

IV.6. L'ESAME SPETTROACUSTICO «AMBULATORIALE»

G. Migliori

L'esame spettroacustico della voce può essere oggi agevolmente realizzata nell'ambulatorio specialistico ORL-foniatrico, grazie a prodotti software «low-cost».

Questi software di analisi vocale (Dr. Speech – Multi-Speech – SoundScope) permettono, in ambiente Windows o Macintosh, di catturare ed analizzare il campione vocale (DSP – digital signal processing) nell'ambito di un PC multimediale.

La nostra esperienza si è realizzata con il «Multi-Speech 3700» della Kay Electronics Corp.

Non è indispensabile che il PC sul quale installare il software abbia un processore particolarmente potente; la Kay richiede un Pentium maggiore di 266 MHz con almeno 16 MB di memoria Ram, nettamente inferiore all'attuale standard medio di un comune PC multimediale (Pentium 3 a 700 MHz con 64 MB di memoria Ram).

Come tutti i programmi basati su un hardware multimediale anche quelli in oggetto sono alquanto limitati dalle specifiche performance della scheda sonora, soprattutto nei riguardi della qualità del rapporto segnale/rumore (SNR: signal-to-noise ratio), abitualmente 40-60 dB contro gli 86 dB del CSL 4300B.

La maggior parte delle schede sonore abitualmente installate sui computer multimediali non hanno particolari funzioni per «l'audio input» che è generalmente alquanto semplice e quasi sempre descritto nelle sue caratteristiche senza dettaglio. È pertanto indispensabile, per effettuare un corretto esame spettroacustico, installare sul proprio PC schede sonore «professionali» che possano migliorare le caratteristiche dell'acquisizione audio, supportando adeguatamente un microfono ad alta sensibilità.

PRELIEVO DEL SEGNALE

È ottimale che la cattura del segnale vocale avvenga in una cabina silente; ove non fosse disponibile è raccomandabile che il rumore di fondo dell'ambulatorio non superi i 40 dB per non inficiare la successiva analisi del segnale vocale.

Il microfono professionale consigliato dalla Kay per il Multi-Speech è il «AKG Acoustic Model C-410» posizionato a 5 cm dalle labbra (per evitare interferenze del rumore ambientale) con una angolazione di 45° (per evitare perturbazioni del flusso aereo).

L'uso di un preamplificatore microfonic esterno protegge il basso livello del segnale microfonic dal rumore del computer, ottenendo un aumento del rapporto segnale/rumore di circa 6 dB.

Fra i microfoni a condensatore di ultima generazione meritano di essere ricordati, per le specifiche caratteristiche tecniche, il «Rode NTK», lo «Shure KSM-44» e lo «Sennheiser MKH800».

Per uniformare il metodo di campionamento, su una acquisizione vocale di almeno 6 secondi, (/a/ tenuta) vanno presi in considerazione solo i 2 secondi centrali sia per evitare le interferenze dell'attacco e dello stacco vocale che per analizzare almeno 110 cicli di vibrazione delle corde vocali. Campioni vocali di 3 secondi permettono di analizzare circa 200 cicli di vibrazione delle corde vocali anche in presenza di voci molto gravi. L'intensità di emissione della /a/ deve essere eguale alla voce di conversazione, senza variazioni di intensità o frequenza, possibilmente fra i 50 e i 65 dB; infatti emissioni vocaliche eguali o maggiori di 70 dB di intensità, tendono a «saturare» l'oscillogramma alterando la successiva valutazione dei parametri vocali quali jitter, shimmer ed H/N ratio.

Le schede sonore multimediali hanno tre frequenze di campionamento standard (11025 Hz, 22050 Hz e 44100 Hz); con il Multi-Speech possiamo effettuare acquisizione e campionamento a qualsiasi valore fino alla frequenza più alta supportata dalla scheda sonora. Alte frequenze di campionamento comportano campioni vocali di notevoli dimensioni ed occupanti quindi grandi quantità di memoria nell'HD (disco rigido); la capacità di memoria degli HD attuali (30-40 GB) e la diffusa presenza di sistemi di back-up a basso costo (masterizzatori) hanno risolto il problema dell'acquisizione di dati ad alte frequenze di campionamento e della loro successiva archiviazione.

Il «sampling rate» (frequenza di campionamento) ottimale è di 20 kHz (le informazioni linguistiche sono significative fino a circa 8-10 kHz) ma per l'analisi di classi di fonemi come le vocali, l'informazione pregnante è tutta contenuta al di sotto dei 4-5 kHz quindi è sufficiente utilizzare una frequenza di campionamento di 10-12 kHz (Ferrero FE, Accordi M, 1998).

Una volta ottimizzato il sistema di acquisizione (livello del rumore ambientale, caratteristiche del PC, scheda sonora, microfono professionale con preamplificatore), il metodo e la frequenza di campionamento, va effettuata una «normativa personale» sulla apparecchiatura, analizzando da 50 a 100 soggetti eufonici con anamnesi ed obiettività negative per patologia vocale. I valori ottenuti con la «normativa personale» debbono rappresentare il «default» di riferimento del sistema per la successiva valutazione dei casi patologici.

Il Multi-Speech permette anche l'analisi del segnale vocale in «tempo reale» (con il modulo opzionale Real-Time Spectrogram). È possibile ottenere uno spettrogramma (con scala di grigi o a colori) del segnale vocale a «finestra di analisi» (window o frame) lunga o corta, a seconda che si voglia porre in evidenza la componente armonica o quella formantica. Inoltre le applicazioni di base permettono di calcolare la Fo (frequenza fondamentale), il Pich synchronous LPC, il Cepstrum, l'Energy contour (intensità vocale), l'LPC (linear predictive coding), l'FFT (fast Fourier transform) (Figg. 1 e 2). Non manca la possibilità di ottenere i valori di Shim-

te nella comprensione della propria affezione e nella accettazione del trattamento proposto.

BIBLIOGRAFIA

- ¹ Barillari U, Previdero G. *Valutazione dell'indice spettrografici H/N*. Acta Phon Lat 1998;20:216-23.
- ² Ferrero FE, Accordi M, Accordi D, de Filippis C, Staffieri A, et al. *Semeiotica avanzata della funzione vocale*. Acta Phon Lat 1998;20:33-58.
- ³ Ferrero FE, Vages K. *Perturbazioni a breve e lungo termine del segnale verbale*. Atti XXII Conv. AIA 1995.
- ⁴ Karnell MP. *Laryngeal perturbation analysis: minimum length of analysis window*. J Speech Hearing Res 1991;34:544-8.
- ⁵ Multi – Speech Model 3700 – Operation Manual – Kay Elemetrics Corp.