della lista di parole, *roba da mangiare* mi appiglio ad un indizio fonetico, come la vocale iniziale o finale, una consonante uno schema accentuale, la durata della parola ed alla fine in base alla mia conoscenza diretta delle cose da mangiare *indovino* o interpreto la parola ascoltata.

<sup>20</sup> Si pensi ad esempio a trascrivere o comprendere una lista di nomi e di cognomi, nessun contesto o conoscenza giova alla interpretazione. Solo l'informazione veicolata sul piano acustico.

<sup>21</sup> Quanto è stato riportato è quindi di fondamentale importanza, soprattutto in una lingua come l'italiano dove le informazioni sintattiche e morfologiche relative al genere, al numero e al caso (l'italiano è una lingua di tipo flessivo)

Per Ladefoged e Broadbent “l’informazione linguistica trasmessa da una vocale non dipende dai valori assoluti delle sue frequenze formantiche, ma dal rapporto fra le frequenze formantiche di quella vocale e le frequenze formantiche delle altre vocali pronunciate dallo stesso parlante”<sup>22</sup>.

Con questo ovviamente, non vogliamo affermare che l’ascoltatore percepisce ogni singolo segmento acustico, al contrario crediamo che la percezione dei singoli suoni sia influenzata da decisioni che vengono prese a livelli ‘superiori’ d’analisi del segnale. Ma la percezione di singoli suoni influenza o può influenzare la percezione-interpretazioni di singole parole (se so di avere un appuntamento al *parcheggio*, mi basterà ascoltare “ci vediamo al XXXXeggio” e questo non creerà nessun fraintendimento con *Reggio*, o *sorteggio* ecc.)<sup>23</sup>.

Nella percezione e nella identificazione di un atto acustico non può essere selezionata una singola parte. Così come la visione è macroscopica, la percezione deve essere globale.

Questo vuol dire che il segnale sonoro registrato non può essere ascoltato per piccoli pezzi ma globalmente. È assolutamente fuorviante andare a cercare singole vocali in un segnale rumoroso e che non abbia caratteristiche acustiche ben definite. Le vocali sono realizzate all’interno di un contesto, sono coarticolate e molta informazione si trova al di fuori di loro stesse, proprio nelle fasi di transizione, nelle metastasi o nelle catastasi consonantiche.

## **Il rumore**

In quanto finora scritto, riveste grande importanza il rumore<sup>24</sup>. Il rumore può essere considerato come qualsiasi fenomeno acustico, generalmente irregolare, casuale e non musicale, costituito da sovrapposizione disordinata di un numero indefinito di onde elementari. Esso è caratterizzato dalla sua intensità (espressa in decibel) e dallo spettro delle frequenze degli elementi che lo compongono.

Quanto detto finora va benissimo per una definizione in termini acustici o fisici ma in linguistica il rumore deve essere inteso come qualsiasi disturbo che comprometta l’intelligibilità di un segnale acustico (o di una registrazione di un segnale acustico).

## **IL RAPPORTO SEGNALE E RUMORE**

Il rapporto segnale-rumore è il rapporto tra la massima ampiezza utile di un segnale e l’ampiezza del rumore presente. Esso è di difficile misura, in quanto fluttua nel tempo in funzione del segnale sonoro. Può essere definito Signal/Noise Ratio (S/N o SNR) o Rapporto Segnale/Rumore (S/R). L’ammontare del disturbo<sup>25</sup> sulla parola può essere valutato in diversi modi: a) *livello di interferenza di disturbo* (espresso in dB); b) *indice di comprensibilità*<sup>26</sup>. L’intelligibilità della parola anche secondo quanto scrive Lehmann<sup>27</sup> con

---

vengono demandate proprio alle vocali. È differente affermare di aver ascoltato ‘l’ho ammazzato’ da l’ha ammazzato’. Liberman et al 1957, Fry et al. 1962, Studdert-kennedy et al. 1963; Stevens et al. 1963 da Robins

<sup>22</sup> Cfr Ladefoged e Broadbent 1980 pag 98

<sup>23</sup> Quello che può accadere è che in base a quanto riportato la percezione di un nome o di un toponimo o comunque di una porzione di segnale svincolata dal contesto, possa essere percepita e quindi ricostruita in base a delle illusioni percettive, all’ascolto di una sequenza *acustica* di alcune consonanti.

<sup>24</sup> Il suono, al contrario, può essere definito come un segnale sonoro composto da un preciso numero di frequenze fisse e ben definite, una somma di onde semplici sinusoidali aventi caratteristiche di periodicità

In fonetica per ogni frequenza acustica, esiste un’intensità sotto la quale il suono non è percepibile (soglia di udibilità) e una oltre la quale provoca dolore (soglia del dolore).

<sup>25</sup> Concetto di disturbo acustico: il disturbo acustico si identifica con l’immissione sonora, cioè nell’aggiunta di vibrazioni sonore a quelle preesistenti nell’ambiente o come si suol dire di fondo. Perché si possa parlare di disturbo acustico non è necessario che il livello sia elevato da essere patologico, ma basta che sia uguale o superiore al livello del rumore di fondo.

<sup>26</sup> questo indice è dato dal contesto e dalla capacità di ricostruire

<sup>27</sup> in Elements de phusio- et de psychoacoustique Parigi 1969. Littler The physics of the ear Londra 1965.

presenza di rumore bianco o comunque di rumore a spettro continuo, è massima con un rapporto S/R > 10 dB e invece è nulla per un rapporto < - 20 dB<sup>28</sup>.

L'intelligibilità dipende direttamente dall'intensità del rumore di fondo, dalla riverberazione e dalla modalità di riflessione (cfr. Cocchi Farina Vezzali 1980) Le controversie riguardano il contributo, il peso di ogni singolo parametro.

Molti sono gli esperimenti, che provano la dipendenza dell'intelligibilità dal tempo di riverberazione<sup>29</sup> o la relazione tra lo spettro del rumore e l'intelligibilità<sup>30</sup>. Viene anche calcolato un indice di Articolazione in base al quale viene definita l'intelligibilità di un segnale<sup>31</sup>. Questi studi sono stati sviluppati anche sull'italiano. Con il calcolo grafico dell'I.A. di Lehmann si arriva, secondo questi studi, ad una ottima approssimazione del livello di intelligibilità, per una conversazione che si svolge all'interno di un ambiente insonorizzato secondo le norme ISO e con la produzione all'esterno di un rumore.

Noi riteniamo che l'intelligibilità così misurata non sia altro che una misura quantitativa della qualità del sistema di trasmissione impiegato, e che l'intelligibilità del parlato debba invece essere la somma dell'informazione fonetica e dell'informazione recata dal contesto linguistico ed extralinguistico.

Inoltre finora abbiamo riportato una definizione *acustica* del rapporto Segnale/Rumore. Riteniamo che *linguisticamente* il rumore debba essere inteso come qualsiasi disturbo regolare o meno, sovrapposto al segnale ritenuto portatore di informazione e che ne comprometta l'intelligibilità. Abbiamo volutamente parlato di segnale e non di atto linguistico in quanto riteniamo che sia il soggetto ascoltatore a decidere di volta in volta quale sia il segnale-informazione e quale il segnale-disturbo. Quindi un suono è tale se ha un carattere informativo, di contro un rumore sarà qualunque segnale indesiderato che disturba, compromette, e maschera la percezione del suono-informazione. Il rumore è quel segnale indesiderato privo di qualsiasi contenuto informativo utile per l'ascoltatore. Il problema sta nel definire quale suono sia <<informazione>> e quale invece sia <<rumore>>. Un brano di musica classica (che sicuramente non è un rumore dal punto di vista fisico-acustico) sovrapposto ad una voce può essere considerato rumore se l'informazione è veicolata dalla voce. Al contrario a teatro, durante una rappresentazione lirica, una voce può essere considerata rumore se la percezione è concentrata sulla musica<sup>32</sup>.

Anche due voci sovrapposte sono l'esempio di un segnale e di un rumore. Dipende quale delle due voci si ritiene carica di contenuto informativo.

Alla luce di quanto detto il rapporto Segnale/Rumore potrebbe essere riproposto come rapporto di intelligibilità, cioè un rapporto tra l'intensità del suono-informazione (qualunque esso sia) e l'intensità del suono-disturbo (qualunque esso sia) ottenendo quindi SI/SD (segnale informazione/segnale disturbo) o più semplicemente rapporto I/D. Questo rapporto così formulato risolverebbe problemi annosi come le intercettazioni in automobile con la radio a tutto volume. Utilizzando il rapporto di Informazione/disturbo (I/D) questo sarebbe sicuramente al di sotto di 0 dB. Un esempio potrebbe essere l'analisi del segnale intercettato in una automobile all'interno del p.p. 1097/03 RG Corte di Appello di Palermo nei confronti di Bruno Contrada. Detto segnale, ad una analisi di S/N tradizionale risulta avere i seguenti rapporti:

Tempo in sec.	S/N in dB
37.34	13
38.18	-6.91

Quindi secondo la scala di Lehmann il segnale è caratterizzato da una intelligibilità superiore al 95%.

Ad una analisi approfondita, si rileva che il segnale è sovrapposto ad una canzone napoletana prodotta dalla radio all'interno dell'automobile (cfr. i grafici in appendice). L'intelligibilità è molto bassa (lo si evince

<sup>28</sup> Cfr. Dubno Dirks and Margon 1984, Gordon-Salant and Fitzgibbons 1995, Sueli Aparecida Caporali, Cox and Mc Daniel 1984 e Nakatani and Dukers 1973.

<sup>29</sup> Knudsen e Harris 1950 *Acoustical designing in architecture*- Wiley N.Y.

<sup>30</sup> Kryter K.D. *Methods for the calculation and use of the articulation index* JASA 1989, 34-1964

<sup>31</sup> Cfr Cocchi Farina Vezzali 1980:1 <<alle diverse bande di frequenza in ottava o terzi di ottava, i vari livelli di pressione di picco, opportunamente definiti vengono depurati dai corrispondenti livelli di rumore di fondo. I contributi in tal modo determinati vengono moltiplicati per opportuni fattori di peso, tabulati in funzione della frequenza e infine sommati per determinare il valore dell'indice di Articolazione Globale>>.

<sup>32</sup> IBBA \*\*\*:103 “ il suono emesso dal motore di un veicolo possiede un contenuto di informazione utile per il conducente che, dal suo ascolto può giudicare del più o meno perfetto funzionamento del motore stesso; pertanto egli non ne sarà disturbato. Al contrario un osservatore esterno [...]“.

anche dalla enorme presenza di <<parole incomprensibili>> presenti all'interno della trascrizione). Effettuando il calcolo I/D i valori risultano essere i seguenti:

Tempo in sec.	S/N in dB	I/D in dB
37.34	13	<b>5.54</b>
38.18	-6.91	<b>-6.91</b>

Confrontando ora i nuovi risultati ottenuti e la scala di Lehmann si evince che l'intelligibilità ha un valore molto basso (pari allo 0% nel caso del valore negativo) e il grande numero di incomprensibili presenti nella trascrizione ne è la prova tangibile.

Quando si ascoltano contemporaneamente due segnali, una televisione ed un parlatore, la voce di un amico in una sala piena di gente e di rumori, l'ascoltatore considera una fonte rumore e l'altra informazione, quindi attiva un processo di selezione attiva, cioè quella capacità di selezionare un input tra tutti quelli che pervengono, ignorando gli altri.

### ***La competenza linguistica, la formazione di nuove professionalità in aiuto alle forze dell'ordine e alla giustizia***

È invalsa l'idea che chiunque possa dare giudizi o opinioni sulla lingua che parliamo.

La trascrizione, per esempio, da sempre è stata considerata una attività talmente semplice da non richiedere studi approfonditi o specializzazioni proprie. Molta gente infatti ne fa il proprio principale lavoro e mezzo di sostentamento<sup>33</sup> nonostante non abbia seguito alcuno studio o percorso di studi sul linguaggio o sulla lingua. Insomma la trascrizione sia di una riunione condominiale sia di un sequestro di persona, in Italia può essere effettuata da chiunque e senza la richiesta di alcun titolo di Studi o alcuna competenza precisa. Lo stesso accade in Germania ed in altri stati europei: <<in Germany only a fraction are professional linguists, let alone academically trained linguists or actually working as linguists in universities. A lot of so-called "experts" are people from other sciences, from law and forensic scientist hand writing experts philologists of one denomination of another, and, of course, people without any academic training whatsoever, for example police officers, lay assessors etc. they have to deal with language data in forensic context as part of their job.>>

In realtà trascrivere una comunicazione orale comporta una enormità di problemi che solo un esperto riesce ad intuire ed a controllare.

Il parlato è caratterizzato da canali paralleli di comunicazione come quello gestuale e intonativo, è fortemente condizionato dal rumore di canale, di contesto. Il trascrittore, dal canto suo, si trova in una posizione di estraneità rispetto la situazione comunicativa, non riceve completamente l'atto comunicativo nella sua interezza ma solo il canale sonoro ed inoltre deve tradurre in caratteri (grafemi) e trasporre su carta una comunicazione nata per essere orale composta di pause innalzamenti di voce, cambi di intonazione, cambi di accento, cambio di stile (ironico, scherzoso, canzonatorio, serio) e sentimenti come ira, felicità, ilarità sdegno, avendo a disposizione per tutti questi fatti sovrasegmentali solo qualche segno di punteggiatura (interpunzione).

La trascrizione deve essere pertanto affrontata come un serio problema prima teorico ed in seguito pratico. Tale problema infatti non è facilmente risolvibile con l'aggiunta di un commento o di una interpretazione. Deve essere la migliore approssimazione della realtà. Una approssimazione scientificamente accettata, esplicitamente e rigorosamente standardizzata senza però essere talmente tecnica da poter essere usata da pochi esperti come per esempio la trascrizione fonetica.<sup>34</sup>

---

<sup>33</sup> Le trascrizioni oggi, vengono richieste per riunioni di: consigli di amministrazione; consigli comunali; ed in ogni riunione dove sia necessario un verbale scritto, vengono richieste per le udienze nelle aule di Tribunale e nei casi di intercettazione e di captazione non dichiarata.

<sup>34</sup> La trascrizione fonetica è usata solo da specialisti della materia come fonetisti linguisti, o dialettologi. Essa serve per rendere scritto un qualsiasi enunciato orale e per far sì che lo stesso enunciato possa essere letto da diversi esperti esattamente come è stato prodotto.



La Bellucci 1986 scrive << nell'intercettazione ambientale, il perito si trova nell'insolita esperienza – che non esito a definire “eccitante” per un sociolinguista- di partecipare “da fantasma” alla situazione >><sup>35</sup> ma è un fantasma cieco che perde tutti gli elementi visivi. Il trascrittore si trova all'esterno del circuito comunicativo, del contesto e del cotesto della conversazione, non percepisce la comunicazione fatta di gesti, di occhiate, di posture del corpo, di movimenti della testa, delle mani e non condivide le conoscenze pregresse degli attori della conversazione.

Il trascrittore percepisce solo una parte della complessa rete della comunicazione: la parte acustica. La comunicazione in presenza però è una rete complessa di canali paralleli, è un evento che abbraccia tutti coloro che vi partecipano e che devono condividere alcune conoscenze sull'ambiente, sulla loro posizione spazio-temporale, sulla loro posizione sociale e questo prima che l'evento avvenga.<sup>36</sup>

Nella trascrizione, qualsiasi parola si trasciva, questa deve conformarsi non solo al modello del segnale percepito ma anche, e forse soprattutto, ad un modello di alternative (legate al contesto) presenti nel vocabolario e nel lessico del soggetto ascoltatore- trascrittore. Le parole possono quindi essere adattate soprattutto se chi trascrive possiede un setaccio fonologico differente (per esempio parla un altro dialetto) e il problema non è solo lessicale ma di frammenti fonetici. È il modello che svolge il ruolo decisivo nelle trascrizioni e nella risposta che deve essere la più probabile possibile. Così il trascrittore oscilla tra un segnale acustico e un modello che ha nella propria testa, e cerca (coscientemente o inconscientemente) di accoppiare frammenti fonetici a probabili risposte secondo il **PROPRIO** modello. La risposta è determinata principalmente dalla scelta personale tra una gamma di alternative considerate possibili<sup>37</sup>.

È quindi misurabile la ricostruzione linguistica di un trascrittore inesperto? È possibile una trascrizione neutra e senza interpretazione?

Un segnale system-oriented influenza inconscientemente la ricostruzione, l'interpretazione e quindi la comprensione del segnale vocale. Solo un segnale output-oriented cioè programmato per essere solo ed esclusivamente verbale, veicola tutta l'informazione nel canale acustico e quindi è privo di interpretazione. Tutto il resto è frutto di un compromesso tra il segnale acustico, le conoscenze pregresse, le aspettative e le inferenze. E quindi possibile una trascrizione neutra? Senza interpretazione?

Come si pone quindi un soggetto ascoltatore-trascrittore quando si trova di fronte ad un segnale da trascrivere e su un argomento a lui noto con molte conoscenze condivise e quando invece si trova di fronte ad un segnale su un argomento a lui sconosciuto? Ha lo stesso atteggiamento di fronte ad un segnale acusticamente pulito e quindi totalmente intelligibile (il canale acustico lo aiuta molto nella percezione) o di fronte ad un segnale con del rumore sovrapposto? Nel secondo caso il trascrittore si basa interamente sulla propria interpretazione dei frammenti fonetici del segnale e adatta su questa base non troppo attendibile il contesto? La scelta sarà anche vincolata dalla ricchezza del suo vocabolario rispetto a quel preciso campo semantico? Cosa succede con i toponimi, con i nomi, con i cognomi o con i numeri? Le domande devono trovare risposta nella competenza del Perito Trascrittore. La garanzia per tutte le parti in causa è legata alla identificazione di una figura professionale preparata, competente e scientificamente riconosciuta.

---

<sup>35</sup> Bellucci 1986 Nota 11 pag 7

<sup>36</sup> Bellucci 1986:8 <<un non-linguista [...] è nei fatti, ancora abituato a credere nell'oggettività del ‘significato di frase’ mentre i vertiginosi sviluppi della linguistica e delle sue subaree disciplinari hanno ormai dimostrato in modo inoppugnabile quanto sia pericoloso separare un segmento verbale dal cotesto e contesto in cui è parte [ ...]>>. Inoltre si vedano a questo proposito una infinità di intercettazioni dove l'informazione ed il significato frase è tutt'altro rispetto alla informazione veicolata nella conversazione: “mi porti le scarpe nuove”, “vieni ad aggiustare il cancello che non funziona?”

<sup>37</sup> Un esempio banale riguardo il setaccio, potrebbe essere quello che relativo al gusto. Quando anni fa comparve sulle nostre tavole il frutto Kiwi, dopo il primo assaggio ognuno si è chiesto che gusto avesse. La risposta più normale dovrebbe essere: di kiwi, il kiwi ha il gusto di Kiwi, come la mela ha il gusto di mela ecc. Eppure il gusto Kiwi all'epoca non faceva parte del nostro modello, non rientrava nel nostro inventario dei gusti possibili e così bisognava catalogare ed etichettare il nuovo gusto con le informazioni in nostro possesso e in base al nostro setaccio, il kiwi assumeva un gusto incrociato tra la banana e il pomodoro (gusti a noi già noti).

## La situazione in Italia

In Italia quando il Magistrato chiede una perizia linguistica o di trascrizione, la fa sempre, o nella maggior parte dei casi, non sulla base dell'ascolto del flusso sonoro delle registrazioni ma piuttosto sul significato frasale che si deduce dai verbali redatti dagli operatori di PG (Polizia Giudiziaria), o su indicazione degli stessi operatori. Nella maggior parte dei casi le informative, le trascrizioni in fase investigativa sono parziali e riportano solo alcuni significati a-contestuali incriminanti come per esempio "si sente –passami la pistola"<sup>38c</sup>. Tale frase senza una precisa ed esauriente definizione del contesto situazionale, del contesto verbale e dei rumori di sottofondo risulta essere sicuramente incriminante, ma con una valenza giuridica decisamente ambigua. Basta infatti che il Consulente di parte della Difesa, dimostri che la frase prodotta sia solo il commento di una trasmissione televisiva per far cadere il peso della frase incriminante ma e soprattutto tutte le altre occorrenze presenti all'interno della stessa trascrizione<sup>39</sup>. Una trascrizione parziale mina alle fondamenta tutto il concetto di comunicazione verbale multidimensionale.

Con questo non si vuole affermare che l'operatore di PG non possiede una competenza acquisita, ma spesso l'intuizione dell'ascoltatore scaturita dall'ascolto complessivo del flusso sonoro, porta a mettere in rilievo una frase o una porzione di frase, l'intuizione può essere scaturita da un rumore, da un'intonazione, da una inflessione particolare nel parlare, o nel rispondere, ma l'operatore non ha gli strumenti per mettere su carta questa intuizione, non ha la competenza per tradurre in *scritto* una interazione orale, quindi sarà costretto a limitarsi al significato molto superficiale delle parole o della frase.

Nella filiera della trascrizione per consulenza, il primo problema che si incontra è quello relativo alla prassi penale che prevede un filtro tra le registrazioni di una scena reale e i magistrati. Detto filtro identificato negli operatori di PG, e forze dell'ordine, dovrebbe avere competenze linguistiche, pragmatiche, dialettologiche e fonetiche<sup>40</sup>.

Le intercettazioni telefoniche sono delle particolari interazioni orali. Hanno un inizio ben preciso, segnalato da uno squillo, e una fine spesso segnalata da un saluto e da una chiusura dell'apparecchio. La comunicazione si svolge tra due persone (o con scambi di interlocutori segnalati da frasi ben precise come "te lo passo"), nella maggior parte dei casi si rispetta il turno (cioè vi è poca sovrapposizione di voci), la distanza dalla fonte sonora alla cornetta (cioè al microfono) è costante, i rumori sono situati alle spalle dell'interlocutore e lontani dalla conversazione, le informazioni sono tutte affidate al canale e soprattutto alla nostra fiducia nel canale (per esempio il nostro modo di interagire è differente al telefono fisso rispetto al telefono cellulare)<sup>41</sup>.

Durante una comunicazione in presenza, o frontale (intercettazione ambientale), i canali si moltiplicano, le interazioni sono più complesse, i rumori assumono nuovi significati, completano spesso lo scambio verbale, si trovano **tra** la fonte sonora e il microfono, i turni non vengono rispettati gli interlocutori sono due o anche di più, la distanza tra la fonte sonora ed il microfono non è mai la stessa, inoltre una conversazione in

---

<sup>38</sup> Altri esempi banali ma riscontrati in alcuni verbali sono "o chimma te viju mortu ammazzatu" it. *potessi vederti morto ammazzato*, che in alcuni dialetti è solo una interazione, o "t'ammazzera" it. *ti ammazzerei*, come preoccupazione per una cosa fatta o "fatti sta sparata dai" it *di ciò che devi dire ma molto velocemente*.

<sup>39</sup> Anche la non uniformità crea delle inesattezze e debolezze interne alla trascrizione. Per esempio i puntini che vengono utilizzati, hanno il valore di una parola non compresa, di una bestemmia non trascritta, di una pausa più o meno lunga, di una inspirazione, di una parolaccia, di un innalzamento di intensità quindi prima di una parola importante di una esitazione di una censura di una sovrapposizione di voci che rende incomprensibile il segnale, o di un rumore che si sovrappone, di una parte non interessante o ritenuta tale ecc. Altro fattore molto importante è la traduzione letterale dal dialetto all'italiano. Spesso le trascrizioni vengono riportate proprio in italiano perdendo così di quell'originalità. In questo caso l'operatore è un doppio filtro, decide cosa è importante e cosa non lo è ed inoltre interpreta un dialetto che a volte è il proprio, a volte no. È una prassi che un MERIDIONALE trascriva tutto il parlato del sud anche se i dialetti meridionali tra di loro sono molto differenti.

<sup>40</sup> Non si considera qui il filtro tra il magistrato e il delegato alla intercettazione. Spesso le microspie sono insufficienti, il luogo scelto per il posizionamento inadatto ecc. Tutto ciò porta già in origine a trattare con del materiale registrato di non ottima qualità.

<sup>41</sup> Nel caso della telefonia fissa, il fatto che ci sia un filo a collegare gli apparecchi, o l'esperienza nelle innumerevoli conversazioni già avute rende più tranquilli gli interlocutori. Nel caso della telefonia cellulare, gli abbassamenti di intensità, i rumori improvvisi, l'esperienza di molte conversazioni interrotte da cadute di linea porta gli interlocutori a fare abbondante uso della funzione fatica della lingua con interazioni tipo "ci sei?, mi ascolti?, sei li?" oppure alzando notevolmente il volume della voce.

presenza non ha mai un inizio ed una fine ben definite.

Un Magistrato attento e scrupoloso potrà chiedere la trascrizione ad un esperto linguista di tutto il materiale intercettato. Così facendo, il trascrittore si troverà di fronte al contesto globale ma anche e spesso ad una sequela di ore e ore di inutili conversazioni parallele. Ciò farà calare l'attenzione su tutta la trascrizione, comporterà un inutile e costoso impiego di tempo da parte di un esperto.

Ma se una telefonata ha una durata sicuramente inferiore, ha un preciso inizio ed una precisa fine, una intercettazione ambientale può essere segnalata da un numero progressivo? Il brogliaccio o il verbale che accompagna il materiale sonoro registrato, spesso riporta solo l'oggetto principale di ogni progressivo, ma ancora oggi, pur con l'ausilio di tecnologie molto avanzate, in una intercettazione ambientale il progressivo viene inserito ogni 59 minuti di registrazione. Cioché un progressivo, non identifica una precisa conversazione, ma un tempo preciso registrato, all'interno del quale può essere contenuto parte di una più ampia conversazione, o al contrario molteplici conversazioni.

In qualche modo chi ha verbalizzato, il filtro tra la registrazione ed il magistrato, ha già fatto una gerarchia per ordine di importanza. È necessario quindi che detto filtro abbia la competenza linguistica, dialettologica, pragmatica e fonetica adeguata.

La trasformazione del parlato in scritto è una operazione tecnicamente molto complessa e tutt'altro che neutra. La ricerca scientifica ha fatto passi da gigante in questo settore, ma tutto ciò non ha neppure scalfito l'opinione comune che vuole che un trascrittore abbia solo pazienza, una cuffia e un registratore.

Una interazione orale non è basata solo ed esclusivamente sul significato delle singole parole, ma bensì su un insieme di tratti e di caratteristiche che compongono il segnale sonoro. Per esempio i tratti sovrasgmentali sono importantissimi in una conversazione. Durante il 1992 scrive: "Molti di quelli che vengono chiamati i *tratti prosodici* del parlato (l'intonazione, il volume, l'altezza, la velocità del parlato) sono usati come segnali di chiave d'interpretazione e cioè per far sapere agli altri che quello che stiamo dicendo deve essere interpretato come serio, scherzoso, ironico, importante, ordinario, ecc. Anche gli aspetti cosiddetti 'paralinguistici', come ad esempio il bisbigliare, il fare la 'voce grossa', il falsetto, sono spesso usati come segnali che marcano una particolare chiave interpretativa. Questo tipo di tratti che accompagnano la produzione di quello che viene a volte chiamato il 'testo verbale' forma una vera e propria 'comunicazione parallela', tutt'altro che secondaria, visto che può modificare o annullare quello che sembra il significato letterale del 'testo'<sup>42</sup>"; e ancora "Le persone coinvolte in un'interazione hanno dei modi per far sapere agli altri partecipanti (siano essi presenti o assenti) la 'chiave' di lettura o d'interpretazione di quello che dicono. Molto spesso la chiave interpretativa per una data espressione linguistica viene data da comportamenti non verbali, che vanno quindi presi in considerazione nell'analisi del parlato: una scrollata di spalle, una smorfia, un'alzata delle sopracciglia, uno strizzare d'occhio, un gesto con le mani"<sup>43</sup>.

È evidente quanto sia necessaria una competenza tecnica per portare tutta l'informazione non verbale su carta o su un verbale.

Una ulteriore ma fondamentale distinzione deve essere effettuata tra i trascrittori di udienza e i periti trascrittori. Questa distinzione viene effettuata raramente dai Tribunali stessi tant'è che spesso i trascrittori di udienza effettuano consulenze e perizie come periti trascrittori. È tutta questa incertezza sui ruoli, sulle competenze, sulle professionalità che rende inattendibile e poco oggettiva la *prova voce registrata*.

Non raggiungendo standard elevati di competenza linguistica nella trascrizione di intercettazioni ambientali, lo stesso valore di prova della trascrizione, da una parte verrà demandato all'abilità delle parti, quindi alle possibilità economiche<sup>44</sup>, alla competenza della difesa o dell'accusa e dall'altra la cattiva competenza rende sempre più debole e basso il valore della prova stessa.

Una non chiara metodologia e una non chiara figura professionale, nel tempo renderà inaffidabile tutte le trascrizioni, sia quelle buone che quelle cattive.

Il CPP e l'inesistenza di un Albo di fronte ad una illegalità così tecnologicamente avanzata perde la partita in partenza.

La trascrizione di intercettazioni ambientali ha una grande potenzialità, è un sottile filo di accertamento dei fatti, può dare un grande aiuto nella ricostruzione ma purtroppo sono poca cosa se non cambiano le

---

<sup>42</sup> Cfr. § 2.6. p. 58.

<sup>43</sup> Cfr. § 2.6, p. 58.

<sup>44</sup> Non per nulla la maggior parte dei linguisti universitari vengono contattati e nominati dalla difesa, È superfluo sottolineare che più alto è il livello di criminalità più disponibilità economiche si hanno per ingaggiare ottimi consulenti.

condizioni in cui molte Procure operano. Il mancato finanziamento e la mancata formazione degli operatori del settore rende sempre più spesso invalidata la trascrizione ed il lavoro operoso di molti.

Così come diventa prova un bossolo nelle perizie balistiche, nel caso delle intercettazioni diventa prova una interazione linguistica, una comunicazione verbale che come tale deve essere studiata ed analizzata da un esperto in comunicazione verbale.

Dovremmo formare operatori addetti alle intercettazioni, creare corsi di formazione per operatori di PG e figure professionali specifiche.

## Appendice

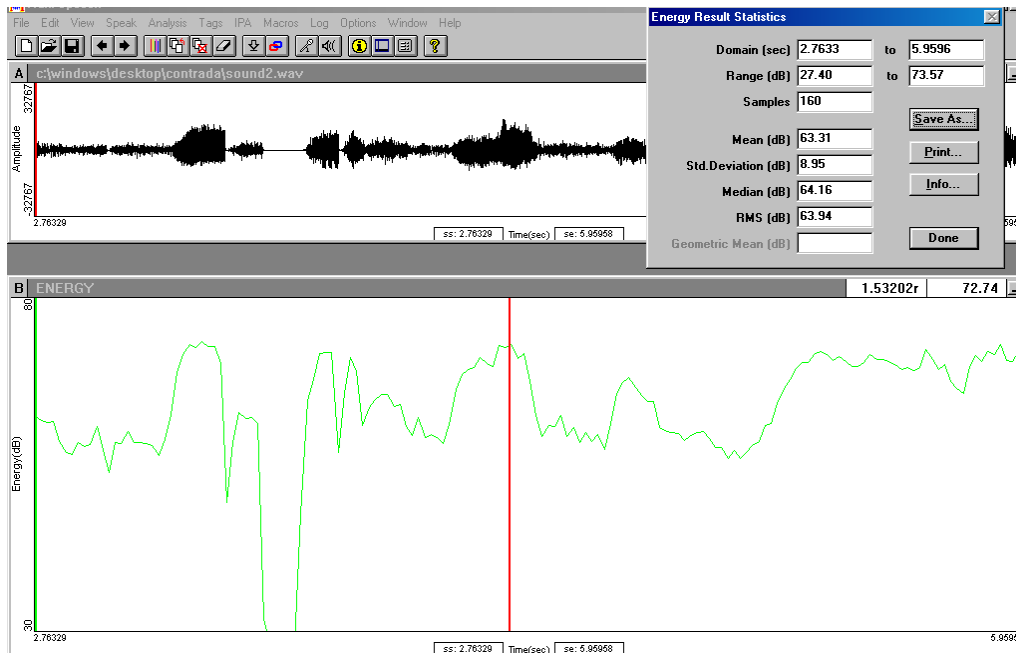


Figura 1 Il grafico mostra l'andamento dell'energia di una porzione di segnale relativo alla musica + il canto di sottofondo.

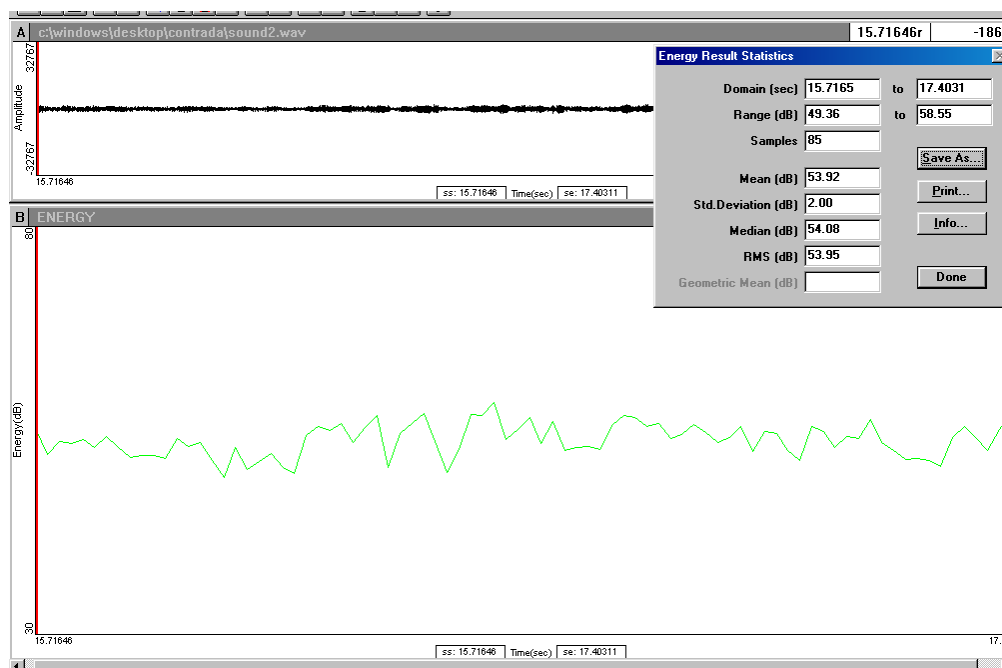


Figura 2 Il grafico mostra l'andamento dell'energia della musica (solo una parte strumentale) di sottofondo.

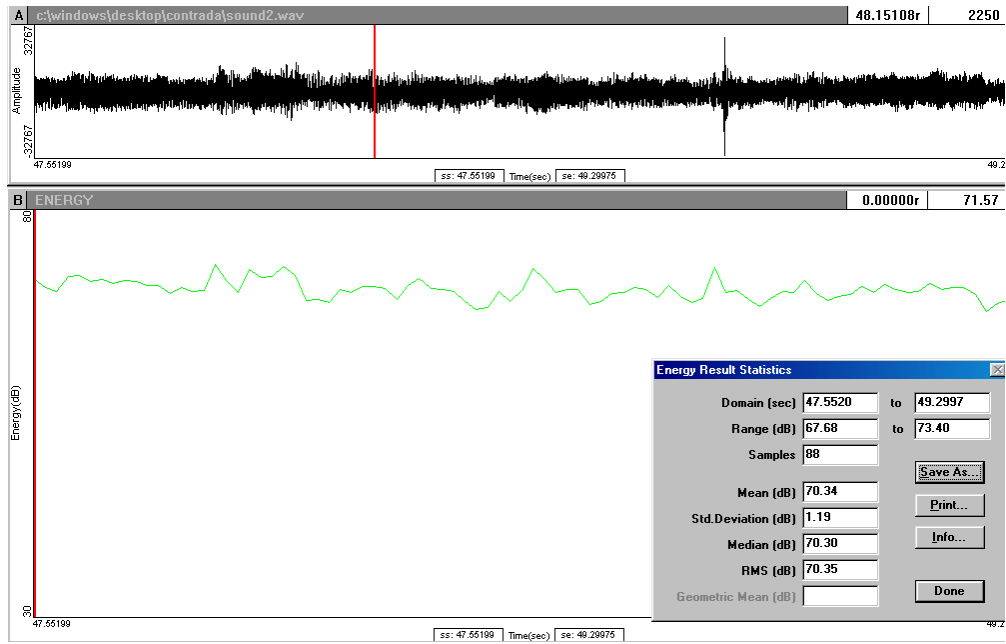


Figura 3 Il grafico mostra l'andamento dell'energia durante la produzione di un acuto del cantante

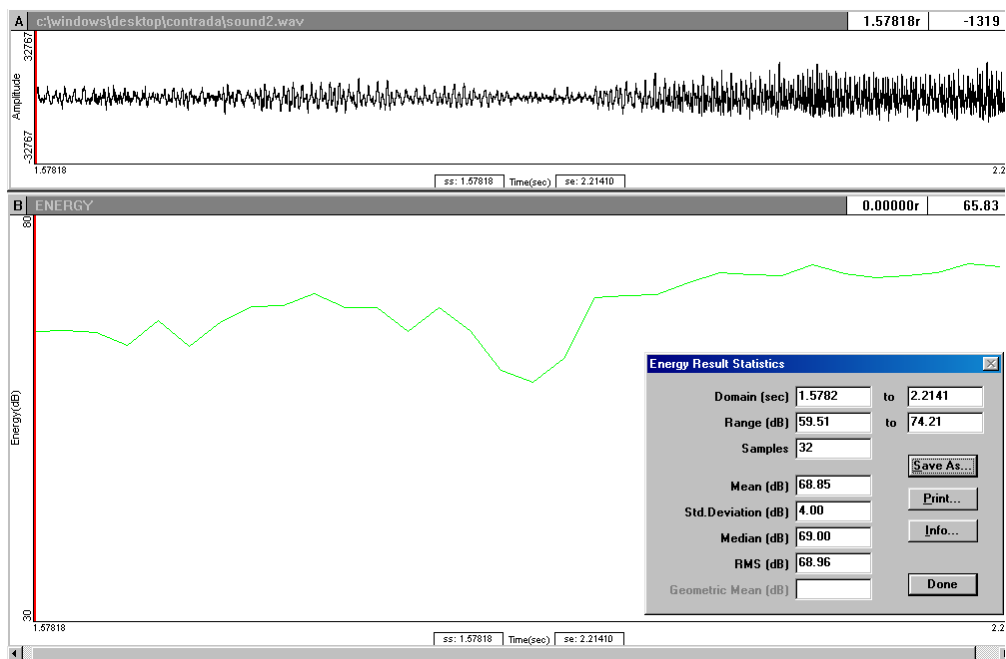


Figura 4 Il grafico mostra l'andamento dell'energia di Uomo 1