



# **LINEE GUIDA PER LE ANALISI SUL MATERIALE SONORO INTERCETTATO**

a cura dell'Osservatorio sulla Linguistica Forense (OLF)

Associazione Italiana Scienze della Voce (AISV)

## Sommario

Introduzione.....	2
Acquisizione, registrazione e conservazione del documento sonoro .....	2
1. Registrazione, originalità e conservazione del segnale .....	2
1.1 Requisiti tecnici minimi del file originale .....	3
1.2 Catena di custodia.....	4
2. Analisi preliminare del segnale .....	5
3. Trascrizione di intercettazioni .....	5
3.1 Parole/frasi contestate (disputed utterances) .....	7
3.2 Relazione tecnica.....	8
Caratterizzazione e identificazione del parlante .....	10
4. Caratterizzazione del parlante (o analisi linguistica).....	10
5. Comparazione forense della voce.....	11
5.1 Metodi di comparazione .....	12
5.2 Integrazione dei metodi .....	14
5.3 Il rapporto di verosimiglianza (Likelihood Ratio, LR),.....	14
5.4 La trasparenza è più importante della conclusione.....	15
5.5 Relazione tecnica.....	16
Miglioramento del segnale audio .....	18
6. Il filtraggio.....	18
6.1 Limiti del miglioramento del segnale .....	18
6.2 Relazione tecnica.....	19
7. Bibliografia.....	21

## Introduzione

Le presenti linee guida sono rivolte al mondo del diritto, e in particolare ad avvocati, magistrati e operatori giudiziari che, a vario titolo, si trovano a interagire con analisi fonetiche, linguistiche e forensi.

La loro finalità è offrire un quadro chiaro, aggiornato e scientificamente fondato delle competenze, dei metodi e dei limiti che caratterizzano il lavoro degli specialisti del settore, così da favorire un dialogo più consapevole tra comunità scientifica e comunità giuridica.

AISV (Associazione Italiana di Scienze della Voce) e OLF (Osservatorio sulla Linguistica Forense) hanno ritenuto necessario formalizzare queste linee guida per rispondere a un'esigenza ormai evidente: **colmare la distanza informativa** che spesso separa il sapere scientifico dalle pratiche giudiziarie, e prevenire fraintendimenti sull'uso, sul valore probatorio e sulle condizioni di applicabilità delle analisi fonetiche e linguistiche in ambito forense.

L'obiettivo non è quello di definire standard rigidi o prescrittivi, ma di fornire criteri condivisi, riconosciuti dalla comunità scientifica, che possano orientare magistrati e avvocati nella valutazione delle consulenze tecniche, nella scelta dei professionisti e nella corretta interpretazione dei risultati.

Le linee guida non hanno carattere vincolante e non intendono sostituire né la discrezionalità del giudice né le norme processuali vigenti. Esse rappresentano invece un riferimento autorevole, fondato sulle migliori pratiche scientifiche, utile a garantire trasparenza metodologica, qualità delle consulenze e tutela delle parti coinvolte nei procedimenti giudiziari.

Le presenti linee guida costituiscono un documento dinamico, soggetto a revisione periodica da parte dell'OLF e di AISV. Al fine di garantire l'aggiornamento scientifico e metodologico, la revisione avviene con cadenza **massima semestrale** (o secondo il periodo che il coordinamento riterrà opportuno). Ogni modifica, integrazione o aggiornamento sarà riportata in una **Revision History**, allegata in appendice o resa disponibile in forma separata, così da assicurare trasparenza sull'evoluzione del documento e sulle versioni ufficialmente approvate.

## Acquisizione, registrazione e conservazione del documento sonoro

### 1. Registrazione, originalità e conservazione del segnale

Secondo un orientamento consolidato della giurisprudenza di legittimità, la prova è costituita dalla registrazione sonora quale fonte primaria; la trascrizione, così come ogni altra forma di rappresentazione del contenuto, non può sostituire l'ascolto diretto<sup>1</sup>. Ne consegue che, prima di qualsiasi attività di analisi, è necessario procedere alla verifica dell'originalità, dell'integrità e della qualità tecnica del dato audio, a prescindere dalle modalità con cui esso è stato acquisito o conservato. L'esperto audio-forense<sup>2</sup> è pertanto chiamato ad accertare che il materiale sia idoneo alle diverse tipologie di esame applicabili al segnale — non limitate alla sola trascrizione — e che presenti caratteristiche tali da consentirne un utilizzo affidabile in ambito processuale.

I file audio, relativi alle intercettazioni e destinati all'utilizzo probatorio, dovrebbero essere prodotti, gestiti e consegnati in conformità alle *Linee Guida sulla formazione, gestione e conservazione dei documenti informatici* (sezione Formati di file e riversamento, Allegato 2) emanate da AgID (Agenzia

---

<sup>1</sup> Il principio di diritto per il quale la trascrizione è un "mero ausilio", mentre la prova è costituita dalla registrazione è affermato, ex multis, da Cassazione, Sez. IV Pen., 23.9.2021, n. 37613 e da Cass., Sez. I Pen., 3.5.2019, n. 41632. In merito al principio di diritto per il quale la registrazione è l'unica fonte primaria di prova, vedi Cass. Pen., Sez. V, 6 aprile 2011, n. 15006. Il principio di diritto per il quale la trascrizione non può sostituire l'ascolto diretto del supporto originale è affermato, tra le altre, da Cass., Sez. VI Pen., 4.2.2025, n. 15772 e da Cass., Sez. II Pen., 17.3.2023, n. 27089.

<sup>2</sup> In queste Linee guida, con l'espressione "esperto audio-forense" si intende designare, in senso ampio e funzionale, il perito quale ausiliario del giudice nel processo penale, i consulenti tecnici delle parti, nonché gli operatori di polizia giudiziaria, ossia tutti i soggetti che, a vario titolo, intervengono sul segnale sonoro.

per l'Italia Digitale)<sup>3</sup>, privilegiando formati **aperti, non compressi, stabili, ampiamente documentati** e idonei alla conservazione a lungo termine.

Esempi di formati rispondenti a tali criteri sono il formato **WAV** (PCM non compresso), **FLAC** (Free Lossless Audio Codec) e **AIFF** (Audio Interchange File Format). Non risultano invece idonei ai fini conservativi e peritali i formati con perdita di informazione (lossy), quali **MP3, AAC, AMR, OGG** (in modalità lossy), **WMA**, nonché, più in generale, i formati fortemente compressi o proprietari, chiusi e non adeguatamente documentati. Tali formati possono essere eventualmente impiegati per finalità operative o come copie di lavoro, ma non dovrebbero essere assunti quale supporto principale ai fini probatori in ambito peritale.

### ***1.1 Requisiti tecnici minimi del file originale***

Per garantire la qualità del materiale audio destinato all'analisi forense, si raccomanda che i file originali rispettino i seguenti parametri tecnici minimi:

#### **Frequenza di campionamento**

- $\geq 44.1$  kHz per registrazioni ambientali
- $\geq 16$  kHz per registrazioni telefoniche (banda telefonica standard)

#### **Profondità in bit**

- $\geq 16$  bit (preferibile 24 bit per nuove acquisizioni, ove tecnicamente possibile)

#### **Integrità del file**

- Calcolo e consegna dell'hash SHA-256 del file originale
- L'hash deve essere riportato nella documentazione di repertazione e nella relazione peritale

Per garantire coerenza metodologica, confrontabilità dei materiali e piena replicabilità delle analisi fonetico-acustiche, l'OLF raccomanda fortemente l'adozione, da parte delle ditte incaricate delle intercettazioni e delle Procure della Repubblica, di formati tecnici di registrazione uniformi su tutto il territorio nazionale. Questa prescrizione riprende il principio già vigente nell'epoca delle registrazioni analogiche, quando:

- Le *intercettazioni telefoniche* venivano registrate ad una velocità di 2.38 cm/s,
- E le *intercettazioni ambientali* venivano registrate ad una velocità di 4.75 cm/s,

garantendo così omogeneità tecnica, comparabilità e affidabilità probatoria.

Scientificamente e tecnicamente, l'uniformità dei parametri di registrazione è necessaria perché:

1. **Assicura la confrontabilità dei segnali.** Registrazioni prodotte con parametri diversi (frequenza di campionamento<sup>4</sup>, bit depth, compressione) non sono confrontabili né tra loro né con eventuali campioni di riferimento.

---

<sup>3</sup> Il documento è reperibile al link [https://www.agid.gov.it/sites/agid/files/2024-05/linee\\_guida\\_sul\\_documento\\_informatico.pdf](https://www.agid.gov.it/sites/agid/files/2024-05/linee_guida_sul_documento_informatico.pdf)

<sup>4</sup> Ad esempio, l'adozione di frequenze di campionamento uniformi assolverebbe oggi alla stessa funzione di standardizzazione tecnica che un tempo era affidata alle velocità di registrazione

2. **Garantisce la replicabilità dell'analisi.** La fonetica forense richiede misure acustiche precise (formanti, F0, durata, ecc.). Parametri diversi alterano tali misure, compromettendo la validità scientifica.
3. **Evita disparità territoriali.** Uffici di Procura diversi non possono seguire livelli di qualità differenti: la prova deve essere uguale per tutti, indipendentemente dal luogo in cui è stata raccolta o dal fornitore (ditta assegnataria dell'intercettazione).
4. **Previene errori giudiziari.** RegISTRAZIONI di bassa qualità aumentano il rischio di incomprensioni, trascrizioni errate, interpretazioni arbitrarie, impossibilità di effettuare analisi fonetiche.
5. **Rende possibile la validazione dei sistemi automatici.** I metodi basati su LR (Likelihood Ratio) richiedono dataset omogenei.

La necessità di uniformare i formati tecnici di registrazione trova fondamento nella normativa vigente. L'*Archivio Unico delle Intercettazioni*, introdotto dal D.L. 30 dicembre 2019 n. 161, convertito nella Legge 8 febbraio 2020 n. 7, e successivamente integrato dalla Riforma Cartabia prevede che: *“Presso ogni Procura della Repubblica è istituito un archivio riservato per la conservazione dei verbali e delle registrazioni delle intercettazioni, accessibile esclusivamente al pubblico ministero e al giudice.”* e che *“Le modalità di conservazione devono garantire sicurezza, integrità, immutabilità e tracciabilità dei dati.”* Questi principi impongono che i file audio siano omogenei nei formati e nei parametri tecnici; le Procure possano gestirli in modo uniforme; la conservazione digitale sia standardizzata; la catena di custodia sia verificabile e non alterabile e infine non siano previste differenze tecniche dovute al fornitore (ditta privata che procede all'intercettazione).

La *Riforma Cartabia (D.Lgs. 150/2022)* ha introdotto una disciplina più rigorosa sulla gestione dei dati processuali, sugli obblighi di sicurezza, di integrità e tracciabilità, rafforzando la responsabilità delle Procure nella gestione delle intercettazioni e richiedendo una maggiore attenzione alla conservazione digitale certificata.

L'uniformità dei formati e la corretta catena di custodia sono quindi condizioni necessarie per l'interoperabilità con l'Archivio Unico, la conservazione conforme agli standard AgID, la tutela dell'integrità probatoria, la parità delle parti nel processo, la tutela del diritto di difesa e i principi del giusto processo (art. 111 Cost.). Pertanto, l'Archivio Unico richiede standard tecnici uniformi per garantire una corretta interoperabilità tra Procure, ditte fornitrici e sistemi di conservazione. L'uniformità tecnica garantisce — come avveniva con le registrazioni analogiche effettuate direttamente in Procura dagli operatori di polizia giudiziaria — una qualità costante del materiale, confronti affidabili, standard nazionali omogenei e l'assenza di discrepanze tra fornitori. Si tratta di condizioni essenziali affinché il materiale audio mantenga le caratteristiche di coerenza e affidabilità richieste a una prova in ambito giudiziario.

## **1.2 Catena di custodia**

Considerato che la prova dal punto di vista giuridico è tradizionalmente la cd. “bobina”, oggi il file audio, la sua originalità deve essere documentata e affidabile. Quindi molta importanza assume la documentazione della catena di custodia. Ogni passaggio di mano del file (autorità giudiziaria, polizia giudiziaria, periti, consulenti) deve essere documentato in modo puntuale, indicando data, ora, soggetto e motivazione. Il file originale non deve essere mai elaborato o modificato: tutte le attività analitiche devono essere svolte esclusivamente su copie di lavoro derivate. La catena di custodia deve essere continua, verificabile e non interrotta, in conformità ai principi di integrità e genuinità della prova. Le copie di lavoro devono essere chiaramente etichettate come tali (nome file, versione, data di creazione, autore della copia), ogni trasformazione (conversione di formato, normalizzazione,

filtraggio, segmentazione) deve essere documentata in modo completo, così da garantire la piena replicabilità del processo analitico. Le copie di lavoro non possono sostituire il file originale ai fini probatori.

## 2. Analisi preliminare del segnale

Prima di procedere a qualsiasi attività di trascrizione, interpretazione o analisi fonetico-acustica, si consiglia di effettuare una valutazione preliminare della qualità del segnale. Tale valutazione dovrebbe essere documentata nella relazione tecnica del perito, del consulente ovvero nella nota tecnica degli operatori di polizia giudiziaria che procedono alla trascrizione e dovrebbe includere l'analisi del rapporto segnale/rumore (SNR), la stima dell'intelligibilità e della riconoscibilità dei segmenti vocalici e consonantici.

Dette analisi dovrebbero essere condotte almeno sui segmenti di segnale che sono rilevanti per il trattamento del caso, e dovrebbero essere corredate da una chiara indicazione del timing, cioè della collocazione dei segmenti sulla linea temporale della traccia audio. Ciò è particolarmente rilevante nelle intercettazioni ambientali, dove la qualità della registrazione non è uniforme lungo tutto il file: all'interno della stessa registrazione possono infatti coesistere segmenti con elevata intelligibilità e valori di SNR molto alti, accanto a porzioni in cui il segnale è quasi assente o completamente mascherato dal rumore.

Al fine di consentire la comparazione delle eventuali misure effettuate dalle parti, le valutazioni del rapporto **SNR** devono essere eseguite secondo una delle procedure riconosciute e comunemente accettate dalla comunità scientifica (Romito 2005a; Romito 2010; Paoloni & Zavattaro 2007; Beritelli et al. 2010; Villalba et al. 2016). Il valore ottenuto deve essere riportato esplicitamente nella relazione tecnica, indicando il metodo utilizzato, il software impiegato, i parametri di analisi e le eventuali criticità riscontrate.

A seguito della valutazione del rapporto SNR, si applicano i seguenti criteri:

SNR < 6 dB

- Le misure acustiche (formanti, VOT, durata, intensità) **non sono affidabili**.
- Le conclusioni fonetiche basate su tali misure **non possono essere utilizzate a fini probatori**.

SNR ≥ 6 dB

- L'analisi fonetica è possibile, purché siano soddisfatti anche i criteri minimi di trascrivibilità.

Una registrazione dovrebbe essere considerata non trascrivibile nella sua interezza se più del 30% della durata complessiva delle porzioni di audio in cui vi sia effettivamente una emissione di voce (quindi escludendo, ad esempio, lunghe pause in cui vi è soltanto rumore di fondo) non soddisfa i criteri sopra indicati, in linea con quanto previsto dall'art. 268, comma 7 c.p.p.<sup>5</sup>, che richiede la produzione di una trascrizione *integrale* e in forma *intelligibile*, secondo le garanzie proprie dell'espletamento delle perizie (art. 220 c.p.p.).

## 3. Trascrizione di intercettazioni

---

<sup>5</sup> «Il giudice dispone la trascrizione **integrale** delle registrazioni ovvero la stampa in forma **intelligibile** delle informazioni contenute nei flussi di comunicazioni informatiche o telematiche da acquisire, osservando le forme, i modi e le garanzie previsti per l'espletamento delle perizie [art. 220]. Le trascrizioni o le stampe sono inserite nel fascicolo per il dibattimento.» (Art. 268, comma 7 c.p.p.)

Le intercettazioni rappresentano una delle principali fonti probatorie nel processo penale italiano e si distinguono in due categorie tecnicamente e metodologicamente differenti:

- **Intercettazioni telefoniche**, caratterizzate da banda limitata (circa 300–3400 Hz), assenza di contesto visivo, numero di parlatori noto e forte dipendenza dalla qualità della linea;
- **Intercettazioni ambientali**, caratterizzate da maggiore variabilità acustica, presenza di rumori di fondo, distanza variabile dal microfono, sovrapposizioni di parlanti e dinamiche internazionali complesse.

Questa distinzione non è meramente tecnica: incide in modo diretto sulla trascrivibilità, sulla qualità del segnale, sulla ricostruzione del contenuto comunicativo e sulla responsabilità epistemica del trascrittore. Per tali ragioni, le relazioni di trascrizione non possono essere identiche né seguire le stesse procedure: ciascuna tipologia di intercettazione richiede criteri specifici, strumenti adeguati e un diverso livello di cautela interpretativa.

Come evidenziato dalla letteratura scientifica (Orletti 2016; Cirillo 2019; Romito 2021), la comunicazione orale è un fenomeno intrinsecamente multimodale, costruito attraverso canali paralleli e sovrapposti che concorrono alla produzione del significato. Una conversazione intercettata non è mai una semplice sequenza di parole: è un intreccio di prosodia, intonazione, pause, silenzi, sovrapposizioni, segnali non verbali, dinamiche di turn-taking (cd. presa di turno di parola) e relazioni situazionali. Ridurre la trascrizione alla sola componente verbale significa alterare la natura del parlato, perdere informazioni rilevanti e compromettere la corretta ricostruzione dell'interazione, con conseguenze dirette sulla verità processuale.

La linguistica forense — e i numerosi errori giudiziari documentati nel tempo — ha mostrato come elementi apparentemente marginali — un silenzio prolungato, un'esitazione, un cambio di intonazione, una risata, un sospiro, una sovrapposizione o un allungamento vocalico — possano costituire mosse comunicative decisive, talvolta più informative delle parole stesse. Il parlato, inoltre, è un fenomeno evanescente: si realizza nel tempo e scompare immediatamente dopo la sua produzione. La registrazione ne cattura solo una parte, spesso in condizioni acustiche difficili e la trascrizione ne restituisce una porzione ancora più ridotta. Per questo motivo, come sottolineato dagli studi sull'analisi della conversazione, la trascrizione non è mai una copia del parlato, ma una sua rappresentazione selettiva.

Da tale principio discendono vincoli precisi riguardo la trascrizione che deve includere pause, silenzi, sovrapposizioni, segnali non verbali e fenomeni prosodici; deve distinguere ciò che è certo da ciò che è dubbio; deve essere contestualizzata, riproducibile e verificabile; e deve essere sempre accompagnata da una relazione tecnica che espliciti criteri, limiti e scelte operative. Solo in questo modo la trascrizione può rispettare la natura del parlato e garantire un contributo affidabile al processo.

Sulla base di tali premesse, valgono i seguenti principi generali:

- **Codice da utilizzare.** La trascrizione deve essere redatta esclusivamente nella lingua o nel dialetto effettivamente ascoltato, senza normalizzazioni arbitrarie. Un'eventuale interpretazione in italiano può essere fornita solo come nota esplicita separata e chiaramente distinta dalla trascrizione fondata sul segnale, che rimane sempre la versione primaria. La riconduzione all'italiano standard è ammessa soltanto su esplicita richiesta del giudice ovvero della parte processuale che ha conferito l'incarico al proprio consulente e comunque come versione secondaria e subordinata.
- **Trascrizione integrale.** La trascrizione deve rappresentare tutti i canali comunicativi rilevanti, inclusi quelli non verbali quando pertinenti, e rispettare la struttura interazionale della conversazione. Ciò implica riportare turni, sovrapposizioni, pause, silenzi e fenomeni

sequenziali, poiché il significato degli enunciati dipende anche dalla loro collocazione nella dinamica conversazionale. Ignorare tali elementi significa perdere informazioni decisive e alterare la natura del parlato. Il principio di integralità della trascrizione impone la raccomandazione di non procedere, contrariamente a quanto accade nella prassi delle trascrizioni di polizia giudiziaria, all’inserimento di “omissis” in corrispondenza delle parti non trascritte perché ritenute irrilevanti ai fini del procedimento. Tale prassi può incidere significativamente sull’accertamento del fatto e, conseguentemente, sulla corretta amministrazione della giustizia. Essa introduce bias cognitivi ed epistemici suscettibili di compromettere le fasi investigativa, difensiva e giudiziaria. L’omissione di intere sequenze conversazionali può generare una fallacia cognitiva, consistente nella ricerca inconsapevole di significati sulla base dei soli segmenti trascritti. In tale prospettiva, chi legge la trascrizione può erroneamente interpretare la conversazione come criptica o allusiva a contenuti illeciti, quando tale impressione deriva unicamente da una rappresentazione incompleta e distorta della sequenza originaria. Inoltre gli “omissis” determinano l’impossibilità di rilevare la presenza di enciclopedie condivise, ossia quell’insieme di conoscenze che si dà per scontato il proprio interlocutore possieda. Gli “omissis” possono occultare scambi rilevanti, quali richieste di chiarimento, accentuando artificialmente il grado di “implicita” dell’interazione e inducendo a sovrastimare la condivisione di conoscenze tra gli interlocutori. Ne deriva l’impossibilità di determinare con precisione ciò che ciascun partecipante sa o non sa, con evidenti ricadute sull’interpretazione complessiva del dialogo.

- **Attendibilità della trascrizione.** La produzione di trascrizioni “pulite”, che omettano esitazioni, pause, sovrapposizioni, allungamenti o segnali non verbali presenti nel segnale, comporta un’alterazione della realtà comunicativa e non rispecchia le condizioni effettive della registrazione. È quindi necessario distinguere in modo trasparente ciò che è certo da ciò che è dubbio, adottando convenzioni esplicite per indicare segmenti incomprensibili, ipotesi interpretative o porzioni non udibili. La trascrizione deve mantenere criteri di verificabilità e replicabilità, fondandosi su procedure esplicite e documentate, e non su intuizioni personali o inferenze legate al contesto investigativo. Non rientra nelle buone pratiche colmare lacune percettive ricorrendo a conoscenze pregresse sul caso, sui parlanti o sul reato: ogni ricostruzione deve derivare esclusivamente dal segnale acustico disponibile.

### 3.1 Parole/frasi contestate (*disputed utterances*)

L’analisi di parole o frasi “controverse” (cd. “contestate”) è una delle attività più delicate della linguistica forense, poiché anche una singola unità lessicale può modificare in modo significativo il senso dell’interazione e incidere sull’accertamento processuale. Per questo motivo, l’identificazione di segmenti controversi deve essere condotta con rigore scientifico, evitando inferenze arbitrarie, ricostruzioni basate sul contesto investigativo o interpretazioni non fondate sul segnale acustico.

In coerenza con i principi della fonetica forense, dell’analisi della conversazione e della psicologia della percezione, è necessario seguire alcune prescrizioni fondamentali: calcolare il rapporto segnale/rumore (SNR) del tratto contestato; valutare la qualità acustica del segmento, considerando rumori, sovrapposizioni, distanza dal microfono e banda limitata; stimare l’intelligibilità; verificare la presenza di distorsioni, compressioni o tagli; e documentare ogni elemento nella relazione tecnica che accompagna la trascrizione.

Il lettore deve essere messo nelle condizioni di comprendere che le parole o le frasi contestate, quando affette da rumori, distorsioni o altre criticità, sono il risultato di un processo analitico e presentano quindi un margine di incertezza. In questi casi, ciò che si legge non può essere considerato certo. È spesso più semplice stabilire ciò che non è presente nel segnale piuttosto che affermare con sicurezza ciò che vi è effettivamente contenuto. È necessario evitare con fermezza la trappola delle inferenze basate sul contesto. Questo tipo di inferenze porta infatti, spesso in modo inconsapevole, a ricostruire

parole sulla base del contesto investigativo, a inferire significati a partire dalla conoscenza del caso, a completare segmenti incomprensibili seguendo aspettative narrative o, ancora, a utilizzare informazioni extralinguistiche per colmare lacune percettive. Proprio per questo motivo la trascrizione — e, più in generale, qualsiasi ricostruzione — deve derivare esclusivamente dal segnale acustico, senza aggiunte, integrazioni o interpretazioni che non trovino riscontro diretto nell’audio. Nel dubbio non si deve scegliere l’ipotesi più coerente con il contesto, o più probabile e più “logica”.

### 3.2 Relazione tecnica

Alla trascrizione deve essere obbligatoriamente allegata una relazione tecnica, che costituisce parte integrante e inscindibile del lavoro trascrittivo. La relazione ha la funzione di documentare in modo trasparente tutte le scelte analitiche, esplicitare criteri, limiti e incertezze (udibile, probabile, interpretato, indeterminabile), garantire replicabilità e verificabilità, rendere chiaro il grado di affidabilità della trascrizione.

La relazione tecnica non costituisce un allegato accessorio, ma un documento tecnico esplicativo essenziale: senza di essa la trascrizione non può essere considerata scientificamente fondata né metodologicamente verificabile. A tale scopo dovrebbe includere i seguenti elementi:

- **informazioni relative alla qualità del segnale**, tra cui valori del rapporto segnale/rumore (SNR); stima dell’intelligibilità; descrizione delle criticità acustiche (rumori, sovrapposizioni, distanza dal microfono, distorsioni, banda limitata, ecc.); indicazione delle parti del segnale che presentano degrado significativo.
- **Informazioni pregresse utilizzate per la trascrizione di:** nomi propri, toponimi, termini specialistici, varianti dialettali o socio-linguistiche. Tali informazioni devono essere dichiarate esplicitamente per evitare inferenze inconsapevoli o bias percettivi.
- **Descrizione del metodo di analisi adottato**, quindi criteri di ascolto, numero di ascolti, uso di cuffie professionali o monitor da studio, eventuale ricorso a strumenti di analisi acustica (sonogrammi, formanti, spettrogrammi), criteri di annotazione dei fenomeni prosodici e non verbali.
- **Le indicazioni delle porzioni non trascrivibili** devono riportare una precisa motivazione tecnica: SNR insufficiente, sovrapposizioni non risolvibili, rumori coprenti, parlato sussurrato o fuori banda, segmenti acusticamente indeterminabili. È importante che la non trascrivibilità sia **motivata**, non semplicemente dichiarata.
- **Elenco dei software e dei parametri utilizzati**, includendo: nome del software, versione, parametri di visualizzazione (ampiezza, scala, finestra, FFT, ecc.), eventuali plug-in o strumenti aggiuntivi.
- **Legenda delle convenzioni utilizzate**, che costituisce parte integrante del documento e ne garantisce la leggibilità, la trasparenza e la replicabilità. Le convenzioni devono essere coerenti con la letteratura scientifica sull’analisi della conversazione e con i principi della linguistica forense, e devono permettere di rappresentare la multimodalità del parlato, la struttura interazionale, le incertezze percettive, i fenomeni prosodici, i segnali non verbali, le sovrapposizioni e i turni. La legenda deve essere riportata all’inizio della trascrizione o in un allegato immediatamente accessibile. Si veda ad esempio il Vademecum per il trattamento delle fonti orali edito dal Ministero della Cultura. Direzione generale, soprattutto riguardo ai contributi relativi a ‘Trascrivere i documenti forensi’ di Romito L., ‘Trascrivere il dialetto’ di Abete G., e ‘Trascrivere la conversazione’ di Cirillo L.
- **Temporizzazione:** la trascrizione deve essere accompagnata da una temporizzazione accurata, che consenta di localizzare con precisione ogni segmento del parlato all’interno del file audio. È uno strumento essenziale per garantire la *verificabilità* della trascrizione; permettere alle parti di ritrovare rapidamente i passaggi rilevanti; documentare la *struttura temporale* dell’interazione; rendere possibile la *replicabilità* dell’analisi da parte di altri

esperti. Il formato da riportare è *hh:mm:ss*, l'intervallo massimo tra due marcatori non dovrebbe superare 1 minuto. Intervalli più brevi sono raccomandati in presenza di parlato sovrapposto, rumori coprenti, tratti acusticamente complessi, segmenti contestati, passaggi rilevanti ai fini processuali. È consigliabile inserire marcatori aggiuntivi in corrispondenza di cambi di turno, pause lunghe, eventi non verbali rilevanti, segmenti incomprensibili o indeterminabili. Anche le porzioni incomprensibili o non trascritte, così come le pause lunghe, devono essere accuratamente segnalate e temporizzate.

## Caratterizzazione e identificazione del parlante

### 4. Caratterizzazione del parlante (o analisi linguistica)

La caratterizzazione del parlante consiste nella descrizione delle proprietà linguistiche, fonetiche, prosodiche e sociolinguistiche che emergono dal segnale audio. Non coincide con l'identificazione del parlante, ma rappresenta un'attività preliminare e distinta, finalizzata a delineare tratti osservabili senza attribuire identità. La caratterizzazione deve essere condotta con un approccio scientifico, evitando inferenze arbitrarie, stereotipi linguistici o conclusioni non supportate dal segnale. Questa si effettua attraverso due livelli di analisi, che dovrebbero essere integrati quando possibile:

#### A. Analisi uditivo-percettiva

Basata sull'ascolto esperto, dei fenomeni prosodici, segmentali come variabili sintattiche, morfologiche, fonologiche, fonetiche e dialettali e paralinguistici<sup>6</sup>.

L'analisi percettiva deve essere sempre documentata e non può essere l'unico metodo utilizzato quando il segnale consente misure acustiche.

#### B. Analisi acustico-fonetica

Comprende misure delle formanti, analisi della F<sub>0</sub>, durata di alcuni segmenti ritenuti significativi, eventuale presenza di VOT (Voice Onset Time o "tempo di attacco della sonorità") o altri fenomeni fonologici, andamento prosodico e caratteristiche della voce (voce soffiata, creaky voice, tensione, ecc.).

L'analisi acustica è possibile solo quando il segnale presenta SNR  $\geq 6$  dB e qualità sufficiente.

In conclusione, la caratterizzazione del parlante (o della comunità linguistica di appartenenza) può includere:

- **Tratti sociolinguistici osservabili** — varietà dialettale, registro, stile comunicativo, marcatori di appartenenza a una comunità linguistica specifica;
- **Tratti fonetici segmentali** — realizzazione di vocali e consonanti, inclusi esiti dialettali rilevanti per l'identificazione areale;
- **Tratti prosodici** — intonazione, ritmo, accento, velocità articolatoria e pattern melodici ricorrenti;
- **Tratti paralinguistici** — qualità vocale (voce tesa, soffiata, creaky voice), intensità, postura fonatoria;
- **Caratteristiche del parlato** — chiarezza, articolazione, esitazioni, marcatori discorsivi, fenomeni di riduzione o iperarticolazione.

Le descrizioni dovrebbero basarsi esclusivamente su ciò che è effettivamente presente nel segnale acustico. Non è invece possibile ricavare con certezza elementi come l'età precisa del parlante, il livello di istruzione, la professione, uno stato emotivo definito, le intenzioni comunicative o valutazioni di tipo psicologico o comportamentale, poiché tali inferenze non trovano un fondamento scientifico nel segnale stesso. Osservazioni di questo tipo possono essere formulate solo evidenziando chiaramente che si tratta di interpretazioni e non di dati ricavabili direttamente dall'audio. Esse,

---

<sup>6</sup> A titolo esemplificativo, tra le variabili linguistiche rilevanti nell'analisi percettiva finalizzata a individuare la comunità linguistica di appartenenza del parlante, possono essere considerate alcune caratteristiche delle varietà dialettali calabresi. Sul piano sintattico, l'uso dei tempi passati mostra una chiara distribuzione areale: nelle zone settentrionali, con sostrato latino, prevalgono forme come «l'hai vistu», con impiego dell'imperfetto o del passato prossimo; nelle aree meridionali, con sostrato greco, è invece più frequente il passato remoto («u vittu») per «io l'ho visto». Dal punto di vista morfologico, il pronome possessivo può posporre nella fascia ionica («patrimma», «sorma») e anteporsi nella fascia tirrenica («me patri», «me soru») per «mio padre» e «mia sorella». Tra le variabili fonetiche, l'esito della -LL- latina presenta realizzazioni diverse anche a breve distanza geografica: /ll/ a Lamezia Terme, /dd/ a Catanzaro, /j/ a Roccella Jonica, /r/ a Badolato e /k/ a San Luca, tra le altre.

inoltre, dovrebbero essere collocate, eventualmente, in nota o in un'apposita sezione separata, e mai inserite nella medesima stringa del parlato trascritto, neppure mediante parentesi quadre, trattini, barre, segni grafici o altri marcatori di separazione. La presenza di annotazioni personali all'interno della stessa stringa trascritta rischia infatti di attribuire a tali osservazioni la medesima affidabilità tecnica della trascrizione, producendo una indebita assimilazione tra dato acustico rilevabile, attività tecnica di trascrizione e valutazione interpretativa e soggettiva del trascrittore.

## 5. Comparazione forense della voce

La comparazione forense della voce è un'attività altamente specialistica che mira a valutare se due campioni vocali — uno di riferimento e uno contestato — possano essere attribuiti allo stesso parlante oppure a parlanti diversi. Si tratta di un'operazione complessa, che richiede competenze avanzate di fonetica, acustica, linguistica e statistica.

La comparazione della voce non si configura come un riconoscimento soggettivo né come un'attività percettiva basata sull'intuizione, ma come una procedura scientifica che richiede criteri rigorosi, documentati e replicabili, analogamente a quanto avviene in altre discipline forensi, come l'analisi del DNA, le indagini di paternità o le comparazioni balistiche. In senso contrario, si segnala tuttavia la recente Sentenza di Cassazione Penale, sezione II, del 13 marzo 2025, numero 10310, che ha ritenuto utilizzabili — ai fini identificativi — le dichiarazioni degli ufficiali o agenti di polizia giudiziaria che affermino di riconoscere (ndr. anche solo percettivamente) con sicurezza la voce dell'imputato, senza necessità di perizia fonica, quando la loro deposizione sia ritenuta attendibile dal giudice. È inoltre necessario riconoscere la differenza tra caratterizzazione del parlante (che ha lo scopo di descrivere i tratti linguistici e fonetici del parlante) e comparazione forense della voce (che ha lo scopo di valutare la compatibilità tra due campioni). La caratterizzazione descrive. La comparazione valuta una relazione probabilistica tra due segnali<sup>7</sup>. Sottolineiamo la differenza tra caratterizzazione del parlante (che ha lo scopo di descrivere i tratti linguistici e fonetici del parlante) e comparazione forense della voce (che ha lo scopo di valutare la compatibilità tra due campioni). La caratterizzazione descrive. La comparazione valuta una relazione probabilistica tra due segnali.

Ancor più della trascrizione, la comparazione della voce è influenzata da numerosi fattori come:

- la qualità del segnale (SNR, banda, rumori, sovrapposizioni);
- la variabilità intra-parlante (stato emotivo, salute, stile comunicativo, contesto, età);
- la variabilità inter-parlante (somiglianze tra parlanti della stessa area linguistica);
- le differenze di canale (telefono vs ambiente, microfoni diversi, compressioni diverse);
- la durata dei campioni da analizzare;
- la presenza di parlato spontaneo vs letto;
- gli effetti del dialetto o del code-switching.

Questi fattori dovrebbero essere esplicitati nella relazione tecnica e considerati nella valutazione finale.

---

<sup>7</sup> <<Il Collegio intende dare seguito al principio di diritto secondo cui le dichiarazioni degli ufficiali e agenti di polizia giudiziaria che abbiano asserito di aver riconosciuto la voce dell'imputato sono legittimamente utilizzabili ai fini dell'identificazione degli interlocutori coinvolti in conversazioni intercettate nel corso delle indagini preliminari, quando è ritenuta attendibile la deposizione di colui che afferma di identificarlo con sicurezza, senza che sia necessario espletare perizia fonica [...]>> Cass. pen., sez. II, 13 marzo 2025, n. 10310.

In particolare, l'omogeneità del codice linguistico adottato dai parlanti da comparare può assurgere a criterio preliminare circa l'opportunità o meno di procedere con la comparazione.

Allo stato dell'arte attuale, l'OLF concorda pienamente con la letteratura scientifica internazionale, nell'affermare la natura probabilistica, e non deterministica, del problema della comparazione forense della voce<sup>8</sup>. Per questa ragione, l'OLF richiama l'adozione del paradigma bayesiano del rapporto di verosimiglianza (Likelihood Ratio, LR) quale criterio scientificamente fondato. In tal senso, la finalità della comparazione è valutare quanto sia più probabile osservare le caratteristiche del parlato contestato se il parlante è la stessa persona (cioè nell'ipotesi accusatoria), rispetto a se il parlante è una persona diversa (cioè nell'ipotesi difensiva). Questa valutazione può essere espressa in forma descrittiva o argomentativa, ma in entrambi i casi deve essere supportata dal rapporto di verosimiglianza tra le due ipotesi in competizione, che quantifica la forza dell'evidenza a favore dell'una o dell'altra ipotesi.

### **5.1 Metodi di comparazione**

Nella comparazione forense della voce si suole distinguere tra metodi soggettivi e metodi oggettivi questi ultimi a loro volta distinguibili in metodi acustico-fonetici o semi-automatici e metodi automatici. Si tratta di approcci profondamente diversi tra loro, non solo per gli strumenti utilizzati, ma soprattutto per il tipo di informazione che producono e per il grado di affidabilità che possono garantire. La scelta del metodo non è mai libera e dipende dalla qualità del segnale, dalla banda passante, dalla durata utile del parlato e dalla comparabilità dei campioni. Ogni metodo ha un ruolo preciso, dei limiti intrinseci e delle condizioni di applicabilità che dovrebbero essere rispettate.

#### **5.1.1 Metodi soggettivi (uditivi / psicoacustici)**

I metodi soggettivi si basano sulla percezione uditiva umana. L'esperto — o un gruppo di ascoltatori — confronta direttamente le due voci e formula un giudizio di somiglianza. Questo confronto può avvenire in due modi: attraverso dei test basati sulla memoria a breve termine, quando il campione di ascolto è numeroso, oppure basati sulla memoria a lungo termine, quando l'analista è esperto.

La percezione umana, tuttavia, è estremamente vulnerabile ai bias cognitivi: aspettative investigative, familiarità con il caso, differenze individuali nelle capacità percettive, variabilità nella sensibilità alle caratteristiche vocali. Per questo motivo si consiglia di utilizzare solo il metodo basato sulla memoria a breve termine con un alto numero di ascoltatori *inconsapevoli*. Alcune voci risultano più facilmente riconoscibili, altre sono molto simili tra loro, e non tutti gli ascoltatori hanno la stessa abilità nel discriminare tratti fonetici sottili. Per questo motivo, i metodi soggettivi non possono mai costituire l'unica base della comparazione quando il segnale consente un'analisi acustica. Il loro ruolo è di supporto: possono orientare l'attenzione verso tratti rilevanti, ma non possono produrre conclusioni deterministiche né sostituire metodi più robusti (riguardo l'applicazione del metodo si veda Romito (2010)).

#### **5.1.2 Metodi semiautomatici (acustico-fonetici / parametrici)**

I metodi semiautomatici rappresentano il cuore della fonetica forense tradizionale. Si basano sull'estrazione di parametri acustici misurabili — formanti, F0, durate, VOT, transizioni — che descrivono in modo oggettivo il comportamento vocale dei parlanti. Questi metodi sono trasparenti, replicabili e scientificamente fondati, ma richiedono condizioni tecniche precise.

---

<sup>8</sup> V. altresì le risoluzioni di IAFPA – International Association for Forensic Phonetics and Acoustics, e di ENFSI – European Network of Forensic Science Institutes.

I metodi semiautomatici sono preferibili ogni volta che il segnale lo consente, perché permettono di fondare la comparazione su dati misurabili e verificabili. Tuttavia, non sono applicabili a campioni troppo brevi o troppo degradati, né su parlato sussurrato o camuffato.

La comparazione forense della voce può essere avviata solo dopo aver verificato che il materiale audio soddisfa i requisiti minimi necessari per un'analisi fonetico-acustica affidabile. Questa fase preliminare è obbligatoria e costituisce il fondamento metodologico dell'intera procedura comparativa. Per ciascun campione (contestato e di riferimento) devono essere valutati i parametri acustici fondamentali che determinano la possibilità stessa di effettuare una comparazione forense della voce. La letteratura internazionale (IAFPA Code of Practice; ENFSI Guidelines for Forensic Speaker Comparison; Rose 2002; Nolan 1983; Gold & French 2011) converge nell'indicare tre requisiti minimi per la comparazione forense della voce:

#### A. Rapporto segnale/rumore (SNR)

Il SNR è il parametro più importante per stabilire se sia possibile effettuare misure acustiche affidabili.

- **SNR < 6 dB** → **non è possibile effettuare misure acustiche affidabili.** Il segnale è troppo degradato per estrarre formanti, F0 o parametri temporali.
- **SNR ≥ 6 dB** → **analisi possibile ma con prudenza.**
- **SNR ≥ 10 dB** → **requisito minimo raccomandato dalla letteratura** per ottenere misure acustiche stabili e ripetibili.

#### B. Banda passante

La banda passante determina quali informazioni acustiche sono effettivamente disponibili per l'analisi. Poiché la banda telefonica standard (300–3400 Hz) limita in modo significativo l'accesso alle formanti e alla F0, è necessario verificare che la banda passante della registrazione superi i 3000 Hz. Solo in presenza di una banda passante più ampia è possibile ottenere misure acustiche stabili e utilizzare parametri discriminanti in modo affidabile.

#### C. Durata utile del parlato e numero di parametri misurabili

Questi due aspetti costituiscono un unico requisito minimo, poiché la durata utile determina il numero di misure acustiche effettivamente estraibili (ovviamente in base alla velocità di eloquio). La durata utile deve escludere rumori coprenti, sovrapposizioni, segmenti fuori banda, parlato sussurrato, distorsioni e porzioni non trascrivibili o incomprensibili.

La letteratura suggerisce che, con una durata utile inferiore a 10 secondi, l'analisi acustica non sia affidabile. Il parametro "durata" è fondamentale per garantire la possibilità di ottenere almeno 5 misure complete (F0, F1, F2, F3) per ogni vocale.

Oltre ai requisiti minimi, è necessario considerare alcuni fattori tecnici legati alla modalità di intercettazione che influenzano la qualità del segnale e, di conseguenza, l'affidabilità del risultato. Tra questi, la *distanza dal microfono*, che incide sull'intensità della voce, sulla qualità timbrica e sul rapporto diretto/riverberato: segmenti prodotti troppo lontano dal microfono possono risultare inutilizzabili.

Un ulteriore elemento riguarda la *compressione* e la *codifica* della registrazione. È noto che formati compressi come AMR, MP3, AAC comportano una significativa perdita di informazioni acustiche, riducendo la qualità e la stabilità delle misure. Per questo motivo è preferibile operare su segnali in formato WAV/PCM, che preservano integralmente il contenuto acustico.

## Sintesi dei requisiti minimi

Parametro	Requisito minimo	Motivazione
Compressione	Preferibile WAV/PCM	Ridotta perdita di informazione
Intensità della voce	Utile e adatta alla misurazione	Stabilità delle misure acustiche
SNR	$\geq 10$ dB	Stabilità delle misure acustiche
Banda passante	$> 3000$ Hz	Accesso alle formanti e alla F0
Durata utile	$\geq 10$ secondi	Sufficienza statistica
Misure per vocale	$\geq 5$	Robustezza del metodo parametrico

Se uno o più di questi requisiti minimi non sono soddisfatti, la comparazione forense della voce non può essere effettuata. In tal caso, il perito, il consulente o l'operatore di polizia giudiziaria deve dichiarare l'impossibilità tecnica dell'analisi.

### 5.1.3 Metodi automatici (modelli statistici e machine learning)

I metodi automatici rappresentano l'evoluzione più recente della comparazione forense della voce. Si basano su modelli statistici complessi — GMM-UBM, i-vectors, x-vectors, PLDA, reti neurali profonde, IA — capaci di analizzare grandi quantità di parametri e di produrre, quando possibile, un Likelihood Ratio (LR).

Questi sistemi richiedono condizioni molto rigorose: campioni comparabili per canale, durata utile sufficiente (che può variare da dieci a oltre trenta secondi a seconda del modello), e soprattutto dataset di addestramento adeguati e rappresentativi della varietà linguistica dei parlanti. Sono metodi potenti, ma sensibili alle differenze di canale, alla compressione e alla qualità del segnale. Inoltre, la loro natura talvolta opaca — la cosiddetta “black box” — impone cautela nell'interpretazione dei risultati<sup>9</sup>. I metodi automatici possono integrare l'analisi acustica, ma non possono sostituirla completamente quando il segnale è complesso o degradato. Il loro uso richiede trasparenza, documentazione e una chiara esplicitazione dei limiti.

## 5.2 Integrazione dei metodi

La comparazione forense della voce non può basarsi su un solo metodo. Ogni approccio — soggettivo, semiautomatico, automatico — illumina una parte diversa del fenomeno vocale, e solo la loro integrazione permette di costruire una valutazione equilibrata, trasparente e replicabile. L'analisi linguistica può restringere il campo di azione, l'analisi percettiva può orientare l'attenzione, l'analisi acustica fornisce parametri misurabili, i metodi automatici aggiungono una prospettiva statistica. La forza della comparazione sta nella convergenza dei metodi, non nell'affidamento esclusivo a uno di essi.

### 5.3 Il rapporto di verosimiglianza (Likelihood Ratio, LR),

La comparazione forense della voce, per essere scientificamente fondata e metodologicamente corretta, si suggerisce che venga condotta all'interno di un quadro probabilistico esplicito.

<sup>9</sup> Per “black box” si intende, appunto, la natura potenzialmente non trasparente di alcuni sistemi automatici, nei quali il risultato finale viene prodotto attraverso elaborazioni statistiche o computazionali complesse, non sempre pienamente interpretabili e verificabili dall'operatore umano.

L'Osservatorio sulla Linguistica Forense raccomanda di adottare come riferimento il modello di Bayes e il rapporto di verosimiglianza (Likelihood Ratio, LR), applicati attraverso metodi semiautomatici e/o automatici, secondo quanto stabilito dalla letteratura internazionale più autorevole (Morrison 2010; Morrison & Zhang 2023). Questo approccio non costituisce una preferenza metodologica, ma viene considerato la soluzione epistemicamente più adeguata poiché un modello probabilistico consente di valutare in modo trasparente la forza dell'evidenza vocale. L'idea centrale è che la comparazione non si limiti a descrivere somiglianze o differenze tra due voci, ma miri a quantificare quanto le caratteristiche osservate siano più compatibili con l'ipotesi che i due campioni provengano dalla stessa persona (Hp) rispetto all'ipotesi che provengano da persone diverse (Hd). Il LR è lo strumento che consente questa valutazione: non stabilisce identità, ma esprime la forza dell'evidenza. Per questo motivo, il LR viene generalmente considerato il cardine dell'intera metodologia di comparazione della voce. È opportuno dichiarare esplicitamente le ipotesi adottate nel modello bayesiano, incluse, a titolo esemplificativo: l'ipotesi del Pubblico Ministero (Hp), l'ipotesi della Difesa (Hd) e la popolazione di riferimento, cioè il gruppo di parlanti rispetto al quale viene valutata la probabilità di osservare quelle caratteristiche vocali. Questi elementi non rappresentano semplici dettagli tecnici ma costituiscono la condizione affinché il LR sia interpretabile. Un LR calcolato senza esplicitare Hp, Hd e popolazione di riferimento è metodologicamente incompleto e non può essere assunto dal giudice e dalle altre parti processuali come tecnicamente affidabile. In sintesi, la metodologia suggerita dall'OLF si fonda su un principio semplice ma decisivo: la comparazione della voce beneficia di un approccio probabilistico, trasparente e replicabile. Il LR è lo strumento che consente di trasformare osservazioni fonetiche in una valutazione scientificamente interpretabile e si propone come elemento centrale dell'intero processo<sup>10</sup>.

Una volta calcolato il Likelihood Ratio, il compito dell'esperto non è semplicemente “dare un numero”, ma spiegare che cosa significa quel numero per la valutazione dell'evidenza vocale. Il LR, infatti, non stabilisce se due voci appartengano alla stessa persona: indica quanto le caratteristiche osservate siano più compatibili con una delle due ipotesi in gioco. L'idea è intuitiva: se un certo tratto vocale è molto comune nella popolazione, la sua presenza non aiuta a distinguere tra Hp e Hd; se invece è raro, la sua presenza può fornire un'evidenza più forte. Il LR serve proprio a misurare questa differenza di compatibilità.

È importante sottolineare che il LR non dice nulla sulla probabilità che l'imputato sia responsabile o meno del fatto di reato addebitato. Il suo ruolo è molto più limitato e molto più preciso: riguarda solo la forza dell'evidenza vocale e non può considerarsi risolutivo rispetto a una specifica vicenda processuale. La valutazione complessiva dell'intero apparato probatorio spetterà al giudice, che valuterà il LR integrandolo con tutte le altre evidenze probatorie, nel contraddittorio delle parti.

#### ***5.4 La trasparenza è più importante della conclusione.***

La comparazione forense della voce richiede un equilibrio costante tra ciò che è tecnicamente possibile e ciò che è scientificamente giustificabile. Anche quando si dispone di modelli probabilistici solidi e di strumenti affidabili, il risultato finale dipende dalla qualità del materiale di partenza e dalla trasparenza dell'intero processo analitico.

Un primo limite riguarda il segnale: rumori di fondo, sovrapposizioni, compressioni aggressive, banda telefonica ristretta o parlato troppo breve possono impedire l'estrazione di parametri affidabili. In queste condizioni non è sufficiente “fare del proprio meglio”: è necessario riconoscere quando il

---

<sup>10</sup> Il metodo basato sul LR è raccomandato da: Association of Forensic Science Providers (UK, Ireland); Royal Statistical Society; ENFSI (European Network of Forensic Science Institutes); National Institute of Forensic Science (Australia–NZ); American Statistical Association; Forensic Science Regulator (England & Wales); IAFPA (International Association for Forensic Phonetics & Acoustics); SWGDE (Scientific Working Group on Digital Evidence).

materiale non soddisfa i requisiti minimi e dichiarare l'impossibilità tecnica di procedere o la necessità di interpretare il risultato con cautela.

Un secondo limite è la **variabilità intra-parlante**, cioè il fatto che la stessa persona può produrre voci molto diverse a seconda del contesto, dello stato emotivo, della salute, della distanza dal microfono o del tipo di interazione. Questa variabilità può rendere difficile distinguere tra differenze dovute alla situazione e differenze dovute a parlanti diversi.

A ciò si aggiunge la **variabilità inter-parlante**: in alcune comunità linguistiche, soprattutto quando i parlanti condividono area geografica o varietà dialettale, le voci possono risultare sorprendentemente simili. Questo rende essenziale definire correttamente la popolazione di riferimento e interpretare correttamente il valore di Likelihood Ratio.

Accanto ai limiti del materiale, esiste un limite più generale legato alla responsabilità epistemica dell'esperto. L'analisi deve basarsi esclusivamente su ciò che il segnale permette di osservare, evitando di colmare lacune con inferenze tratte dal contesto investigativo o da aspettative esterne. Quando l'evidenza è debole, anche la conclusione deve essere debole; quando è inconcludente, la conclusione deve riportare tale constatazione.

Perché il risultato sia affidabile, non basta applicare un modello bayesiano o calcolare un LR: l'intero processo deve essere trasparente e validato. La trasparenza riguarda la chiarezza con cui vengono presentati i passaggi dell'analisi: selezione dei segmenti vocali, parametri misurati, strumenti utilizzati, ipotesi del modello (Hp, Hd e popolazione di riferimento) e limiti del materiale. La validazione riguarda invece la solidità del metodo: uno strumento è affidabile solo se ha dimostrato di funzionare per la lingua, la qualità del segnale e la popolazione coinvolte. A questo scopo è utile includere nella relazione i valori di convalida, come il grafico di Tippett e il Cllr (Cost Log-Likelihood Ratio), per individuare eventuali problemi di calibrazione<sup>11</sup>.

Il requisito della trasparenza include anche lo strumento informatico, cioè tipicamente i software utilizzati per l'elaborazione dei dati e per la comparazione. Il software dovrebbe essere:

- noto nella letteratura scientifica di riferimento
- corredato da documentazione che ne comprovi l'affidabilità in casi reali di comparazione analoghi al caso concreto di volta in volta affrontato
- utilizzabile autonomamente anche dalle altre parti del processo, nei termini previsti dalla licenza d'uso; ove la licenza d'uso non sia disponibile ovvero non consenta un utilizzo autonomo del software, la parte interessata dovrebbe poter ottenere dalla parte che ne ha la disponibilità l'accesso allo stesso, al fine di utilizzarlo sotto la sua supervisione.

La forza dell'evidenza non dipende solo dal valore numerico del LR, ma dalla qualità del percorso che porta a quel risultato. Rendere visibile questo percorso, riconoscere i limiti del materiale non indebolisce l'affidabilità del risultato: ne costituisce la solidità scientifica.

### 5.5 Relazione tecnica

La comparazione forense della voce non si esaurisce nell'analisi tecnica o nel calcolo del LR: richiede una documentazione accurata e completa di ogni passaggio. La documentazione non è un adempimento formale, ma una parte essenziale della metodologia, perché permette a chi legge — giudice, avvocati, altri esperti — di comprendere come si è arrivati al risultato e di verificarne la

---

<sup>11</sup> Il grafico di Tippett mostra, tramite due curve cumulative, la distribuzione dei log-LR per confronti *same-speaker* e *different-speaker*, permettendo di valutare visivamente la separazione tra le due condizioni. Una buona separazione indica un sistema discriminante, mentre sovrapposizioni o forme anomale delle curve segnalano problemi di calibratura, bias o eccessiva variabilità. Il **Cllr** è l'indice numerico che misura l'efficienza complessiva del sistema LR: valori prossimi a 0 indicano un sistema ben calibrato e informativo, valori vicini a 1 un sistema non informativo, valori superiori a 1 un sistema fuorviante. Le sue componenti, **Cllr min** (discriminazione) e **Cllr cal** (calibratura), permettono di distinguere le due dimensioni di prestazione. Utilizzati insieme, grafico di Tippett e Cllr forniscono una valutazione completa, visiva e numerica, della qualità di un sistema di confronto vocale.

correttezza. Per questo motivo, l'OLF richiede che ogni relazione tecnica includa una serie di elementi fondamentali, presentati in modo chiaro e trasparente.

Il primo elemento riguarda *la descrizione del materiale analizzato*. L'esperto deve indicare quali file ha ricevuto, in quale formato, con quali caratteristiche tecniche e in quali condizioni sono stati prodotti. È importante specificare eventuali criticità già presenti nel materiale, come rumori, sovrapposizioni o compressioni, perché queste condizioni influenzano direttamente la qualità dell'analisi.

Il secondo elemento è *la descrizione delle operazioni preliminari*. L'esperto deve spiegare come ha digitalizzato, etichettato e preparato i file, quali porzioni sono state selezionate e quali sono state escluse, e per quali motivi. Questa parte permette di comprendere il percorso che ha portato alla scelta dei segmenti utilizzati per l'analisi.

Il terzo elemento riguarda *la verifica dei requisiti minimi*. L'esperto deve documentare in modo esplicito se il materiale soddisfa i requisiti necessari per una comparazione affidabile: rapporto segnale/rumore, banda passante, durata utile, assenza di sovrapposizioni e così via. Se uno o più requisiti non sono soddisfatti, l'esperto deve dichiararlo e spiegare le conseguenze sulla possibilità di procedere.

Il quarto elemento è *la descrizione dei metodi utilizzati*. L'esperto deve indicare se ha utilizzato metodi semiautomatici, automatici o entrambi, spiegando in modo semplice quali parametri sono stati misurati o quali modelli sono stati applicati. Non è necessario entrare nei dettagli tecnici in questa sede, ma è fondamentale rendere comprensibile il percorso metodologico.

Il quinto elemento riguarda *il calcolo del Likelihood Ratio*. L'esperto deve riportare il valore del LR, spiegare quali ipotesi sono state adottate ( $H_p$ ,  $H_d$ ) e quale popolazione di riferimento è stata utilizzata. Questa parte deve essere chiara e leggibile anche per chi non ha competenze statistiche.

Il sesto elemento è *la presentazione dei limiti*. Ogni analisi ha dei limiti: qualità del segnale, durata del parlato, variabilità dei parlanti, differenze di canale. L'esperto deve dichiarare questi limiti in modo trasparente, spiegando come possono aver influenzato il risultato.

Infine, la documentazione deve includere *la conclusione finale*, senza espressioni categoriali e nel pieno rispetto del modello probabilistico.

In sintesi, la documentazione obbligatoria non è un semplice allegato tecnico, ma una parte integrante della comparazione forense della voce. È ciò che permette al giudice e alle parti processuali di comprendere la portata dell'evidenza, di valutarne la forza e di collocarla correttamente nel quadro probatorio complessivo. La trasparenza del processo è la garanzia più solida della qualità del risultato.

## Miglioramento del segnale audio

### 6. Il filtraggio

Il miglioramento del segnale è una fase delicata della comparazione forense della voce. Non si tratta di “modificare” l’audio, né di “ripulirlo” in senso estetico, ma di applicare interventi minimi e controllati che permettano di rendere il materiale più leggibile dal punto di vista acustico, senza alterare le caratteristiche vocali del parlante. L’obiettivo è facilitare l’analisi, migliorare l’intelligibilità, non trasformare il segnale.

La prima regola è che ogni intervento deve essere documentato e giustificato. L’esperto deve sempre conservare una copia del file originale, non modificato, e deve indicare con precisione quali operazioni sono state effettuate, con quali parametri e per quali finalità. La trasparenza è essenziale ed è parte integrante della validità del risultato: un miglioramento non documentato è, a tutti gli effetti, una manipolazione.

Un aspetto importante riguarda la selezione delle porzioni utili. In molti casi, il miglioramento del segnale non consiste nel “modificare” l’audio, ma nello scegliere quali parti analizzare e quali escludere. Rumori coprenti, sovrapposizioni, distorsioni o segmenti troppo brevi devono essere lasciati fuori dall’analisi, perché nessun miglioramento può rendere affidabile l’analisi acustica della voce contenuta in tali segmenti. La selezione è una forma di miglioramento: non altera il segnale, ma migliora la qualità dell’analisi.

#### 6.1 Limiti del miglioramento del segnale

Il miglioramento del segnale, per quanto utile, non può trasformare un materiale inadeguato in uno idoneo. Vi sono limiti tecnici e scientifici che non possono essere superati senza compromettere l’integrità della voce e, di conseguenza, la validità della comparazione. È quindi essenziale attenersi a principi fondamentali che aiutano l’esperto a distinguere quando un intervento è praticabile, quando comporta rischi e quando è preferibile evitarlo.

Il primo limite riguarda la natura stessa del segnale. Se il parlato è troppo breve, troppo degradato, coperto da rumori intensi o prodotto in condizioni estreme (come distanza eccessiva dal microfono, distorsione o compressione aggressiva), nessun miglioramento può renderlo adatto a un’analisi affidabile. In questi casi, l’esperto deve dichiarare l’impossibilità tecnica di procedere, senza tentare interventi che rischiano di introdurre artefatti o alterazioni.

Un secondo limite riguarda la non ricostruibilità del parlato. Non è possibile — né lecito — tentare di “recuperare” parti di parlato coperte da rumore, sovrapposizioni o distorsioni. Qualsiasi algoritmo che *prometta* di “ricostruire” la voce originale, di “estrarre” la voce da un rumore intenso o di “separare” parlanti sovrapposti introduce inevitabilmente elementi artificiali. L’esperto deve evitare questi strumenti e limitarsi a selezionare le porzioni realmente utilizzabili.

Nel caso della comparazione, è necessario rispettare il principio della coerenza tra i campioni. Il miglioramento del segnale non deve mai creare differenze artificiali tra la voce anonima e la voce di riferimento. Se un intervento viene applicato a un campione, deve essere applicato — quando possibile — anche all’altro, per evitare che la comparazione sia influenzata da differenze introdotte dal processo tecnico e non dai parlanti.

Il miglioramento del segnale è uno strumento utile ma limitato: non può sostituire la qualità del materiale originale, né ricostruire ciò che non esiste, né superare i limiti intrinseci del segnale. Il suo ruolo è rigoroso e circoscritto, ovvero rendere analizzabile ciò che è già idoneo senza alterare la voce del parlante. Per questo motivo è vietato utilizzare file filtrati come unica base per conclusioni probatorie, mentre è consigliabile fondare la trascrizione e l’analisi fonetica sul file originale, ricorrendo al filtrato solo come supporto ausiliario. È inoltre opportuno dichiarare esplicitamente ogni limite tecnico o metodologico che possa influire sull’interpretazione del segnale filtrato.

## **6.2 Relazione tecnica**

Qualora si proceda al filtraggio, il perito deve documentare in modo completo e replicabile l'intera procedura; pertanto, è consigliabile descrivere dettagliatamente il software utilizzato, la versione del software, il tipo di filtro applicato, i parametri impostati, l'ordine dei filtri e le eventuali normalizzazioni o equalizzazioni. Inoltre, è necessario allegare alla relazione il file filtrato, il file residuo (ciò che il filtro ha rimosso) e una valutazione tecnica degli effetti del filtraggio.

## 7. Conclusioni

Le presenti linee guida intendono fornire un quadro metodologico condiviso per il trattamento, la trascrizione, l'analisi e la comparazione del materiale sonoro intercettato, nella consapevolezza che il dato audio costituisce una fonte probatoria complessa, fragile e non riducibile alla sola dimensione verbale.

La registrazione sonora deve essere preservata nella sua integrità, documentata nella catena di custodia, valutata nella qualità tecnica e analizzata secondo criteri trasparenti, verificabili e replicabili. La trascrizione, la caratterizzazione del parlante, la comparazione forense della voce e l'eventuale miglioramento del segnale non possono fondarsi su intuizioni percettive, inferenze investigative o ricostruzioni narrative, ma devono derivare esclusivamente da ciò che il segnale consente scientificamente di osservare.

In questa prospettiva, il compito dell'esperto non è trasformare l'incerto in certo, né colmare le lacune dell'audio con ipotesi coerenti con il contesto processuale, ma indicare con precisione ciò che è udibile, ciò che è probabile, ciò che è dubbio e ciò che resta tecnicamente indeterminabile.

La qualità dell'accertamento non dipende soltanto dal risultato finale, ma dalla chiarezza del percorso che conduce a quel risultato: formati adeguati, parametri tecnici dichiarati, criteri di trascrizione espliciti, indicazione dei limiti, uso prudente dei metodi comparativi, adozione di modelli probabilistici quando si procede alla comparazione vocale e documentazione completa delle operazioni svolte. Tali cautele non rappresentano un irrigidimento formalistico dell'attività peritale, ma una garanzia essenziale per il processo: tutelano l'affidabilità della prova, il contraddittorio tra le parti, la comprensibilità delle consulenze e la corretta valutazione giudiziaria dell'evidenza intercettata.

L'Osservatorio sulla Linguistica Forense, ribadisce pertanto che ogni analisi sul materiale intercettato deve essere guidata da un principio di responsabilità epistemica: affermare solo ciò che il dato consente di affermare, dichiarare i limiti quando il dato non consente conclusioni affidabili, rendere sempre controllabile il metodo adottato. Solo in questo modo la linguistica e la fonetica forense possono offrire al processo penale un contributo realmente scientifico, rispettoso delle garanzie, delle parti e della verità processuale.

## 8. Bibliografia

**Nota bibliografica** La bibliografia riportata di seguito include sia i riferimenti citati nel testo sia una serie di opere di base considerate essenziali per chi voglia orientarsi nella letteratura internazionale sulla fonetica forense, sulla comparazione della voce e sui metodi probabilistici applicati alle scienze forensi.

Abete G., (2023) ‘Trascrivere il dialetto’, Tavolo permanente per le fonti orali. Vademecum per il trattamento delle fonti orali. Ministero della Cultura. Direzione generale

Beritelli F., Casale S., Grasso R., Spadaccini A. (2010). *Performance Evaluation of SNR Estimation Methods in Forensic Speaker Recognition*. In: 2010 Fourth International Conference on Emerging Security Information, Systems and Technologies, Venice, Italy, pp. 88–92.

Cirillo L., (2023) ‘Trascrivere la conversazione’, Tavolo permanente per le fonti orali. Vademecum per il trattamento delle fonti orali. Ministero della Cultura. Direzione generale

Cooke M., King S., Garnier M., Aubanel V. (2014). *The Listening Talker: A Review of Human and Algorithmic Context-Induced Modifications of Speech*. *Computer Speech & Language*, 28(2): 543–571.

ENFSI (2015). *Guidelines for Forensic Speaker Comparison*. European Network of Forensic Science Institutes, Speech and Audio Analysis Working Group.

Garayzábal Heinze, E., Queralt Estévez, S., & Reigosa Riveiros, M. (2019). *Fundamentos de la lingüística forense*. Madrid: Editorial Síntesis.

Gold E., French P. (2011). International Practices in Forensic Speaker Comparison. *International Journal of Speech, Language and the Law*, 18(2): 293–307.

Grimaldi M. (1995). *Aspetti pragmatolinguistici, caratterizzazione del parlante e attività investigativa*. In: Fedi F., Paoloni A. (a c. di), *Caratterizzazione del Parlatore*, Atti delle 6e Giornate di Studio del GFS (A.I.A.), Roma, Fondazione Ugo Bordoni, Vol. XXIII, pp. 109–120.

IAFPA (2019). *Code of Practice*. International Association for Forensic Phonetics and Acoustics.

Jessen M. (2008). *Forensic Phonetics*. *Language and Linguistics Compass*, 2/4: 671–711.

Jessen M. (2018). *Forensic Voice Comparison*. In: Visconti J. (ed.), *Handbook of Communication in the Legal Sphere*, Berlin: De Gruyter, pp. 219–255.

Jessen M. (2021). *Speaker Profiling and Forensic Voice Comparison: The Auditory Acoustic Approach*. In: Coulthard M., May A., Sousa-Silva R. (eds.), *The Routledge Handbook of Forensic Linguistics* (2nd ed.), London: Routledge, pp. 382–399.

Loakes D. (2024). *Automatic Speech Recognition and the Transcription of Indistinct Forensic Audio: How Do the New Generation of Systems Fare?* *Frontiers in Communication*, 9:1281407.

McMenamin, G. R. (2002). *Introducción a la lingüística forense: Un libro de curso*. Bogotá: Universidad Nacional de Colombia.

Miller G.A. (1951). *Language and Communication*. New York: McGraw-Hill.

- Morrison G.S. (2009). *Forensic Voice Comparison and the Paradigm Shift*. Science & Justice, 49: 298–308.
- Morrison G.S. (2010). *Forensic Voice Comparison*. In: Freckelton I., Selby H. (eds.), *Expert Evidence*, Sydney: Thomson Reuters, cap. 99.
- Morrison G.S. (2014). *Distinguishing Between Forensic Science and Forensic Pseudoscience: Testing of Validity and Reliability, and Approaches to Forensic Voice Comparison*. Science & Justice, 54: 245–256.
- Morrison G.S., Zhang C. (2023). *Forensic Voice Comparison: Overview*. In: Houck M., Wilson L., Eldridge H., Lewis S., Lothridge K., Reedy P. (eds.), *Encyclopedia of Forensic Sciences* (3rd ed.), vol. 2, Elsevier, pp. 737–750.
- Natarajan S., Abdul Rahman Al-Haddad S., Arif Ahmad F., Kamil R., Hassan M.K., Azrad S., et al. (2025). *Deep Neural Networks for Speech Enhancement and Speech Recognition: A Systematic Review*. Ain Shams Engineering Journal, 16(7): 103405.
- Nolan F. (1983). *The Phonetic Bases of Speaker Recognition*. Cambridge University Press, Cambridge.
- Nolan F. (2005). *Forensic Speaker Identification and the Phonetic Description of Voice Quality*. In: Hardcastle W.J., Beck J. (eds.), *A Figure of Speech: A Festschrift for John Laver*, Mahwah, NJ: Erlbaum, pp. 385–411.
- O’Shaughnessy D. (2024). *Speech Enhancement—A Review of Modern Methods*. IEEE Transactions on Human-Machine Systems, 1(1): 110–120.
- Orletti F. (2016). *La trascrizione delle intercettazioni telefoniche ed ambientali: un esercizio di analisi della conversazione applicata*. In: Gatta F. (ed.), *Parlare insieme. Studi per Daniela Zorzi*, Bologna: Bononia University Press, pp. 49–64.
- Orletti F. (2019). *La trascrizione delle intercettazioni telefoniche e ambientali*. Quaderni del Dipartimento di Linguistica, Università della Calabria, 27: 349–365.
- Paoloni A. (2006). *Limiti alla trascrizione giudiziaria*. In: Atti del 3° Convegno Nazionale AISV, CNR Padova, pp. 215–221.
- Paoloni A., Zavattaro D. (2007). *Intercettazioni telefoniche e ambientali. Metodi, limiti e sviluppi nella trascrizione e verbalizzazione*. Torino: Centro Scientifico Editore.
- Romito L. (2005a). *Il contesto, l’intelligibilità, il rapporto segnale-rumore*. In: Cosi P. (a cura di), *Misura dei parametri*, Torriana: EDK, pp. 539–566.
- Romito L. (2005b). *La competenza linguistica nelle trascrizioni forensi: l’intelligibilità, l’oggettività e il rapporto segnale/rumore*. Detective and Crime, pp. 52–63.
- Romito L. (2010), *Manuale di fonetica articolatoria, acustica e forense*, Centro editoriale Università della Calabria.

- Romito L. (2021). *La trascrizione in ambito forense*. In: Bernini G. (a cura di), *Superare l'evanescenza del parlato*, L&F – Linguistica e Filologia, Università di Bergamo.
- Romito L. (2026), *La perizia fonica e gli errori giudiziari: l'importanza della linguistica forense*, Melazzo, R. & Meluzzi, C. (a cura di), *Appunti di linguistica forense*, Milano: Milano University Press, coll. Consonanze.
- Romito L., (2013), *Manuale di Linguistica Forense*. Roma: Bulzoni Editore
- Romito L., (2026), *Il processo è fatto di parole: l'urgenza di una professionalità certificata*, in *La voce della prova: il perito trascrittore tra giustizia e intelligence*, a cura di L. Romito e M. Caligiuri, Rubbettino editore.
- Romito L., (2026), *La linguistica forense tra scienza e pseudoscienza e tra ricerca e applicazione*, in *La voce della prova: il perito trascrittore tra giustizia e intelligence*, a cura di L. Romito e M. Caligiuri, Rubbettino editore.
- Romito L., Frontera M. (2017). *La trascrizione forense di intercettazioni ambientali: una proposta di metodologia procedurale*. *Quaderni di Linguistica*, 5: 105–120.
- Romito, L. (2023), 'Trascrivere i documenti forensi', Tavolo permanente per le fonti orali. *Vademecum per il trattamento delle fonti orali*. Ministero della Cultura. Direzione generale
- Romito, L. (2023), *La Linguistica Forense: Prefazione*, *La Linguistica Forense in prospettiva multidisciplinare - Forensic Linguistics in a multidisciplinary perspective*, a cura di C. Meluzzi e S. Cenceschi, studi AISV, Vol. 10 - Collana peer reviewed (pp. 9-27), Officinaventuno (MI).
- Romito, L. e Vigè, M.D. (2019), Un software per la gestione delle trascrizioni in ambito processuale: un database di documenti sonori, in L. Romito (a cura di) *Quaderni del Dipartimento di Linguistica, 30 anni di Laboratorio di Fonetica*, 27 (numero speciale), Università della Calabria, ISSN 1128-7675.
- Romito, L., (2017), Chapter Three: A training program for expert forensic transcribers, in F. Orletti e L. Mariottini (a cura di) *Forensic Communication in Theory and Practice A Study of Discourse Analysis and Transcription*, Cambridge: Cambridge University Press, pp. 47-62, ISBN (10) 1-4438-9569-5; ISBN (13) 978-1-4438-9569-9.
- Romito, L., Tarasi, A., Ciardullo, M.A. e Graziano, E., (2017), Un modello per l'annotazione di fatti prosodici nelle trascrizioni forensi, in L. Romito & M. Frontera (a cura di) *La scrittura all'ombra della parola - Quaderni di Linguistica*, 5: 139-155, Milano: Officinaventuno, ISBN: 978-889765-17-0, ISSN: 2239-6748.
- Roods G.M., James S.H. (2006). *In-flight Communication*. In: Rainford D.J., Gradwell D.P. (eds.), *Ernsting's Aviation Medicine* (4th ed.), London: Edward Arnold, pp. 385–394.
- Rose P. (2002). *Forensic Speaker Identification*. Taylor & Francis, London.
- Rose P., Morrison G.S. (2009). *A Response to the UK Position Statement on Forensic Speaker Comparison*. *International Journal of Speech, Language and the Law*, 16: 139–163.
- San Segundo E., Foulkes P., French P., Harrison P., Hughes V., Kavanagh C. (2019). *The Use of the Vocal Profile Analysis for Speaker Characterization: Methodological Proposals*. *Journal of the International Phonetic Association*, 49(3): 353–380.

Schilling N., Marsters A. (2015). *Unmasking Identity: Speaker Profiling for Forensic Linguistic Purposes*. *Annual Review of Applied Linguistics*, 35: 195–214.

Trumper J. (2018). *Le telefonate. Le perizie foniche originali sulla strage di Peteano e sul caso Moro*. Padova: Cleup.

Villalba J. et al. (2016). *Analysis of Speech Quality Measures for the Task of Estimating the Reliability of Speaker Verification Decisions*. *Speech Communication*, 78: 42–61.

Yang F., Wang Z., Li J., Xia R., Yan Y. (2020). *Improving Generative Adversarial Networks for Speech Enhancement Through Regularization of Latent Representations*. *Speech Communication*, 118: 1–9.

Yu D., Gong Y., Picheny M.A., et al. (2023). *Twenty-Five Years of Evolution in Speech and Language Processing*. *IEEE Signal Processing Magazine*, 40(5): 27–39.

Zhang T., Sun Q., Zhang Z., Geng Y. (2026). *NRSRNet: Speech Enhancement Network Based on Noise Reduction and Restoration Module Under Extremely Low SNR Conditions*. *Applied Acoustics*, 241.

Le presenti Linee Guida sono state sottoscritte da associazioni scientifiche, enti di ricerca, organismi professionali e osservatori che riconoscono la necessità di promuovere trasparenza, rigore metodologico e uniformità operativa nelle attività di trascrizione, analisi fonetica e comparazione vocale in ambito forense. L'adesione rappresenta un impegno condiviso verso la qualità scientifica e la tutela delle garanzie processuali.

L'Unione delle Camere Penali Italiane (UCPI), primo firmatario del presente documento, ha partecipato ai lavori anche attraverso uno dei suoi membri, che è contestualmente componente dell'Osservatorio sulla Linguistica Forense. Tale doppia appartenenza ha favorito un dialogo diretto tra comunità scientifica e comunità professionale, contribuendo alla redazione condivisa delle Linee Guida.



Commissione sulla Linguistica Giudiziaria dell'Unione delle Camere Penali