

Lorenzo Cioni
Multi-Speech e ESPS: tecniche di scripting per l'analisi acustica
(*work in progress*)

1. INTRODUZIONE

Scopo del presente lavoro è quello di illustrare l'utilità e l'uso di due tipologie di tecniche di scripting, che diremo rispettivamente tecniche sincrone e tecniche asincrone, per l'esecuzione di operazioni ripetitive per l'analisi acustica di segnali di parlato.

Nei paragrafi che seguono vengono inizialmente svolte alcune considerazioni sulla necessità di standardizzare l'esecuzione di sequenze di operazioni, si passa quindi alla definizione delle tecniche di scripting asincrone e sincrone per terminare con la presentazione di due case study relativi ai programmi Multi-Speech™ e ESPS™.

ESPS sta per "Entropic Signal Processing System", software package che rappresenta uno strumento potente e versatile per l'elaborazione dei segnali in ambiente Unix. ESPS può essere visto come una estensione di Unix ed è composto da più di duecento programmi utilizzabili dall'utente singolarmente o all'interno di script.

Multi-Speech è un prodotto software utilizzabile per l'acquisizione, l'analisi acustica e il playback di segnali di parlato e mette a disposizione degli utenti un ricco set di operazioni che possono essere raggruppate all'interno di script che, in ambiente Multi-Speech, sono denominati macro.

La differenza sostanziale, dal punto di vista del presente lavoro, risiede nelle differenze concettuali che esistono fra gli script di ESPS e le macro di Multi-Speech dal momento che i primi sono una estensione delle seconde: in prima approssimazione si può affermare che se una macro è una lista di comandi che permette agli utenti di ripetere sempre nello stesso ordine una serie di operazioni uno script in più ha strutture di controllo e istruzioni per l'accesso ai file.

2 L'ESECUZIONE DI OPERAZIONI RIPETITIVE PER L'ANALISI ACUSTICA

Le operazioni che, di solito, sono eseguite sui file di parlato sono, in genere, operazioni di analisi (calcolo dello spettro e dello spettrogramma, estrazione di F0 e analisi LPC) o di calcolo di durate nel dominio del tempo (VOT e Rise Time, ad esempio) eseguite su file o porzioni di file individuate da etichette posizionate dall'utente in modo da evidenziare le sezioni dei file in cui sono localizzati i fenomeni acustici di interesse. L'esecuzione di tali operazioni su grossi corpora di dati per la validazione di teorie linguistiche richiede l'adozione di cautele procedurali che sono tanto più gravose quanto più ampia è la mole dei dati da sottoporre ad analisi.

Una macro o uno script sono, in prima approssimazione, dei file di comandi ovvero dei file di testo contenenti successioni di comandi che vengono eseguiti sempre nello stesso ordine e con gli stessi valori dei parametri significativi.

L'adozione di macro e di script consente, pertanto, la ripetizione di successioni di comandi sempre nello stesso ordine. In tal modo si ha la possibilità di analizzare senza difficoltà grossi corpora di dati e di implementare protocolli di ricerca che, applicati a tali corpora, danno origine a dei risultati che non risentono dell'ordine di esecuzione delle singole operazioni. Si ha così la certezza di non aver introdotto variazioni accidentali nei parametri utilizzati per l'analisi o nella successione delle operazioni, variazioni che possono

falsare i risultati delle analisi portando, al limite, alla validazione erronea di teorie linguistiche.

3 TECNICHE DI SCRIPTING ASINCRONE E SINCRONE

Le tecniche di scripting¹ utilizzabili per l'esecuzione di operazioni di cui al punto precedente possono essere classificate come asincrone oppure sincrone.

Nel caso di tecniche sincrone l'esecuzione delle operazioni contenute nello script o nella macro è guidata da azioni dell'utente che esegue o lo script o la macro dall'interno di un programma per il DSP (per Multi-Speech è il programma stesso, nel caso di ESPS è possibile usare sia EnSig² sia waves+, [5]). L'utente, in questo caso, imposta lo spazio di lavoro, carica uno o più file necessari, setta su di essi dei marker che individuano istanti temporali ed usa tali marker per eseguire operazioni che producono nuovi file e operazioni che li visualizzano arricchendo lo spazio di lavoro ed eventualmente creando risultati numerici che possono essere salvati in file di testo.

Nel caso di tecniche asincrone l'utente esegue uno o più script in modo disgiunto dal programma per il DSP ottenendo, a partire da file di solito contenenti porzioni di parlato, file il cui contenuto è il risultato di operazioni di analisi e che possono essere visualizzati, nel caso specifico, sia con comandi di ESPS sia all'interno di EnSig e/o waves+. Nel caso di waves+ ([7]) un possibile protocollo che può essere utilizzato è il seguente: l'utente visualizza un file di parlato, posiziona su di esso etichette che vengono salvate in file di testo e alle quali corrispondono marker temporali e poi usa tali marker (e gli associati valori di posizione) in script ad hoc per l'esecuzione di operazioni su quel file di parlato (e altri file etichettati in modo simile).

4 CASE STUDY 1: MULTI-SPEECH

4.1 Breve descrizione del programma

Multi-Speech ([6]) è un prodotto software in ambiente Windows per l'acquisizione, l'analisi, la visualizzazione e l'ascolto di segnali di parlato. Il programma è caratterizzato da un ricco set di comandi accessibili sia attraverso menù sia attraverso parole chiave utilizzabili sia sulla linea di comando del programma stesso sia all'interno di macro.

Oltre ai menù, il programma è caratterizzato da un set di pulsanti posizionati all'interno di una toolbar (vedi figura 1), pulsanti che consentono di accedere rapidamente ai comandi di utilizzo più comune.

Il programma interagisce con la scheda audio per l'acquisizione e l'ascolto di segnali di parlato e il settaggio dei parametri per l'acquisizione fra i quali la frequenza di campionamento e la durata massima del file da acquisire.

¹ Indicheremo nel seguito con la dizione di "tecniche di scripting" le tecniche necessarie per la stesura di script e di macro. La distinzione sarà chiarita, di volta in volta, dal contesto.

² EnSig è, in pratica, una interfaccia tramite la quale l'utente può accedere sia a waves+ sia a ESPS sia ai programmi per la gestione della scheda audio sul sistema ospite mentre xwaves è il programma che fornisce l'interfaccia grafica per l'accesso a waves+.

Multi-Speech possiede un ricco set di tool per l'analisi che però, se acceduti da menù, presuppongono che il segnale da analizzare (in tutto o in parte) sia stato visualizzato in una finestra vuota.

I menù di Multi-Speech contengono comandi raggruppati in base alla funzione svolta. Si hanno pertanto i menù "File" (per l'accesso ai file di parlato compresa l'acquisizione di nuovi segnali), "Edit" (per l'editing della forma d'onda), "View" (per la visualizzazione dei dati sullo schermo), "Speak" (per la riproduzione di un segnale visualizzato), "Analysis" (per l'esecuzione sui segnali di operazioni di analisi), "Tags" (per l'inserimento di puntatori nei file), "IPA" (per la gestione dei simboli IPA nelle trascrizioni), "Macros" (per la gestione delle macro) e altri ("Log", "Options", "Window" e "Help").

I comandi contenuti nei singoli menù corrispondono a sequenze di comandi inseribili dall'utente nelle macro per ottenere risultati equivalenti.

Ad esempio il comando "New (Record)" del menù "File" può essere emulato in una macro utilizzando la successione di comandi seguente:

```
PURGE =  
CAPTURE  
SOURCE =
```

Il primo di tali comandi pulisce la finestra attiva dal suo contenuto, il secondo consente la registrazione di un nuovo segnale e il terzo lo definisce come il segnale attivo per ulteriori elaborazioni.

Sostituendo al comando CAPTURE il comando LOAD si ha una emulazione del comando "Open (Load)" sempre del menù "File".

In modo analogo si può attivare una finestra A (con il comando USE A) e copiare in essa il contenuto di un'altra (COPY B 0 *) oppure copiare una porzione, individuata da due marker di inizio e fine selezione (SS e SE rispettivamente), del segnale visualizzato in una finestra (ad esempio la B) in un'altra, resa attiva con un comando di USE, con il comando COPY B SS SE.

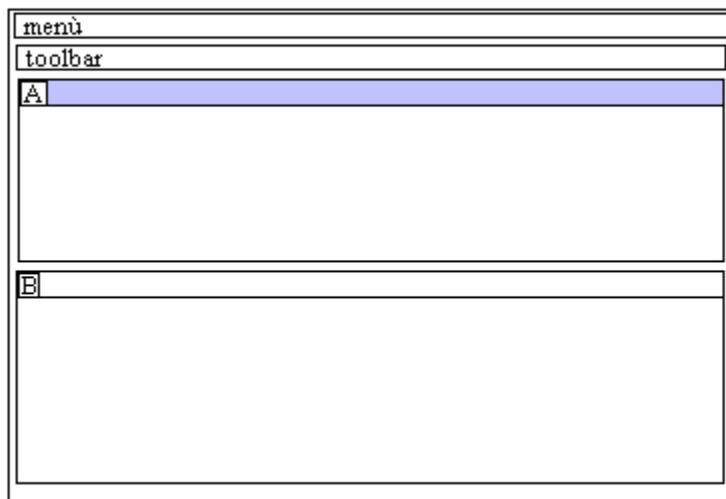


Figura 1

Dagli esempi dovrebbe essere evidente come ai comandi dei menù corrispondano comandi singoli o gruppi di comandi utilizzabili all'interno delle macro, comandi che fanno uso di simboli per l'accesso ai dati quali:

0 * per individuare la totalità dei dati contenuti in un file;

< > per individuare i dati visualizzati;

SS SE per individuare i dati selezionati

! per individuare la finestra attiva corrente

e altri per i quali si rimanda a [6].

Altri esempi significativi sono quelli relativi ai comandi del menù "View" che possono essere emulati quasi tutti con il comando SHOW seguito da vari parametri che fanno riferimento alle diverse porzioni del segnale da visualizzare e quelli relativi al menù "Speak" che possono essere emulati quasi tutti con il comando SPEAK seguito da vari parametri che fanno riferimento alle diverse porzioni del segnale da ascoltare.

Sempre con riferimento alle equivalenze fra comandi dei menù e comandi utilizzabili all'interno di macro si fa notare come la selezione con il mouse di una finestra per renderla attiva equivalga all'uso del comando USE. Con riferimento alla figura 1 la selezione della finestra etichettata come "A" la si può ottenere sia con il mouse sia, all'interno di una macro, con il comando USE A.

La successione seguente:

USE A

LOAD

USE B

PITCH A 0 *

permette di caricare (mediante una finestra di dialogo classica dell'ambiente Windows) un file di speech nella finestra etichettata "A" di cui calcola il pitch che viene visualizzato nella finestra etichettata "B". Tali comandi emulano la successione seguente "click su A", "File" -> "Open (Load)", "click su B", "Analysis -> Pitch Contour -> All data". In modo analogo la successione

USE A

LOAD

USE B

SPG 0 *

(che emula una successione analoga alla precedente di comandi con il mouse) permette il calcolo dello spettrogramma a banda larga della forma d'onda presente nella finestra etichettata "A", spettrogramma che viene visualizzato nella finestra etichettata "B".

Il menù "Analysis" presenta comandi che consentono diversi tipi di analisi e ai quali corrispondono comandi che possono essere inseriti nelle macro per svolgere compiti analoghi in modo ripetitivo e controllato.

Nel menù "Analysis" si hanno comandi:

- per l'analisi LPC;
- per l'analisi spettrale FFT;
- per l'analisi LTA (long term analysis);
- per l'analisi cepstrale;

- per il calcolo dello spettrogramma;
- per il calcolo e la visualizzazione delle formanti;
- per l'estrazione del pitch;
- per il calcolo dell'andamento dell'energia e
- per il calcolo delle statistiche di sonorità.

I parametri utilizzati dai comandi di analisi sono settabili utilizzando le voci del menù "Options" cui corrispondono comandi del tipo "SET <oggetto>" inseribili dall'utente all'interno di macro e che consentono l'accesso a finestre di dialogo per il settaggio dei valori dei parametri relativi ad <oggetto>.

Ad esempio il menù presenta le voci "Capture..." e "Waveform..." cui corrispondono i comandi "SET CAPTURE" e "SET SIGNAL" mediante i quali l'utente può accedere dall'interno di una macro a finestre di dialogo, rispettivamente, per la scelta dei valori dei parametri di registrazione e dei valori dei parametri per la visualizzazione della forma d'onda. Altri comandi che corrispondono a voci presenti nel menù "Options" sono quelli per il settaggio dei valori dei parametri per l'analisi ("SET LPC", "SET FFT" e altri) e per il settaggio dei parametri dell'editing ("SET COPY", "SET APPEND" e altri) per i quali si rimanda a [6].

Multi-Speech consente l'esecuzione dei comandi sia sulla linea di comando (invocabile utilizzando la barra spaziatrice) sia all'interno di macro la cui gestione avviene tramite il menù "Macros". Tale menù consente la creazione in tempo reale delle macro ma l'utente può creare una macro utilizzando un qualunque editore di testo per memorizzare i comandi che vuole inserire nella macro in un file in formato testo.

Il menù "Macros" contiene un certo numero di voci a ciascuna delle quali, come di consueto, corrisponde un comando utilizzabile sia dalla linea di comando sia all'interno di una macro. Tali comandi sono relativi o alla creazione ed esecuzione delle macro o alla configurazione da utilizzare per l'esecuzione delle macro.

I primi sono del tipo "MACRO <parametro>" (dove <parametro> può assumere i valori "LEARN", "DISPLAY", "UNDO", "SAVE", "EXIT" e "RUN" per i quali si rimanda a [6]) mentre i secondi consentono il salvataggio della configurazione corrente perché sia utilizzabile da una macro ("WINDOW SAVEALL WORKAREA" e "SET TO"), il ripristino di una configurazione ("SET FROM" e "WINDOW RESET WORKAREA") o il ripristino della situazione precedente l'esecuzione di una macro ("SET FROM.NOSAVE.INI").

4.2 L'uso delle macro

Multi-Speech permette di usare esclusivamente tecniche sincrone dal momento che le macro sono eseguibili solo dall'interno del programma e non sono previsti strumenti né per l'accesso ad insiemi di file individuati da pattern né per l'accesso a porzioni di file individuate da marker inseriti dall'utente in altri file (vedi oltre).

La creazione di una macro richiede da parte dell'utente per prima cosa la creazione di un file di configurazione che contiene sia la descrizione dello spazio di lavoro sia i valori dei parametri utilizzati dai comandi di acquisizione, di analisi e di visualizzazione. Le macro possono essere poi create sia usando la modalità di LEARN di Multi-Speech sia inserendo comandi in file di testo utilizzando un qualunque editor. La prima modalità presenta una serie di limitazioni ma è utile in fase di progettazione e di apprendimento.

Per illustrare alcune caratteristiche delle macro di Multi-Speech riporto qui di seguito tre macro da me scritte e usate da Silvia Calamai per alcuni suoi lavori di analisi del vocalismo tonico e atono nella varietà pisana di italiano ([1]). Le tre macro, eseguibili sia dalla linea di comando di Multi-Speech usando il comando MACRO RUN sia dalla voce Run del menù Macros, usano il file MASTER.INI come file di configurazione per impostare sia lo spazio di lavoro sia i parametri per l'analisi. La prima macro (detta master) consente all'utente di caricare un file di parlato (LOAD), ascoltarlo (SPEAK) e selezionarne una porzione su cui potrà eseguire in successione le due macro slaves (slave1 e slave2).

```
* Macro slaves:
* slave1
* slave2
SET FROM.NOSAVE C:\Kaylab\ini\user\MASTER.INI
CLOSE ALL
WINDOW RESET WORKAREA
USE A
MESSAGE Carica un file di speech
LOAD
SOURCE =
INFO
SPEAK A 0 *
TITLE Speech
MESSAGE Ora scegli una porzione di segnale e poi esegui la macro successiva quando hai
selezionato la porzione desiderata
SET FROM.NOSAVE INI
```

La seconda macro (slave1, vedi qui di seguito) copia in una nuova window la porzione di segnale selezionata dall'utente (USE B e COPY A SS SE) e su di essa esegue il calcolo dello spettrogramma, (SPG B 0 *), l'estrazione della F0 (PITCH B 0 *) e sia il calcolo sia il tracciamento (in rosso, SET WINDOW.PEN RED) delle formanti (FMT B 0 *).

<pre>* Macro slave1 SET FROM.NOSAVE C:\Kaylab\ini\user\MASTER.INI USE B COPY A SS SE SPEAK B 0 * TITLE Speech Selection USE C SPG B 0 * USE D PITCH B 0 * USE B LINK C LINK D USE C</pre>	<pre>* Macro slave2 SET FROM.NOSAVE C:\Kaylab\ini\user\MASTER.INI SOURCE B WINDOW LPC != RESULT SET FROM.NOSAVE INI</pre>
---	---

```
SET WINDOW.PEN RED
FMT B 0 *
SET FROM.NOSAVE INI
```

La terza macro (slave2), infine, esegue un'analisi LPC (LPC ! =) sulla porzione di file selezionata (SOURCE B) e visualizza il risultato su una nuova finestra (WINDOW) oltre a visualizzare (RESULT) i risultati di tutte le analisi svolte in una finestra di dialogo che ne consente il salvataggio in un file di testo, utile per elaborazioni successive.

Da un esame anche sommario di tali macro risulta evidente che la gestione della esecuzione è a totale carico dell'utente che deve localizzare e visualizzare un file di parlato, selezionarne la porzione desiderata e poi passare alla esecuzione delle macro slave1 e slave2 nell'ordine, dal momento che la macro slave2 fa uso della finestra in cui la macro slave1 copia la sezione del file di parlato, sezione su cui viene eseguita l'analisi LPC (vedi i comandi SOURCE B e LPC ! =).

Qui di seguito si riporta un'altra macro tratta da [6], cui sono state apportate delle modifiche, macro che illustra altre caratteristiche del linguaggio delle macro di Multi-Speech.

La macro carica un file di configurazione (SET FROM FMT.INI), chiude tutte le finestre aperte (CLOSE ALL) e crea le finestre che sono necessarie per le operazioni che saranno eseguite nel seguito (WINDOW RESET WORKAREA). L'utente può impostare la struttura dell'area di lavoro necessaria per le operazioni che vuole svolgere (e che sarà richiamata automaticamente dai comandi suddetti) semplicemente creando le finestre con i comandi del menù "Window", posizionandole e dimensionandole a mano e usando il comando "Save a New Alternate Configuration File" del menù "File" per creare il file .INI che verrà usato all'inizio della macro per ricreare la struttura dell'area di lavoro. tale file consente di impostare i parametri che saranno usati dai comandi presenti nella macro.

Fatto questo la macro attiva la finestra A e in essa carica un file di parlato che viene fatto ascoltare all'utente per un controllo. A questo punto la macro attiva la finestra B ed esegue una estrazione del pitch sul file contenuto nella finestra A e visualizza il risultato nella finestra attiva. In modo analogo viene attivata la finestra C in cui viene visualizzato il risultato del calcolo dell'energia eseguito sul file contenuto nella finestra A (con i simboli < e > si indica la porzione di file visualizzata).

```
SET FROM FMT.INI
CLOSE ALL
WINDOW RESET WORKAREA
USE A
LOAD
SPEAK A 0 *
USE B
PITCH A 0 *
USE C
ENERGY A <>
SET FROM.NOSAVE.INI
```

A questo punto volendo poter selezionare porzioni del segnale visualizzato nella finestra A è necessario definire una macro slave ed eseguirla dopo aver effettuato la porzione del segnale su cui si vuole affinare l'analisi. Ad esempio si può avere:

```
SET FROM FMT.INI
WINDOW
USE D
SPG A SS SE
WINDOW
USE E
SET SPG.LENGTH 512
SPG A SS SE
USE A
LINK B
LINK C
LINK D
LINK E
SET FROM.NOSAVE.INI
```

In tal modo l'utente opera la selezione ed eseguendo la macro slave ottiene un'analisi spettrale a banda larga (SPG A SS SE) ed una a banda stretta (SET SPG.LENGTH 512 e SPG A SS SE). Gli ultimi comandi attivano la finestra A e sincronizzano i cursori presenti su tutte le finestre fin qui definite.

5 CASE STUDY 2: ESPS ([5], [7])

5.1 Breve introduzione

ESPS è un package software utilizzabile per il signal processing in ambiente Unix composto da un certo numero di programmi accessibili dalla linea di comando di una shell di Unix, dall'interno di script e tramite sia EnSig sia waves+.

EnSig rappresenta l'interfaccia attraverso la quale l'utente interagisce con waves+, con il software per la gestione della scheda audio sul sistema ospite e con i programmi di ESPS. Lo scopo principale di EnSig è quello di facilitare l'analisi, l'elaborazione e la visualizzazione dei segnali di parlato. A tal fine EnSig fornisce dei menù e dei pulsanti per interagire con i menù di ESPS e di waves+. Date queste premesse nel seguito EnSig non sarà più preso in considerazione.

waves+ invece è a rigore un insieme di programmi per l'analisi del parlato che consente la visualizzazione, la misurazione, l'annotazione e l'analisi di file di parlato. Di tali programmi il più importante è xwaves. Per semplicità nel seguito considereremo waves+ e xwaves come sinonimi avendo ben presente che, invece, xwaves è solo l'interfaccia grafica di una serie di programmi utilizzati da xwaves come tool esterni per l'esecuzione di funzioni di elaborazione dei segnali quali il calcolo degli spettrogrammi e l'estrazione del pitch.

I programmi utilizzati da xwaves sono veri e propri comandi Unix-like che fanno parte del package ESPS.

xwaves può essere mandato in esecuzione semplicemente digitando “xwaves &” sulla linea di comando di una shell di Unix ed è caratterizzato da un pannello di controllo principale e da un pannello di controllo ausiliario, oltre ad altri pannelli la cui trattazione esula dalle presenti note. Mediante il primo è possibile accedere ai file di input da analizzare e ai programmi ausiliari (detti attachment, vedi oltre), impostare il file di output e inserire comandi. Mediante il secondo è possibile definire nuove operazioni da aggiungere ai menù, personalizzare i menù aggiungendo operazioni predefinite, personalizzare le azioni eseguibili con i tasti del mouse e molto altro ancora.

Un elemento saliente del pannello di controllo principale è il campo “COMMAND (or @file)” che consente l'immissione di un comando di xwaves o del nome di un file di comandi. Il comando immesso può essere un comando globale e pertanto diretto a xwaves oppure un comando locale e diretto ad un oggetto visualizzato da xwaves. Nel primo caso è sufficiente digitare il nome del comando mentre nel secondo il nome deve essere preceduto dal nome dell'oggetto cui il comando è indirizzato oppure dal carattere “.” se il comando è rivolto all'oggetto correntemente visualizzato. L'inserimento del nome di un file di comandi preceduto dal carattere “@” fa sì che l'esecuzione inizi dal primo comando del file. L'esecuzione dei comandi immessi può essere sospesa e ripresa mediante due pulsanti (PAUSE e CONTINUE) presenti sul pannello di controllo principale.

Il pannello di controllo ausiliario contiene un vasto insieme di pulsanti alcuni dei quali consentono:

- l'aggiunta di nuove funzioni alle liste di menù (“Add_op...”);
- la visualizzazione e la modifica delle configurazioni dei comandi e di xwaves (“waves+profile...”);
- la modifica dei menù utilizzando menù standard e/ funzioni predefinite (“Menù Changes...”) e
- la personalizzazione delle azioni associabili ai singoli pulsanti del mouse (“Mouse Bindings...”).

5.2 Le finestre di visualizzazione dei dati

Utilizzando il pannello di controllo principale l'utente può, specificando un pattern in corrispondenza del campo “INPUT file:”, elencare i file di un certo tipo contenuti nella directory di lavoro³, file su cui può eseguire operazioni di visualizzazione, di ascolto e di analisi della struttura interna.

Ad esempio, specificando il pattern “*.wav” l'utente ottiene un elenco di tutti i file di speech il cui nome presenta l'estensione “.wav” e pertanto in formato “wav” contenuti nella directory di lavoro. Selezionando uno dei nomi l'utente può visualizzare il file corrispondente in una window ad hoc detta “waveform window” (“spectrogram window” nel caso l'utente visualizzi file il cui contenuto è uno spettrogramma, a banda stretta o a banda larga a seconda dei casi).

All'interno di una “waveform window” è possibile visualizzare una forma d'onda, settare due marker verticali che permettono di individuare una porzione della forma d'onda visualizzato e richiamare un menù che consente l'esecuzione di operazioni su tale forma d'onda. All'interno di una “waveform window” è, inoltre, possibile visualizzare

³³ Si fa notare che solo EnSig fornisce operazioni di file browsing mentre xwaves consente di lavorare direttamente solo sui file contenuti nella directory di lavoro.

sovrapposte e in colori diversi più forme d'onda in modo da poter confrontare sia gli andamenti sia i valori mentre, nel caso di una "spectrogram window" è possibile sovrapporre allo spettrogramma i tracciati stimati delle formanti calcolati usando il comando "formant".

Fra le operazioni eseguibili su una "waveform window" ci sono di default quelle per il calcolo e la visualizzazione degli spettrogrammi a banda stretta e a banda larga sia di tutta la forma d'onda sia di sue porzioni. Per il calcolo e la visualizzazione degli spettrogrammi sono usati valori dei parametri contenuti in due file di configurazione editabili e customizzabili dall'utente. Di default per il calcolo degli spettrogrammi viene usato il programma "sgram" che esegue il calcolo degli spettrogrammi usando l'algoritmo FFT ma è possibile aggiungere operazioni ad hoc di calcolo degli spettrogrammi utilizzando il comando "add_op".

5.3 Comandi predefiniti o built-in

Le finestre di visualizzazione dei dati hanno associati dei menù che possono essere richiamati utilizzando il pulsante destro del mouse e che possono essere modificati dinamicamente utilizzando comandi di xwaves immessi tramite il campo "COMMAND (or @file)" del pannello principale di xwaves oppure eseguiti dall'interno di script oppure tramite un altro pannello di controllo. L'aggiunta di voci è possibile utilizzando il comando "add_op" mentre la rimozione è possibile con i comandi "delete_item" e "delete_all_items". Le voci dei menù possono essere anche ridefinite utilizzando i comandi "add_op" e "add_waves". Volendo usare i comandi dei menù all'interno di uno script non è possibile usare il comando "send_xwaves" ma è necessario usare il comando particolare "op". La sintassi da usare è la seguente:

object op file filename op nome-di-comando-built-in

dove "object" individua uno degli oggetti noti al programma.

E' possibile aggiungere comandi all'elenco dei comandi predefiniti associati alle varie tipologie di window selezionando con il pulsante destro del mouse il pulsante "Menu Changes" del "miscellaneous controls panels".

I principali comandi predefiniti consentono di:

- riprodurre il contenuto di un file di speech per intero o per sezioni;
- scorrere la porzione visualizzata del file verso l'inizio o verso la fine;
- eseguire operazioni di zoom sul segnale visualizzato;
- eseguire operazioni di analisi sul file o su sue porzioni;
- di estrarre una parte del segnale e memorizzarla in un altro file;
- di allineare e ridimensionare più window contenenti forme d'onda di segnali di speech oppure spettrogrammi oppure altri tipo di segnali (ad esempio la curva del pitch);
- di cancellare una parte del segnale e
- di inserire il contenuto di un altro file nella posizione individuata dal cursore mobile.

Le operazioni di analisi di solito contenute nell'elenco dei comandi predefiniti consentono il calcolo degli spettrogrammi a banda larga e a banda stretta della porzione di segnale individuata dai due marker verticali e utilizzando i valori dei parametri come settati nel file .wave_pro presente nella home page dell'utente.

5.4 Tecniche di scripting

Nel caso di ESPS⁴ ([2]) l'utente ha a disposizione tecniche di scripting sia sincrone sia asincrone. Le prime ([7]) prevedono che l'utente esegua singoli comandi o script attraverso il pannello di controllo principale di waves+, comandi che possono essere diretti sia al programma (comandi globali) sia ad un oggetto visualizzato dal programma (comandi puntuali). Le seconde prevedono l'uso di script immessi sulla linea di comando di una shell di Unix.

Esempi di tecniche del primo tipo sono i seguenti ([7]):

waves make file DgmtA01F.wav	comando globale che carica il file indicato e lo visualizza in una window di segnale;
. activate file DgmtA01F.wav.out op set id F0	comando puntuale che consente la sola visualizzazione del grafico della F0

Un esempio delle tecniche del secondo tipo è dato dallo script seguente ([4]):

```
#!/bin/csh
foreach f (*.wav)
  get_f0 $f $f.f0
end
```

Tale script può essere eseguito come un qualunque comando di Unix e consente l'estrazione del pitch (comando `get_f0`) da un certo numero di file di parlato contenuti nella directory di lavoro e i cui nomi sono individuati da `$f`. I file prodotti dallo script hanno nomi `$f.f0`. Una caratteristica di tale script è quella di essere completamente autonomo rispetto a waves+ ed EnSig pur facendo uso di un comando di ESPS (`get_f0`).

Volendo far interagire tale script con waves+ si può fare uso del comando "send_xwaves" che implementa l'interazione con waves+ sia dalla linea di comando (ad esempio è possibile far visualizzare da waves+ un file di parlato con il comando "send_xwaves make name speech file speech.wav") sia dall'interno di script di Unix.

La versione seguente dello script precedente ([4]) calcola il pitch per ciascun file e poi visualizza in una finestra il file di parlato e in un'altra il pitch e altre grandezze (ovvero l'energia RMS, la curva di probabilità sordo/sonoro e il valore di picco dell'autocorrelazione):

```
#!/bin/csh
foreach f (*.wav)
  get_f0 $f $f.f0
  send_xwaves make name speech file $f
  send_xwaves make name speech file $f.f0
  send_xwaves pause
  if ( $status != 0 ) then
    exit
  endif
end
```

⁴ Simbologie e convenzioni tipiche del Sistema Operativo Unix usate nel paragrafo sono date per familiari al lettore.

```

send_xwaves kill
end

```

L'esecuzione è resa sincrona (ovvero dipendente da azioni dell'utente dall'interno di waves+) dalla presenza del comando "send_xwaves pause": tale comando permette all'utente di analizzare con calma il contenuto di una coppia di finestre per poi passare alla successiva premendo il pulsante CONTINUE presente sul pannello di controllo di waves+. Scopo del comando "send_xwaves kill" è quello di rimuovere le finestre non più utili per impedire un inutile affollamento dello schermo.

Come ultimo esempio illustrativo delle potenzialità di ESPS si riporta qui di seguito il testo di uno script tratto da [7] e il cui compito principale è quello di semiautomatizzare l'etichettatura di un gruppo di file di parlato dei quali provvede a calcolare anche lo spettrogramma.

```

#!/bin/csh -f
send_xwaves set middle_op 'blow up; function'
send_xwaves set ref_size 3 ref_step 2.5 ref_start 0.0
send_xwaves attach function xlabel
foreach file (*)
    set bn=`basename $file`
    send_xwaves make name $bn file $file loc_x 0 loc_y 150 height 200
    send_xwaves $bn spectrogram loc_x 0 loc_y 375
    send_xwaves send make name $bn file file.1.lab color 125
    send_xwaves send make name $bn file file.2.lab color 118
    send_xwaves send activate name $bn fields 1 3 5
    send_xwaves pause
    if ( $status != 0 ) then
        exit
    endif
    send_xwaves kill
end

```

Lo script suddetto permette l'elaborazione di un certo numero di file i cui nomi sono passati sulla linea di comando come parametri. Di ciascun file viene visualizzata la forma d'onda (send_xwaves make name \$bn file \$file loc_x 0 loc_y 150 height 200) e viene calcolato e visualizzato lo spettrogramma (send_xwaves \$bn spectrogram loc_x 0 loc_y 375).

Fatto questo, lo script associa a ciascun file di parlato (e all'associato spettrogramma) due file di etichette (con i comandi send_xwaves send make name \$bn file file.1.lab color 125 e send_xwaves send make name \$bn file file.2.lab color 118) sfruttando il fatto che prima di iniziare a lavorare sui singoli file l'utente ha aperto il programma di etichettatura (send_xwaves attach function xlabel) e si mette in attesa di un'azione dell'utente (comando send_xwaves pause) prima di ripulire lo schermo e passare al file successivo. In questo modo l'utente può lavorare a suo piacimento su ciascun file per passare al successivo solo quando ritiene di aver terminato, avendo la possibilità di riprendere il lavoro di etichettatura in un secondo tempo, semplicemente rieseguendo lo script.

Lo script illustra altre due caratteristiche molto utili di ESPS: la possibilità di aggiungere comandi personalizzati ai menù di waves+ e la possibilità di interagire con programmi ausiliari, i cosiddetti “attachments”.

La prima possibilità è illustrata dai comandi “send_xwaves set middle_op ‘blow up; function’” e “send_xwaves set ref_size 3 ref_step 2.5 ref_start 0.0” che consentono all’utente di definirsi un comando attivabile con il pulsante centrale del mouse per ingrandire una porzione del segnale attorno alla posizione corrente del cursore. Tale comando rende più agevole il posizionamento di marker etichettati sulla forma d’onda del segnale.

La seconda possibilità è illustrata dal comando “send_xwaves attach function xlabel” con il quale viene richiamato il programma “xlabel” ([7]) che consente di visualizzare ed editare fino a 32 file di etichette contemporaneamente. Altri attachment predefiniti sono “xspectrum” (che consente di eseguire sui file visualizzati da waves+ una power spectrum analysis parametrica), xmarks (che consente di assegnare temporizzazioni a sequenze predefinite di etichette) e “xchart” (una estensione di “xlabel”) ma l’utente può scrivere attachment personalizzati usando il linguaggio Tcl/Tk e alcune funzioni ad hoc per interagire con waves+. Per ulteriori dettagli si rimanda ai manuali citati in Bibliografia, in particolare a [7].

CONCLUSIONI

Il presente lavoro contiene una breve introduzione a problematiche vaste e complesse su cui ho in programma di tornare in un futuro prossimo con la stesura sia di macro per Multi-Speech sia di script per ESPS.

RINGRAZIAMENTI

Vorrei ringraziare Barbara Gili Fivela e Silvia Calamai per avermi dato, a vario titolo e in tempi diversi, l’occasione di avvicinarmi a queste problematiche e Franco Ferrero per una epifania avuta leggendo il suo prezioso articolo ([3]).

Bibliografia

- [1] Calamai S., Vocali atone e toniche a Pisa, *Atti delle XII Giornate di Studio del Gruppo di Fonetica Sperimentale*, Macerata, 13-15 Dicembre 2001, in corso di stampa.
- [2] ESPS Programmer’s Manual, Entropic Inc., 1998.
- [3] Ferrero F., User’s Manual of C:\CSL50\MACROS\FORMANTI.FRA [-> ALT-R], *Quaderni del Centro di Studio per le Ricerche di Fonetica*, Volume XIV, 1995:395-402.
- [4] Gili Fivela B., *Script di ESPS*, strumenti di lavoro, Laboratorio di Linguistica, Scuola Normale Superiore.
- [5] *Introducing esps/waves+*, Entropic Inc., 1998.
- [6] *Multi-Speech, Software Instruction Manual*, Kay Elemetrics Corp, 1999.
- [7] *waves+ Manual*, Entropic Inc., 1998.