

TRA ABDUZIONE E “SERENDIPITY”: ANCORA SULLA STIMA (AUTOMATICA) DELLE FORMANTI

Mauro Falcone

Fondazione Ugo Bordoni

RIASSUNTO

Semberebbe, è forse lo è realmente, un argomento obsoleto. Dovremmo pertanto trovare molte informazioni a riguardo in letteratura. Se non fosse che cercando “stima delle formanti” con Google il risultato è uno zero tondo, tondo. Proviamo con “misura delle formanti” ottenendo un desolante due. Allarghiamo lo spazio e cerchiamo solo “delle formanti”. Si trova ben poco; più che altro informazioni di carattere musicale o nel caso peggiore assurdità del tipo “*Anche la voce ha delle formanti (5 secondo il modello più utilizzato)...*”. Presi da sconforto torniamo a consultare i sacri testi (Biblioteca-Internet uno a zero). È gioco forza partire dal lavoro di Ferrero, ed in particolare dal suo ultimo lavoro [1] le cui conclusioni sono per noi distillato di saggezza, di ineguagliabile esperienza, ma soprattutto di umiltà (che è poi l’unico atteggiamento che un ricercatore può avere nel suo lavoro). Ferrero stabilisce, in pratica, che non è possibile attualmente evidenziare un metodo automatico per la caratterizzazione spettrale delle vocali. Manca infatti una “chiarezza” nella metodologia; ma a nostro avviso manca “chiarezza” anche nelle procedure non automatiche, e questo ci porta a riconsiderare alla base il problema ed in particolare quello della stima delle formanti. Come prima cosa non vi è certo univocità sulla definizione di “*formante*” [2]; come secondo punto poniamoci obiettivi “*semplici*” e nello spirito di ottenere il risultato migliore secondo strategie collaborative (*) [3].

Il presente lavoro introduce una prima versione del programma FORA che, sulla base delle tecniche LPC e CEPSTRUM, opera una serie di analisi al fine di trovare quei punti in cui si ottengono risultati “concordanti” utilizzando metodologie, modelli e vincoli in comune. I vincoli possono essere “rilassati” attraverso un file di comandi esterno. Il software viene qui utilizzato sul database VIPS, ed una prima elaborazione dei risultati ottenuti viene proposta al fine di poter valutare possibili utilizzi nel riconoscimento del parlatore utilizzando sistemi già in uso comune in ambito forense. Pur utilizzando strumenti e procedure “chiare” (nella terminologia di Ferrero) riteniamo che questi risultati non possano che successivamente essere utilizzati in ragionamenti abduktivivi [4] [5]. È chiaro infatti che induzione e deduzione possono esserci utili nella implementazione algoritmica piuttosto che, almeno parzialmente, nella fusione di tecniche. Tuttavia, come si intuisce anche dalle operazioni analitiche effettuate in [1], questo non risulta sufficiente. Dalla risultanza di più analisi/risultati dobbiamo abduire delle nuove informazioni o, a volte, correggere le misure già operate realizzando così una specie di circolo virtuoso (ovvero una validazione a posteriori [1]). In questo lavoro si propone una metodologia “chiara” per la stima automatica delle formanti, con le limitazioni sopra delineate. I risultati così ottenibili devono essere considerata la base informativa per ulteriori analisi che, tra abduzione e “serendipity”, possano finalmente aprire uno spiraglio a quelle tecniche automatiche auspiccate, ma giustamente e scetticamente giudicate come non realizzabili in [1].

Il software utilizza un file di comando in XML di facile preparazione; il sorgente è libero. Per quanto riguarda il formato dei file di risultati siamo ancora in una fase di “work in progress”.

BIGLIOGRAFIA

- [1] Franco Ferrero, “*Problemi spettroacustici di classificazione e di misurazioni delle vocali: un confronto*”, Atti delle 7° Giornata di Studio del G.F.S., Napoli 14-15 Novembre 1996
- [2] Mauro Falcone, Andrea Paoloni, “*Formants for clean and telephone corrupted speech*”, COST232 Final report, 1994
- [3] <http://nobelprize.org/economics/laureates/1994/nash-autobio.html>
- [4] François-Marie Arouet (Voltaire), “*Zadig*”, Einaudi (tascabile, serie bilingue), 1997
- [5] Pek van Andel, Danièle Bourcier, “*Serendipity and abduction in proofs, presumptions and emergine laws*”.

SESSIONE

- 8 Elaborazione del Segnale VOCE
- 13 Caratterizzazione e Riconoscimento del Parlato

(*) http://it.wikipedia.org/wiki/Equilibrio_di_Nash

Secondo Adam Smith, un gruppo ottiene il massimo risultato quando ogni componente del gruppo fa ciò che è meglio per sé stesso: "l'ambizione individuale serve al bene comune", e di conseguenza "il risultato migliore si ottiene quando ogni componente del gruppo fa ciò che è meglio per sé". L'intuizione di Nash, lo porterà ad intuire che tale enunciato è incompleto: "il risultato migliore si ottiene quando ogni componente del gruppo fa ciò che è meglio per sé e per il gruppo, secondo la teoria delle dinamiche dominanti".