

IDENTIFICAZIONE AUTOMATICA DELLA PROMINENZA PROSODICA NEL LINGUAGGIO PARLATO

Fabio Tamburini

Università di Bologna

RIASSUNTO

Questo contributo presenterà uno studio che si inserisce nel filone delle indagini nell'ambito dei fenomeni prosodici e della loro identificazione con metodi automatici, proponendosi di indagare le complesse relazioni che intercorrono, a diversi livelli, tra il fenomeno percettivo della prominente prosodica e i fenomeni di tipo acustico che possono essere desunti dalla componente fonetico/acustica degli enunciati, in particolare nel parlato continuo. La prominente risulta essere uno dei fenomeni prosodici maggiormente interessanti e anche uno di quelli maggiormente studiati, proprio per l'importanza che riveste nell'ambito dei processi comunicativi umani. Nonostante la notevole quantità di studi nel settore, soprattutto dal punto di vista linguistico, il problema di un'accurata identificazione del fenomeno con metodi automatici risulta ancora estremamente attuale e tutt'altro che risolto.

Questo lavoro, fortemente basato su informazioni di tipo fonetico/acustico, tenta di costruire, visti anche i lavori sviluppati finora in questa direzione, un modello che consenta l'identificazione della prominente, permettendo la creazione di sistemi automatici di etichettatura del fenomeno. Riteniamo opportuno che tale indagine utilizzi come uniche informazioni i parametri derivabili direttamente, anche se in modo estremamente articolato, dall'espressione sonora dell'enunciato. Il modello proposto e l'algoritmo che implementerà il riconoscimento della prominente non si basano su fonti di informazioni alternative, quali trascrizioni degli enunciati, sia ortografiche che fonetiche, risorse linguistiche etichettate dal punto di vista fonetico, fonologico, prosodico, e nemmeno risorse che contengono informazioni di tipo segmentale sugli enunciati. L'unica informazione fornita all'algoritmo di annotazione dovrà essere la digitalizzazione dell'enunciato (*waveform*).

Il fenomeno percettivo della prominente è supportato, dal punto di vista prosodico, da due fenomeni: il *pitch accent* e lo *stress*. Dal punto di vista fonetico/acustico i *pitch accent* risultano essere connessi con movimenti nel profilo della frequenza fondamentale (F0) e con l'energia globale all'interno del nucleo sillabico, mentre lo *stress* presenta una forte correlazione con la durata temporale del nucleo sillabico e la quantità di energia in specifiche bande di frequenza (*mid-to-high-frequency emphasis*).

Con questo lavoro intendiamo mostrare come con un'attenta misurazione di questi parametri e identificando opportune relazioni tra essi e i fenomeni prosodici oggetto di questo studio, sia possibile costruire un sistema automatico capace di identificare le sillabe prominenti nel parlato continuo con livelli di accordo tra sistema e annotatori umani confrontabili con quelli ottenibili da esperti del settore utilizzando unicamente informazioni ricavabili dalla *waveform* dell'enunciato.

Gli esperimenti che descriveremo sono principalmente basati sull'inglese americano, ma sono in corso studi su altre lingue europee e asiatiche.