

Le perizie foniche in ambito forense

Francesco Beritelli

Dipartimento di Ingegneria Elettrica Elettronica e Informatica
DIEEI - Università di Catania

Argomenti del seminario

- La fonetica forense
- La biometria e i parametri prestazionali
- Sistema fonatorio e caratteristiche del segnale vocale
- Frequenza fondamentale e formanti
- Metodi di identificazione del parlante
- Confronto delle tecniche e strumenti
- IDEM
- Degradazione della qualità del segnale fonico
- Conclusioni

La fonetica forense

Trascrizione di voci

Identificazione del parlante

Caratterizzazione del parlante

Genuinità di una registrazione

Restauro del segnale vocale

Identificazione di un suono

Modificazioni di una voce

Analisi di segnali audio

Riconoscimento biometrico

Con il termine “riconoscimento biometrico” si fa riferimento alle tecniche automatiche per riconoscimento dell’identità di un individuo basate sull’uso di caratteristiche fisiologiche o comportamentali distintive. La modalità del riconoscimento può essere in termini di :

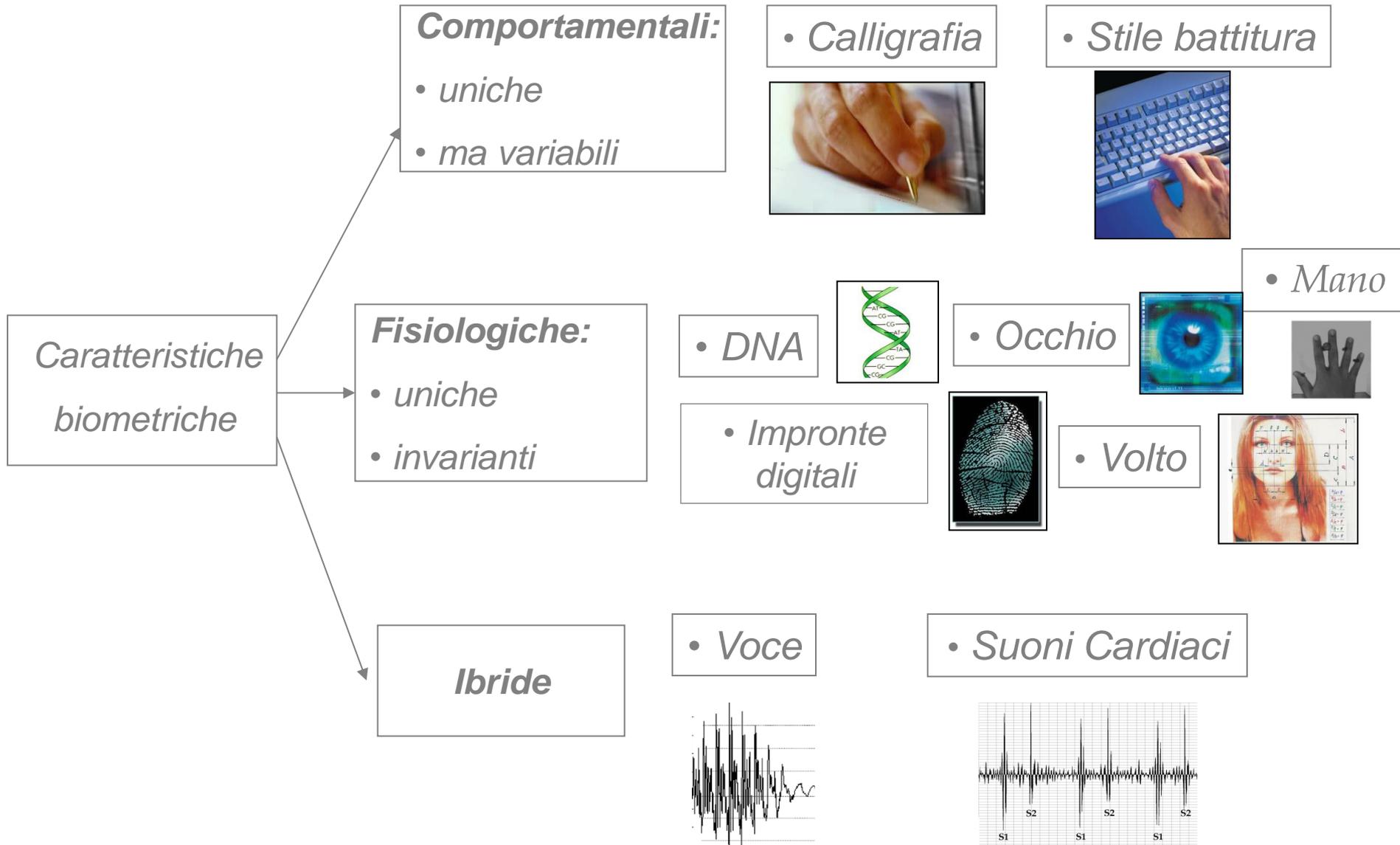
Verifica (Autenticazione): si dichiara l’identità

Confronto uno a uno al fine di determinare se l’identità dichiarata dall’utente è vera o no

Identificazione: l’utente non dichiara l’identità

Confronto uno a molti al fine di stabilire l’identità dell’individuo

Indici biometrici



Errori nei sistemi biometrici

- **False Match** (nel riconoscimento positivo chiamato spesso **False Acceptance** o **Falsa Attribuzione**)
 - misurazioni biometriche di persone diverse vengono erroneamente considerate come appartenenti alla stessa persona
- **False Non-Match** (nel riconoscimento positivo chiamato spesso **False Rejection** o **Falsa Esclusione**)
 - misurazioni biometriche della stessa persona vengono erroneamente attribuite a persone diverse

Ipotesi:

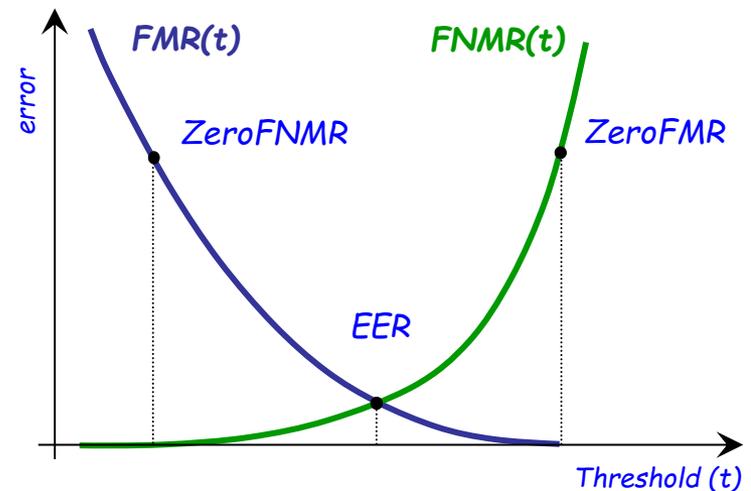
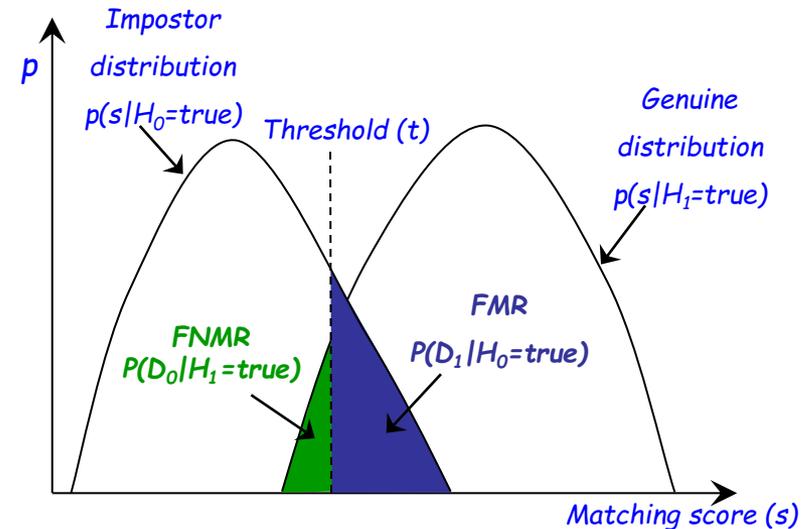
H_0 : persona diversa

H_1 : stessa persona

Decisioni possibili:

D_0 : persone diverse

D_1 : stessa persona



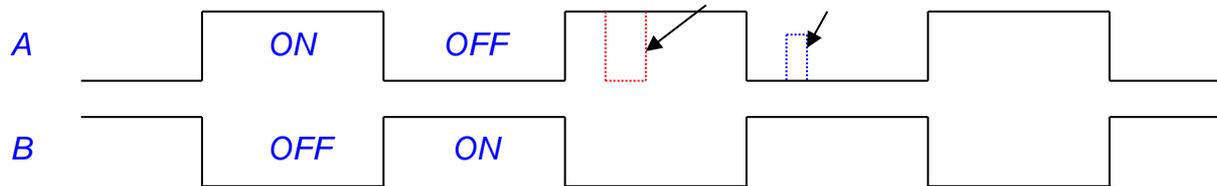
Attendibilità degli indici biometrici

ERRORE	Falsa Esclusione (FE)	Falsa Attribuzione (FA)
IMPRONTE VOCALI (voce naturale Hi-Fi)	1 % (1 su 100)	0,1 % (1 su 1000)
IMPRONTE VOCALI (segnale telefonico)	3 % (3 su 100)	0.4% (4 su 1000)
IMPRONTE DIGITALI	0,5 % (5 su 1000)	0,001 % (1 su 100.000)
FIRMA	0,2 % (2 su 1000)	0,6 % (6 su 1000)
RETINA	2,8 % (2,8 su 100)	0%

Caratteristiche del segnale vocale

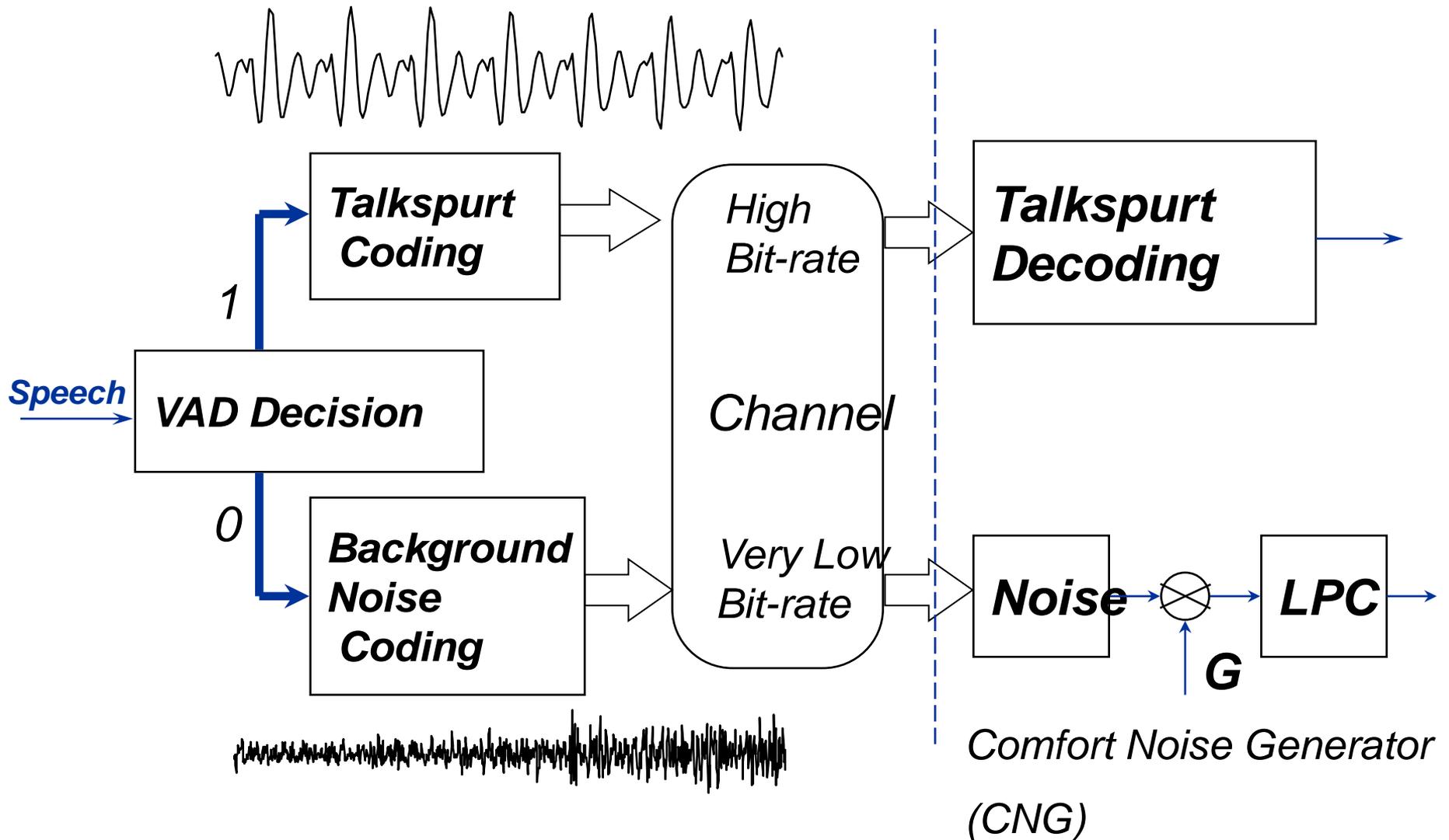
- Voce naturale: banda 20-10000 Hz
- Segnale telefonico: banda standardizzata ITU-T:
 - Narrowband (20÷4000 Hz, qualità telefonica)
 - Wideband (20÷7000 Hz, qualità audioconferenza)
- Media nulla, distribuzione non uniforme
- Segnale non stazionario con correlazione a breve e lungo termine
- Struttura “on-off” nel tempo con il 40 % di attività vocale e circa il 60 % di pause di silenzio
- Suoni vocalizzati, non vocalizzati e misti

Struttura ON-OFF di una conversazione

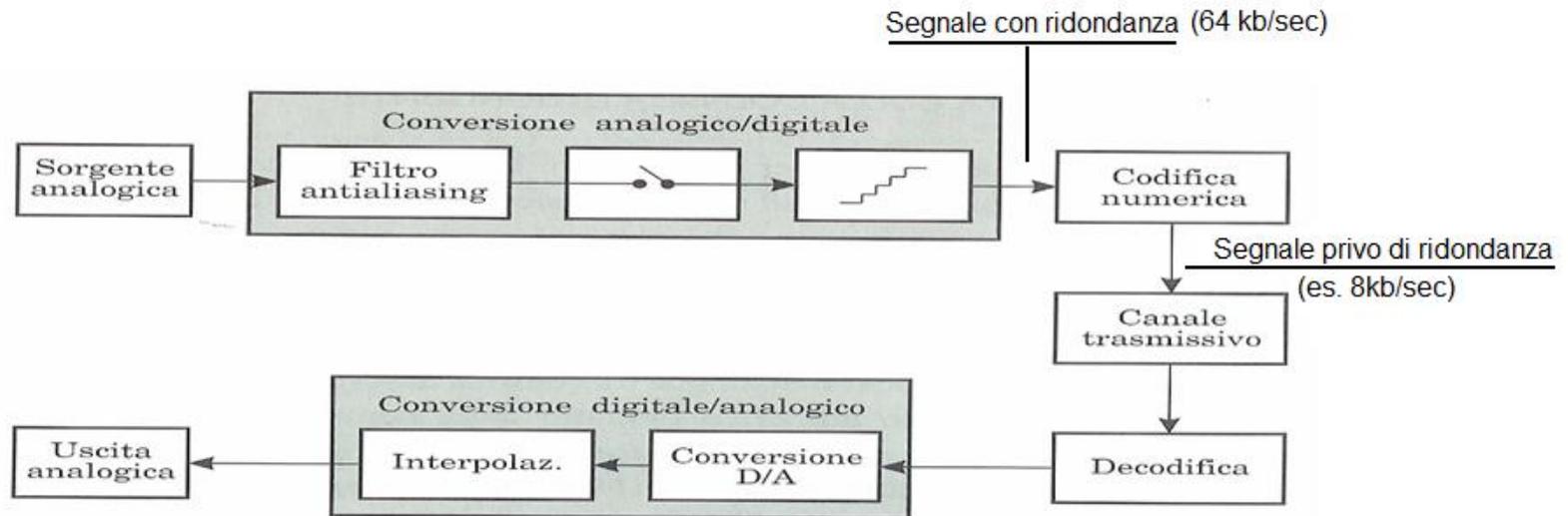


- *Utilizzo di un codec CBR + VAD (source driven)*
- *Codifica CBR dei tratti di ON (Talkspurt)*
- *Trasmissione periodica dei SID nei tratti di OFF (silenzio o rumore ambientale)*
- *Sintesi dei tratti di OFF in Rx tramite un CFG*
- *Prestazioni dipendono dal rumore ambientale*

Voice Activity Detection (VAD)



Compressione/Decompressione del segnale vocale



COding + DECo ding = CODEC

- *bit rate $r_b = f_c * b$,*
- *La frequenza di campionamento f_c (NB o WB) e la risoluzione b determinano la qualità del segnale originale non compresso*
- *Fattore di compressione = input bit rate/output bit rate (es. $64/8=8$)*

Standard ITU-T G.711 Log PCM

- Banda telefonica 0÷4kHz
- Frequenza di campionamento $F_c=8\text{kHz}$
- Distribuzione non uniforme
- Compressione della dinamica (legge A o μ)
- Bitrate: $r_b=b \cdot F_c=b \cdot 8\text{kHz}$ (b n. bit di quantizzaz.)
- 12 bit lineari/campione \rightarrow 8 bit logaritmi/campione
- 96kbit/sec (b=12) \rightarrow 64kbit/sec (b=8)
- Qualità MOS 4.3
- Standard per l'accesso base ISDN

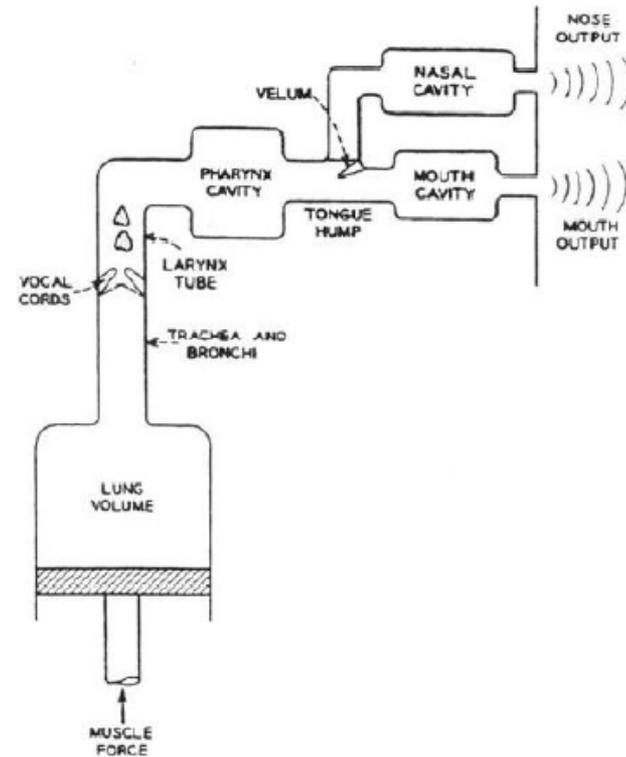
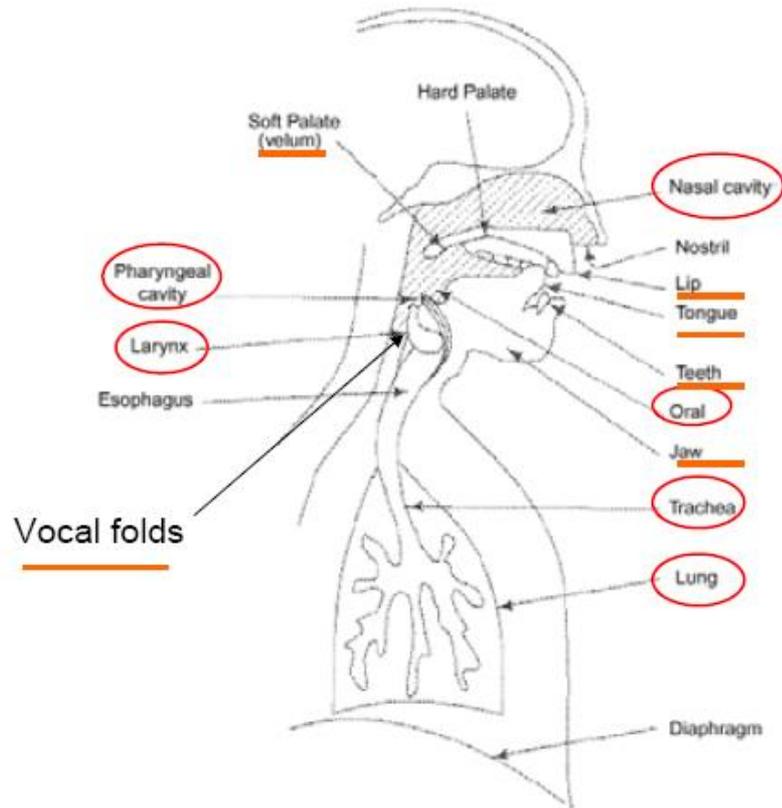
Standard di codifica vocale ITU-T

Standard	Description	Bit rate	Mos (Ber=0)
ITU-T G.711	Log PCM	64 kbit/s	4.3
ITU-T G.723.1	Dual rate speech coder for multimedia applications (MP-MLQ/ACELP)	6.3/ 5.3 kbit/s	3.8 / 3.6
ITU-T G.726	Adaptive Differential Pulse Code Modulation (ADPCM); contains obsolete Rec. G.721, G.723	16,24, 32 ,40 kbit/s	4.0
ITU-T G.727	5-, 4-, 3-, and 2 bits per sample embedded Adaptive Differential Pulse Code Modulation (ADPCM)	16,24,32,40 kbit/s	N.D.
ITU-T G.728	Low Delay CELP (LD-CELP)	16 kbit/s	4.2
ITU-T G.729	Conjugate Structure Algebraic CELP (CS-ACELP) Annex A:Reduced complexity algorithm at 8 kbit/s Annex D:Low-Rate extension at 6.4 kbit/s AnnexE : High-Rate extension at 11.8 kbit/s	8 kbit/s (6.4, 11.8 kbit/s)	3.9
ITU-T G.722	Wideband (7 kHz) audio codec by Subband ADPCM (SB-ADPCM)	64 (56, 48) kbit/s	N.D.

Standard di codifica vocale ETSI

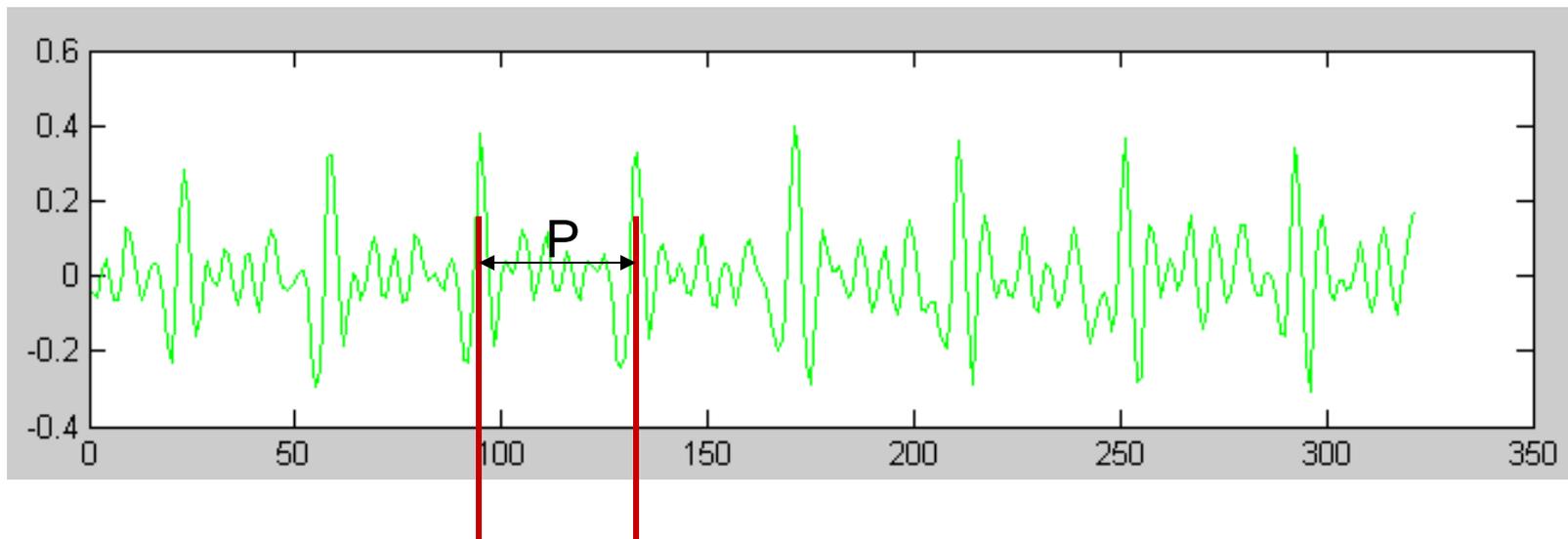
Standard	Description	Bit rate	Mos (Ber=0)
ETSI GSM 06.10	Full Rate (FR) speech transcoding (RPE-LTP:Regular Pulse Excitation-Long Term Prediction)	13 kbit/s	3.7
ETSI GSM 06.20	Half Rate (HR) speech transcoding (VSELP:Vector sum Excited Linear Prediction)	5.6 kbit/s	3.5
ETSI GSM 06.60	Enhanced Full Rate (EFR) speech transcoding (ACELP:Algebraic CELP)	12.2 kbit/s	3.9
ETSI GSM (AMR)	Used in UMTS	4.75 - 12.2 kbit/s	3.4 - 3.9
ETSI GSM (AMR-WB) = ITU-T G.722.2 WB	Used in UMTS	6.60 – 23.85 kbit/s	3.5 – 4.2

Sistema fonatorio umano e schema a blocchi



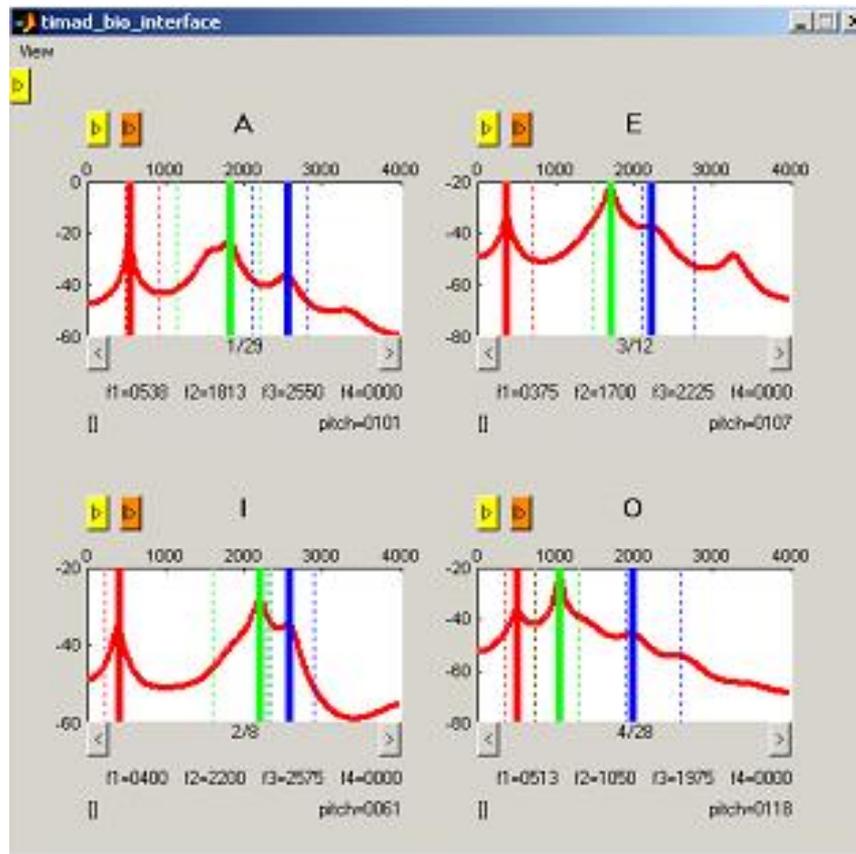
La frequenza fondamentale

La *frequenza fondamentale* o *pitch* è l'inverso del periodo vibrazione delle corde vocali durante l'emissione di un suono vocalizzato



Le formanti

- Sono le frequenze dello spettro vocale in cui è massima l'energia
- Sono rappresentative delle caratteristiche fisiche del tratto vocale
- Si estraggono tramite l'analisi LPC o CEPSTRUM



La voce è un suono complesso perché è dato dalla combinazione di tre effetti:

1. La vibrazione delle corde vocali

(genera la Frequenza Fondamentale)

2. Il rumore prodotto nella fonazione

3. Il transito attraverso il tratto vocalico

(genera le Frequenze Formanti)

Modello del sistema fonatorio

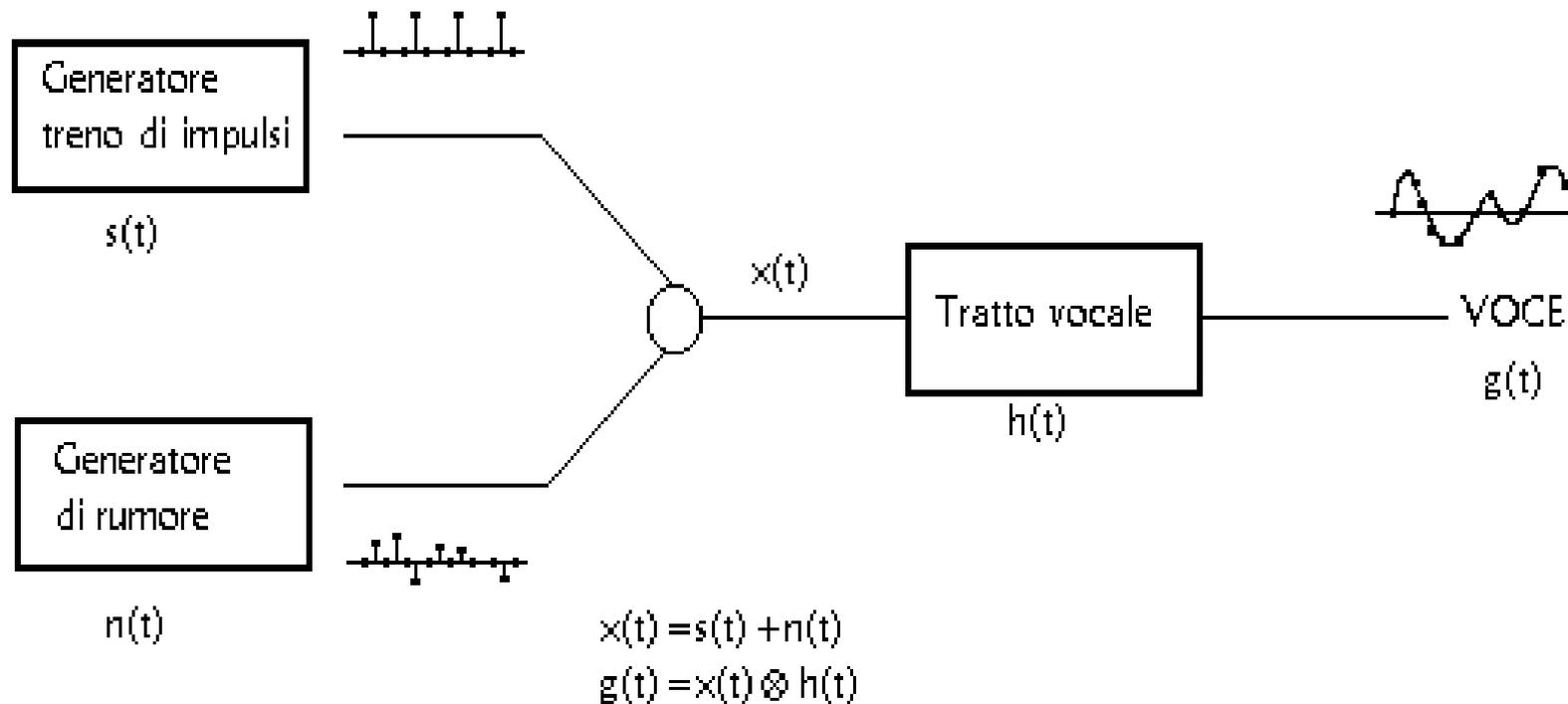
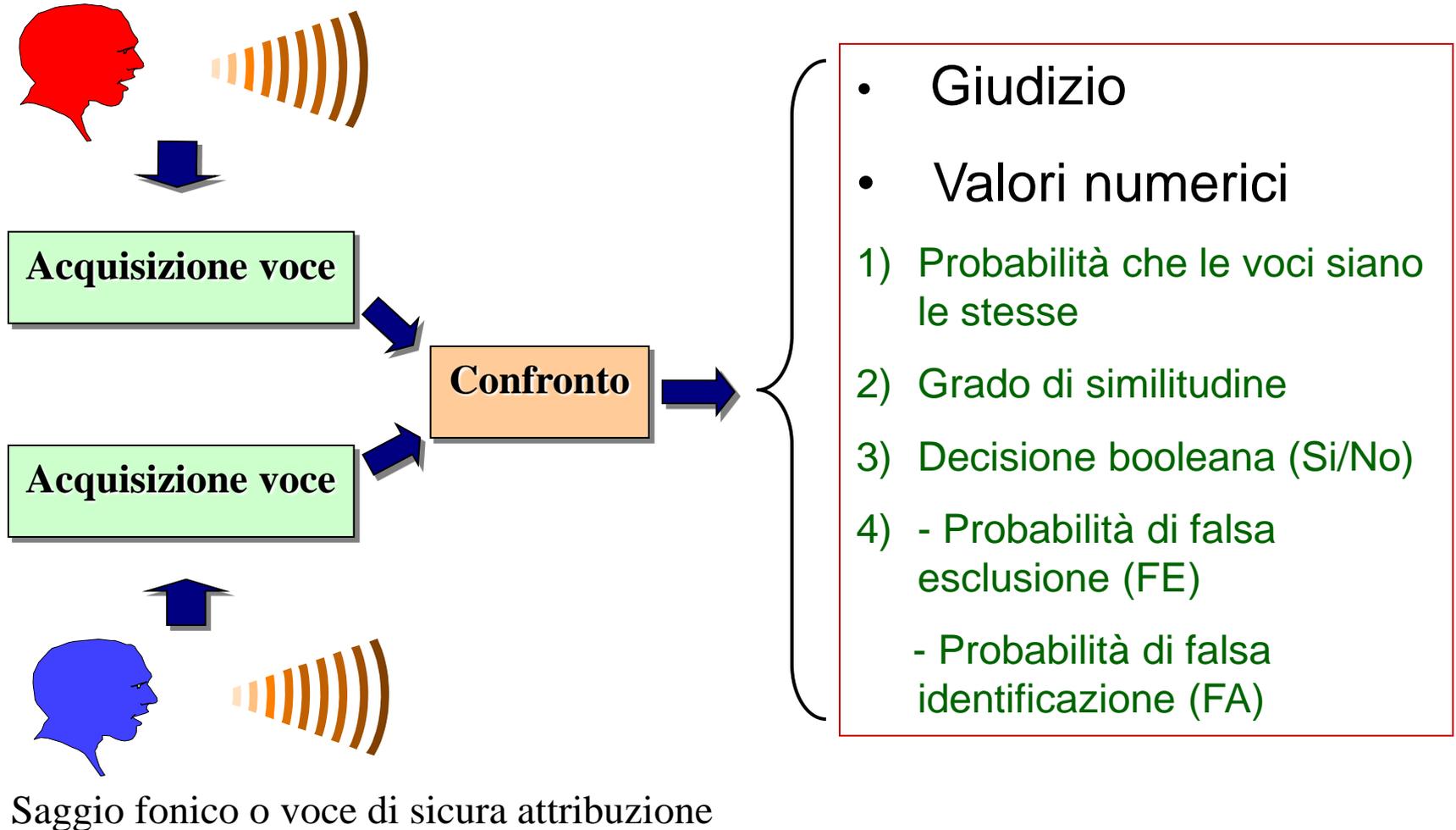


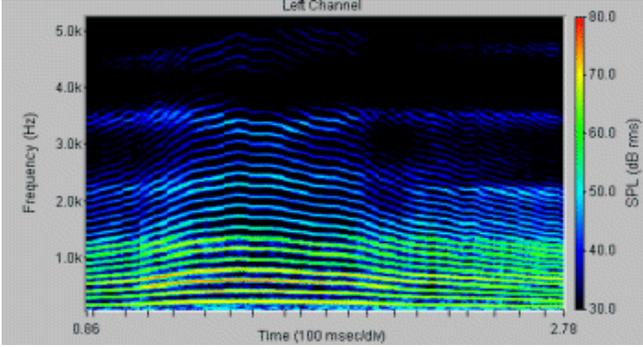
FIG. 1- MODELLO DI GENERAZIONE DELLA VOCE

Comparazione fonica

Parlatore ignoto (voce intercettata)



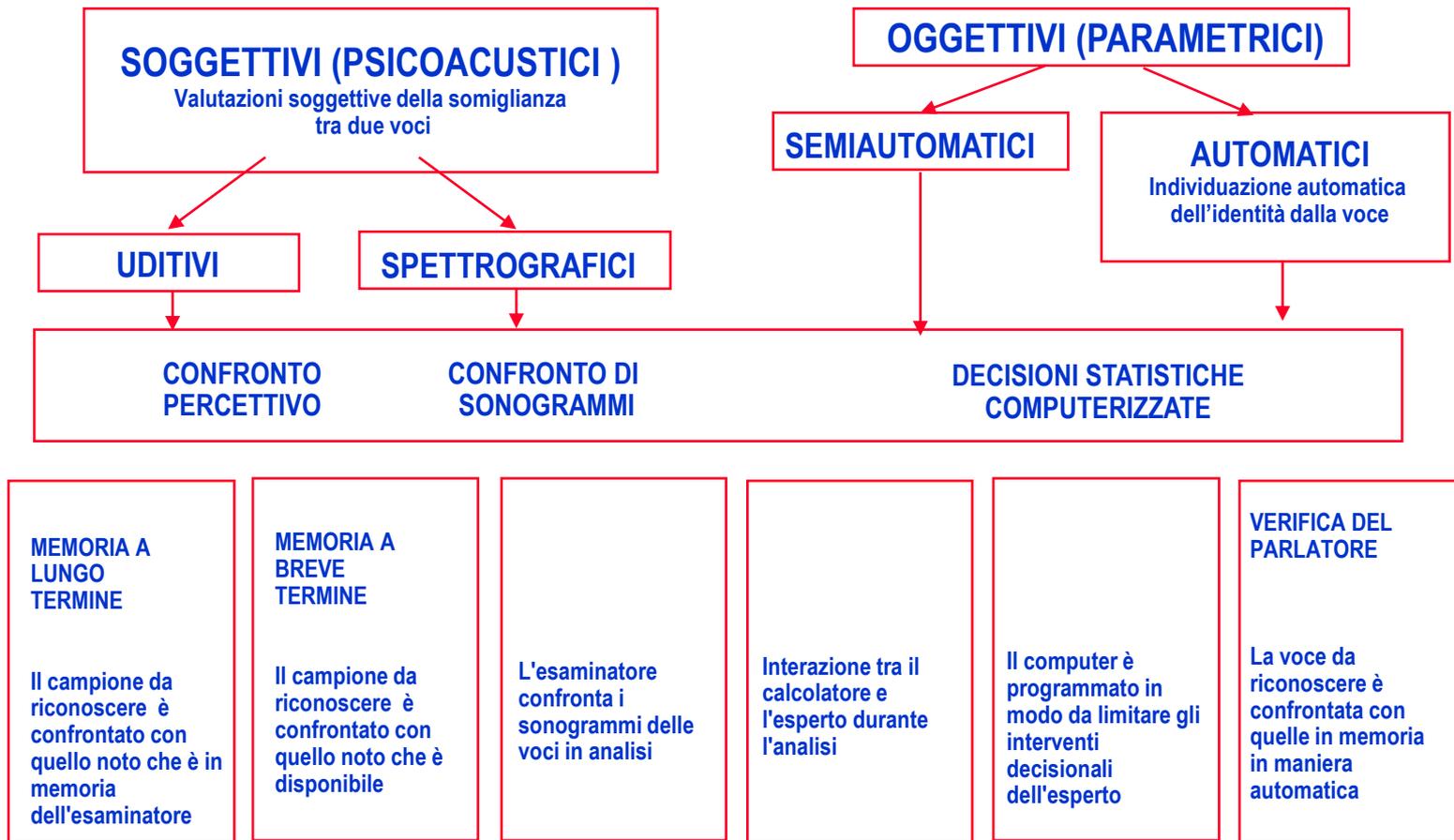
Metodi di identificazione

<i>SOGGETTIVI</i>	<i>OGGETTIVI</i>
- Confronto uditivo	Metodo parametrico
- Fonetica acustica	
- Metodo sonografico 	

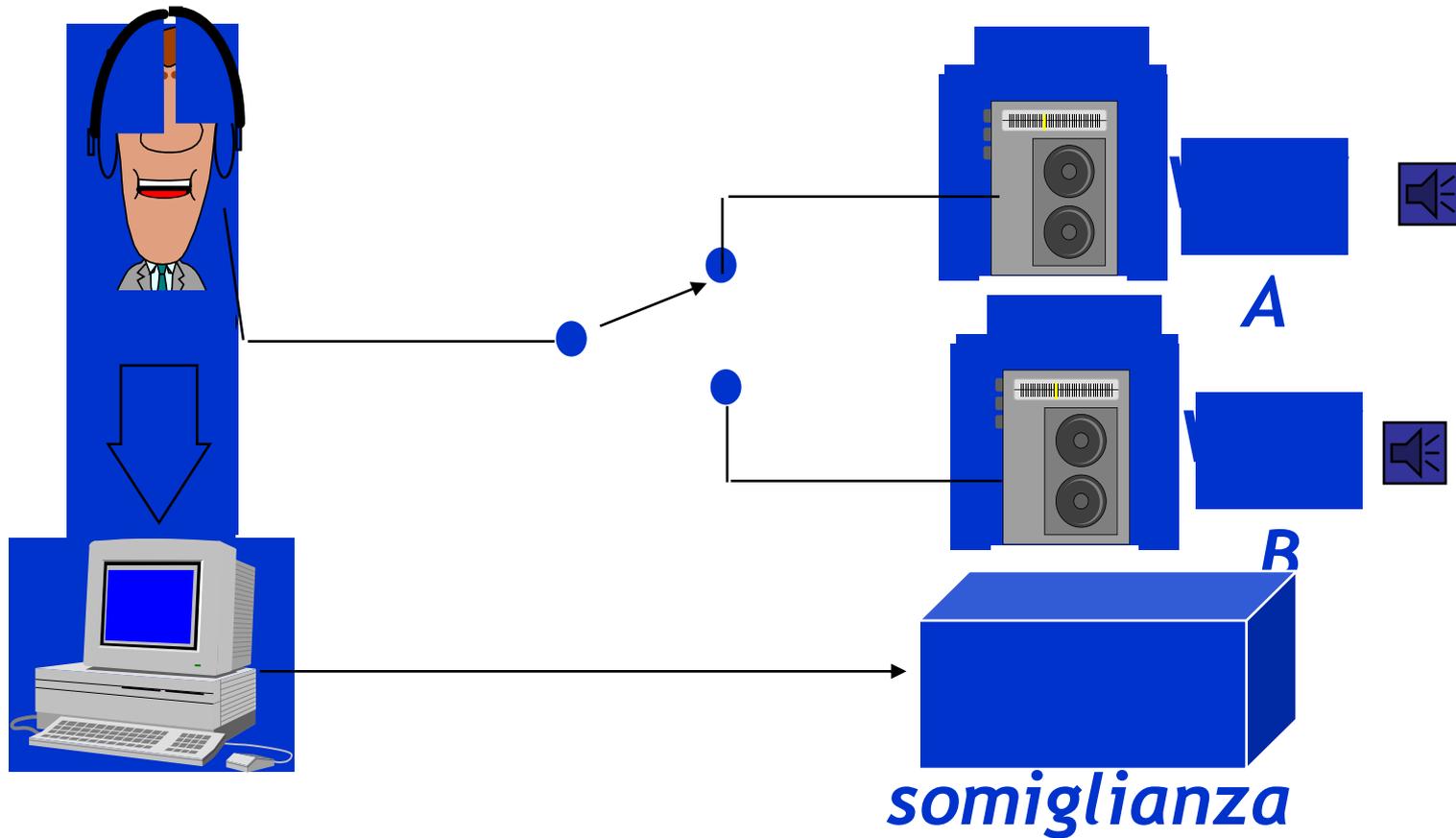
Metodi di identificazione

METODI	VANTAGGI	SVANTAGGI
<i>SOGGETTIVI</i>	<ul style="list-style-type: none">- Più semplici	<ul style="list-style-type: none">- Risultato opinabile- Grado di similitudine- Sensibili a disturbi
<i>OGGETTIVI</i>	<ul style="list-style-type: none">- Approccio statistico- Automatico o semiautomatico- Calcola FE e FA	<ul style="list-style-type: none">- Numerosità dei dati- Sensibili ai disturbi- Più complessi

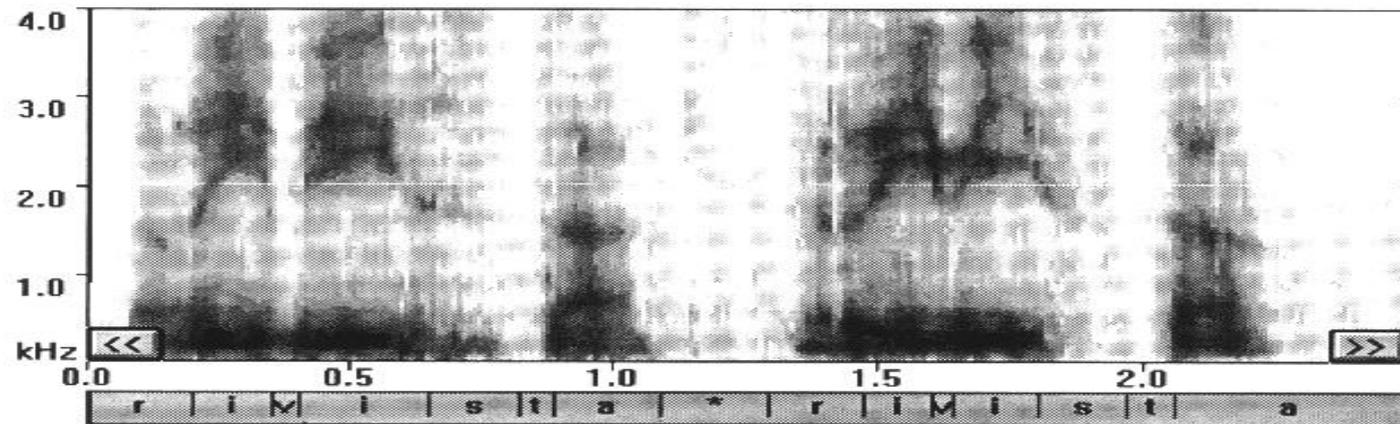
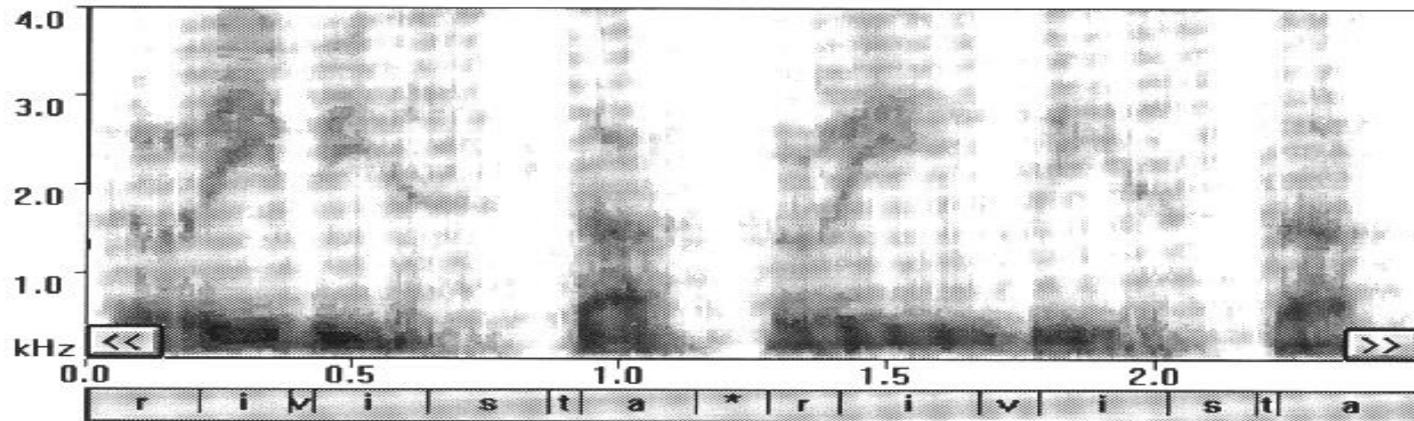
Metodi di identificazione del parlante



Confronto uditivo



Confronto di sonogrammi





International Association for Forensic Phonetics and Acoustics

Home

News

Journal

Conferences

Officers & Committees

Code of Practice

Resolutions

Language ID

Voiceprints

Constitution

Membership

Contact

Resolutions

IAFPA Resolution - Voiceprints

The following resolution was passed at the IAFPA annual conference, Plymouth, UK, 24th July, 2007.

IAFPA dissociates itself from the approach to forensic speech comparison known as the "voiceprint" or "voicegram" method in the sense described in Tosi (1979).

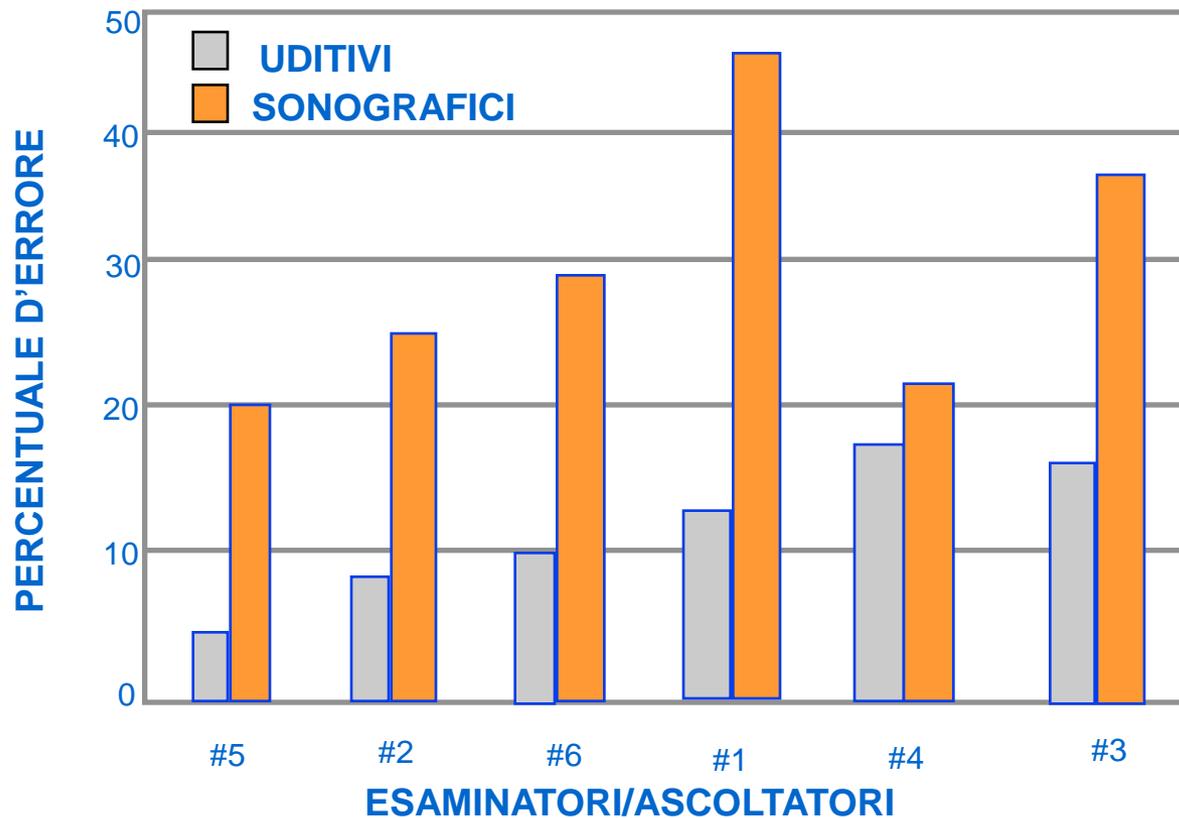
This approach to forensic speaker identification involves the holistic, i.e., non-analytic, comparison of speech spectrograms in the absence of interpretation based on understanding of how spectrographic patterns relate to acoustic reflexes of articulatory events and vocal tract configurations.

The Association considers this approach to be without scientific foundation, and it should not be used in forensic casework.

Reference

Tosi, O.I. (1979) Voice Identification: Theory and Legal Applications. Baltimore: University Park Press.

Confronto tra metodi sonografici e uditivi



Metodo Linguistico-Fonetico

- ✓ Storicamente è stato il primo ad apparire
- ✓ Richiede un notevole background di carattere fonetico
- ✓ E' fortemente dipendente dalla lingua
- ✓ Richiede grandi quantità di materiale audio

Metodo Linguistico-Fonetico

Le analisi si dividono in:

- 1. Prove d'ascolto*
- 2. Esami fonetici*
- 3. Esami lessicali*
- 4. Esami prosodici*

1. Prove d'ascolto

Sono basate sulla memoria a breve termine

Si costruisce un file, contenente sia le frasi anonime che le medesime pronunciate dal sospettato, poi si procede all'ascolto, ripetendo l'operazione in tempi diversi

2. Esami fonetici

Si ricercano gli elementi caratterizzanti che differiscono dalla pronuncia corretta nazionale

ES. TRE → tré, trè, tce...

ES. NOVE → nòve, nóve...

L'insieme dei risultati darà indicazioni sulla regione linguistica di provenienza del parlatore, tuttavia il riscontro di particolari difetti fornisce importanti elementi di riconoscimento individuale. Per questo scopo le consonanti più sensibili sono:

'R', 'S', 'C', 'G', 'Z'

3. Esami lessicali

Lingua, Dialecto

Forniscono indicazioni sulla regione linguistica di provenienza del parlatore

Terminologia specifica (gergo)

Padronanza della lingua

Aiutano ad individuare l'ambiente socio-culturale e scolastico-professionale

Terminologia preferenziale (ripetizioni)

Terminologia personale

Forniscono informazioni legate alle abitudini individuali

4. Esami prosodici

Inflessione/cadenza

Danno indicazioni sulla regione linguistica di provenienza del parlatore

Pause (vuote e piene) - Lunghezza consonanti plosive (p,t,k)

Sono elementi individuali misurabili strumentalmente

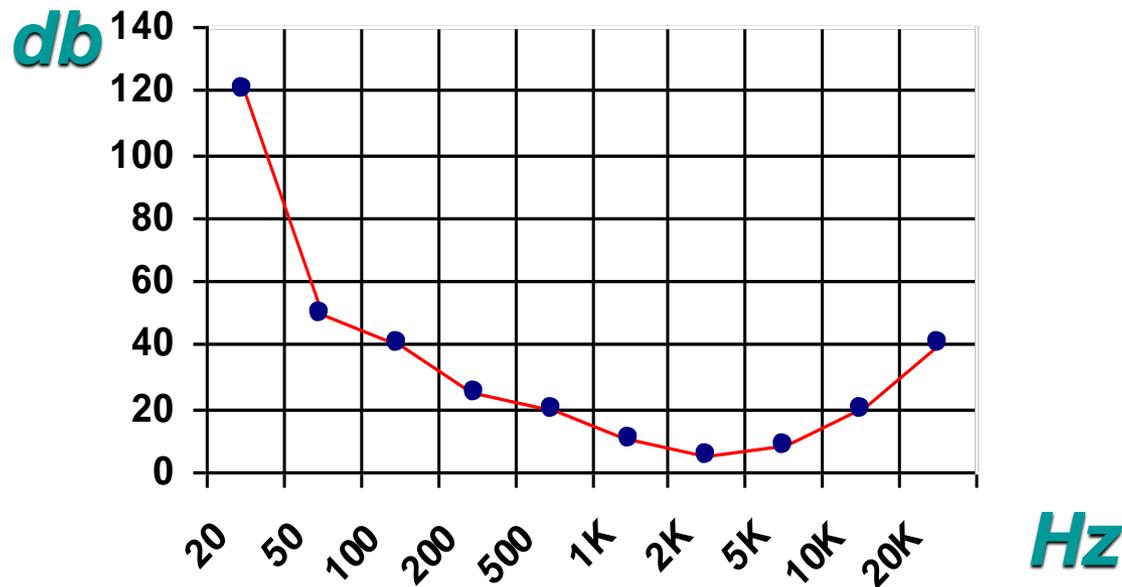
Velocità di locuzione

E' il parametro più utile a fini forensi, per misurabilità, stabilità e trattabilità statistica

Metodo Linguistico-Fonetico: considerazioni

- ✓ *Alcuni parametri sono 'robusti' rispetto agli effetti del canale e al rapporto segnale/rumore*
- ✓ *Richiede giorni per un esame completo*
- ✓ *E' complicato creare database*
- ✓ *Le risposte sono generalmente qualitative, le valutazioni soggettive*

L'orecchio è uno strumento attendibile?



*GRAFICO DELLA SENSIBILITA'
DELL'ORECCHIO UMANO*

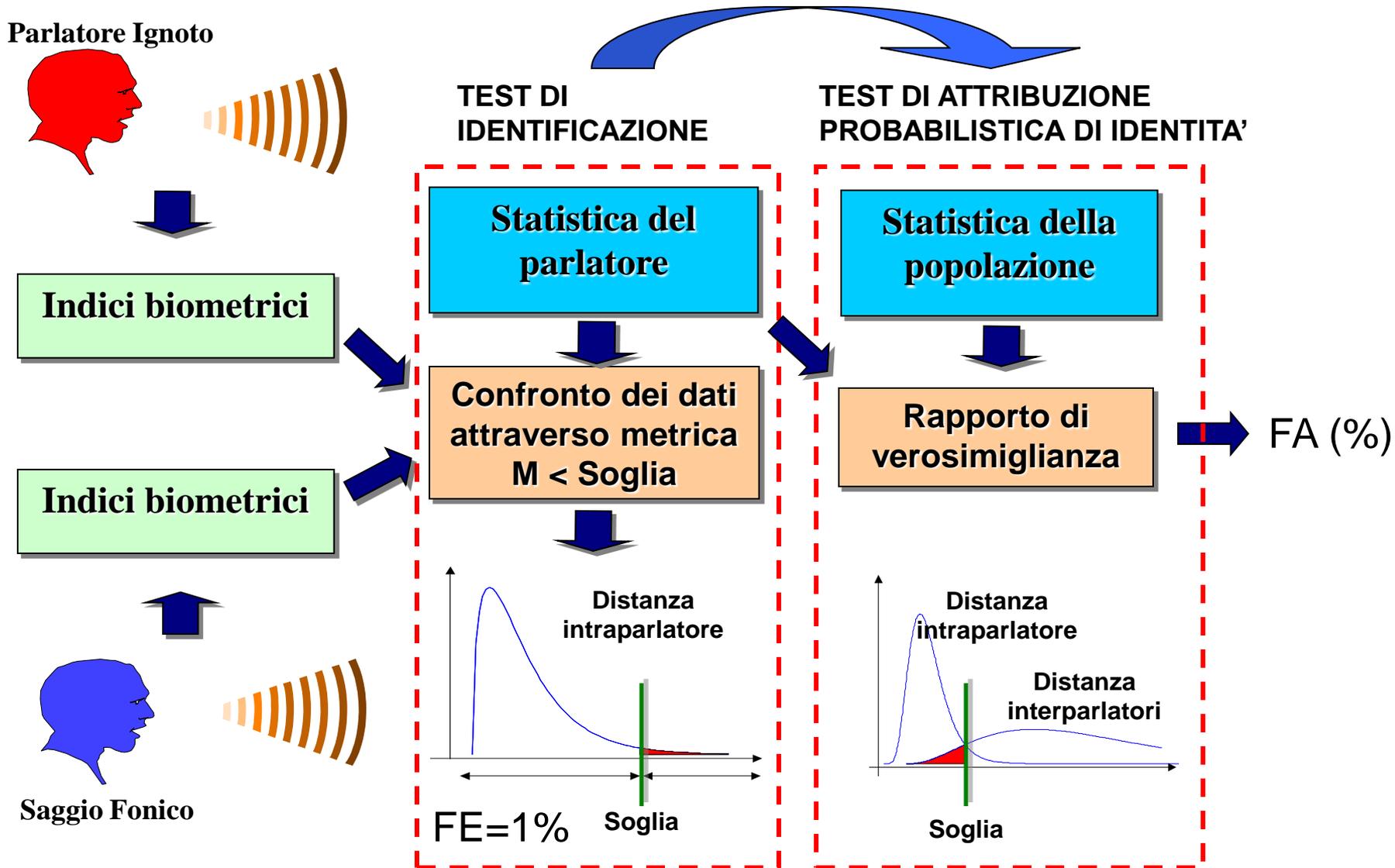
La percezione è utile?

1. *Tutti abbiamo forme di percezione*
2. *La percezione solitamente è PARZIALE*
3. *La verità può essere AMBIGUA*
4. *La verità può essere UNIVOCA ma la nostra percezione porta a conclusioni ERRATE*

Decisioni oggettive (in ambito forense)

- L'identificazione del parlatore con metodi uditivi o con i sonogrammi non consente di valutare l'affidabilità raggiunta.
- La principale fonte di controversia è la soggettività delle decisioni prese dall'esperto.

Metodo parametrico



Proprietà dei parametri biometrici

PER UNA BUONA IDENTIFICAZIONE:

- PICCOLE VARIABILITA' INTRAPARLATORE
- GRANDE VARIABILITA' INTERPARLATORE
- POCO SENSIBILI ALLE DEGRADAZIONI

PARAMETRI ACUSTICI PRINCIPALMENTE USATI:

- LA FREQUENZA FONDAMENTALE
- LE FORMANTI (IN GENERE LE PRIME TRE)

Metodo SEMI-AUTOMATICO

Sistemi computerizzati che richiedono l'interazione uomo-macchina. Alcune dei task sono automatici:

- *Stima rapporto s/n*
- *Estrazione di parametri (pitch, formanti...)*
- *Modelli matematici intra-interparlatore (GMM, Distances, Matrix...)*
- *Analisi statistiche e decisionali (Tests)*

Metodo SEMI-AUTOMATICO

Attività 'manuali' lasciate all'operatore:

- *Selezione dei tratti da misurare (editing)*
- *Controllo delle misure*
(correttezza, selezione delle vocali, deleting...)
- *Selezione dei parametri decisionali*
(livello di confidenza, variabili e database...)
- *Controllo finale*

Metodi Semiautomatici - considerazioni

- ✓ *Lavorano nel dominio delle frequenze in condizioni di stazionarietà (i.e. F0 e Formanti)*
- ✓ *Le formanti sono robuste rispetto al canale*
- ✓ *E' possibile una ricca trattazione statistica (Lr o P.F.A./P.F.R.)*
- ✓ *Sono metodi accettati in dibattito*

Metodi semiautomatici europei

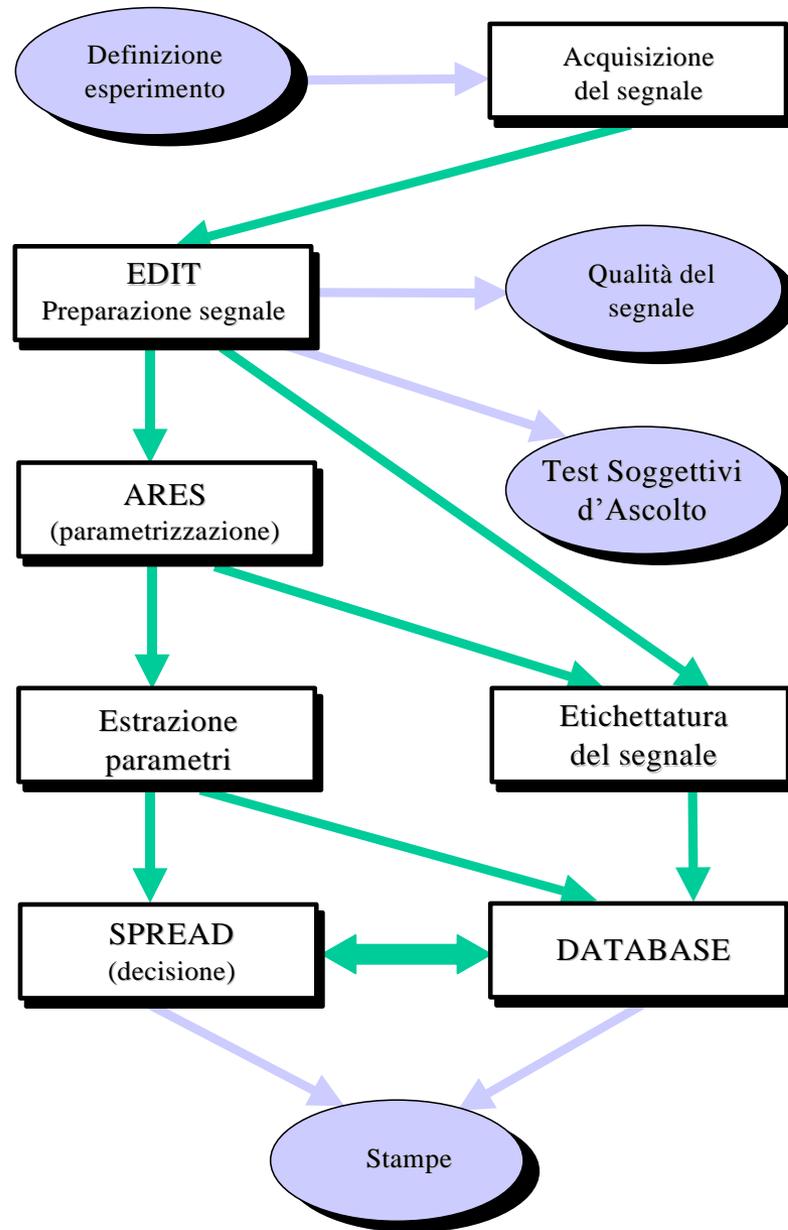
IDEM	ITALY (Carabinieri)
SMART	ITALY (Police) (+FRANCE and SPAIN -Police-)
SIVE	LITHUANIA, LATVIA, ESTONIA CHECZ REP., POLAND, FINLAND
DIALECT -Phonexi-	RUSSIA, BELORUSSIA, UKRAINE, KAZAKHSTAN, UZBEKISTAN

IDEM

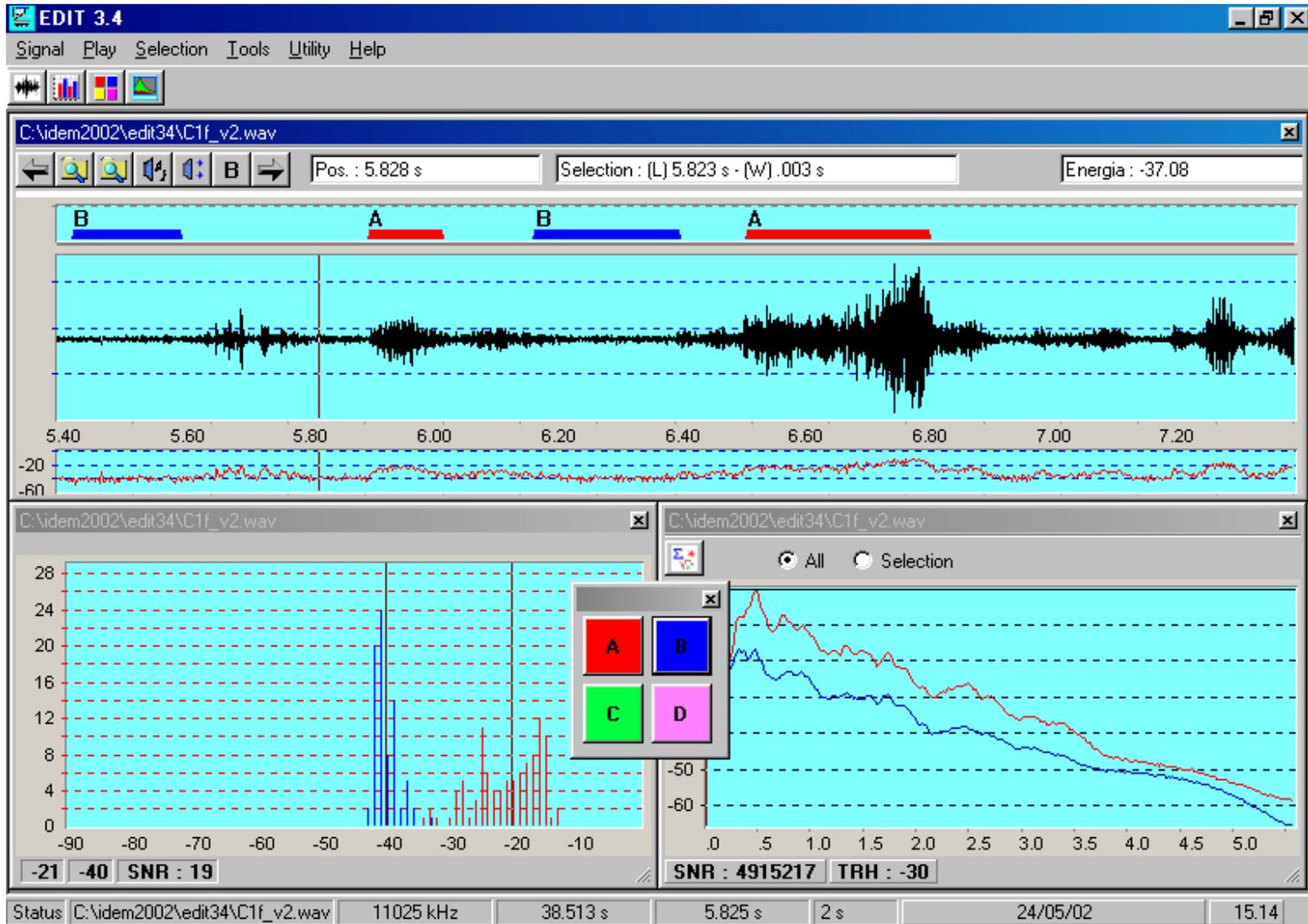
Fondazione

Ugo

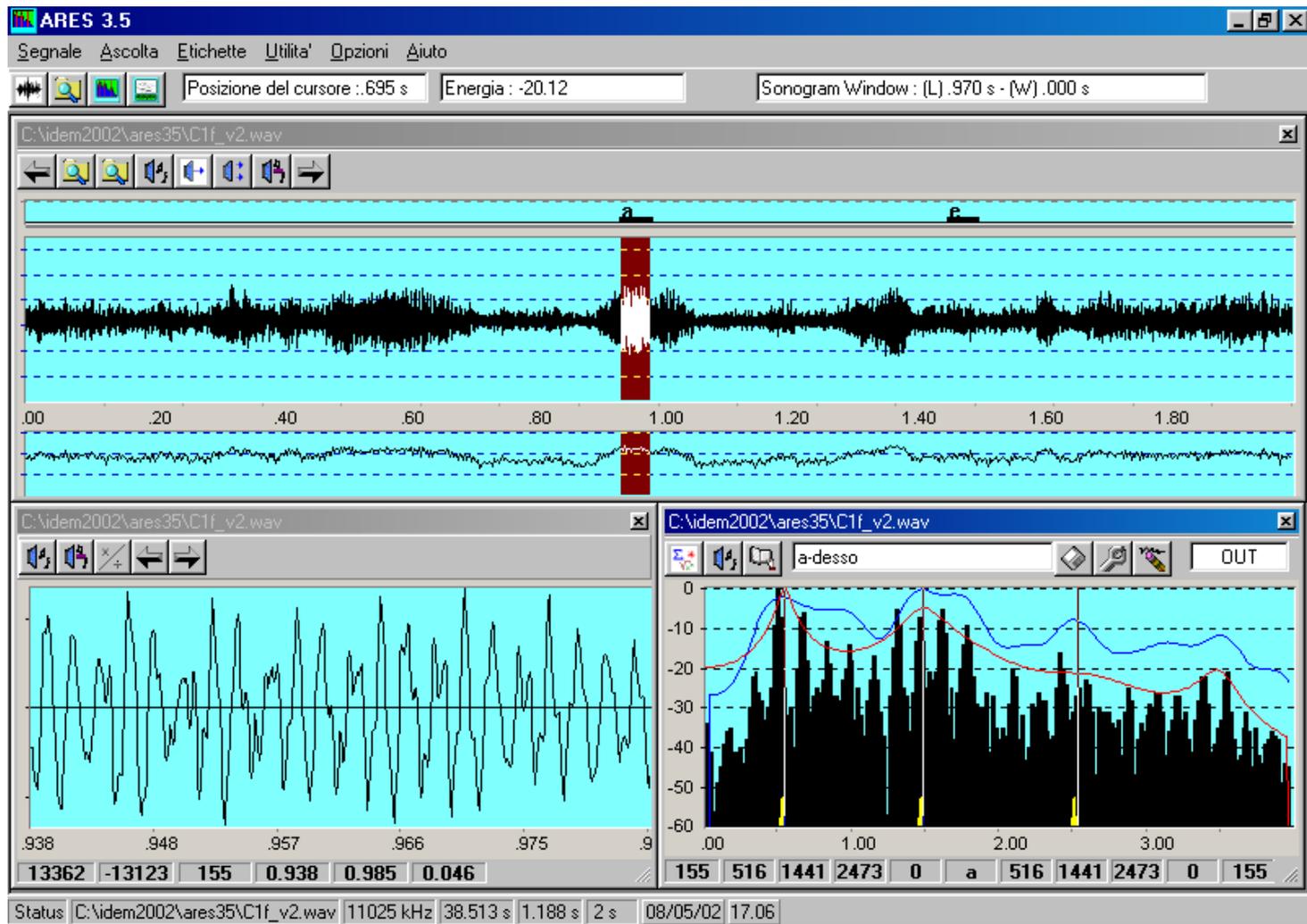
Bordoni



Modulo EDIT



Modulo ARES: segmentazione ed estrazione delle formanti



Test di verifica etichettatura

The screenshot shows a software window titled "ETICHETTE". It contains a table with 10 columns: SEQ, VOCALE, PAROLA, POSIZIONE, FF0, FF1, FF2, FF3, and FF4. The table has three rows of data. Below the table are two main control panels: "Ordinamento" (Ordering) with radio buttons for "Posizione", "Sequenza", and "Testo"; and "Ascolta" (Listen) with a slider from 0 s to 1 s (set to 0.5), buttons for "Ascolta singola" and "Ascolta in sequenza", and checkboxes for vowels "a", "e", "i", "o", and "u". At the bottom are buttons for "Aggiungi", "Cancella", "Ripristina", "Aggiorna", and "Chiudi", along with navigation arrows.

	SEQ	VOCALE	PAROLA	POSIZIONE	FF0	FF1	FF2	FF3	FF4
▶	1	a	a-desso	0.938	155	600	1450	2300	0
	2	o	t-o-mo	1.333	235	430	1054	2452	0
	3	e	p-e-r	1.588	119	452	1634	2430	0

Ordinamento

- Posizione
- Sequenza
- Testo

Ascolta

0 s. 0.5 1 s.

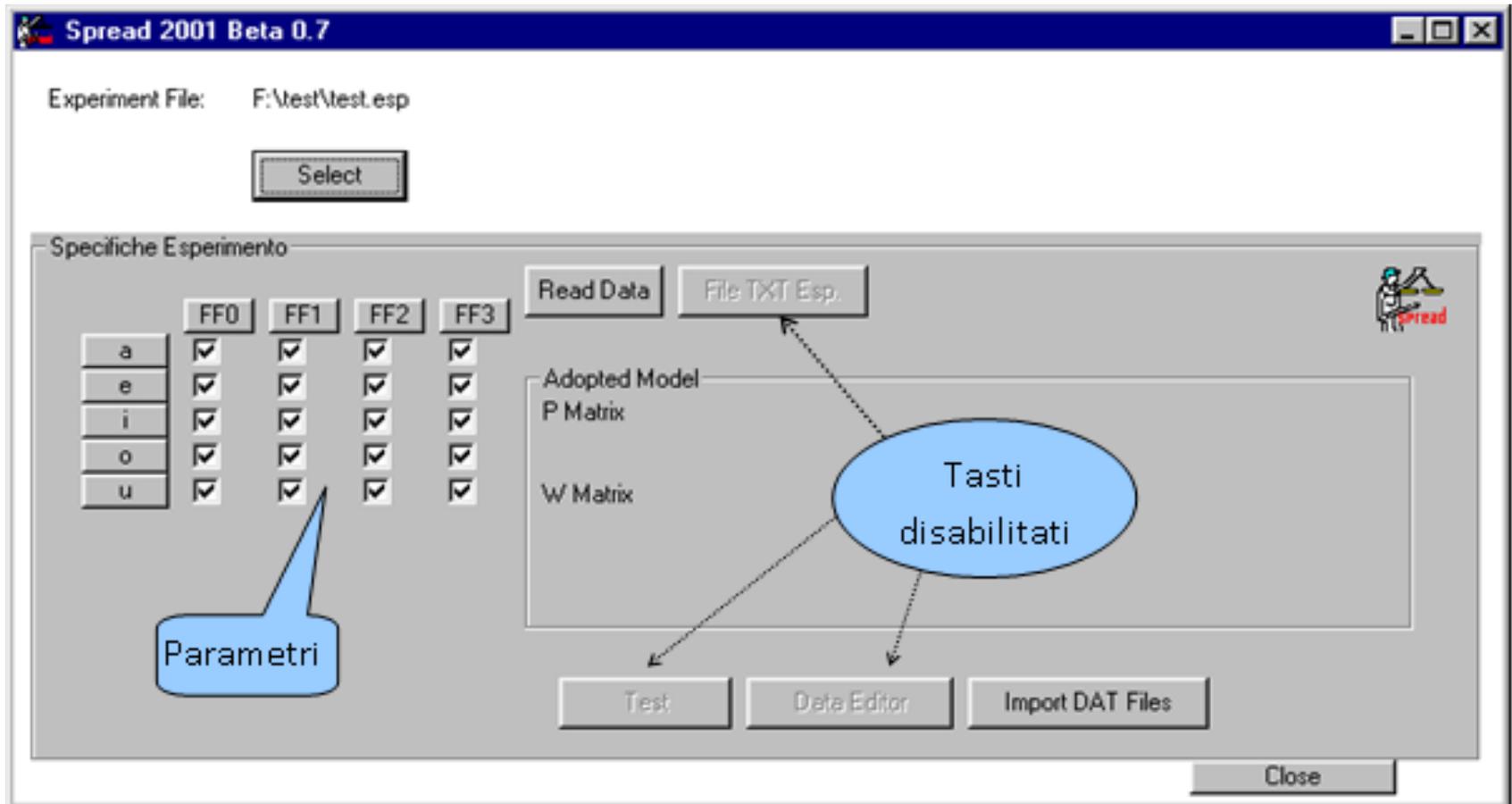
Ascolta singola

Ascolta in sequenza

- a
- e
- i
- o
- u

Aggiungi Cancella Ripristina **Aggiorna** Chiudi

Modulo SPREAD



Analisi e filtraggio dei dati

Criteri di Ricerca

Parlatore FF0 <= >= N.SD

File FF1

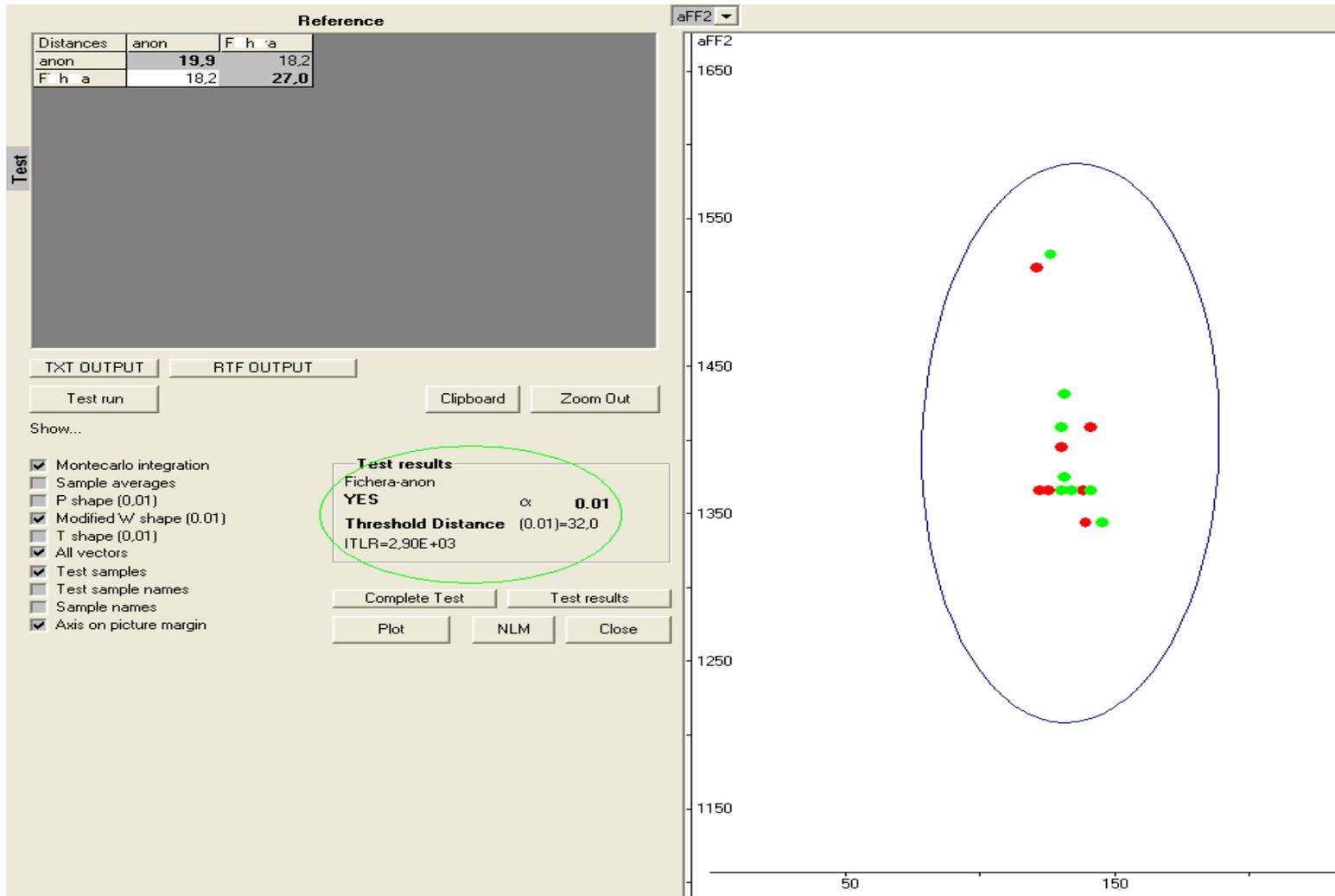
Fono FF2

Frase FF3

Sort Discendente

	parlatore	file	frase	fonema	#0	#1	#2	#3
▶	crl_int1	F:\test\PE	buonas-e-ra	e	160	449	1640	2656
	crl_int1	F:\test\PE	ciare-lli	e	161	488	1679	2636
	crl_int1	F:\test\PE	s-o-no	o	164	488	1289	2734
	crl_int1	F:\test\PE	lu-i-gi	i	142	351	2187	2519
	crl_int1	F:\test\PE	ciare-lli	e	152	527	1699	2636
	crl_int1	F:\test\PE	-e-h	e	166	527	1738	2734
	crl_int1	F:\test\PE	stam-a-tina	a	198	683	1406	2792
	crl_int1	F:\test\PE	cerc-a-va	a	156	605	1464	2480
	crl_int1	F:\test\PE	l-o-	o	172	507	1191	2675
	crl_int1	F:\test\PE	s-o-	o	162	468	957	2578
	crl_int1	F:\test\PE	radiom-o-bil	o	164	566	1035	2617
	crl_int1	F:\test\PE	-a-	a	135	605	1347	2675
	crl_int1	F:\test\PE	t-e-rno	e	128	527	1679	2714
	crl_int1	F:\test\PE	m-o-ro	o	148	527	1074	2675
	crl_int1	F:\test\PE	mor-o-	o	135	527	1171	2753
	crl_int1	F:\test\PE	trentan-o-ve	o	128	488	1054	2597
	crl_int1	F:\test\PE	pesc-a-ra	a	139	605	1386	2617
	crl_int1	F:\test\PE	vic-o-	o	196	507	1171	2695
	crl_int1	F:\test\PE	m-o-ro	o	151	546	1015	2656
	crl_int1	F:\test\PE	mor-o-	o	127	507	1054	2636

Test di compatibilità e calcolo dell'errore



Metodi Totalmente automatici

Rappresentano le più moderne opportunità

Tutti i passi sono automatici:

- estrazione dei parametri*
- trattamento statistico e catalogazione*
- calcolo della likelihood ratio*

L'operatore deve solo:

- inserire i tratti di segnale nel sistema*
- scegliere il database di riferimento*

Metodi Totalmente automatici

Il segnale, una volta inserito nel sistema, viene processato secondo il seguente schema:

- ✓ *realizzazione di una FFT ogni 10 ms, con finestra di 25-30 ms;*
- ✓ *estrazione automatica dei 15-30 coefficienti (MFCC...) per singola 'frame';*
- ✓ *catalogazione delle misure;*
- ✓ *realizzazione del modello intra-speaker;*
- ✓ *calcolo della likelihood ratio finale*

Metodi Totalmente automatici

Lavorano nel campo delle frequenze e del tempo (parametri MFCC e Δ MFCC)

Caratteristiche principali:

-velocità

-sensibilità al canale

-sensibilità al rapporto segnale/rumore

-discreta robustezza sulla lingua

Rapporto segnale/rumore - vincoli

Linguistic/ Phonetic	~ 10 dB
Semi Automatic	~ 12 dB
Full Automatic	~ 17 dB

Metodi automatici europei

LVIS (Loquendo)	<i>in fase di test presso il RaCIS</i>
IDENTIVOX	SPAGNA (Guardia Civil)
'Meuwly'	Svizzera (Zurigo-Losanna)
'SIS'	FRANCIA (Gendarmeria)
Trawl	RUSSIA

Loquendo Voice Investigation System (LVIS)

- ✓ *Speaker separation*
- ✓ *Language identification*
- ✓ *Gender detection*
- ✓ *Speaker identification*

*Loquendo SpA (ex spin-off TILab - Divisione
Tecnologie Vocali - oggi Nuance)*

Loquendo Voice Investigation System (LVIS)

Log Likelihood Ratio Calculation: for describing the likelihood of a voice matching a specific subject versus a reference population;

Tippett Plots: a visual validation of the Log Likelihood Ratio, clearly showing the reliability of the results;

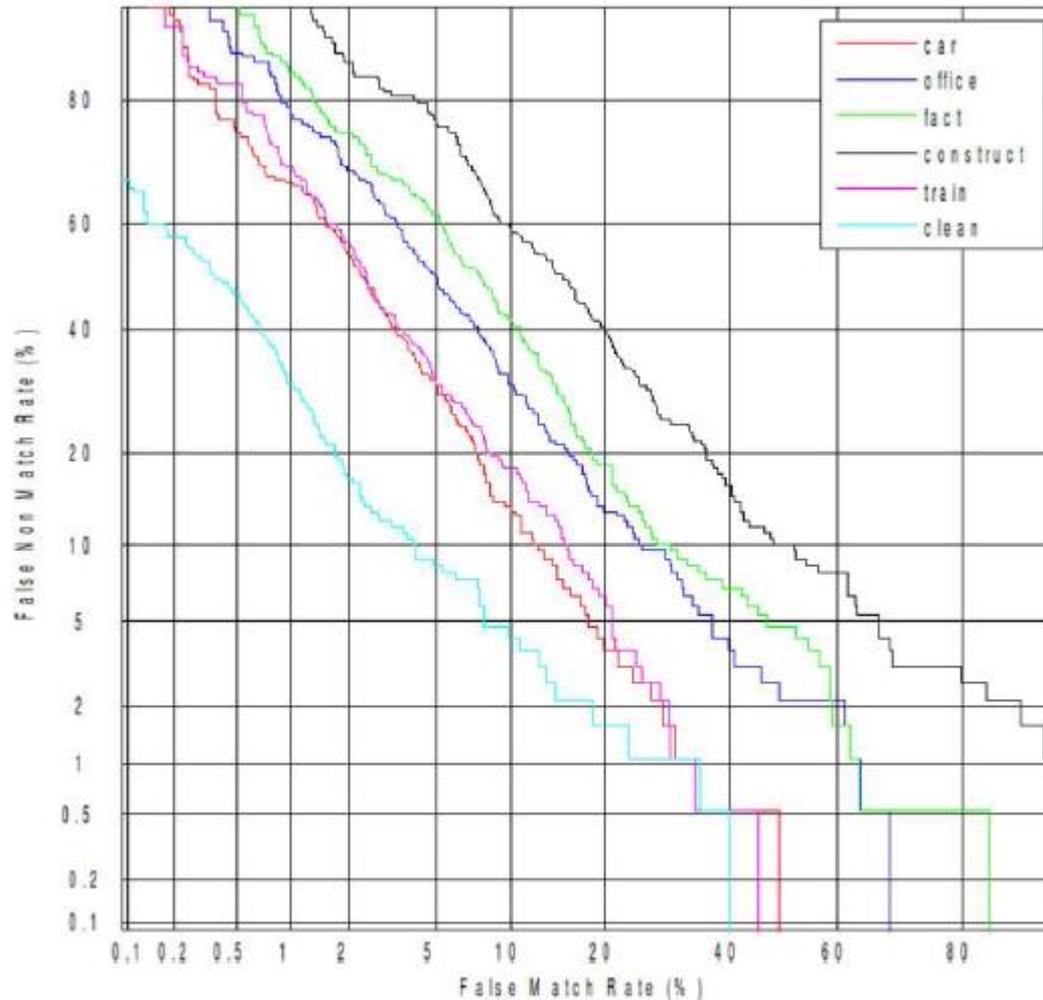
Reference Population Selection: More than 65 gender, channel, and language reference populations are now built and selectable by the operator;

Flexible Channel Normalization: operators can supply recordings from non-standard collection methods to recalibrate the system.

ALIZE/SpkDET

- ✓ Piattaforma open source
- ✓ Sviluppata inizialmente con fondi del Ministero della Ricerca scientifica francese
- ✓ Altri fondi LIA e BioSecure
- ✓ French National Research Agency: progetti BIOBIMO e MISTRAL
- ✓ Modelli statistici GMM/UBM (Gaussian Mixture Model / Universal Background Model) utilizzati per la stima della densità di probabilità multidimensionale e, quindi, poi per il calcolo di un matching score tramite un Likelihood test
- ✓ Parametri: 19MFCC + 19 Δ MFCC + E + Δ E

ALIZE/SpkDET: performance



(b) 10 dB

Vincoli sulla durata

Linguistic/ Phonetic	??
Semi Automatic	~ 12"
Full Automatic	~ 30" per l'anonimo ~ 2' per il sospettato

Robustezza sul canale

Linguistic/ Phonetic	+++
Semi Automatic	++/+++
Full Automatic	-

Robustezza sulla lingua (operatore non madrelingua)

Linguistic/ Phonetic	-/+ (dipende dal grado di con.)
Semi Automatic	+ / ++ (dipende dalla lingua)
Full Automatic	++ / +++

Velocità (comparazione a 2 voci)

Linguistic/ Phonetic	~ 1-2 Days
Semi Automatic	~ 1 Day
Full Automatic	~ 10 min

Parametri utilizzati

Linguistic/ Phonetic	Fonemi, semantica, prosodia, respiro... (‘imparati’ / caratteristiche time- dependent / frequency dependent)
Semi Automatic	F0, Formanti (Frequency domain)
Full Automatic	F0, MFCC, DMFCC, DDMFCC (Tutti i domini)

Risposte

Linguistic/ Phonetic	Scala verbale – Soggettiva/qualitativa
Semi Automatic	P.F.A./P.F.R. o Lr
Full Automatic	Lr / Tippet Plot

Performance

Linguistic/ Phonetic	Alta ?? (non esiste una reale validazione)
Semi Automatic	EER < 1 %
Full Automatic	EER 1%-15% (dipende dal tipo di canale)

Casi rigettati per i limiti del sistema (il segnale non rispetta i vincoli)

Linguistic/ Phonetic	1% - 5%
Semi Automatic	~ 20%
Full Automatic	30% - 60%

E' accettato in dibattimento ?

Linguistic/ Phonetic	generalmente sì <i>(non in U.S.A. - caso Daubert)</i>
Semi Automatic	Sì
Full Automatic	?? (Sono da considerarsi alla stregua di AFIS e IBIS, cioè utili per per analisi preventive)

Conclusioni

- ✓ *Nessuno dei 3 metodi può definirsi 'il migliore' in senso assoluto, avendo differenti limiti*
- ✓ *La richiesta dominante delle Scienze Forensi è 'fornire risposte quantitative'*
- ✓ *I sistemi linguistico-fonetico, pur qualitativi, sono però finora accettati quasi ovunque*
- ✓ *I metodi Semi-automatici appaiono i più versatili, potendo combinarsi con le variabili time-dependent usate dai fonetisti*

Conclusioni

L'accertamento completo dovrebbe essere condotto con la combinazione di tutte le tecniche (il 'metodo dei metodi') in modo da poter garantire applicabilità e performance superiori in qualsiasi condizione.

Caso: comparazione di
voce alterata

Il caso

Svizzera, Canton Ticino

Un ragazzo uccide la nonna con un lama da 30 cm e aggredisce il padre. Viene arrestato immediatamente.

Nella camera del ragazzo viene trovato un video di 6 minuti nel quale un soggetto maschile, totalmente incappucciato, con una lama in mano, legge un discorso delirante, con voce alterata, nel quale spiega che bisogna compiere 'grandi azioni' prima di suicidarsi.

Il P.M. vuole sapere se la voce è dell'assassino oppure sia un'azione di plagio sul ragazzo

Il caso

La Polizia Scientifica del Canton Ticino non opera comparazioni di voce

L'Istituto Centrale di Zurigo non è in grado di trattare il caso per via della lingua e perché, utilizzando sistemi automatici, la voce alterata rappresenta un limite.

Tuttavia la voce naturale del ragazzo, all'ascolto, appare molto simile.

Metodo di lavoro

- ✓ *Effettuazione del saggio fonico secondo le modalità standard*
- ✓ *Approccio linguistico (laddove possibile)*
- ✓ *Approccio strumentale (laddove possibile)*

Approccio linguistico

Il parlatore anonimo si esprime con accento ticinese (simile al lombardo), mostrando una chiara 'r' uvulare.

Nel messaggio, tuttavia, si evidenziano 6 parole pronunciate con accento differente da quello locale:

*dèa – Atèna – appartèngo – sènza – intèndere -
sèmpre*

La dimostrazione che l'accento fosse effettivamente diverso nella popolazione ticinese è stata ottenuta intervistando una quindicina di persone del luogo, invitate a leggere frasi contenenti queste parole.

Il saggio fonico

Si effettua in carcere, in modalità ortofonica, e si propone al sospettato, dopo un colloquio informale, di ripetere frasi contenenti le parole evidenziate, per verificarne le caratteristiche fonetiche, nonché di rileggere il messaggio pronunciato nel video

Approccio strumentale

Dalla voce del sospettato e dell'anonimo sono stati estratti 7-8 campioni per ogni vocale e da questi sono state misurate le formanti.

Poiché la F_0 (frequenza fondamentale) è indubbiamente modificata dall'alterazione volontaria, non è stata presa in considerazione (nel saggio la voce era naturale)

*L'esito del test è stato sorprendentemente **positivo.***

Reference

iff1

Distances	Anonimo	NICOLAS
Anonimo	18,0	10,3
NICOLAS	10,3	21,1

Test

TXT OUTPUT RTF OUTPUT

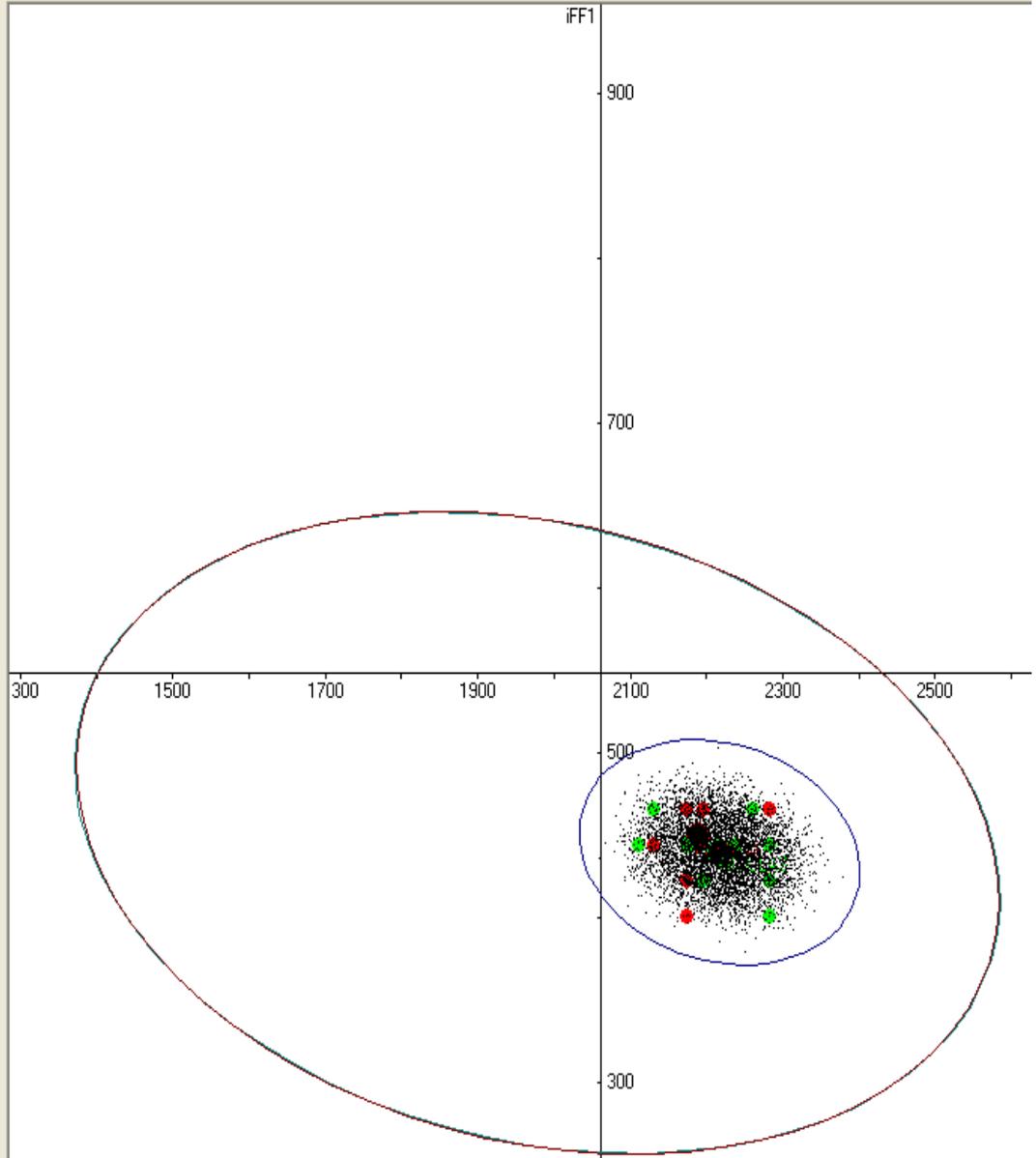
Test run Clipboard Zoom Out

Show...

- Montecarlo integration
- Sample averages
- P shape (0,01)
- Modified W shape (0,01)
- T shape (0,01)
- All vectors
- Test samples
- Test sample names
- Sample names
- Axis on picture margin

Test results
NICOLAS-Anonimo
YES α **0.01**
Threshold Distance (0,01)=26,2
ITLR=2,01E+03

Complete Test Test results
Plot NLM Close



Attendibilità delle impronte vocali

- **L'esame comparativo tramite campione di voce non è in grado di fornire un risultato certo ($FA=10^{-9}$) !**

ERRORE	IMPRONTE VOCALI (segnale telefonico di buona qualità)
Falsa Esclusione (FE)	3 % (3 su 100)
Falsa Attribuzione (FA)	0.4% (4 su 1000)

L'attendibilità (certezza dell'attribuzione) dipende da:

- Durata del segnale fonico (> 15 sec.)
- Numero e bontà dei parametri biometrici
- Qualità del campione di voce (SNR > 10 dB)
- Dalla voce in esame (comune o particolare)

Degradazioni del segnale fonico

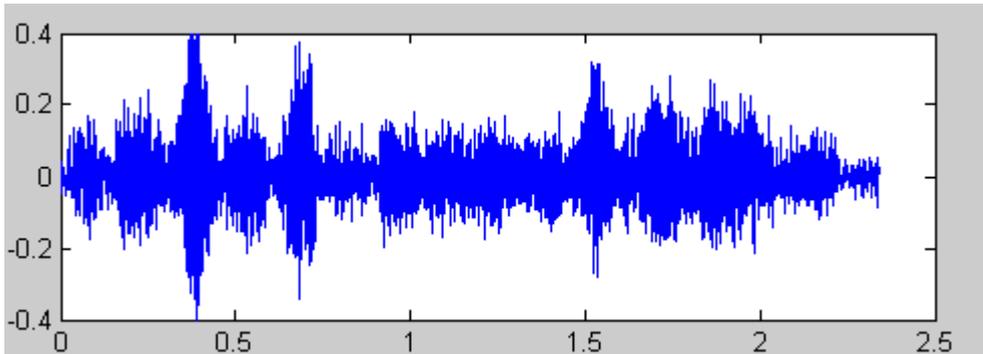
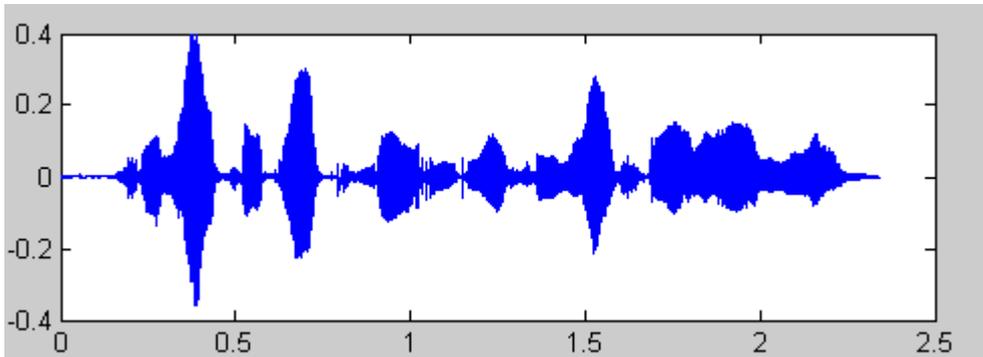
- APPARATI DI REGISTRAZIONE ++
- RUMORE (AMBIENTALE o di CANALE) +++++
- EFFETTO “LOMBARD” +++
- MICROVARIABILITA' TEMPORALE DELLA VOCE +
- RIVERBERI +++++
- REGISTRATORE MAGNETICO +
- ALGORITMI DI COMPRESSIONE +
- ALGORITMI DI SOPPRESSIONE DEL SILENZIO +
- IL CANALE DI COMUNICAZIONE +++

Riverbero

Riflessione dell'onda sonora da parte di un ostacolo posto davanti alla fonte sonora. Il riverbero ha aspetti negativi, come il rischio di mascheramento delle sillabe del parlato

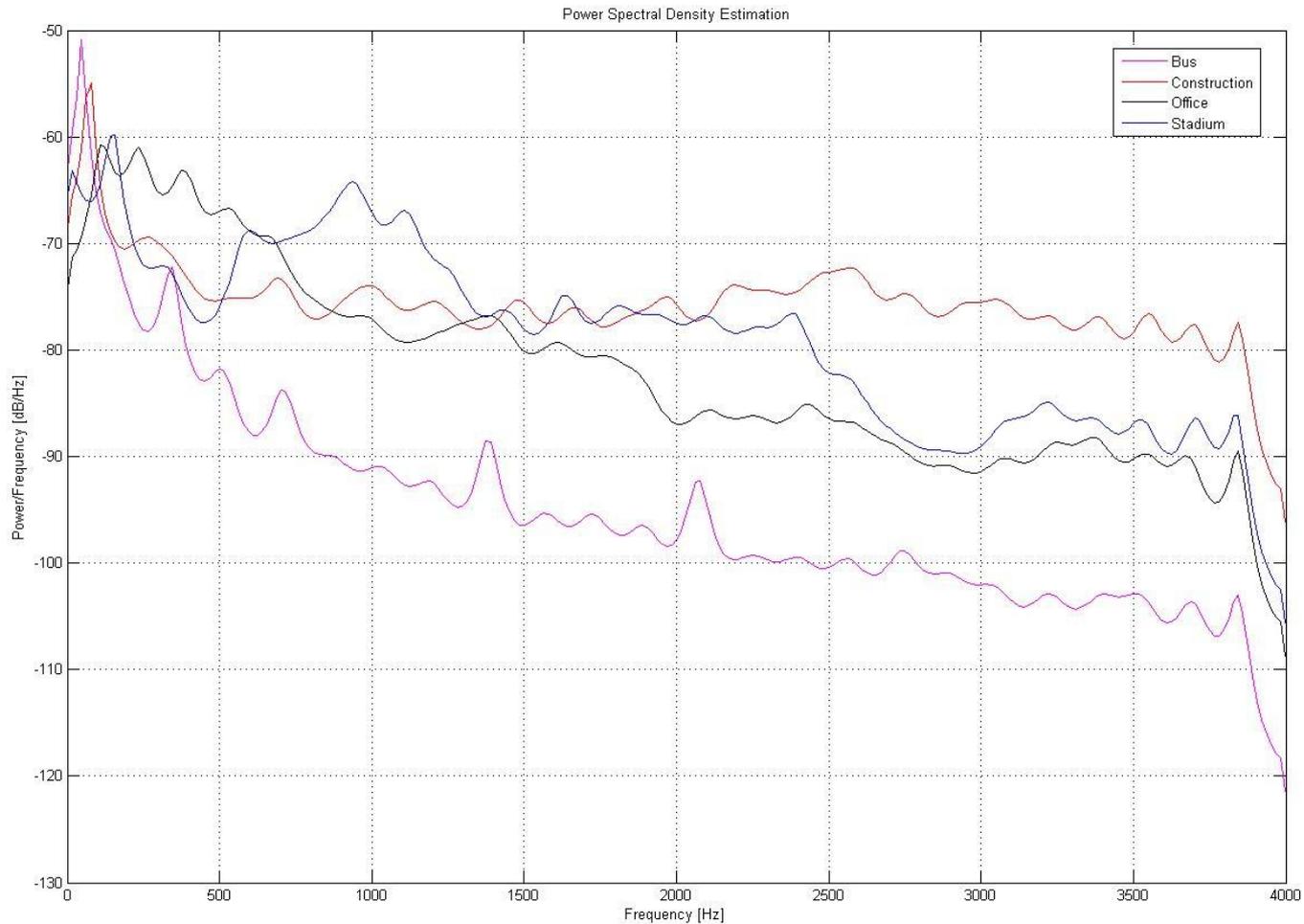


Effetto del rumore ambientale



SNR {
Clean
20 dB
15 dB
10 dB
05 dB
00 dB

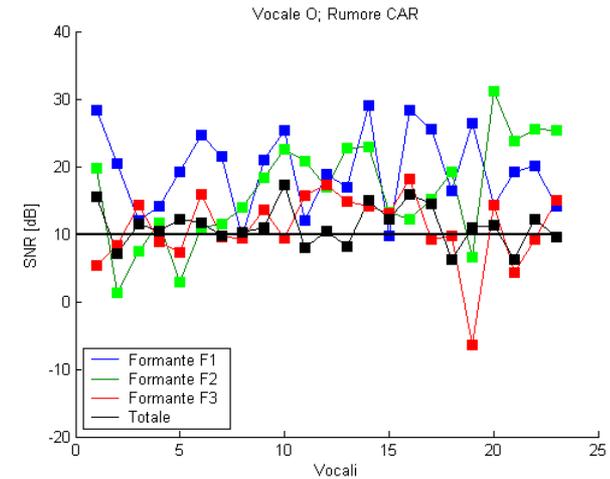
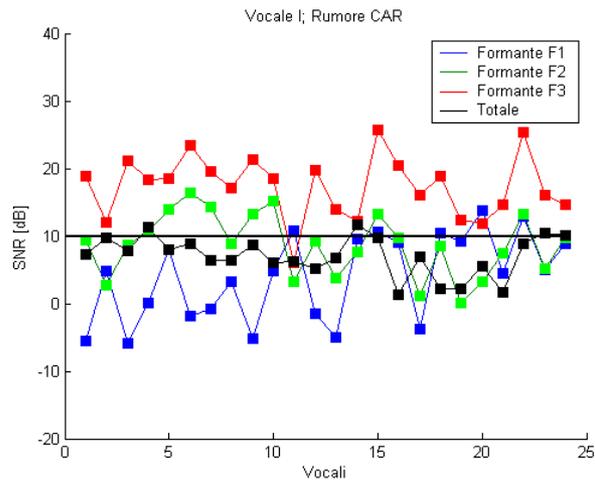
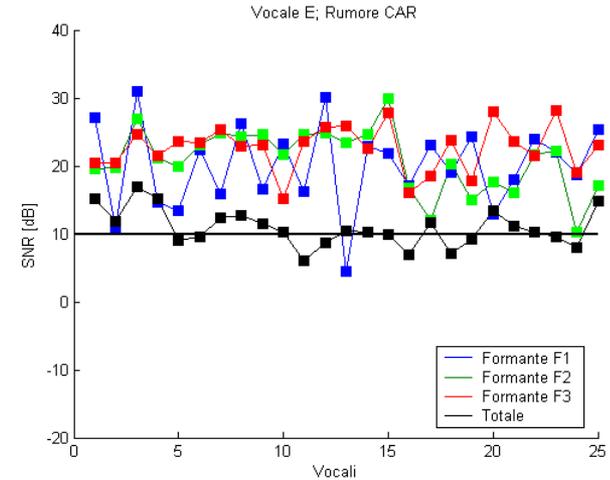
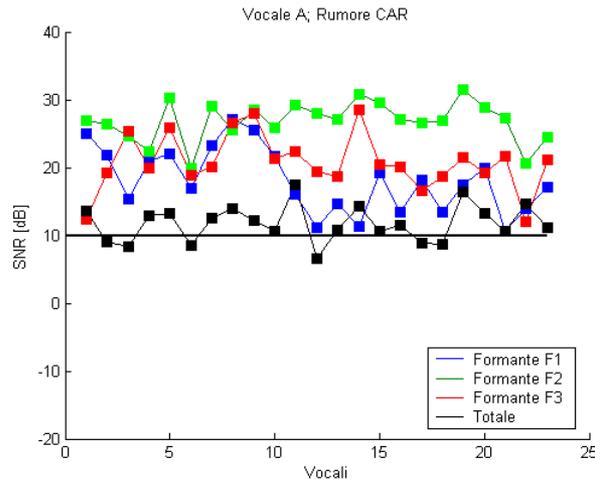
Caratterizzazione del rumore ambientale



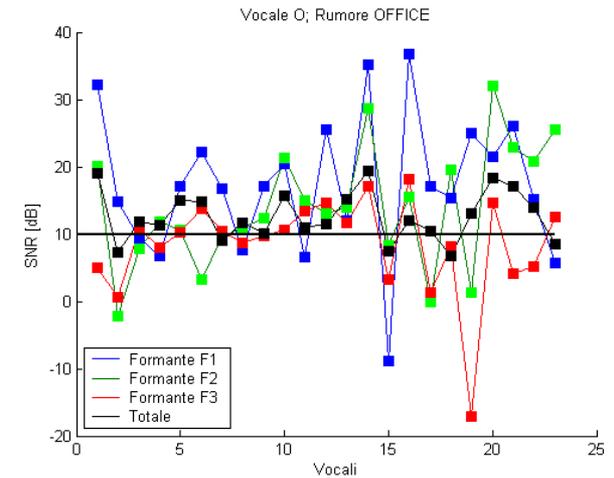
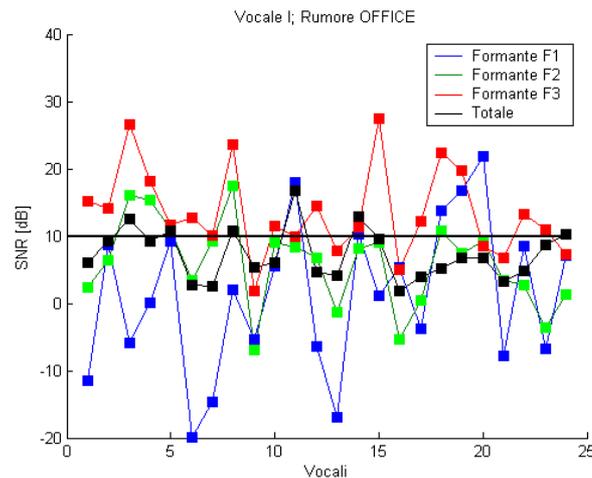
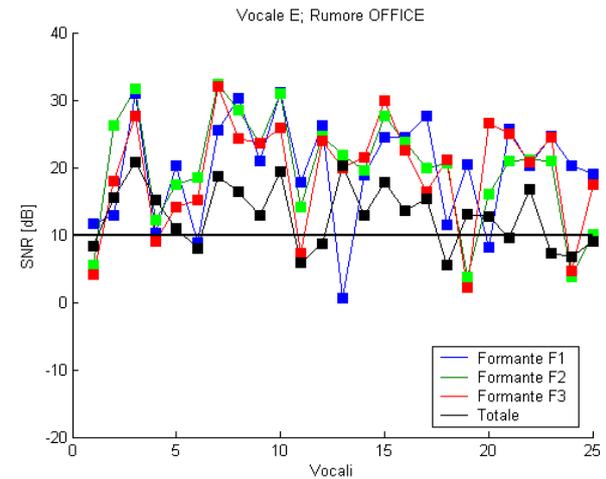
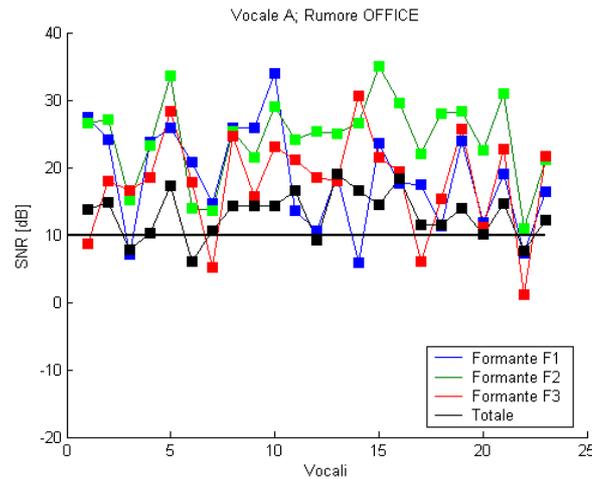
Sottobande indici biometrici

Vocale \ Formante	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
A	478÷752	1154÷2700	2141÷3700
E	343÷700	1480÷2170	2191÷2750
I	220÷434	1746÷2450	2451÷3100
O	379÷715	720÷1202	2000÷2800

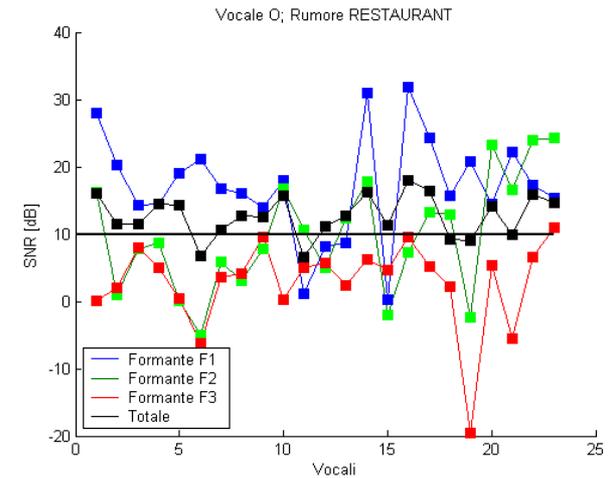
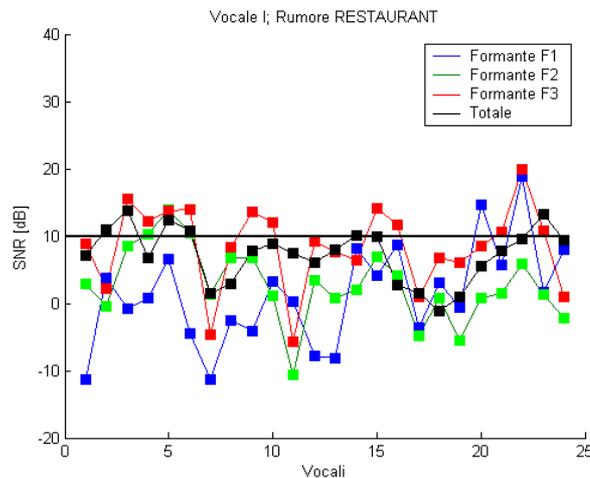
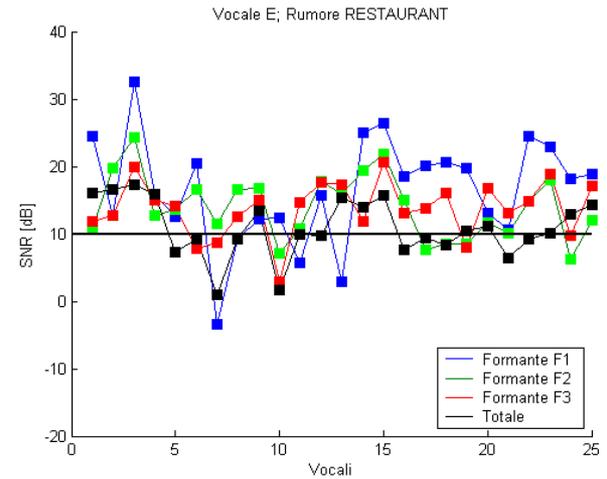
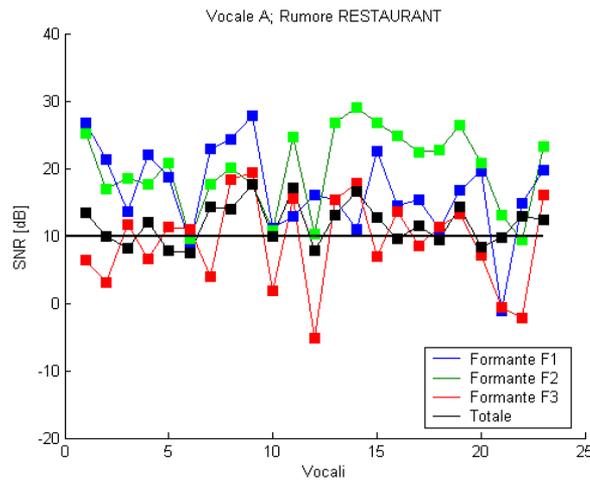
SNR nelle sottobande F1, F2, F3: esempi con rumore CAR 10 dB



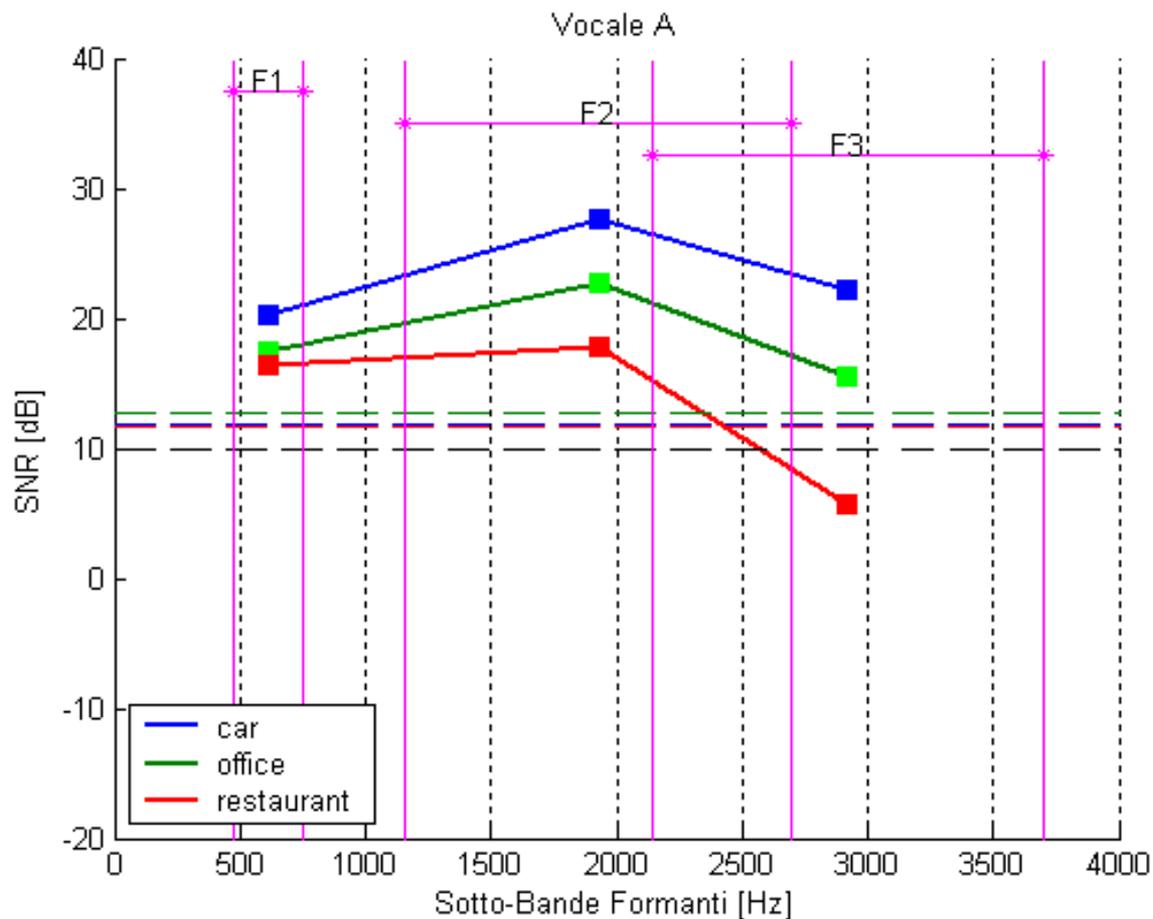
SNR nelle sottobande F1, F2, F3: esempi con rumore OFFICE 10 dB



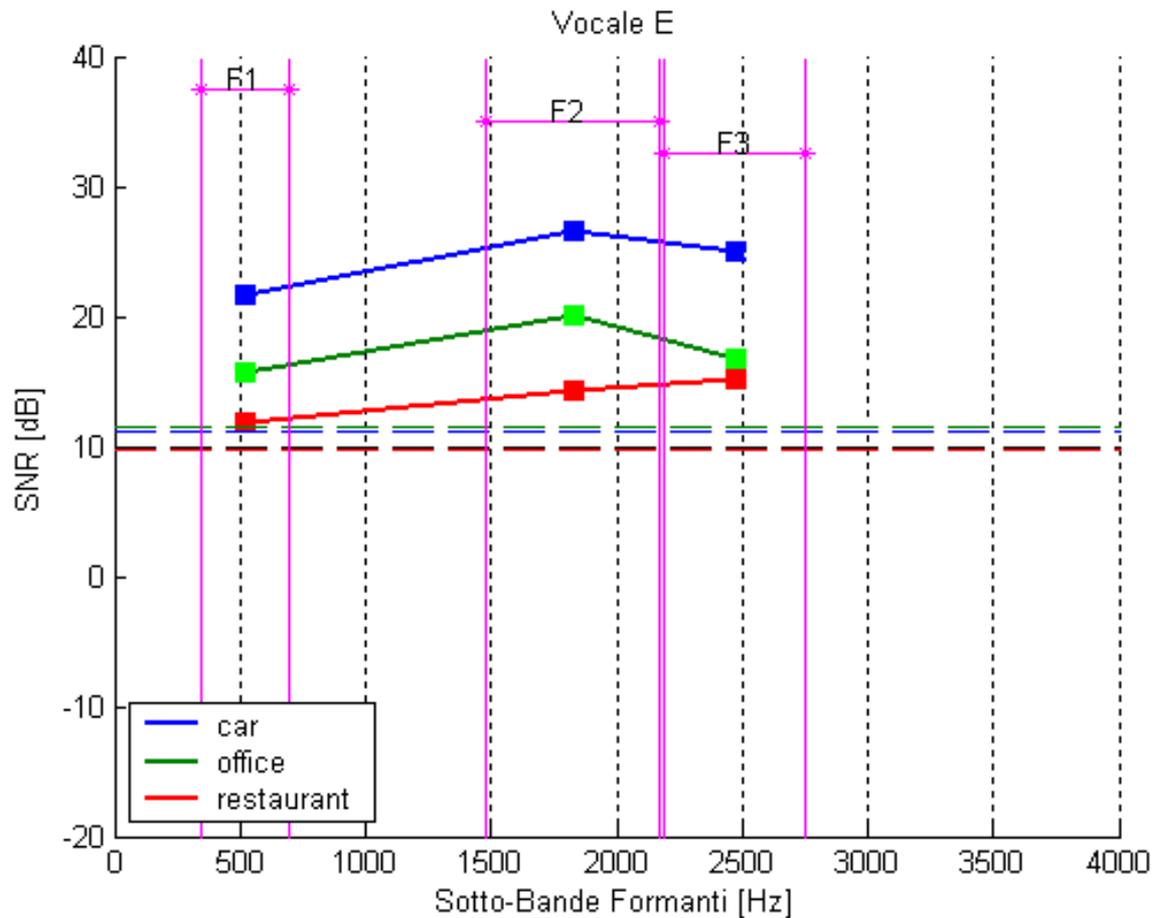
SNR nelle sottobande F1, F2, F3: esempi con rumore REST 10 dB



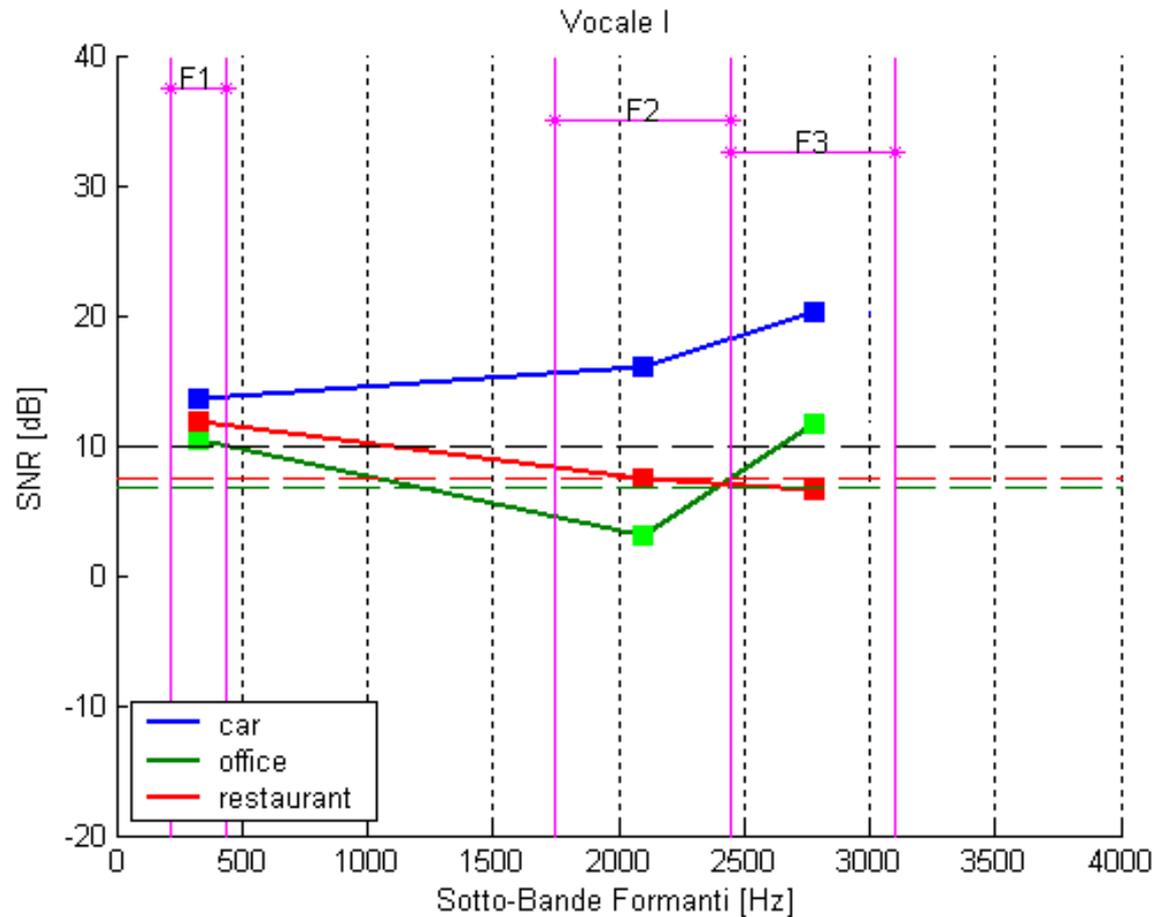
SNR nelle sottobande F1, F2, F3: medie vocale A



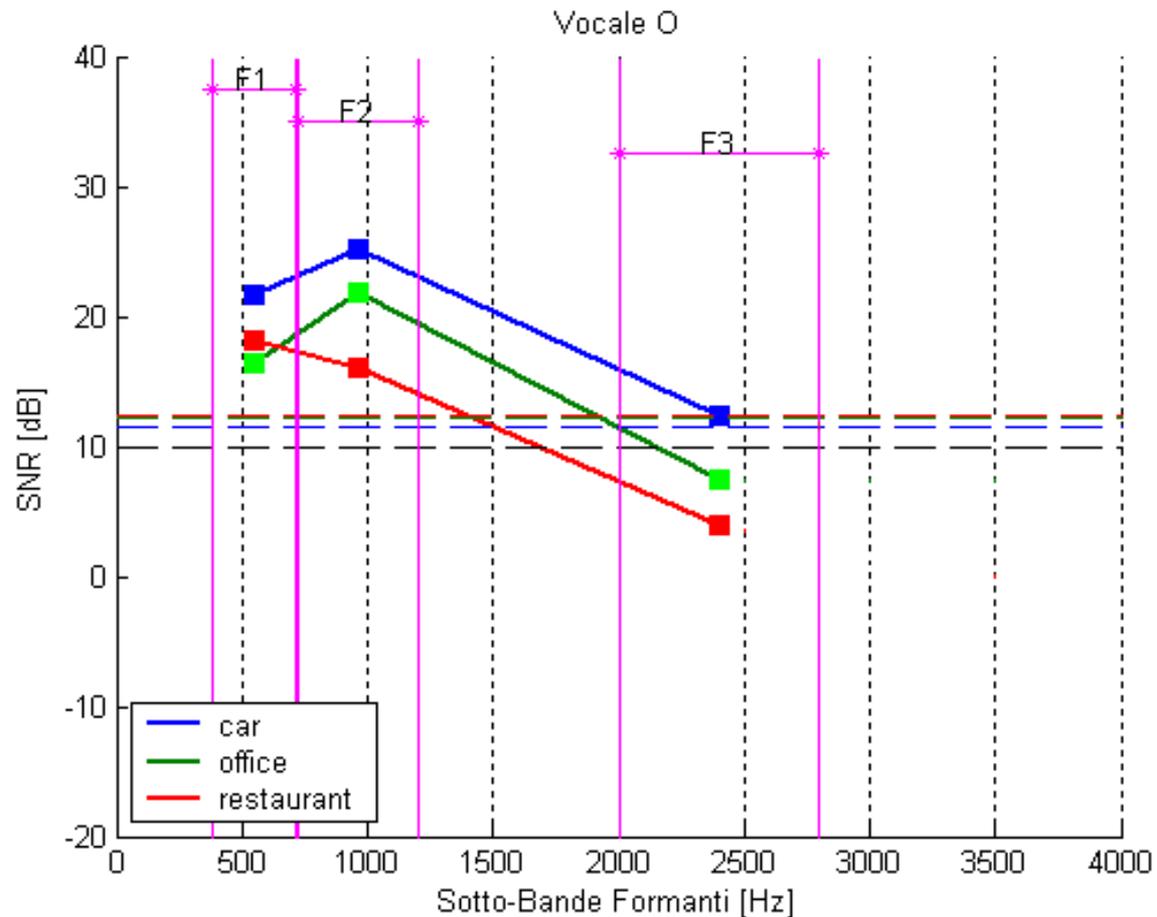
SNR nelle sottobande F1, F2, F3: medie vocale E



SNR nelle sottobande F1, F2, F3: medie vocale I



SNR nelle sottobande F1, F2, F3: medie vocale O



Acquisizione del saggio fonico

- Il soggetto deve essere iscritto nel registro degli indagati
- L'avvocato deve essere invitato a presenziare
- Le apparecchiature di acquisizione devono essere simili a quelle utilizzate per registrare la voce anonima
- Preventivamente bisogna stilare un testo contenente le frasi pronunciate dall'anonimo

Acquisizione del saggio fonico

- FASE 1: Colloquio introduttivo con richiesta generalità
- FASE 2: L'indagato deve ripetere le frasi contenute nel testo, sotto dettatura
- FASE 3: Colloquio informale con l'indagato

Al termine: Redazione del verbale

Nomina del perito

(art. 221 c.p.p.)

1. “Il Giudice nomina il perito scegliendolo tra gli iscritti negli appositi albi o tra persone fornite di particolare competenza nella specifica disciplina...”
2. “Il Giudice affida l’espletamento della perizia a più persone quando le indagini e le valutazioni risultano di notevole complessità ovvero richiedono distinte conoscenze in differenti discipline”.
3. “Il perito ha l’obbligo di prestare il suo ufficio, salvo che ricorra uno dei motivi di astensione previsti dall’articolo 36”.

Conclusioni

➤ LE PERIZIE BASATE SU METODI PARAMETRICI:

- ✓ PRESENTANO UNA MAGGIORE COMPLESSITA'
- ✓ RICHIEDONO COMPETENZE SPECIFICHE SU DIVERSI AMBITI DISCIPLINARI
- ✓ PIU' SENSIBILI ALLE DEGRADAZIONI DEL SEGNALE
- ✓ RICHIEDONO UNA DURATA MINIMA DI SEGNALE
- ✓ CALCOLANO IN MANIERA OGGETTIVA LA PROBABILITA' DI IDENTITA' ATTRAVERSO TEST STATISTICI

... anche la Biometria su Cloud .. ?!



**"I'm sorry, but someone else with
that identity is already here."**