

LA PROMINENZA IN ITALIANO: DEMARCAZIONE PIU' CHE CULMINAZIONE

Gloria Gagliardi
Università di Firenze
gloria.gagliardi@unifi.it

Edoardo Lombardi Vallauri
Università Roma Tre
lombardi@uniroma3.it

Fabio Tamburini
Università di Bologna
fabio.tamburini@unibo.it

1. INTRODUZIONE

L'espressione della Struttura Informativa (SI) negli enunciati linguistici è in gran parte affidata a pattern acustici. Il livello più strettamente associato ad essi è quello a cui prevalentemente ci si riferisce con i termini di "Tema-Rema", "Topic-Focus" o "Topic-Comment", per i quali adottiamo le definizioni proposte da Cresti (1992; 2000) e Lombardi Vallauri (2001; 2009), fondate su quali parti dell'enunciato ne veicolano la forza illocutiva. In questo quadro, il Focus è *la parte dell'enunciato che veicola la forza illocutiva e realizza lo scopo informativo dell'enunciato stesso*. Il Topic invece è *la parte dell'enunciato che non ha forza illocutiva e la cui funzione è di permettere la comprensione del Focus all'interno del discorso*.

Queste definizioni coincidono nella sostanza con i concetti di Topic e Focus (Tema-Rema, Topic-Comment) adoperati da molta della letteratura sui correlati acustici della SI (Halliday, 1989; Ladd, 1978; 1996; Pierrehumbert, 1987; Selkirk, 1984 e, più vicino alla nostra analisi, Avesani, 2000; Avesani, Vayra, 2004; Avesani, *et al.* 2007; Breen, *et al.* 2010; D'Imperio, 2002b; Féry, Krifka, 2008; Frascarelli, 2000; 2004; Frascarelli, Hinterrhölzl, 2007, ecc.).

Nel presente studio si sono esaminati enunciati appartenenti a due corpora di italiano parlato, individuando le categorie di Topic e di Focus secondo due principali criteri:

- Primo, la valutazione (in base alla percezione dei pattern acustici e all'applicazione di test di negazione) di quali parti dell'enunciato veicolano forza illocutiva e sono quindi responsabili dell'atto linguistico che viene compiuto; cioè del fatto che l'enunciato sia un'asserzione, una domanda, una richiesta, un comando o qualsiasi altro atto pragmaticamente rilevante (v. Cresti, 2000 per una lista di oltre 80 atti illocutivi).

- Secondo, l'esame del contesto precedente per stabilire quale informazione sia da considerare attiva (Chafe, 1987; 1992) nel momento in cui viene prodotto l'enunciato, cioè Data, e perciò meno probabilmente in Focus; e quale informazione si possa considerare inattiva, cioè Nuova, e quindi più probabilmente in Focus.

Si sono presi in esame solo i tre tipi di SI più frequenti negli enunciati dei corpora, cioè Focus Esteso (a tutto l'enunciato), Topic-Focus e Focus-Appendice¹ (cioè costruzioni con un Focus Ristretto a sinistra dell'enunciato).

Alcuni lavori sull'argomento studiano direttamente le relazioni fra la SI e i fenomeni fonetici, mentre altri introducono un livello intermedio di natura fonologica (ad es. Ladd, 1996; Pierrehumbert, 1987 e tutti gli studi che adottano il sistema di trascrizione ToBI (Beckman, *et al.* 2005). In questa seconda prospettiva le categorie fonologiche sono fatte derivare dai parametri acustici considerando soprattutto l'intonazione, cioè i profili di F0.

¹ Con il termine "Appendice" intendiamo un Topic collocato a Destra del Focus.

La maggior parte degli studi sui correlati prosodici della SI (in particolare del Focus) per l'italiano sono stati condotti entro il paradigma della fonologia Autosegmentale Metrica (AM). Al momento attuale il panorama delle varietà diatopiche esaminate non può dirsi completo, e la quasi totalità dei lavori ha riguardato solo il parlato letto (non il parlato spontaneo o semi-spontaneo). La Tabella 1 sintetizza gli andamenti tonali degli enunciati assertivi individuati dagli studiosi per le varietà che questo contributo condivide; considerando il margine di variabilità riscontrato nei comportamenti dei locutori nella realizzazione degli accenti (Gili Fivela, 2006; Avesani e Vayra, 2004) tali pitch-accent sono da intendersi solo come pattern tipici.

	Focus Esteso	Focus Ristretto	Focus Contrastivo
Roma (Frascarelli, 2004)	H+L*	H*, H*+L	H*, H*+L
Firenze (Avesani, Vayra, 2004)	H+L*	H+L*	L+H*, (L+H)* H+H*
Napoli (D'Imperio, 2002b)	H+L*	L+H*	L+H*

Tabella 1: Profili intonativi tipici di enunciati assertivi nell'ambito degli studi AM.

Come mostra la tabella, l'accento nucleare varia sia in dipendenza della portata del Focus, sia secondo la dimensione diatopica: in particolare, mentre per la varietà Fiorentina ad essere marcata intonativamente è la contrastività, per le varietà Romana e Napoletana è la portata del focus ad essere associata a pitch-accent diversi. Non è ancora chiaro se tali difformità siano interamente imputabili alla variazione diatopica oppure siano legate alle caratteristiche di trascrizione di ToBI. La notazione, da un lato, sembra non riuscire a rendere conto di differenze melodiche chiaramente percepite dai parlanti: ad esempio, nonostante i locutori siano in grado di identificare la provenienza geografica di un parlante solo sulla base dell'intonazione, il broad focus delle assertive è rappresentato mediante lo stesso pitch-accent (Marotta, 2008). Dall'altro lato, sembra esistere un problema di agreement nella descrizione degli accenti. Se l'accordo può dirsi infatti consistente nell'identificazione degli edge tone e dei pitch-accent, è basso nella classificazione dei pitch-accent (Pitrelli et al., 1994; Syrdal & McGorg, 2000). A questo proposito risultano spesso problematici non solo l'allineamento (D'Imperio, 2002a; Gili Fivela, 2002), ma anche l'identificazione dei target tonali, in particolar modo nei plateau, in cui un unico massimo o minimo non possono essere agevolmente identificati (D'Imperio, 2002a). Le informazioni su *scaling* e *slope* sono sottostimate, sebbene potenzialmente distintive (Gili Fivela 2002).

Come suggerito in alcuni studi classici (ad esempio Ladd, 1996) e confermato in ricerche più recenti (Breen, et al. 2010; Lee, Yu, 2010), un item in focus potrebbe coinvolgere una combinazione complessa di tratti acustici differenti, vale a dire durata, pitch e intensità, e non può pertanto essere analizzato solo mediante il profilo intonativo. Per queste ragioni proveremo a indagare la correlazione tra elementi focalizzati e caratteristiche fonetiche considerando il concetto di prominza prosodica come un insieme complesso di tratti acustici, combinati in modo articolato. L'identificazione automatica dei livelli di prominza è indubbiamente un task complesso che richiede un'analisi più approfondita.

2. IDENTIFICAZIONE AUTOMATICA DELLA PROMINENZA

Facendo riferimento agli studi di (Couper-Kuhlen, 1986; Jensen, 2004; Kohler, 2006; Mertens, 1991; Terken, 1991), è possibile definire il concetto di prominenzza prosodica come *un fenomeno percettivo, di natura continua, che consente di enfatizzare alcune unità linguistiche di tipo segmentale rispetto al contesto che le circonda, ed è supportato da una complessa interazione di parametri di tipo prosodico e fonetico/acustico.*

Dei numerosissimi lavori in questo settore sembra opportuno fare riferimento primariamente al lavoro di (Kohler, 2005), per la chiarezza, la lucidità e il rigore metodologico con cui descrive i fenomeni coinvolti. Dai lavori di Kohler emergono chiaramente due attori precisi, a livello linguistico-prosodico, in grado di supportare il fenomeno della prominenzza frasale (o *sentence accent*): i *pitch accent* e i *force accent*. Il primo (*pitch accent*) risulta coincidere pressoché totalmente col concetto omonimo introdotto da Bolinger (1958) ed essenzialmente legato a variazioni, o a specifiche configurazioni, nel profilo della frequenza fondamentale (F0), mentre il secondo (*force accent*) risulta essere un fenomeno completamente indipendente dalla componente intonativa degli enunciati e intimamente legato a fenomeni acustici di altro tipo, per esempio l'intensità e la durata delle unità segmentali.

I due fenomeni sembrano giocare entrambi un ruolo preminente nel supportare la prominenzza percepita a livello di enunciato, in linea con ciò che sostengono alcuni studiosi (si veda ad esempio il lavoro di Ladd, 1996), ma anziché in un'ottica antagonista o gerarchica in un'ottica di interazione e rinforzo reciproco.

Una delle sfide più rilevanti nell'identificazione del livello di prominenzza sillabica riguarda la determinazione dell'influenza che i vari parametri esercitano sulla percezione della prominenzza, in particolare le escursioni della frequenza fondamentale F0, la durata delle unità sillabiche, misure di intensità e anche le aspettative dell'ascoltatore. A livello acustico, numerosi studi (Sluijter, van Heuven, 1996; 1997; Anastakos et al. 1995; Bagshaw, 1994; Heldner, 2003; Streefkerk, 1996) suggeriscono, anche in una prospettiva interlinguistica, una dipendenza tra i *force accent* e parametri come la durata e l'enfasi spettrale (*spectral emphasis*, *spectral tilt* o *spectral balance*), mentre i *pitch accent* sarebbero supportati prevalentemente da movimenti o specifiche configurazioni nel profilo di F0 e dall'intensità globale all'interno dell'unità segmentale di riferimento. Uno degli autori ha condotto alcuni esperimenti che hanno suffragato l'esistenza di tali relazioni in riferimento ad alcune lingue (Tamburini, 2003; 2005; 2006).

Queste considerazioni puramente qualitative si possono trasformare in legami quantitativi definendo una funzione che sia in grado di assegnare livelli continui di prominenzza ai nuclei sillabici utilizzando unicamente informazioni di tipo acustico:

$$Prom^i = W_{FA} \cdot [SpEmph_{SPLH-SPL}^i \cdot dur^i] + W_{PA} \cdot [en_{ov}^i \cdot (A_{event}^i(at_M, at_m) \cdot D_{event}^i(at_M, at_m))]$$

dove $SpEmph_{SPLH-SPL}^i$ riguarda una misura di enfasi spettrale, dur^i è la durata del nucleo sillabico, en_{ov}^i è l'energia globale del nucleo e A_{event}^i , D_{event}^i sono parametri derivati dal modello intonativo TILT (Taylor, 2000) calcolati in funzione della tipologia di allineamento scelto per i massimi – at_M – e i minimi – at_m – presenti nel profilo. Tutti questi parametri sono riferiti al generico nucleo sillabico i all'interno dell'enunciato. Si veda la Tabella 2 per alcuni dettagli sul calcolo di questi parametri.

La struttura della funzione $Prom$, sebbene sembri scelta arbitrariamente, riflette in realtà le relazioni tra i parametri che abbiamo descritto e, in particolare, la somma dei due contri-

buti esprime matematicamente la visione di rinforzo reciproco che attribuiamo alle due tipologie accentuali considerate.

Parametro	Descrizione
Durata del Nucleo (dur)	Durata temporale del nucleo sillabico normalizzata considerando la media e la varianza delle durate dei nuclei all'interno dell'enunciato ($z-score$), e calcolata, all'interno di questo lavoro, utilizzando le segmentazioni manuali disponibili per i corpora considerati.
Enfasi Spettrale ($SpEmph_{SPLH-SPL}$)	Parametro SPLH-SPL (Fant, <i>et al.</i> 2000), normalizzato come nel caso precedente ($z-score$).
Configurazioni del Pitch	Rappresentazione del profilo intonativo dell'enunciato utilizzando il modello TILT (Taylor, 2000) a partire da un profilo del pitch calcolato utilizzando il programma <i>ESPS get f0</i> (Talkin, 1995).
Intensità Globale (en_{ov})	Energia RMS calcolata nella banda di frequenza 50-5000 Hz, normalizzata anch'essa considerando la media e la varianza delle intensità all'interno dell'enunciato ($z-score$).

Tabella 2: Parametri acustici utilizzati nell'algoritmo automatico per l'identificazione della prominenza prosodica.

Il nucleo della funzione *Prom* contiene nove parametri. Cinque di essi possono essere visti come parametri acustici in grado di supportare il fenomeno della prominenza da un punto di vista cross-linguistico ($SpEmph_{SPLH-SPL}$, dur , en_{ov} , A_{event} e D_{event}), mentre gli altri quattro, rappresentati dal vettore $\mathbf{W} = (W_{FA}, W_{PA}, at_M, at_m)$, possono essere visti come specifici di una determinata lingua. Nel nostro modello, W_{FA} e W_{PA} pesano il contributo delle due tipologie di accenti, mentre at_M e at_m modellizzano le differenti possibilità di allineamento tra i *pitch accents* e i nuclei sillabici nelle varie lingue (si veda la Figura 1).

Tutti i parametri coinvolti nel calcolo della funzione *Prom* sono normalizzati all'interno dell'enunciato, quindi i contributi dei differenti locutori e dei differenti intervalli numerici dovrebbero essere stati fattorizzati. In tutti gli esperimenti che presenteremo è stato utilizzato l'insieme di parametri $\mathbf{W} = (1.0, 1.0, 2, 2)$, che risulta essere il più opportuno per l'italiano (Tamburini, 2009).

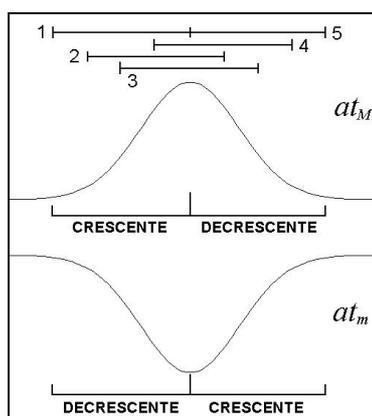


Figura 1: Parametri per l'allineamento tra i pitch accents e i nuclei sillabici.

3. ESPERIMENTI

I due esperimenti che presenteremo sono volti all'identificazione di regolarità tra la posizione e il livello della Prominenza Principale, identificata utilizzando l'algoritmo automatico presentato nella sezione precedente, e la struttura informativa dell'enunciato così come è stata classificata da un annotatore esperto.

Il primo esperimento presenta uno studio pilota su un corpus piuttosto limitato dell'italiano parlato a Roma. Il secondo esperimento ha lo scopo di verificare i risultati ottenuti nel primo sulla stessa varietà, ma considerando un corpus differente, e di estendere l'analisi a due ulteriori varietà, l'italiano parlato a Firenze e a Napoli. L'annotatore ha identificato manualmente l'unità di Focus e le altre unità, se presenti, di Topic e Appendice. Ha inoltre determinato l'estensione del Focus e la sua possibile contrastività.

Negli esperimenti considereremo solo tre tipologie di enunciati che possono essere classificati: (a) TOPIC | FOCUS; (b) FOCUS ESTESO; (c) FOCUS | APPENDICE, FOCUS RISTRETTO, FOCUS CONTRASTIVO. Gli enunciati contenenti riprogrammazioni, esitazioni o disfluenze sono stati esclusi dallo studio, almeno in questa prima fase.

(a) TOPIC FOCUS							
Varietà-Corpus	Prominenza Principale sulla...						Nessuna Prom. Principale
	UsT	UsF	UsA	sIT	sIF	sIA	
Roma-Bonvino	18	1	-	0	1	-	3
Roma-CLIPS	12	3	-	1	0	-	3
Firenze-CLIPS	24	1	-	0	1	-	7
Napoli-CLIPS	8	0	-	2	1	-	2
(b) FOCUS ESTESO							
Varietà-Corpus	Prominenza Principale sulla...						Nessuna Prom. Principale
	UsT	UsF	UsA	sIT	sIF	sIA	
Roma-Bonvino	-	4	-	-	0	-	4
Roma-CLIPS	-	4	-	-	6	-	8
Firenze-CLIPS	-	3	-	-	3	-	2
Napoli-CLIPS	-	4	-	-	7	-	6
(c) FOCUS APPENDICE, FOCUS RISTR., FOCUS CONTRAST.							
Varietà-Corpus	Prominenza Principale sulla...						Nessuna Prom. Principale
	UsT	UsF	UsA	sIT	sIF	sIA	
Roma-Bonvino	-	14	0	-	2	0	0
Roma-CLIPS	-	22	1	-	2	0	2
Firenze-CLIPS	-	14	1	-	1	0	2
Napoli-CLIPS	-	25	0	-	6	0	0

Tabella 3: Numero di enunciati divisi per Varietà-Corpus e configurazioni (es. UsT=Ultima sill. del Topic, sIF=sill. Interna del Focus). Alcune combinazioni non sono ammissibili; in questi casi è stato inserito il simbolo '-' nella casella corrispondente.

3.1 Esperimento 1

I dati sono stati estratti dal corpus "Bonvino", una sezione di *Ar.Co.Dip.* (Bonvino, 2005). Il corpus è formato da 12 conversazioni tra locutori provenienti da Roma, omogenei a livello sociale, età, titolo di studio e origine geografica. Da tre delle dodici conversazioni sono stati selezionati 47 enunciati, estratti gli oscillogrammi, ed è stata prodotta la trascrizione

fonetica allineata al fine di identificare i nuclei sillabici necessari alla procedura automatica per l'identificazione della prominza.

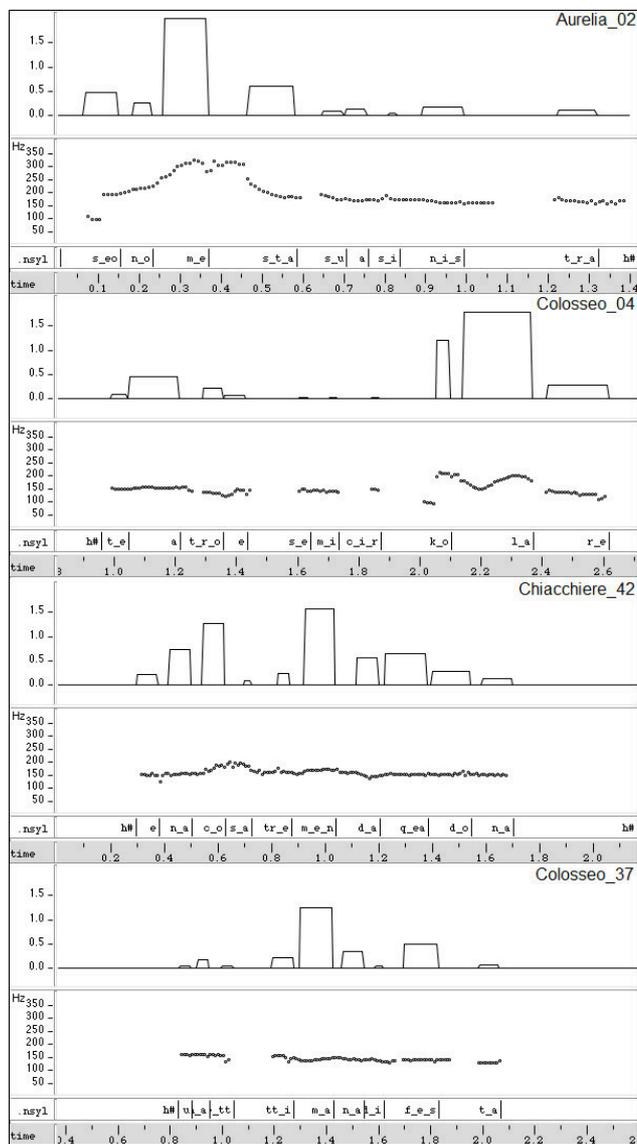


Figura 2: Profili della funzione di prominza – *Prom* – e profili del pitch per alcuni enunciati considerati in questo studio. Aurelia_02: “Secondo me_T | stava sulla sinistra_F”. Colosseo_04: “Il teatro è semicircolare_F”. Chiacchiere_42: “E’ una cosa tremenda_F | quella donna_A”. Colosseo_37: “Una settimana_F | di festa_A”.

3.2 Esperimento 2

I dati per il secondo esperimento sono stati estratti dal subcorpus dialogico del corpus CLIPS (in particolare dalle sezioni riguardanti i map-task e i test delle differenze), corpus che risulta stratificato rispetto alla dimensione diatopica e diafasica (Albano Leoni, 2003). La scelta è quindi caduta su testi relativi alla varietà romana, allo scopo di replicare l'esperimento precedente utilizzando dati diversi, e sulle varietà di italiano parlate a Napoli e Firenze, entrambe particolarmente studiate nell'ambito della fonologia Autosegmentale-Metrica. Sono stati selezionati un totale di 184 enunciati: 64 per la varietà romana, 59 per quella fiorentina e 61 per quella napoletana.

I risultati di entrambi gli esperimenti, visualizzati nella Tabella 3, mostrano regolarità notevoli tra la posizione della Prominenza Principale e la Struttura Informativa degli enunciati. Innanzitutto è possibile notare che, considerando le diverse strutture informative, non emergono differenze rilevanti tra le varietà di italiano considerate nello studio: la distribuzione delle Prominenze Principali sembra seguire schemi molto simili in tutte le coppie varietà-corpus. Inoltre, la posizione della Prominenza Principale tende a collocarsi al confine tra le due unità informative per quanto riguarda le strutture TOPIC | FOCUS e FOCUS | APPENDICE, mentre, nel caso di FOCUS ESTESO, il quadro generale sembra essere piú complesso anche se è possibile riscontrare una marginale tendenza della Prominenza Principale a collocarsi alla fine dell'enunciato. La Figura 2 mostra queste regolarità in riferimento a tre enunciati, presi come esempi di tali comportamenti: Aurelia_02 (TOPIC | FOCUS), Colosseo_04 (FOCUS ESTESO) and Chiacchiere_42 (FOCUS | APPENDICE), tutti estratti dal corpus Bonvino.

E' importante notare che un numero rilevante di Prominenze Principali tra quelle considerate (per esempio 14 enunciati tra i 47 estratti dal corpus "Bonvino") sono supportate principalmente o unicamente da *force accents*, come mostrato dall'enunciato Colosseo_37 nella Figura 2. In questi casi nessun fenomeno di tipo intonativo (*pitch accent*) ha contribuito a supportare quelle prominenze.

Queste regolarità si sono dimostrate altamente rilevanti anche attraverso un test statistico di Fisher.

4. UN'INTERPRETAZIONE "TOPOLOGICA": DEMARCAZIONE PRIMA CHE CULMINAZIONE

I risultati che abbiamo ottenuto non sono certo assoluti. La convergenza fra percezione e misurazione rivela forti (anche molto forti) tendenze, ma non è mai totale. Risultati del tutto netti, in cui i pattern prosodici associati al Topic e al Focus sono perfettamente coerenti, quando si lavora su corpora di parlato spontaneo si ottengono forse solo con procedure *ex post*, cioè se prima si fa la misurazione e poi sulla base di essa si procede all'etichettatura; cioè, se a tutti gli enunciati che alla misurazione presentano lo stesso pattern si dà la stessa etichetta (ad es. Topic-Focus; o Focus Estesio; ecc.). Vale a dire, solo se si adotta un procedimento circolare. Ma se l'etichettatura si fa prima su base percettiva, poi la misurazione è destinata a produrre sempre qualche sorpresa.

Comunque, dai risultati a cui ha condotto l'esperimento che ora si illustrerà, è stato possibile trarre alcune conclusioni interessanti.

Come si è visto in Tabella 3, l'accostamento fra la valutazione percettiva sugli enunciati del corpus e la loro misurazione automatica mediante l'algoritmo che adottiamo ha portato ai seguenti risultati:

<p>Topic-Focus</p> <ul style="list-style-type: none"> - La maggioranza degli enunciati hanno la Prominenza principale all'estremità destra del Topic. - Una minoranza sembrano non distinguere fra le due unità, su cui cadono Prominenze paragonabili. <p>Focus Ristretto (a sinistra)</p> <ul style="list-style-type: none"> - È sempre marcato dalla Prominenza principale dell'enunciato, all'estremità destra del Focus. <p>Focus Esteso</p> <ul style="list-style-type: none"> - Circa metà degli enunciati hanno la Prominenza principale all'estrema destra. - L'altra metà non hanno una Prominenza principale, ma varie prominenze equivalenti.

Cioè, sembra che ad essere segnalati stabilmente dalla Prominenza principale siano solo i costituenti che si trovano alla sinistra dell'enunciato (Topic, o Focus Ristretto), e più precisamente l'estremità destra di tali costituenti. Questo ammette la seguente possibile spiegazione: la funzione primaria della Prominenza principale potrebbe essere *demarcativa*, piuttosto che *culminativa*. In altre parole, il suo primo, immediato effetto potrebbe essere quello di tracciare il confine tra due unità informative, piuttosto che quello di "descrivere" in modo riconoscibile ciascuna di esse.

Questo non significa che diversi tipi di Topic e di Focus non possano essere caratterizzati da diversi e specifici contorni intonativi, che determinino diversi tipi di illocuzione e di funzione pragmatica. Ma la *mera presenza e posizione* della Prominenza Principale (quale risulta dalle misurazioni effettuate) è già sufficiente a segnalare se l'enunciato contiene un confine tra unità informative, e dove esso si trova. E una volta che la Prominenza Principale segnala un confine fra due unità, per riconoscere di quali unità si tratti è sufficiente che il contorno intonativo di quella che si trova a destra segnali se si tratta di un Focus o di un'Appendice.

Gli indizi minimi necessari per rendere riconoscibili da parte del destinatario i confini tra unità informative sono dunque quelli mostrati nella Tabella 4 (PP = Prominenza Principale):

unità di SI	<i>inizio segnalato da:</i>	<i>fine segnalata da:</i>
Topic	inizio dell'enunciato/del contorno intonativo	PP sull'ultima sillaba accentata del Topic
Focus a Destra (dopo un Topic)	PP sull'ultima sillaba accentata del Topic	fine dell'enunciato/del contorno intonativo
Focus Esteso	inizio dell'enunciato/del contorno intonativo	fine dell'enunciato/del contorno intonativo
Focus Ristretto (a Sinistra)	inizio dell'enunciato/del contorno intonativo	PP sull'ultima sillaba accentata del Focus, e inizio di un contorno piatto di Appendice
Appendice	PP sull'ultima sillaba accentata del Focus, e inizio di un contorno piatto di Appendice	fine dell'enunciato

Tabella 4: Indizi minimi per il riconoscimento delle unità informative

Questo fornirebbe una spiegazione piuttosto semplice delle questioni seguenti:

- Perché i Topic sono segnalati più energicamente dei Focus Estesi e dei Focus a Destra che seguono un Topic, benché la rilevanza comunicativa dei Focus sia maggiore di quella dei Topic: la ragione può essere che i Topic, a differenza dei Focus Estesi e a Destra, sono seguiti da un'altra unità di informazione all'interno dello stesso enunciato, e perciò il confine fra le due unità deve essere segnalato.

- Perché anche i Focus Ristretti (a Sinistra) sono segnalati energicamente: la ragione è la stessa, e cioè che anche i Focus a Sinistra sono seguiti da un confine tra unità informative entro l'enunciato.

Dunque, per spiegare come la Prominenza Principale consente (almeno in alcune varietà italiane) il riconoscimento delle unità informative, proponiamo di partire da una spiegazione di natura *squisitamente strutturale*, e più precisamente di natura "topologica"; cioè una spiegazione basata solo sulla *presenza e posizione*, non su aspetti qualitativi della Prominenza e dei contorni intonativi:

Ipotesi topologica sulla prominenza principale

"Ciò che è segnalato dalla Prominenza principale è il confine tra unità informative"

In termini essenziali, l'unica differenza *qualitativa* necessaria per il riconoscimento della Struttura Informativa di un enunciato è quella fra la marcatura di un Topic e quella di un Focus (ristretto) a Sinistra, perché entrambi sono seguiti da un'altra unità. Tale differenza può essere assicurata dai diversi contorni intonativi delle unità che seguono (rispettivamente, dopo un Topic si avrà un Focus a Destra, e dopo un Focus a Sinistra si avrà un'Appendice); oppure (anche con qualche ridondanza) da specifici contorni intonativi che caratterizzino rispettivamente il Topic e il Focus a Sinistra.

L'assenza di una Prominenza principale, o il suo trovarsi sull'ultima sillaba accentata dell'enunciato, segnalano entrambi un Focus Esteso (non preceduto da un Topic), i cui confini a rigore non richiedono una Prominenza principale che li segnali, poiché coincidono con i confini dell'intero enunciato.

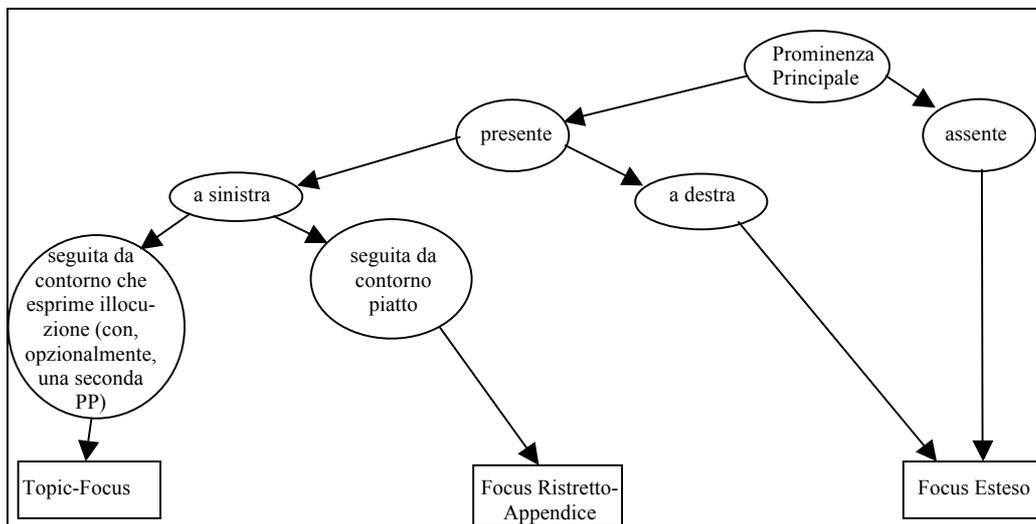
I passi attraverso i quali il destinatario può "computare" la Struttura Informativa di un enunciato orale sono evidenziati nello Schema 1.

In questa interpretazione, i parlanti obbediscono in misura (non) sorprendente alla *Legge del Minimo Sforzo*. Gli unici elementi strettamente necessari sono (a) una Prominenza principale per ogni enunciato, e (b) la differenza tra il contorno "illocutivo" del Focus e il contorno privo di illocuzione di un'Appendice. Ora, poiché i diversi contorni di Focus sono comunque necessari per esprimere le diverse illocuzioni degli enunciati (cioè i diversi atti linguistici), il costo specifico richiesto per segnalare la Struttura Informativa risulta molto basso. Segnalare ciascuna unità informativa con una prominenza culminativa costerebbe più sforzo che segnalare semplicemente i confini, perché:

- distinguere il Topic dal Focus richiederebbe la produzione di due diverse prominenze (una per ciascuna unità) invece di una sola (al confine);

- distinguere il Focus Esteso dal Focus Ristretto richiederebbe due prominenze riconoscibilmente diverse, perché anche i Focus Estesi dovrebbero avere una prominenza "dedicata". Invece, secondo questa interpretazione il linguaggio preferisce funzionare in modo più economico, e cioè marcare solo... l'elemento marcato: cioè il Focus Ristretto.

Questa situazione è ben rappresentata nel corpus, come mostra la Tabella 5. Ma la situazione è più complessa di così, come vedremo nella sezione 4.1.



Schema 1: Passi essenziali per il riconoscimento delle unità di Struttura Informativa.

	Enunciati che corrispondono alla descrizione	Enunciati che non corrispondono alla descrizione
Roma – Bonvino	40 (85.10%)	7 (14.90%)
Roma – Clips	46 (71.88%)	18 (28.12%)
Firenze – Clips	42 (71.19%)	17 (29.81%)
Napoli – Clips	43 (70.49%)	18 (29.50%)
TOTALE	170 (73.59%)	61 (26.41%)

Tabella 5: Risultati previsti e non previsti per le realizzazioni acustiche della SI nel corpus.

4.1. Non alternative discrete, ma un continuum

Come si vede nella Tabella 3, nel corpus una minoranza degli enunciati che sono percepiti come Topic-Focus non hanno una Prominenza principale; e una minoranza di quelli valutati come Focus Estesi hanno una Prominenza principale in posizione interna, non diversa da strutture Topic-Focus.

In altre parole, enunciati acusticamente misurabili come Focus Estesi possono essere percepiti come Topic-Focus, e viceversa. Questo si può spiegare: Topic-Focus e Focus Esteso non sono strutture separate e reciprocamente esclusive, ma piuttosto i due estremi di un *continuum*. Il centro del continuum è occupato da quegli enunciati in cui le unità informative non sono segnalate in maniera netta, e la distinzione fra le due possibili SI rimane sottospecificata o addirittura non specificata.

Insomma, *il parlante non è obbligato a decidere* tra Topic-Focus e Focus Esteso. Alme-

no, non prosodicamente; poiché la disambiguazione può sempre essere affidata a fattori contestuali e pragmatici. Questo è ancora più vero se si considera che il locutore e il destinatario possono valutare in maniera diversa gli indizi prosodici, e il locutore è sempre consapevole di tale possibilità. Pertanto, si regola sempre sulla previsione che la percezione della SI può essere soggetta a un certo grado di indeterminatezza.

Più radicalmente, non c'è ragione di pensare che un contenuto debba necessariamente essere focalizzato o al 100% o allo 0%. Al contrario, ogni contenuto può essere focalizzato in una varietà illimitata di gradi (Daneš, 1967, 1974; Firbas, 1966, 1987, 1989; Sgall 1975; Sgall *et al.* 1973), o anche in una misura che può semplicemente rimanere sottospecificata.

Perciò non c'è da stupirsi se la Prominenza principale non è sempre chiaramente riconoscibile. E' buona norma aspettarsi che alcuni enunciati abbiano uno statuto intermedio tra Topic-Focus e Focus Esteso. E lo statuto informativo di alcune porzioni di informazione, tipicamente "a metà" dell'enunciato, può rimanere incerto.

Insomma, Topic versus Focus non si presenta come un'opposizione di bianco e di nero, ma come una scala di toni di grigio. Questo è ciò che si osserva in enunciati come quelli in Figura 3.

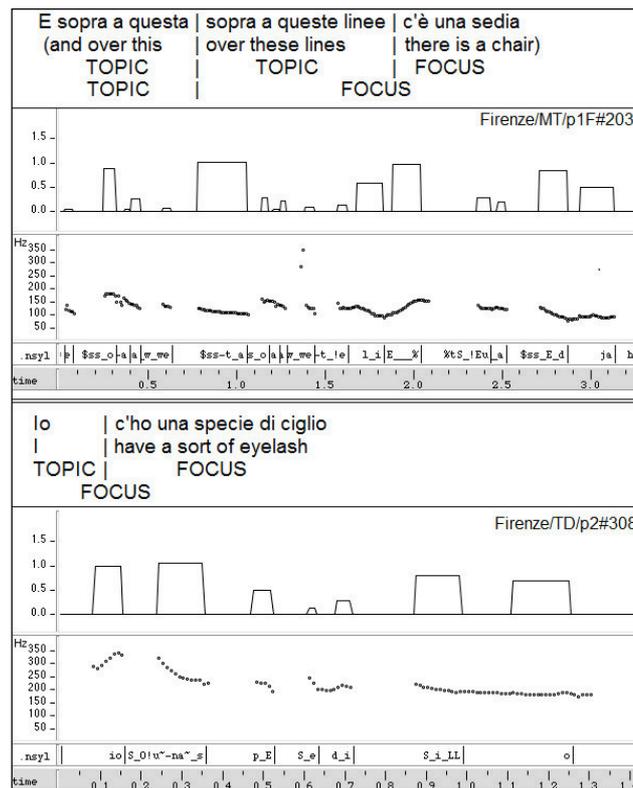


Figura 3: Enunciati sottospecificati per la distinzione fra Topic-Focus e Focus Esteso.

L'assenza di una distinzione netta fra Topic-Focus e Focus Esteso corrisponde al fatto che si tratta di strutture spesso possibili negli stessi contesti, e che spesso non influenzano il seguito del discorso in modo significativamente diverso. Inoltre, si può fare un'osservazione più generale: il fatto che le categorie della SI rimangano sottospecificate negli scambi comunicativi reali non è un problema per gli esiti della comunicazione, come non lo è il fatto che questo ovviamente accada per molti altri aspetti dell'interpretazione semantico-pragmatica degli enunciati.

Per esempio, se dico "Tito ha fermato la macchina", il destinatario può produrre ogni tipo di arricchimento libero per interpretare il mio enunciato, arrivando a interpretazioni anche molto diverse: che Tito è il conducente della macchina, oppure un vigile che ha intimato l'Alt, o un elefante che gli ha attraversato la strada davanti, e così via. Anche informazione di natura meno pragmatica può restare sottospecificata o del tutto inespressa. Per esempio, in molte lingue il tempo verbale può restare non espresso, permettendo diverse interpretazioni (spesso non del tutto disambiguate dal contesto) delle coordinate temporali dell'evento descritto dall'enunciato. Ancora più comunemente, l'identità dei partecipanti a un evento può mancare di espressione esplicita nelle lingue in cui il Soggetto esplicito non è la regola e il Verbo non ha marche morfologiche per la persona. L'esempio giapponese che segue contiene entrambe le ambiguità di cui abbiamo appena parlato.

Tokyo-e ikimasu

Tokyo-a andare

"Io/tu/lui/lei/noi/voi/loro vado/vai/andrò/andremo ecc. a Tokyo"

Ebbene, se nel nostro corpus consideriamo come coerenti con il modello anche tutti i casi in cui la SI rimane sottospecificata tra Topic-Focus e Focus Esteso, otteniamo le nuove percentuali mostrate in Tabella 6:

	Enunciati che corrispondono alla descrizione	Enunciati che non corrispondono alla descrizione
Roma – Bonvino	43 (91.49%)	4 (8.51%)
Roma – Clips	55 (85.94%)	9 (14.06%)
Firenze – Clips	53 (89.83%)	6 (10.17%)
Napoli – Clips	53 (86.89%)	8 (13.11%)
TOTALE	170 (87.88%)	28 (12.12%)

Tabella 6: Risultati previsti e non previsti per le realizzazioni acustiche della SI nel corpus (incluso il continuum tra Topic-Focus e Focus Esteso).

Questo significa che quasi il 90% degli enunciati presentano uno dei seguenti tipi di convergenza fra valutazione percettiva e risultati della misurazione:

- strutture valutate come Topic-Focus, con la Prominenza Principale all'estremità destra del Topic;
- strutture valutate come Focus-Appendice, con la Prominenza Principale all'estremità destra del Focus;

- strutture valutate come Focus Estesì, o senza Prominenza Principale, o con la Prominenza Principale all'estremità destra;
- strutture valutate o come Topic-Focus o come Focus Esteso, senza una Prominenza Principale evidente.

Solo nel 10% dei casi, le misurazioni automatiche danno risultati in cui la Prominenza principale è in posizioni diverse da quelle previste. Questi possono probabilmente considerarsi "rumore" residuo nella procedura: l'esistenza di una minoranza di casi con pattern diversi è da attendersi, perché (a) ragionevolmente devono esserci stati errori umani nella prima fase (determinazione della posizione delle Unità Informative negli enunciati, mediante valutazione soggettiva dei parametri acustici e del contesto), (b) una parte dei dati devono necessariamente risentire dei tipici "difetti" dell'oralità, come produzioni imperfette, cambi di programmazione ed esecuzione, ecc., e (c) l'efficienza dell'algoritmo automatico nell'assegnare livelli di prominenzza alle sillabe non può essere, e non è, del 100%.

5. CONCLUSIONI

Gli esperimenti descritti e la loro possibile interpretazione data qui sopra consentono le seguenti provvisorie conclusioni riguardo alle varietà italiane esaminate:

1. Si può evidenziare un livello astratto e meramente strutturale della Prominenza, topologico e non qualitativo, in cui la sua *mera collocazione* ha la funzione di demarcare il confine tra le unità informative, prima che quella di produrre una culminazione su ciascuna di esse, e una "caratterizzazione" di ciascuna.
Questo aspetto della Prominenza potrebbe già bastare a spiegare i processi attraverso cui i parlanti interpretano la Struttura Informativa degli enunciati nel discorso. Altri tratti, come gli specifici contorni intonativi delle diverse unità informative, in questa funzione potrebbero dunque rappresentare una certa dose di ridondanza.
2. Gli enunciati reali non segnalano sempre in maniera chiara la distribuzione di Topic e Focus. Sul piano acustico, molti rimangono sottospecificati da questo punto di vista. Ciò è vero in modo particolare per la distinzione tra Topic-Focus e Focus Esteso, che spesso non ha effetti rilevanti sulla progressione del dinamismo comunicativo nel discorso successivo.
3. La coerenza di questi risultati con la legge del minimo sforzo, e l'alta percentuale di coincidenza fra valutazioni percettive e misurazione automatica, sembrano convalidare in maniera molto soddisfacente l'algoritmo adoperato per questa analisi.

6. BIBLIOGRAFIA

- Albano Leoni, F. (2003), Tre progetti per l'italiano parlato, in Atti del XXXIV Congresso SLI, Firenze, 675–683.
- Avesani, C. (2000), Costruzioni marcate e non marcate in italiano. Il ruolo dell'intonazione., in Atti delle X giornate di studio del GFS, Il parlante e la sua lingua (D. Locchi, A. Giannini & M. Pettorino, editors), Napoli, 1–14.
- Avesani, C. & Vayra, M. (2004), Focus ristretto e focus contrastivo in italiano, in Il Parlato Italiano, Atti del Convegno Nazionale (F. Albano Leoni, F. Cutugno, M. Pettorino & R. Savy, editors), Napoli, 1–20.

- Avesani, C., Vayra, M., Zmarich, C., Paggiaro, R. & Sperandio, D. (2007), Le basi articolatorie della prominza accentuale in italiano, in *Atti del III convegno AISV* (V. Giordani, V. Bruseghini & P. Cosi, editors), Trento, 1–22.
- Bagshaw, P. (1994), *Automatic prosodic analysis for computer-aided pronunciation teaching*. PhD thesis, University of Edinburgh, UK.
- Beckman, M.E., Hirshberg, J. & Shattuck-Hufnagel, S. (2005), The original ToBI system and the evolution of the ToBI framework, in *Prosodic models and transcription: Towards prosodic typology* (S. Jun, editor), Oxford: Oxford University Press, 9–54.
- Bolinger, D. (1958), A theory of pitch-accent in English, *Word*, 14, 109–149.
- Bonvino, E. (2005), *Le sujet postverbal. Une étude sur l'italien parlé*, Paris: Ophrys.
- Breen, M., Fedorenko, E., Wagner, M. & Gibson, E. (2010), Acoustic correlates of information structure, *Language and Cognitive Processes*, 25, 1044–1098.
- Chafe, W. (1987), Cognitive Constraints on Information Flow, in *Coherence and Grounding in Discourse* (R.S. Tomlin, editor), Benjamins, 21–51.
- Chafe, W. (1992), Information Flow in Speaking and Writing, in *The Linguistics of Literacy* (P. Downing, S.D. Lima & M. Noonan, editors), Benjamins, 17–29.
- Couper-Kuhlen, E. (1986), *English prosody*, London: Edward Arnold.
- Cresti, E. (1992), Le unità d'informazione e la teoria degli atti linguistici, in *Atti del XXIV Congresso SLI* (G. Gobber, editor), Bulzoni, 501–529.
- Cresti, E. (2000), *Corpus di italiano parlato*, Firenze: Accademia della Crusca.
- Daneš, F. (1967), Order of Elements and Sentence Intonation, in *Studies to Honor Roman Jakobson*, The Hague-Paris: Mouton, 499–512.
- Daneš, F. (1974), Functional Sentence Perspective and the Organization of the Text, in *Papers on Functional Sentence Perspective* (F. Daneš, editor), Prague: Academia /The Hague: Mouton, 106–128.
- D'Imperio, M. (2002a), Language-specific and universal constraints on tonal alignment: the nature of targets and anchors, in *Proceedings of Speech Prosody 2002*, Aix-en-Provence, France, 101–106.
- D'Imperio, M. (2002b), Italian Intonation: An overview and some questions, *Probus*, 14, 37–69.
- Fant, G., Kruckenberg, A. & Liljencrants, J. (2000), Acoustic-phonetic Analysis of Prominence in Swedish, in *Intonation* (A. Botinis, editor), Kluwer, 55–86.
- Féry, C. & Krifka, M. (2008), Information structure. Notional distinctions, ways of expression, in *Unity and diversity of languages* (P. van Sterkenburg, editor), Benjamins, 123–136.
- Firbas, J. (1966), On Defining the Theme in Functional Sentence Analysis, *Travaux Linguistiques de Prague*, 1, 267–280.
- Firbas, J. (1987), On the Delimitation of the Theme in Functional Sentence Perspective, in *Functionalism in linguistics* (R. Dirven & V. Fried, editors), Amsterdam-Philadelphia: Benjamins, 137–156.

- Firbas, J. (1989), Degrees of communicative dynamism and degrees of prosodic prominence (weight), *Brno Studies In English*, 18, 21–66.
- Frascarelli, M. (2000), The Syntax-Phonology Interface in Focus and Topic Constructions in Italian, *Studies in Natural Language and Linguistic Theory*, 50, Kluwer.
- Frascarelli, M. (2004), L'interpretazione del Focus e la portata degli operatori sintattici, in *Il Parlato Italiano, Atti del Convegno Nazionale* (F. Albano Leoni, F. Cutugno, M. Pettorino, R. Savy, editors), B06, Napoli.
- Frascarelli, M. & Hinterhölzl, R. (2007), Types of Topics in German and Italian, in *On Information Structure, Meaning and Form* (S. Winkler & K. Schwabe, editors), Benjamins, 87–116.
- Gili Fivela, B. (2006), Tonal alignment in two Pisa Italian peak accents, in *Proceedings of Speech Prosody 2002, Aix-en-Provence, France*, 339–342.
- Halliday, M.A.K. (1989), *Spoken and Written Language*, Oxford: Oxford University Press.
- Heldner, M. (2003), On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish, *Journal of Phonetics*, 31, 39–62.
- Jensen, C. (2004), *Stress and Accent*, Phd thesis, University of Copenhagen.
- Kohler, K.J. (2005), Form and Function of Non-Pitch Accents, in *Prosodic Patterns of German Spontaneous Speech*, AIPUK, 35a, 97–123.
- Kohler, K.J. (2006), What is emphasis and how is it coded? In *Proceedings of Speech Prosody 2006, Dresden, Germany*, 748–751.
- Ladd, D.R. (1978), *The Structure of Intonational Meaning*, Bloomington: Indiana University Press.
- Ladd, D.R. (1996), *Intonational Phonology*, Cambridge: Cambridge University Press.
- Lee, Y. & Xu, Y. (2010), Phonetic Realization of Contrastive Focus in Korean, in *Proceeding of Speech Prosody 2010, Chicago*, paper 033.
- Lombardi Vallauri, E. (2001), La teoria come separatrice di fatti di livello diverso, L'esempio della struttura informativa dell'enunciato, in *Atti del XXXIII Congresso SLI, Napoli*, 151–173.
- Lombardi Vallauri, E. (2009), *La struttura informativa, Forma e funzione negli enunciati linguistici*, Carocci.
- Marotta, G. (2008), Phonology or non phonology? That is the question (in intonation), *Estudios de Fonética Experimental, Universitat Autònoma de Barcelona*, XVII, 177–206.
- Mertens, P. (1991), Local prominence of acoustic and psychoacoustic functions and perceived stress in French, in *Proceedings of ICPhS'91, Aix-en-Provence*, 218–221.
- Pierrehumbert, J. (1987), *The Phonology and Phonetics of English Intonation* (Ph.D. thesis 1980), Indiana University Linguistics Club.
- Pitrelli, J.F., Beckman, M.E., Hirschberg, J. (1994), Evaluation of Prosodic Transcription Labelling Reliability in the ToBI Framework, in *Proceedings of ICSLP'94, Yokohama*, 123–126.

- Selkirk, E. (1984), *Phonology and Syntax: The Relation between Sound and Structure*, Cambridge, MA: MIT Press.
- Sgall, P. (1975), Conditions of the Use of Sentences and a Semantic Representation of Topic and Focus, in *Formal Semantics of Natural Language* (E. Keenan, editor), Cambridge: Cambridge University Press, 297–312.
- Sgall, P., Hajicová, E. & Benesová, E. (1973), *Topic, Focus and Generative Semantics*, Kronberg Taunus: Scriptor.
- Sluijter, A., & van Heuven, V. (1996), Spectral balance as an acoustic correlate of linguistic stress, *Journal of the Acoustical Society of America*, 100, 2471–2485.
- Streefkerk, B. (1996), Prominent accent and pitch movements, *Institute of Phonetic Sciences Proceedings*, University of Amsterdam, 20, 111–119.
- Syrdal, A. & McGorg, J. (2000), Inter-transcriber reliability of ToBi prosodic labelling, in *Proceedings of ICSLP2000*, Beijing, 235–238.
- Tamburini, F. (2005), Automatic Prominence Identification and Prosodic Typology, in *Proceedings of InterSpeech 2005*, Lisbon, 1813–1816.
- Tamburini, F. (2006), Reliable Prominence Identification in English Spontaneous Speech, in *Proceedings of Speech Prosody 2006*, Dresden, Germany, PS1–9–19.
- Tamburini, F. (2009), Prominenza frasale e tipologia prosodica: un approccio acustico, in *Atti del XL Congresso SLI*, Vercelli, 437–455.
- Taylor, P.A. (2000), Analysis and Synthesis of Intonation using the Tilt Model, *Journal of the Acoustical Society of America*, 107, 1697–1714.
- Talkin, D. (1995), A robust algorithm for pitch tracking (rapt), in *Speech coding and synthesis* (W. Kleijn & K. Paliwal, editors), New York: Elsevier, 495–518.
- Terken, J. (1991), Fundamental Frequency and perceived prominence parameters, *Journal of the Acoustical Society of America*, 87, 1768–1776.