

Progetto CLIPS

Corpora e Lessici di Italiano Parlato e Scritto

W1 – a3

(Definizione Corpora Speciali)

Title: *Definizione del database ortofonico in camera anecoica*

Document No: CLIPS/W1-a3/DCS/002

Document Name: RMDL002

Status: pubblico

Date: 10/3/2006

Authors: *Mauro Falcone* (Fondazione Ugo Bordini), Antonio Barone,
Alessandro Bonomi (Istituto Superiore C.T.I)

Realizzazione del database ortofonico in camera anecoica

Mauro Falcone
(Fondazione Ugo Bordoni)

Antonio Barone, Alessandro Bonomi
(Istituto Superiore C.T.I.)

INDICE

1.0 Introduzione	4
2.0 La camera anecoica	4
2.1 Caratterizzazione acustica della sala	5
3.0 Strumentazione utilizzata e sua disposizione	6
3.1 Alcune considerazioni sul microfono utilizzato	10
3.2 Procedure di calibrazione	10
4.0 Procedure esecutive adottate.....	10
5.0 Reclutamento dei parlatori	12
6.0 Alcune note sulla esecuzione delle sessioni di registrazione	13
7.0 Conclusioni.....	14
Ringraziamenti	15
Bibliografia.....	15
ALLEGATO A – Caratteristica del microfono utilizzato.....	16
ALLEGATO B – Fotografia della camera anecoica	17

1.0 Introduzione

In questo documento si descrive la strumentazione e le procedure utilizzate nella realizzazione del database vocale registrato in camera anecoica da parlatori professionisti nell'ambito del progetto CLIPS. La descrizione e la caratterizzazione dell'ambiente di registrazione (camera anecoica) viene altresì riportata come parte integrante della strumentazione utilizzata.

Le scelte operate nella disposizione e nell'utilizzo della strumentazione vengono dettagliatamente descritte e giustificate riportandone, qualora necessario, le procedure ed i risultati dei test di calibrazione.

Oltre ad una dettagliata descrizione tecnica e strumentale saranno inoltre descritte: le procedure esecutive utilizzate durante le sessioni di registrazione; i problemi incontrati e quindi le soluzioni utilizzate. Saranno infine discusse le esperienze compiute sia in campo squisitamente tecnico, sia in quello organizzativo/esecutivo.

Non saranno invece analizzati né i contenuti del database vocale, né la sua struttura. Per tali informazioni si rimanda il lettore ai relativi documenti (Falcone M., Barone A., Bonomi A., "Definizione e caratterizzazione di un database vocale ortofonico realizzato da parlatori professionisti in camera anecoica"), (Falcone M., Barone A., Bonomi A., "Struttura e contenuto del database vocale ortofonico realizzato da parlatori professionisti in camera anecoica").

2.0 La camera anecoica

Una delle caratteristiche principali del database in questione consiste nel fatto che questo è stato registrato in camera anecoica. Ricordiamo che con questo termine si intende un ambiente trattato acusticamente in modo molto particolare, e tale per cui risulti totalmente "sordo" nel senso di privo di qualsiasi riflessione e isolato da qualsiasi rumore esterno.

La camera anecoica viene solitamente utilizzata nei laboratori delle ditte che producono componenti audio per effettuare misurazioni senza interferenze ambientali sui diffusori e sui microfoni. Nel nostro caso invece è stata utilizzata per registrare un segnale vocale il più fedele possibile a quello "realmente" emesso dall'apparato fonatorio dei parlatori senza tener conto di interferenze dovute alle diverse riflessioni acustiche e/o ad eventuali rumori di fondo. L'acustica di una camera anecoica deve tuttavia considerarsi piuttosto "irreale" in quanto in natura non esistono situazioni che realizzano un'acustica ambientale simile.

La camera anecoica dell'Istituto Superiore C.T.I. è stata costruita nel 1979, ed è posta al terzo livello interrato dell'edificio ministeriale in viale America 201 di Roma. Più precisamente la sala si trova a circa nove metri sotto il livello stradale, e realizza un ambiente cubico di circa 7 metri di lato (spazio utile al limite dei cunei) ovvero una cubatura di circa 343 m³, con piano di calpestio in griglie metalliche amovibili posto a poco meno di un terzo dell'altezza. In questo modo una sorgente sonora posta al centro del piano di calpestio, e ad una altezza media di circa 1,5m si trova approssimativamente al centro della sala. Nella sala sono disponibili connessioni elettriche di alimentazione e per collegamenti audio, mentre non sono presenti sistemi di aria condizionata o altri sistemi meccanici o a motore elettrico.

L'ingresso alla camera anecoica è posto nella adiacente saletta regia, la quale è anche in comunicazione visiva con la prima per mezzo di vetrata che può essere chiusa per mezzo di un portello fonoassorbente. Al fine di isolare al meglio l'ambiente di registrazione, durante tutte le sessioni di registrazione il pannello fonoassorbente è stato lasciato montato impedendo così la comunicazione visiva ma garantendo, semmai necessario, un miglior isolamento da eventuali rumori esterni.

Il sistema di chiusura della camera prevede il posizionamento di un pannello isolante, scorrevole su di un binario del vano di chiusura, e la successiva chiusura del portellone a doppia guarnizione. Poiché l'apertura di detto pannello può risultare difficoltosa dall'interno si rende necessario, per motivi di sicurezza, monitorare continuamente la sala ed inoltre assicurare la presenza di almeno un paio di persone nella adiacente saletta di regia.

2.1 Caratterizzazione acustica della sala

La camera anecoica dell'Istituto Superiore C.T.I. è stata in passato oggetto di diversi studi (Vincenti A., Poscetti L., "Realizzazione di una camera anecoica nella sede dell'I.S.P.T.", Note Recensioni e Notizie, Vol. XXX, N.2, pg.82-91, Giugno 1981); si è voluto comunque effettuare una serie di nuove misure al fine di verificarne le attuali caratteristiche. Poiché le dimensioni e la struttura della sala sono rimaste le stesse nel tempo, in particolare doveva essere controllato il livello e la distribuzione del rumore di fondo della sala: fattore legato anche alla situazione ambientale ed edilizia circostante la sala stessa.

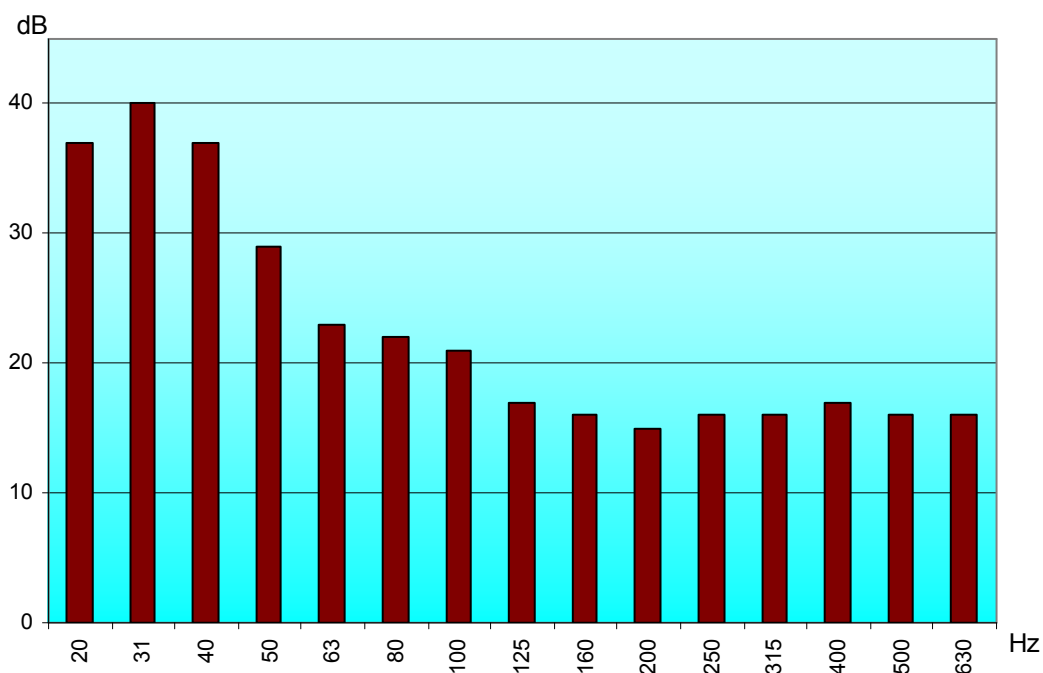


Figura 1 – spettro di rumore della camera anecoica per frequenza inferiori a 700Hz

A tal fine si è proceduto ad effettuare una serie di misure utilizzando un fonometro Brüel & Kjær 2230 con filtro a 1/3 d'ottava 1625. Nel grafico sottostante sono riportati i valori in dB ottenuti. Come si vede è riportata solo la parte in bassa frequenza dello

spettro in quanto il rumore al di sopra dei 500 Hz risulta sempre uguale o inferiore ai 16 dB. Dal grafico si evince anche livelli di rumore “significativi”, ovvero non ottimali per una camera anecoica, sono presenti solo a frequenze inferiori ai 50Hz, ed in particolare si rileva un picco di rumore intorno ai 32Hz pari a 40dB.

Questa “risonanza” alle basse frequenze non è di immediata interpretazione: tra le sue possibili cause vi è ovviamente quella di un problema legato alle risonanze delle strutture dell'edificio in cui la camera anecoica è situata. Essendo la camera realizzata in una struttura muraria cubica di lato di circa 10m ed essendo la velocità del suono circa di 330m/s, detta struttura cubica risulta avere una risonanza a circa 33Hz. In seconda ipotesi non da escludere la possibilità che detto rumore a basse frequenze non possa ricondursi a fenomeni di misurazione. Deve comunque considerarsi che il livello del rumore di fondo rimane di altissima qualità e risponde pienamente ai requisiti aspettati. Da una misura integrale su tutto lo spettro audio si ottiene infatti che il rumore di fondo della camera pesato in curvatura A è pari a 18dB (SPL). Con il termine SPL (Sound Pressure Level) si indica il valore di pressione per unità di superficie, è corrisponde alla normale procedura di misurazione di un livello sonoro. Discorso diverso invece deve farsi per la “pesatura” che al contrario viene operata al fine di “adattare” la misura affinché questa rispecchi in modo efficace una determinata grandezza fisica. La pesatura A è una speciale metodologia di calcolo per cui l'intensità di ogni valore sonoro misurato ad una specificata frequenza viene “pesato” (ovvero considerato più o meno importante) a seconda di quanto i suoni a quella determinata frequenza sono percepiti dall'uomo. Una misurazione del livello sonoro in pesatura A è quindi indice di quanto segnale (nel nostro caso rumore) sia effettivamente percepito all'ascolto. Una misurazione senza pesature (o più comunemente una misura lineare) è invece indice di quanto segnale sia fisicamente presente nell'ambiente.

3.0 Strumentazione utilizzata e sua disposizione

La strumentazione utilizzata nella realizzazione del database può dividersi essenzialmente in strumentazione di registrazione e strumentazione di memorizzazione/archiviazione. Con il termine di strumentazione di registrazione indicheremo tutti quei sistemi atti alla rilevazione, adattamento e memorizzazione (anche se in formato numerico) del segnale audio vero e proprio, ovvero tutta la catena di apparecchi atti all'effettivo scopo di registrazione (vedi tabella 1). I sistemi di memorizzazioni sono invece dei personal computer opportunamente equipaggiati con specifico hardware e/o software atti a memorizzare i segnali audio su opportune memorie di massa.

Apparecchio	Marca e modello	
Microfono	Brüel & Kjær 4009	
Amplificatore microfonico	Brüel & Kjær 4000	
Mixer	Soundcraft B200	
Registratore DAT	Panasonic SV 3700	
Amplificatore per cuffie	Samson Q5	
Cuffie	AKG 141 M	
Monitor audio	Fostex 6301 B	

Tabella 1 – elenco materiale analogico utilizzato nelle registrazioni

Non passeremo in rassegna le specifiche dei diversi dispositivi analogici, vogliamo solo ricordare che trattasi di strumentazione professionale accuratamente selezionata e comunemente utilizzata nei maggiori studi di registrazione. In particolare si deve sottolineare il fatto che, essendo il segnale audio di altissima qualità, si deve necessariamente ricorrere ad un dispositivo di rilevazione del segnale di altissima qualità al fine di non pregiudicare l'intero sistema di registrazione.

Particolare cura è stata posta nel posizionamento del microfono la cui asta di sostegno è stata fissata al soffitto della camera anecoica al fine di evitare il rischio che piccoli urti da parte dello speaker o leggeri movimenti del piano di calpestio, potessero essere trasmessi al microfono.

Per inciso vogliamo sottolineare il fatto che solamente il microfono, l'amplificatore microfonico ed un monitor audio erano disposti all'interno della sala anecoica, mentre tutto i rimanenti apparecchi erano situati nella adiacente saletta di regia. Un opportuno meccanismo garantiva il fatto che il monitor audio all'interno della sala anecoica fosse spento durante le normali fasi di registrazione. Il solo dispositivo attivo dentro la camera anecoica durante le registrazione risultava quindi essere l'amplificatore microfonico che, dovendo essere situato in modo prossimo al microfono, non poteva essere portato al di fuori della sala medesima.

Nella saletta regia erano situati inoltre un personal computer con scheda audio Creative "SoundBlaster Live! Platinum 5.1", e software "Sound Forge 4.0" della ditta Sonic Foundry, ed un notebook Compaq M700.

Il segnale proveniente dal microfono, collegato al suo preamplificatore, veniva inviato al mixer, ed in uscita da questo all'ingresso analogico del DAT. Il mixer operava esclusivamente da adattatore di impedenza e l'allineamento del livello di registrazione veniva impostato esclusivamente attraverso i comandi del DAT.

Sempre attraverso il mixer veniva gestito un canale indipendente per le comunicazioni audio con il parlatore e veniva prelevato il segnale da inviare all'amplificatore delle cuffie per il monitoraggio in regia.

L'uscita digitale (SPDIF) del DAT veniva inviata all'ingresso SPDIF della scheda audio del PC in modo da ottenere due copie contemporanee della registrazione, una delle quali, quella su DAT, poteva servire per recuperare dati eventualmente perduti (copia di backup).

Ad ogni pausa di registrazione (all'incirca ogni 10-15 minuti) i dati acquisiti sul personal computer venivano copiati sul notebook, attraverso una connessione diretta in rete tra i due personal computer.

Nella figura 2 è riportato uno schema riassuntivo della disposizione degli apparati utilizzati nella registrazione.

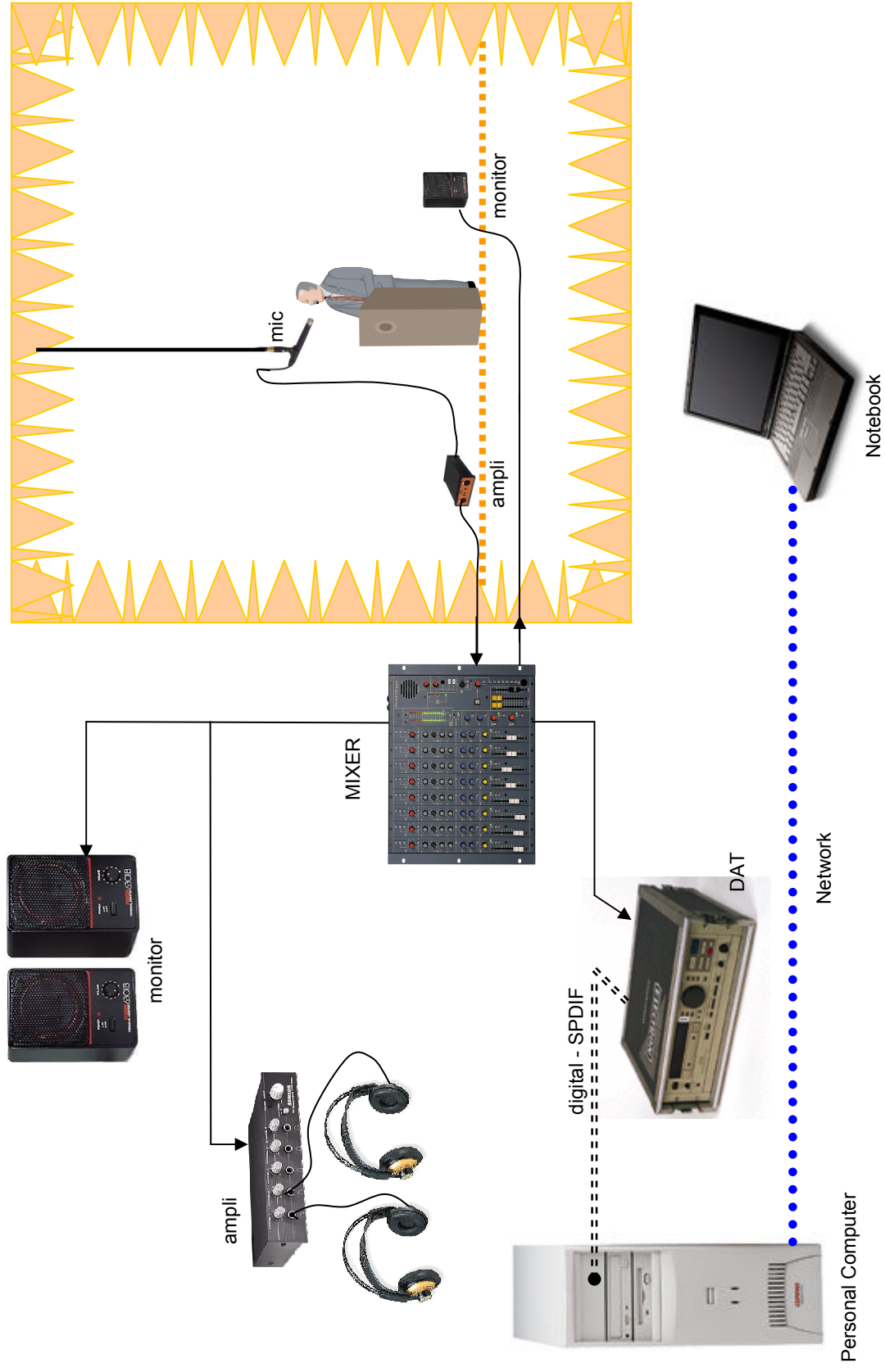


Figura 2 – schema di collegamento degli apparati utilizzati

3.1 Alcune considerazioni sul microfono utilizzato

Il microfono utilizzato è un modello a condensatore di altissima qualità con caratteristica di sensibilità di tipo omnidirezionale. Ad un primo esame l'uso di un microfono omnidirezionale per registrare voce può non sembrare la migliore soluzione o quantomeno può risultare inconsueta. L'esperienza maturata in precedenti acquisizioni audio in camera anecoica, ci hanno portato a ritenere l'impiego di tale microfono molto soddisfacente in considerazione del fatto che in camera anecoica si è in assenza di riverberi e soprattutto perché microfoni omnidirezionali denotano una sensibile diminuzione dell'effetto di prossimità, tipico dei dispositivi con caratteristica del tipo a cardiodide o ipercardiode. Altro aspetto non secondario è il fatto che la sensibilità degli omnidirezionali è buona anche per segnali "fuori asse", e questo permette di limitare le variazioni di livello dovute a spostamenti involontari della testa da parte del parlatore.

3.2 Procedure di calibrazione

Le procedure di calibrazione adottate ad ogni inizio di registrazione sono state di due tipi. La prima prevedeva la registrazione di un tono sinusoidale puro ad una frequenza di 1 kHz emesso da una bocca artificiale Brüel & Kjær con una pressione di 90 dB (SPL), a circa 30cm dal microfono. Tale livello di pressione era misurato con un fonometro Brüel & Kjær 2230. Una volta verificato il livello attraverso il fonometro, con detto segnale si allineavano i livelli di registrazione del DAT e del software Sound Forge in modo da ottenere un livello di circa -15dB su entrambi i loro misuratori di livello. Così calibrati sia il sistema DAT, sia il sistema di acquisizione su personal computer, si poteva esser ragionevolmente sicuri che la dinamica a disposizione ci metteva al riparo da fenomeni di saturazione per i livelli di parlato alla normale distanza tra parlatore e microfono.

La seconda procedura consisteva nella lettura, da parte dello speaker, del testo di calibrazione per mezzo del quale si verificava la bontà del livello immediatamente prima dell'inizio delle sessioni di registrazione. Nel caso in cui si rilevava un livello non soddisfacente si operava esclusivamente sull'attenuatore del preamplificatore microfonico, situato all'interno della camera anecoica, il cui intervento copriva un intervallo di 12dB (da 0dB a -12dB) a passi di 6 dB.

Ad esclusione di un solo parlatore avente un livello di voce leggermente meno potente del normale e che conseguentemente è stato registrato con una attenuazione a -6 dB, per tutti gli altri la regolazione del preamplificatore microfonico è stata impostata a -12 dB.

4.0 Procedure esecutive adottate

Dopo aver definito il materiale e le procedure di presentazione da adottare (Falcone M., Barone A., Bonomi A., *"Definizione e caratterizzazione di un database vocale ortofonico realizzato da parlatori professionisti in camera anecoica"*), e la strumentazione da utilizzare nelle registrazioni, si è passati alla fase di realizzativa disponendo tutto il materiale hardware nella sala regia e nella camera anecoica;

effettuando per diversi giorni consecutivi test di calibrazione e di registrazione al fine di valutare l'affidabilità della soluzione proposta e contemporaneamente per saggiare la qualità del materiale vocale acquisito. Si sono inoltre effettuate una serie di registrazioni pilota simulando una reale sessione di registrazione. Attraverso queste esperienze si sono messe a punto tutta una serie di fattori quali l'illuminazione della camera anecoica, la predisposizione del leggio e la sua illuminazione, la disposizione dei personal computer e le procedure di registrazione nonché le procedure di controllo da utilizzare in tempo reale durante le sessioni di registrazione. A tal fine si è deciso di effettuare un doppio controllo in tempo reale da parte di due diversi ascoltatori che riportavano su apposito modulo tutte le frasi che a loro avviso necessitavano essere ripetute. Questa fase pilota ha permesso di controllare anche le stime fatte sulla durata delle sessioni di registrazione, nonché evidenziare e correggere una serie di piccoli problemi che potevano mostrarsi durante la registrazione.

Solo dopo aver effettuato una serie di prove (della durata approssimativa di una settimana) e avuta la certezza che le scelte operavano in un ambiente di registrazione affidabile e con i requisiti richiesti si è passati alle registrazioni vere e proprie con parlatori professionisti. Poiché per ogni sessione si è stimato un tempo totale di 3-4 ore, si è deciso di realizzare registrazioni di un parlatore al giorno e soltanto di mattina. Le registrazioni sono state effettuate nei dieci giorni lavorativi nel periodo dal 14 al 27 Marzo 2001.

Per ogni sessione di registrazione venivano effettuate le operazioni di calibrazione sopra descritte. Quindi, sebbene tutti i parlatori fossero preventivamente stati avvisati sul tipo di prestazione a loro richiesta, la prima operazione fatta è stata quella di spiegare il lavoro da farsi al parlatore e di rispondere a sue eventuali domande. Quindi si mostrava la sala anecoica al parlatore, facendolo ambientare all'acustica e spiegando i diversi sistemi di comunicazione tra la sala e la regia. Infine si operava il posizionamento del microfono a circa 30cm dalla bocca del parlatore e si controllava la postura del parlatore davanti al leggio dando così anche indicazione per una corretta posizione dei piedi affinché si mantenesse una distanza corretta dal microfono. Tutte le registrazioni sono state effettuate stando in piedi davanti al leggio, costituito da un'asta saldamente ancorata a terra e da un supporto dalle dimensioni atte a contenere comodamente fogli di formato A4 in posizione orizzontale. Una luce di colore rosso situata in prossimità del leggio segnalava (quando accesa) al parlatore che si era in fase di registrazione. Durante le registrazioni in camera anecoica era disponibile dell'acqua naturale nel caso il parlatore avesse necessità di bere.

Tipicamente i parlatori non hanno mostrato particolari problemi di adattamento e comunque, nonostante le ripetute assicurazioni, si è riscontrato, in alcuni soggetti, una "comprensibile" preoccupazione per il fatto di essere ermeticamente chiusi in una camera isolata. Spesso i parlatori richiedevano maggiori informazioni sulle modalità di lettura e di pronuncia da effettuare. Anche al fine di meglio chiarire questo problema si faceva preventivamente ascoltare ai parlatori (ad esclusione del primo parlatore) un campione del materiale registrato dal primo parlatore che, lo ricordiamo, ha effettuato le registrazioni sotto la supervisione di esperto fonetico.

Dopo l'espletamento di tutte le pratiche relative all'istruzione del parlatore, si passava quindi alla fase di registrazione del testo di calibrazione e quindi del corretto posizionamento del valore di attenuazione sul preamplificatore microfonico. Quindi si iniziava la serie di sottosessioni secondo lo schema previsto. Durante le registrazioni il parlatore non veniva mai fermato dalla regia a meno di persistenti e sistematici errori. Le singole frasi che venivano evidenziate come sbagliate o "dubbe" da almeno uno delle due persone che ascoltavano in regia venivano tutte segnate su un opportuno modulo. Al termine della registrazione veniva quindi compilata una lista con tutte le frasi segnalate dagli ascoltatori in regia e si eseguiva una ulteriore sessione di registrazione in cui dette frasi venivano ripetute, anche più volte sino ad ottenere una corretta realizzazione. Solo in questa ultima fase, e nel caso di ripetizioni multiple la regia entrava in contatto con il parlatore per evidenziare i problemi legati alla pronuncia delle frasi in questione.

Al termine di ogni sessione i dati registrati venivano archiviati su CD-ROM. Si avevano quindi i dati registrati su nastro DAT, i file acquisiti su personal computer e copiati su notebook, ed infine ulteriore copia su CD-ROM. Questa diversificazione e molteplicità dei dati garantiva la loro integrità anche a fronte di possibili danneggiamenti dei supporti o di uno o più dei sistemi di registrazione o di memorizzazione del segnale vocale.

5.0 Reclutamento dei parlatori

I parlatori sono stati reclutati attraverso una cooperativa di doppiatori dell'area romana. La maggior parte di loro oltre a effettuare attività di doppiaggio, effettua correntemente attività di attore in ambito teatrale e/o televisivo. Tutti i parlatori hanno effettuato scuola di dizione.

N	Nome	Luogo di nascita	Anno di nascita	Riferimento nel DB
1	Alessia Ranaldi	Roma	1970	F1
2	Marcello Caponi	Roma	1971	M1
3	Pietro Bontempo	Roma	1964	M2
4	Giuseppina De Maria	Genova	1964	F2
5	Cinzia Villari	Libia	1965	F3
6	Giordano Aquilini	Lucca	1967	M3
7	Luciano Melchionna	Roma	1967	M4
8	Maria Carolina Salomè	Roma	1965	F4
9	Alessio Buccolini	Roma	1971	M5
10	Elisabetta Cianchini	Roma	1967	F5

Tabella 2 – Dati anagrafici dei parlatori utilizzati nel database ortofonico

Non è stato possibile effettuare nessun tipo di pre-screening dei parlatori; tuttavia era previsto che, nel caso un parlatore fosse ritenuto "inadatto" al compito richiesto, questo fosse sostituito da altro parlatore. In ogni caso tutti e dieci i parlatori sono stati giudicati idonei, e quindi non è stato necessario ricorrere all'utilizzo di parlatori aggiuntivi. Nella tabella seguente sono riportate alcune indicazioni anagrafiche dei parlatori. Le registrazioni sono state effettuate nell'ordine sopra riportato; sono iniziate il giorno 14/03/2001 e sono terminate il giorno 27/03/2001, per un totale di dieci giorni lavorativi.

6.0 Alcune note sulla esecuzione delle sessioni di registrazione

Al termine della registrazione del database possiamo operare alcune considerazioni sulla base delle esperienze fatte durante le due settimane di lavoro, nonché sull'intera attività di raccolta. In linea generale possiamo dire che sia per quanto riguarda la tempistica delle sessioni, sia per quanto riguarda la fase di spiegazione del compito e di adattamento all'ambiente di registrazione, le stime e le procedure originariamente stabilite, sebbene operate cercando di avere dei margini di garanzia piuttosto larghi, si sono rivelate corrette e certamente non sovrastimate. Infatti l'impatto con l'acustica della camera anecoica, ed anche i fattori psicologici relativi al fatto di rimanere isolati in un ambiente così particolare hanno creato dei problemi che, sebbene previsti, sono risultati piuttosto rilevanti. Un fattore originariamente sottostimato è invece stato quello dell'apprendimento o meglio dell'adeguamento al compito specifico da parte dei parlatori. Trattandosi di parlatori professionisti, questi erano facilmente in grado di adattare la pronuncia, la velocità di eloquio e la prosodia della frase molto facilmente. Tuttavia i parlatori, chi più e chi meno, avevano anche una "personalità" professionale che li portava a voler seguire quella che "a loro giudizio" era la migliore o la corretta modalità di pronuncia. Si poteva quindi avere da parte del parlatore un atteggiamento psicologico negativo in quanto forzato a realizzare delle frasi diversamente da quella che loro (professionisti) ritenevano la "corretta" pronuncia. Anche per quanto riguarda lo stile di parlato si poteva osservare a volte una perdita di concentrazione che portava il parlatore ad assumere uno stile di parlato più "recitato" e troppo "fluidico". Ovviamente quando questi fenomeni venivano individuati dagli ascoltatori in regia si cercava di avvertire il parlatore. Non è comunque risultato facile controllare tali fattori in tempo reale (tale compito è molto più semplice in post produzione ma ovviamente in questo caso non è possibile recuperare gli errori se non operando nuove sessioni di registrazione).

Durante la sessione di registrazione in sala regia erano sempre presenti almeno quattro persone: una addetta al controllo della strumentazione analogica (mixer, DAT, ecc.); due addette all'ascolto (in realtà il segnale era diffuso anche in ambiente, quindi tutti i presenti potevano ascoltare il segnale) ed almeno una persona per sopperire ad altre eventuali necessità in quanto durante la registrazione le postazioni di controllo non potevano mai essere disattese per alcun motivo.

In generale si è riscontrato che il controllo in tempo reale della qualità del segnale risulta un compito molto difficile e non privo di errori. Inoltre bisogna considerare che oltre a frasi "chiaramente" pronunciate in modo diverso da quello richiesto, vi sono una serie di fattori (velocità di eloquio, stato emozionale, coarticolazione, prosodia,

ecc.) che spesso possono semplicemente dare adito a dei “dubbi”, ed il cui controllo richiede una analisi più dettagliata. Lo stesso discorso vale per determinati rumori di fondo che si riesce ad individuare solo in fase di post registrazione. Per questi motivi si ritiene, sulla base delle esperienze fatte, che la “garanzia” di un materiale che rispecchi completamente delle specifiche di qualità molto stringenti come nel nostro caso sia possibile solo con una serie di sessioni e di controlli di qualità ripetute. Ovviamente questo comporta un moltiplicarsi dei costi in termini di lavoro di analisi del materiale e in termini di reclutamento dei parlatori.

7.0 Conclusioni

Nel presente documento si sono esaminate le operazioni, i problemi e le scelte operate durante la preparazione e la realizzazione del database. Quindi, oltre a descrivere gli apparati e le metodologie utilizzate, si è voluto anche presentare una serie di problemi di carattere meno tecnico ma altrettanto importanti e che, in accordo con la esperienza fatta, possono influire profondamente con i risultati della raccolta e sulla qualità del materiale.

La realizzazione del database in camera anecoica con parlatori professionisti ha costituito una parte a sé stante del progetto CLIPS, e non è di conseguenza possibile comparare le particolari esperienze fatte in questo ambito con quelle delle altre raccolte di segnale vocale. Considerata la particolarità del compito, la alta qualità del segnale e la specificità dei parlatori, si è cercato di raggiungere dei risultati che non solo fossero aderenti alle specifiche dettate, ma che in qualche modo potessero considerarsi anche i migliori ottenibili con le risorse disponibili.

Se da un punto squisitamente tecnico è stato possibile raggiungere risultati ottimali, grazie anche alla esperienza acquisita in precedenti raccolte di dati, alla scrupolosa scelta degli apparecchi e ad una attenta e puntale verifica della intera catena di acquisizione del segnale attraverso una ripetuta serie di misure e di calibrazioni; al contrario da un punto di vista organizzativo/esecutivo oltre ad una scrupolosa preparazione e definizione delle modalità esecutive e di simulazioni di registrazioni si è dovuto, di volta in volta, cercare di ottenere i risultati “migliori” da ciascun parlatore. Gli aspetti emotivi, la problematicità di ambientazione alla camera anecoica, la non sempre chiara comprensione dello stile di parlato richiesto, congiuntamente alla inevitabile tendenza a caratterizzare lo stile di parlato da parte del “doppiatore/attore”, hanno sicuramente costituito una serie di problemi che si sono dovuti affrontare lungo il corso delle registrazioni e che si riproponevano in modalità e aspetti diversi da parlatore a parlatore. Conseguentemente anche le soluzioni a tali problemi non necessariamente sono state le stesse per tutti i parlatori.

In conclusione possiamo dire che i risultati ottenuti sono conformi a quanto originariamente aspettato, e che non si sono incontrati significativi problemi nella realizzazione del database.

Da un punto di vista tecnico va evidenziato che alcuni problemi riguardo al fatto che ci si trovava in camera anecoica dove anche il più piccolo rumore (ad esempio tic nervosi come strofinare un’unghia contro l’altra, o aggiustarsi il vestito, ovvero altri rumori dovuti a braccialetti, collane e così via) sono stati originariamente

sottovalutati. Inoltre la possibilità di avere un sistema di registrazione numerico a 24bit avrebbe facilitato e semplificato le operazioni di calibratura garantendo, semmai necessario, una maggiore robustezza della catena di acquisizione.

Da un punto di vista non-tecnico dobbiamo notare come il diverso aspetto emotivo dei parlatori abbia condotto a difficoltà che hanno conseguentemente portato ad una qualità abbastanza diversa da un parlatore all'altro. Anche per quanto riguarda la comprensione del compito, ovvero delle istruzioni e dei suggerimenti da fornire al parlatore, si è riscontrata un'ampia variabilità nonostante i soggetti fossero tutti professionisti e in qualche modo tutti avvezzi alle stesse modalità di lavoro (che però, lo ricordiamo, sono piuttosto diverse da quelle del compito specifico). Il fattore umano è risultato pertanto quello più difficile da trattare e da controllare nella realizzazione del database ortofonico.

Ringraziamenti

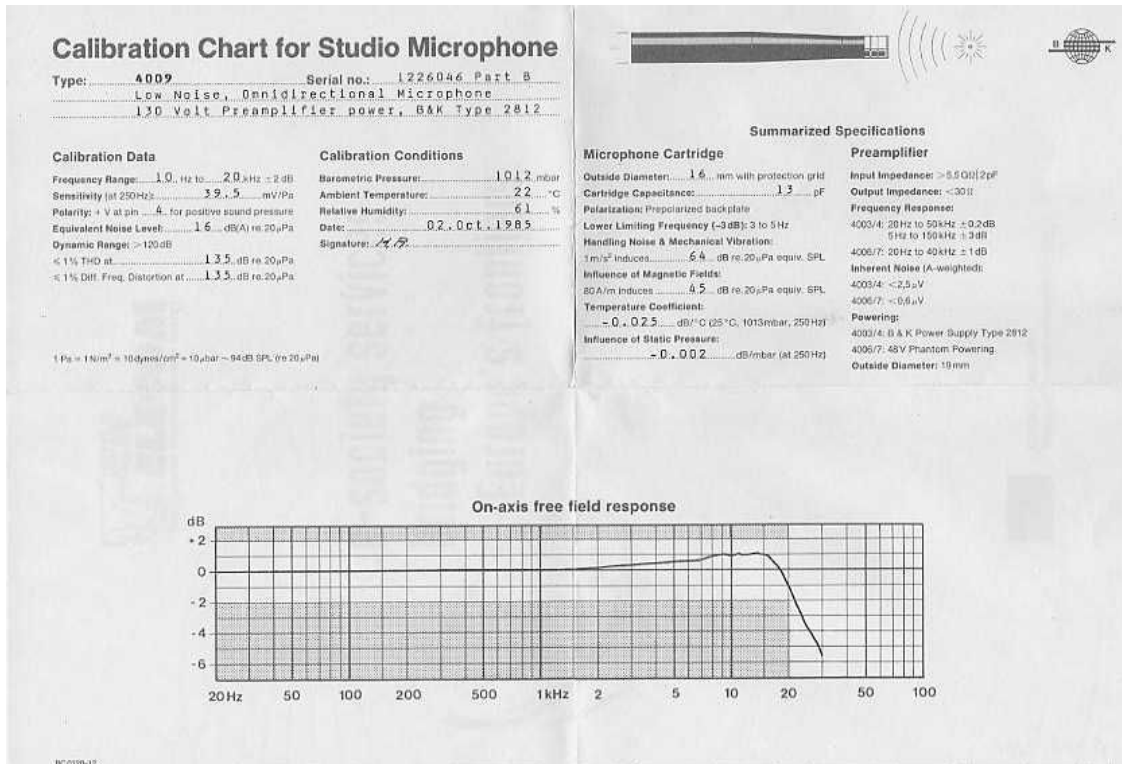
Si ringraziano tutte le persone che nei diversi giorni di esecuzione delle registrazioni, e nella fase di allestimento, hanno assistito e supportato le normali procedure di registrazione, ed in particolare Angelo Ciavardini e Stefano Renzi dell'Istituto Superiore T.C.I.

Si vuole inoltre ringraziare la dott.ssa Renata Savy che ha supervisionato l'inizio delle registrazioni, fornendo peraltro le opportune indicazioni relativamente alla pronuncia fonetica dei testi letti dai parlatori.

Bibliografia

- [1] Falcone M., Barone A., Bonomi A., "*Definizione e caratterizzazione di un database vocale ortofonico realizzato da parlatori professionisti in camera anecoica*", documento CLIPS, 2001
- [2] Vincenti A., Poscetti L., "*Realizzazione di una camera anecoica nella sede dell'I.S.P.T.*", Note Recensioni e Notizie, Vol.XXX, N.2, pg.82-91, Giugno 1981

ALLEGATO A – Caratteristica del microfono utilizzato



ALLEGATO B – Fotografia della camera anecoica

