LE UNITÀ DI RIFERIMENTO PER L'ESPRESSIONE DELLE EMOZIONI: DATI DAL CORPUS C-ORAL-ROM

Massimo Moneglia
LABLITA, Università di Firenze
massimo.moneglia@unifi.it

SOMMARIO

Il lavoro propone un sondaggio sulla manifestazione acustiche delle emozioni nel corpus di parlato spontaneo C-ORAL-ROM (Cresti&Moneglia, 2005) e verifica la pertinenza alla loro annotazione della nozione di enunciato (Cresti,2000). Nel lavoro viene mostrato che espressione emotiva negli enunciati ha forte definizione semantica, ma che la presenza delle emozioni propriamente dette (Damasio, 2000) è sporadica (frequente solo nelle sessioni in cui la stimolazione contestuale è particolarmente forte). Sulla base della descrizione di esempi tipici si evidenzia che le proprietà acustiche con cui l'emozione è espressa sono qualità olistiche degli enunciati e non influiscono sugli indici prosodici che esplicano il valore illocutivo. La ricerca mostra anche che le manifestazioni espressive (enfatiche) che si riscontrano in modo pervasivo nel parlato spontaneo non esprimono propriamente emozioni e parallelamente pertengono, non l'enunciato (unità minima interpretabile pragmaticamente), ma le unità di informazione in cui questo si articola (Cresti, 2000).

1. INTRODUZIONE

La ricerca sulla manifestazione delle emozioni ha la sua primaria fonte di dati nelle realizzazioni spontanee della comunicazione parlata, nelle quali le emozioni si esprimono in natura. Il corpus del parlato romanzo C-ORAL-ROM (Cresti & Moneglia, 2005), che contiene un corpus di italiano parlato spontaneo, soddisfa due condizioni preliminari per utilizzare effettivamente tale fonte di dati:

a) è un corpus accessibile rappresentativo del parlato spontaneo;[1]
b) identifica le unità linguistiche dell'orale alle quali le emozioni sono riferite

I segnali che esprimono le emozioni nella comunicazione parlata non pertengono, ovviamente, la sola modalità vocale, ma riguardano, oltre a questa, le espressioni facciali, la quantità e velocità dei movimenti oculari, il profilo dinamico degli arti e della postura, la gestualità, ovvero tutti caratteri che pertengono la modalità visiva (Scherer,1998). Deve essere premesso quindi che il corpus C-ORAL-ROM, se può essere considerato rappresentativo della varietà degli atti linguistici eseguiti nel parlare quotidiano, non contiene però l'informazione visiva. Lo studio qui proposto sarà quindi limitato alle

[1] L'accesso a C-ORAL-ROM è possibile in due soluzioni: 1) per lo sviluppo di tecnologie del linguaggio e per i laboratori fonetici il corpus è distribuito da European Language Resources Distribution Agency (ELDA) in 9 DVD con dati non-compressi e non criptati; 2) per la larga distribuzione accademica, attraverso Benjamins Publishing Company (Corpus Linguistics Studies 15) in un solo DVD con dati compressi e criptati.

manifestazioni delle emozioni che si esprimono attraverso la voce e possono essere percepite dall'ascolto del solo segnale acustico.[2]

1.1 Il corpus C-ORAL-ROM

C-ORAL-ROM è costituito da 772 campioni, per complessive 121:43 ore. Le strategie adottate per assicurare la rappresentatività del corpus seguono criteri desunti dalla tradizione di studi sociolinguistici (Cfr. Cresti et al. 2002), criteri rispetto ai quali i quattro corpora, italiano, francese, portoghese e spagnolo, sono stati costruiti in modo comparabile. La comparabilità e la rappresentatività è ottenuta attraverso il corpus design presentato in Tavola 1. In particolare, in ogni corpus romanzo, i campioni sono divisi per varietà di lingua. La varietà informale contiene sessioni che variano per contesto sociologico e struttura dell'evento comunicativo, assicurando l'attestazione di molti generi e domini. La varietà formale comprende un campionamento equilibrato tra parlato nei media e in contesto naturale, testimoniando i domini d'uso più tipici del parlato formale.

Language type	***Sociological context***	***Structure of the communication event***
Informal	Family/private Public	Dialogue Multi-dialogue Monologue

Language type	***Channel***	***Typical domain of use***
Formal	**Natural context**	political speech, political debate; preaching; teaching; professional explanation; conference; business; law
Formal	**Media**	talk shows; scientific press; reportage; interviews; sport; news, weather forecast
Informal	**Telephone**	Private conversations; human-machine interactions

Tavola 1: Corpus design

[2] Unitamente alla serie di annotazioni relative ai metadati generali dei testi, contenuti nel corpus, che permettono la corretta comprensione e contestualizzazione dei testi stessi. La serie completa dei metadati C-ORAL-ROM, redatti nello standard IMDI, sono consultabili in formato html nella pagina ufficiale di C-ORAL-ROM http://lablita.dit.unifi.it/coralrom, e attraverso il browser IMDI, nel catalogo delle risorse linguistiche del Max Planck Institute http://corpus1.mpi.nl/IMDI/metadata/IMDI.imdi

Questo lavoro propone un sondaggio sull'espressione delle emozioni nella voce in un campione di 1.269 enunciati tratti da due testi monologici e due dialogici del sottocorpus italiano.[3] In particolare il campione è selezionato all'interno del sottocorpus dei testi di varietà informale in ambiente familiare e privato, sottocorpus che rappresenta il dominio più specifico della lingua parlata ed è costituito da 20.475 enunciati in 83 sessioni, per circa 129.000 parole. La definizione del campione è tesa quindi ad assicurare un livello di rappresentatività sufficiente solo per uno studio pilota.[4]

1.2. L'annotazione per enunciati di C-ORAL-ROM

L'informazione emotiva veicolata dalla "voce" (Anolli e Ciceri, 1995; Magno Caldognetto 2003) è in teoria una qualità correlabile con vari livelli linguistici: la fonazione, la scelta della parola, le unità linguistiche specificate dalla sintassi (clausole), le unità prosodiche, le unità pragmatiche (atto linguistico), la presa di parola (turno dialogico di un parlante). Sembra dunque cruciale evidenziare quali di questi livelli specifichino le correlazioni più pertinenti della manifestazione emozionale.

La scansione del segnale in enunciati e unità tonali è l'annotazione principale del corpus C-ORAL-ROM e in questo lavoro verificheremo che tale annotazione è effettivamente pertinente per lo studio della manifestazione delle emozioni. Avanzeremo anche una ipotesi più forte, ovvero che le emozioni si manifestino principalmente come qualità delle unità linguistiche che hanno un fondamento azionale e che corrispondono per questo a programmazioni prosodiche unitarie: gli enunciati (Cresti,2000). In sintesi si ipotizza che l'espressione emozionale nella voce si realizzi simultaneamente all'atto linguistico e pertenga quindi l'entità linguistica corrispondente all'atto stesso (enunciato). Di questo assunto evidenzieremo la pertinenza e i limiti.

L'enunciato in C-ORAL-ROM è annotato sulla base di proprietà indipendenti dall'espressione dell'emozione. In accordo con la *Teoria della lingua in atto* (Cresti, 2000), l'enunciato è identificato su base prosodica, ed è considerato l'unità comunicativa di riferimento del parlato spontaneo superiore alla parola. L'identificazione avviene attraverso una euristica semplice, basata sull'assunto che i parlanti competenti sono estremamente sensibili alle variazioni prosodiche intenzionali, quali i patterns tonali ('t Hart et al., 1990) che presiedono all'esecuzione dell'atto linguistico e delle unità di informazione che lo compongono. Su questa base si assume in particolare che ad ogni confine prosodico terminale corrisponda un enunciato (Karcevsky, 1931; Crystal, 1975).

L'annotazione dei confini di enunciato e di scansione prosodica è stata operata su base percettiva da annotatori esperti (Moneglia, 2004). Il corpus è stato interamente annotato nella sua trascrizione con i segni di demarcazione degli enunciati (breaks terminali), entro i turni dialogici e con quelli dei principali fenomeni di scansione prosodica interna agli enunciati stessi (breaks non terminali). Il materiale linguistico corrispondente alla scansione

---

[3] ifamdl08, ifamdl19, file ifamnn03, ifammn18. I testi, tutti avente l'autore e persone a lui vicine tra i partecipanti, sono stati preferiti anche per facilitare eventuali prove di laboratorio.

[4] L'ipotesi di un campionamento statistico casuale sull'intero corpus è stata scartata perché non avrebbe dato esiti significativi, a causa del carattere di sporadicità dell'espressione delle emozioni nei corpora di spontaneo, che evidenzieremo. Probabilmente con una tale metodologia non avremmo avuto alcun esempio di enunciato connotato emozionalmente. I testi sono stati quindi selezionati per il significato che le tipologie in questione ricoprono e l'analisi di tutti gli enunciati di ciascuna sessione ha massimizzato la probabilità di incontrare eventi comunicativi connotati in tal senso.

in enunciati è allineato all'onda, per cui il corpus si configura come una base di dati di enunciati (circa 45.000 nell'intero corpus italiano).

Lo schema di annotazione è stato validato sia all'interno del consorzio che ha prodotto il corpus (Moneglia et al. 2005) sia da enti esterni[5] (Danieli et al. 2004). La validazione mostra un notevole accordo sulla rilevanza percettiva dei confini prosodici.

	French	Italian	Portuguese	Spanish
Total Agreement Rate	96,48%	95,21%	98,93%	97,17%
Worst possible Result (baseline)	62,17%	30,84%	37,28%	37,93%
K Index (General)	0,952	0,928	0,980	0,946
K Index (Realistic)	0,776	0,807	0,920	0,827

Tavola 1

La valutazione conferma in particolare l'ipotesi della rilevanza percettiva a livello multilingue dei breaks prosodici, sia terminali che non terminali. Tavola 1 presenta l'accordo sull'annotazione di tutte le posizioni che possono essere sede di un break. L'accordo tra i giudizi dei validatori e il l'annotazione proposta in C-ORAL-ROM è calcolato sia in termini percentuali sia in termini di K statistica (i dati sono riportati per esteso in Moneglia et al. 2005).

1.3. Premesse teoriche per l'annotazione delle emozioni

C-ORAL-ROM annota tutti i contesti linguistici che corrispondono ad un enunciato e la loro scansione interma in unità tonali, ma le qualità funzionali, azionali e emotive delle unità segnalate non sono specificate. Per condurre il sondaggio sono stati adottati due criteri, rispettivamente per la definizione delle qualità illocutorie e funzionali delle unità linguistiche e per la definizione delle emozioni correlate:

a) per l'illocuzione e la struttura funzionale degli enunciati ci riferiamo alla *Teoria dell'Articolazione dell'Informazione* e alla *Classificazione Affettiva dell'Illocuzione* sviluppate in Cresti (2000) e applicate in particolare in Firenzuoli (2003) per lo studio dei correlati prosodici dell'atto linguistico;

b) per l'annotazione delle emozioni ci riferiamo all'opera generale di Damasio (Damasio, 1996; 2000) che consente di collocare le emozioni all'interno di un quadro neuro-psicologico esplicito.

Ciò, come vedremo, è necessario al linguista per distinguere le manifestazioni espressive specificamente emotive da altre manifestazioni espressive, ma non emotive, veicolate dalla voce, come ad esempio i contenuti semantici dell'enfasi. In particolare traiamo da Damasio la distinzione tra Emozione e Sentimento e dal lavoro di Cresti la definizione di Affetto:

EMOZIONE
•collezione di risposte chimiche e neurali che formano una configurazione o "rappresentazione" a livello subcorticale e che interessano un organismo;

[5] Loquendo (Torino) con la supervisione di M. Swerts e J. M. Garrido.

•Le risposte che sono innescate da un'emozione sono automatiche (non sono controllate), sono inconsce, e non possono essere fermate volontariamente dal soggetto.

SENTIMENTO

•esperienza mentale, privata, di un'emozione;
•riguarda non solo un organismo, ma anche un "sé";
•i sentimenti che derivano dalle emozioni sono automatici e non possono essere bloccati.

AFFETTO

•attivazione pulsionale di uno schema motorio di atto relazionale connesso ad una rappresentazione;
•è caratterizzato da intenzionalità (conscia o inconscia);
•riguarda un soggetto (organismo, sé, coscienza) in relazione ad altri soggetti e fonda gli schemi di comportamento e le azioni sociali;

Deriviamo inoltre da Damasio una semplice e collaudata classificazione delle emozioni, sintetizzata qui di seguito, che ci servirà come schema di annotazione nel sondaggio: [6]

EMOZIONI DI FONDO
tensione vs rilassamento
benessere vs malessere
affaticamento vs energia
anticipazione vs timore

EMOZIONI PRIMARIE O UNIVERSALI
gioia
tristezza
paura
rabbia
sorpresa
disgusto

EMOZIONI SOCIALI
imbarazzo
gelosia
colpa
orgoglio

Sulla base di queste premesse teoriche, nel paragrafo 2. verificheremo la congruenza dell'annotazione in enunciati di C-ORAL-ROM con le manifestazioni emozionali riscontrabili nel corpus, e quindi l'utilizzabilità dell'annotazione della risorsa ai fini dello studio dei correlati acustici delle emozioni. Dall'analisi di alcuni esempi di parlato spontaneo emotivo evidenzieremo le proprietà acustiche con cui le emozioni si manifestano

[6] Sono in corsivo neretto le emozioni che saranno esemplificate in questo articolo. Il lettore noterà che le emozioni non hanno una definizione operativa!

e verificheremo l'ipotesi della connessione tra espressione dell'emozione e enunciato a due diversi livelli di compromissione teorica, espressi dalle seguenti proposizioni:

1. La spinta emozionale è correlata all'enunciato, ovvero da un lato ad unità superiori sia alla parola sia all'unità di informazione e dall'altro ad unità inferiori al turno dialogico.
2. Due enunciati diversi non sono mai attivati dalla stessa spinta emozionale.

Presenteremo inoltre una serie di osservazioni tese a facilitare lo studio corpus based delle emozioni: a) la frequenza con cui le emozioni si manifestano negli enunciati del corpus e la tipologia dei testi in cui l'espressione delle emozioni è più frequente; b) le modalità di categorizzazione, ovvero come all'emozione espressa è assegnato un valore semantico.

Vedremo poi che all'interno del cosiddetto "parlato emotivo" è necessario distinguere nettamente tra le modalità con cui nel linguaggio si manifestano le emozioni propriamente dette e le modalità con cui si manifesta, più in generale, l'espressività. In particolare il sondaggio evidenzia una netta distinzione sia quantitativa che qualitativa tra i casi in cui nella voce si ha espressione del contenuto emozionale e una varietà di casi nei quali l'espressione si connette invece a manifestazioni espressive "enfatiche".Vedremo nel paragrafo 3. che l'enfasi veicola contenuti di diversa natura neuropsicologica rispetto all'espressione emotiva propriamente detta, ha diversa semantica e non ha come riferimento linguistico l'enunciato, bensì l'unità d'informazione.

2. L'ESPRESSIONE DELLE EMOZIONI: UNITÀ DI RIFERIMENTO LINGUISTICA E ATTRIBUZIONE DI SIGNIFICATO

L'Espressione delle emozioni emerge in modo significativo nel corpus solo in presenza di forti stimoli situazionali che innescano la reazione emotiva del parlante. Nelle conversazioni *de plano*, che costituiscono la maggior parte delle sessioni registrate, tali stimolazioni contestuali sono ridotte, con una conseguente sporadicità degli enunciati connotati da un punto di vista emotivo. In effetti l'azione verbale reciproca, esercitata dei parlanti nel dialogo, costituisce in genere lo stimolo principale alla verbalizzazione e deve essere sottolineato che difficilmente questa innesca una risposta emozionale marcata, o almeno così avviene nelle interazioni che sono oggetto di registrazione nel pur ampio campionamento C-ORAL-ROM.

Questa caratteristica fa sì che il numero delle chiare espressioni emozionali nel corpus sia esiguo, per cui il primo risultato del sondaggio è che l'espressione delle emozioni in un corpus di parlato spontaneo ha un carattere di sporadicità ed è legata massimamente alla occorrenza nella situazione della verbalizzazione di stimoli extralinguistici significativi (vedi i dati più oltre).

Data questa premessa lo studio delle modalità con cui l'emozione si manifesta nel segnale vocale nel parlato spontaneo assume, più che un valore quantitativo e statistico, un valore qualitativo: veniamo a disporre di dati naturali attendibili per identificare le qualità acustiche differenziali dell'espressione delle emozioni.

In questo paragrafo esamineremo più tipi di espressione naturale delle emozioni tratti dal corpus: a) tre esempi nei quali la manifestazione dell'emozione avviene in risposta ad uno stimolo contestuale forte (situazioni di pericolo durante una sessione di scuola guida); b) un esempio nel quale la manifestazione dell'emozione è connessa all'emergenza, nel

contesto dialogico, di significati che toccano profondamente la sfera emozionale dei parlanti (un ricordo che provoca dispiacere nel parlante durante un racconto di vita).

Nell'analisi di tali contesti ci soffermeremo in particolare su tre punti:

- la possibilità di definire la semantica dell'emozione e la sua conseguente annotabilità
- la pertinenza dell'enunciato in quanto unità linguistica di riferimento per l'espressione dell'emozione
- la descrizione dei caratteri acustici ascrivibili all'emozione, in rapporto ai caratteri degli enunciati con identica forma locutiva e illocutiva, ma non marcati da spinta emotiva

2.1. Espressione dell'emozione di paura-rabbia

Il primo esempio è tratto dalla registrazione di una sessione di scuola guida (ifamdl19). In questa registrazione la verbalizzazione dell'istruttore è condizionata dalla rilevanza degli eventi nella viabilità e in relazione a questa stimolazione, forte e frequente, si riscontrano altrettanto frequenti connotazioni emozionali. A riprova della connessione tra espressione emozionale nel parlato e forte stimolazione contestuale, abbiamo constatato che in questa sessione l'espressione delle emozioni non è sporadica, ma è una caratteristica che riguarda ben 54 dei 376 enunciati (1/7), con connotazioni che registrano più frequentemente *paura*, *rabbia*, *ira* e anche alcuni casi di *ira contenuta*, *tensione* e *tristezza*.

Nell'esempio seguente (nel file ifamdl19-sel1.wav) il contenuto emozionale espresso nella locuzione dell'istruttore cambia in relazione al pericolo esterno, che aumenta tanto più si avvicina uno stop che il guidatore non sembra prepararsi a rispettare.

*MAX: [219] qui / sulla destra / è una round about / credo //$ [220] no //$ [221] c' hai / lo stop //$ [222] vai vai //$ [223] accàpati //$ [224] ferma //$ [225] ferma //$ [226] ferma //$ [227] ferma //$ [228] c' hai lo stop //$
%sit: l'auto si avvicina ad un incrocio
Scarica l'audio

Il caso è particolarmente interessante perché la connotazione emotiva è relativa a una locuzione "ferma" che viene ripetuta più volte (negli enunciati da 224 a 227) a partire da una prima istanza non marcata, tutte le volte con diverso contenuto emozionale.

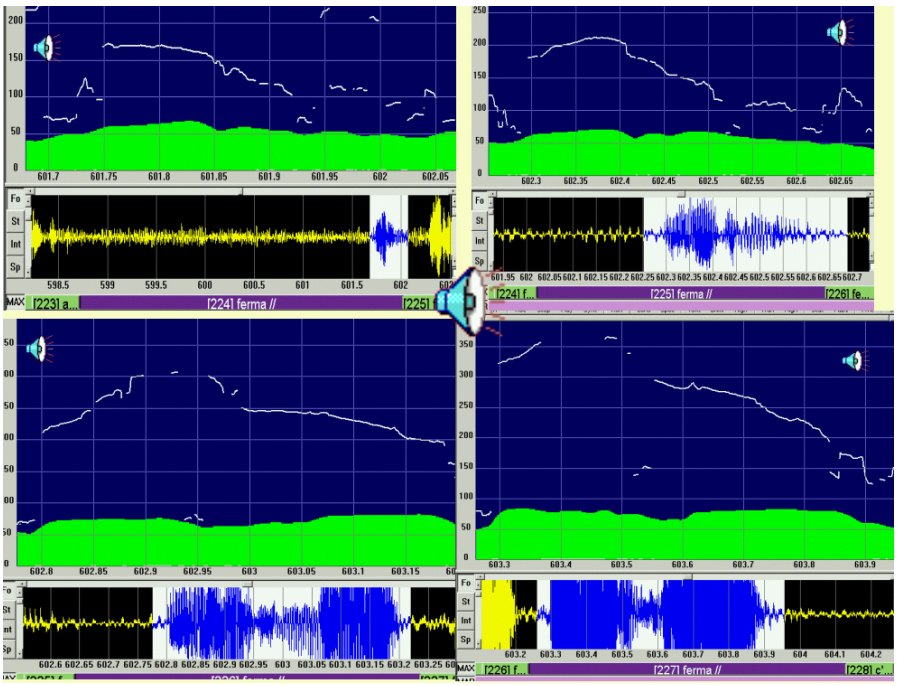

Figura.1

Scarica l'audio

L'identità di locuzione in più atti linguistici contigui consente quindi, in linea di principio, una analisi dei tratti differenziali attribuibili ai diversi contenuti emozionali espressi. La serie di tracciati in Figura 1 mostra le caratteristiche acustiche salienti di F0, Durata e Intensità della serie di enunciati (in sequenza nel file ifamdl19-sel2.wav).

La definizione dell'espressione dell'emozione sulla base della percezione dei caratteri acustici che la manifestano è notevole. Non è difficile raggiungere un accordo tra i parlanti competenti sul fatto che le modificazioni di modalità vocale nella locuzione "ferma" negli enunciati 224, 225, 226, 227 esprimono gradi diversi di *paura* relativamente ai pericoli emergenti nel contesto e, nello stesso tempo, anche *rabbia* relativa all'interlocutore. [7]

Il giudizio, ottenuto sulla base dello schema di annotazione precedentemente proposto, in realtà è disponibile facilmente anche come risposta non vincolata ad una specifica serie di alternative.Il lettore può facilmente replicare tale giudizio ascoltando il turno dialogico.

In altri termini, nonostante lo schema utilizzato sia estremamente sommario, la definizione dell'emozione risulta sufficientemente chiara e l'interpretazione è condivisa. L'ambito di variazioni emotive riscontrate si inscrive con chiarezza tra le emozioni universali *paura-rabbia*, che sono provate dal soggetto e si manifestano nella verbalizzazione, peraltro non attraverso una modulazione volontaria della voce, ma attraverso caratteristiche che, come previsto dalla natura dell'emozione, risultano involontarie.

D'altro canto il valore universale, e quindi immediatamente percepibile dei caratteri che qualificano le emozioni, è sempre stato sottolineato. Mostrare l'uniformità dell'attribuzione di significato alle variazioni acustiche in considerazione conferma tale proprietà generale, che, come vedremo, è anche discriminante l'espressione dell'emozione.

Dato che non vi è altra informazione oltre le modalità di realizzazione vocale che consenta di indurre il contenuto in questione, le variazioni dei parametri vocali appaiono gli unici candidati rilevanti alla percezione e sono quindi immediatamente connesse al valore semantico individuato precedentemente. Dalla prima analisi dell'esempio emerge in

---

[7] Trattandosi di un sondaggio, la procedura di valutazione non è standardizzata. Qui e nei casi seguenti i giudizi sono stati ottenuti da 7 ricercatori del laboratorio LABLITA e ripetuti con almeno due soggetti non appartenenti a nessun gruppo specifico.

particolare che il turno è diviso in enunciati, ognuno dei quali ha una caratteristica emozionale indipendente, ovvero è in corrispondenza biunivoca con una espressione emotiva specifica.

La previsione che l'enunciato è l'unità linguistica pertinente la manifestazione dell'emozione risulta quindi confermata da questo esempio, almeno per quanto riguarda il limite imposto dalle unità di descrizione linguistica superiore all'enunciato: dentro un turno si realizzano infatti più enunciati che non sono vincolati ad una stessa espressione emotiva, ma che sono interessati ognuno da una spinta emotiva specifica.

Dato che gli enunciati di questa serie sono sempre monorematici, però, l'esempio non è significativo da punto di vista delle correlazioni tra espressione delle emozioni e unità linguistiche inferiori all'enunciato (unità tonale, costituente, parola). Enunciato e unità inferiori di parola e di unità tonale, infatti, coincidono.

I parametri della voce interessati dalla variazione emotiva degli enunciati in questione e sono sintetizzati nella Tavola 2. Prendendo come punto di riferimento l'enunciato emozionalmente neutro (riportato sulla prima riga) si evidenziano negli altri tre enunciati variazione positive sia di durata, che di intensità, che di F0 media, come peraltro previsto in generale per i parametri macro-prosodici delle emozioni ad "alta attivazione"(Scherer, 2003; Magno Caldognetto et. al, 1998):

Durata	Intensità tonica	Intensità post tonica	Altezza attacco	Mov. sillaba 1	Mov. sillaba 2
315	40	30	175	1A	
349	48	40-30	210	1A	
436	50	50	300	[1]	[0][A]
691	>50	>50	350	[1]	[0][A]

Tavola 2

Tale aumento è speculare all'aumento della prominenza dello stimolo che determina, nel giudizio degli ascoltatori, un grado maggiore del contenuto emozionale.

All'aumento della tensione connessa all'insorgere dell'emozione di paura-rabbia può essere quindi correlato, in accordo alle descrizioni più note, a variazioni positive simultanee in tutti i parametri di realizzazione vocale.

2.1.1 Espressione dell'emozione e espressione della forza illocutoria

Le emozioni si incrociano, oltre che con l'aspetto locutorio, anche con l'aspetto di valore illocutorio, che nel parlato è espresso fondamentalmente dai movimenti prosodici percettivamente rilevanti dell'enunciato e dalle loro configurazioni (Cresti, 2000; Cresti Firenzuoli 1999; Firenzuoli, 2003). A questo proposito la prima analisi della serie di enunciati in oggetto descrive la variazione emotiva riscontrata come un aumento, via via sempre maggiore, del contenuto emozionale primario di rabbia-paura in enunciati locutivamente identici di uno stesso locutore, aventi uno stesso valore illocutivo (direttivo di ordine).

Questa descrizione non è però coerente con le diversità di forma prosodica negli enunciati in questione riscontrate nell'ultima colonna di Tavola 2. Se i parametri di durata, intensità e altezza di attacco variano quantitativamente, nella seconda coppia di enunciati la descrizione dei movimenti di F0 nei tracciati (riportata sull'ultima colonna della tabella precedente) evidenzia un movimento tipologicamente diverso nella seconda coppia di enunciati rispetto alla prima. I movimenti percettivamente rilevanti nei primi due esempi presentano una ascesa e discesa repentine (movimento 1 A nei termini di annotazione

prosodica IPO), mentre il secondo presenta un movimento su più sillabe costituito da ascesa separata con una tenuta dalla discesa (1 0 A, sempre nella notazione IPO). Tali movimenti veicolano diverso valore illocutorio.

Le informazioni sui valori illocutori dei movimenti in considerazione, desumibili dagli studi precedenti, permettono qui una analisi di dettaglio. Confrontando i profili con la griglia dei rapporti tra forma prosodica e valore illocutorio proposta in Cresti e Firenzuoli (1999) e più estesamente in Firenzuoli (2003), possiamo osservare che mentre il primo movimento corrisponde strettamente alla descrizione della forma tipica dell'*ordine*, il secondo corrisponde piuttosto alla descrizione di un diverso tipo di atto direttivo, ovvero l'atto di *presentazione distale*.

In effetti da una più attenta analisi del valore pragmatico degli enunciati in questione, sorprendentemente,[8] possiamo notare che alla corrispondenza formale tra coppie di enunciati e tipi di movimenti percettivamente rilevanti, corrisponde anche un diverso valore funzionale. Le due coppie esprimono due distinti atti della tipologia direttiva: nei primi due enunciati l'atto illocutorio è caratterizzabile come il tentativo di *indurre un cambiamento nel mondo tramite una azione diretta su un orizzonte conoscitivo e attenzionale comune a parlante e interlocutore* (il contenuto dell'ordine), mentre nei secondi l'atto direttivo tende piuttosto *a mutare l'orizzonte attenzionale dell'interlocutore, che non è più presupposto essere comune* (il contenuto della presentazione distale). In breve, diversamente dagli enunciati della prima coppia che tendono semplicemente a far sì che l'interlocutore fermi la macchina, gli enunciati della seconda coppia sono interpretati come il tentativo di far prendere atto a chi guida che c'è lo stop e che, di conseguenza, si deve fermare.

Dunque, nel passaggio dalla prima alla seconda coppia di enunciati, insieme all'aumento di contenuto emozionale, si realizza anche un cambiamento del valore illocutorio della locuzione, che passa da *ordine* a *presentazione*. Tale cambiamento corrisponde, come previsto dalla teoria motivata indipendentemente nei lavori di Cresti e Firenzuoli, al mutamento dei movimenti percettivamente rilevanti dell'enunciato.

L'esempio è significativo in quanto evidenzia i rapporti tra indici prosodici della forza illocutoria e manifestazione dell'emozione nella voce: l'espressione emozionale non altera la caratteristica prosodica che veicola la forza illocutoria, caratteristica che conserva tutto il suo valore differenziale, nonostante l'estrema marcatezza dell'espressione emozionale stessa. L'indipendenza tra indici illocutori e indici emotivi, dimostrata qui in vivo, è coerente con l'ipotesi che l'intonazione veicola formalmente i significati illocutori, che, in quanto caratteri linguistici, rimangono costanti al variare quantitativo di contenuto emozionale paralinguistico.

2.2 L'aumento di tensione: lo scope degli indici vocali dell'emozione e la sua categorizzazione

L'enunciato, inteso come l'entità linguistica minima iterpretabile pragmaticamente, è o meno anche l'unità linguistica minima a cui riferire l'espressione emozionale? L'espressione di paura e rabbia nella voce nell'esempio precedente sono coerenti questa ipotesi in quanto i confini prosodici che demarcano l'enunciato correlano con le proprietà macro-prosodiche e percettive dell'espressione dell'emozione.

Non è possibile ovviamente fare qui generalizzazioni, ma il sondaggio permette di evidenziare dati empiricamente significativi relativamente a questa ipotesi. L'enunciato

---

[8] La sorpresa riguarda gli informatori, ma il fenomeno è invece previsto teoricamente in quanto la categorizzazione dell'atto illocutorio, come la maggior parte dei processi di elaborazione del senso, non è un processo conscio (Moneglia, 1999).

seguente, tratto dalla stessa sessione, è interessante in quanto, se apparentemente sembra contraddirla, in realtà la convalida.

*MAX: [283] quando lui è andato avanti / vai un po' più avanti //
%sit: in coda alla pompa di benzina per fare rifornimento, MAR si avvicina troppo all'auto che si sta allontanando dalla pompa
Scarica l'audio

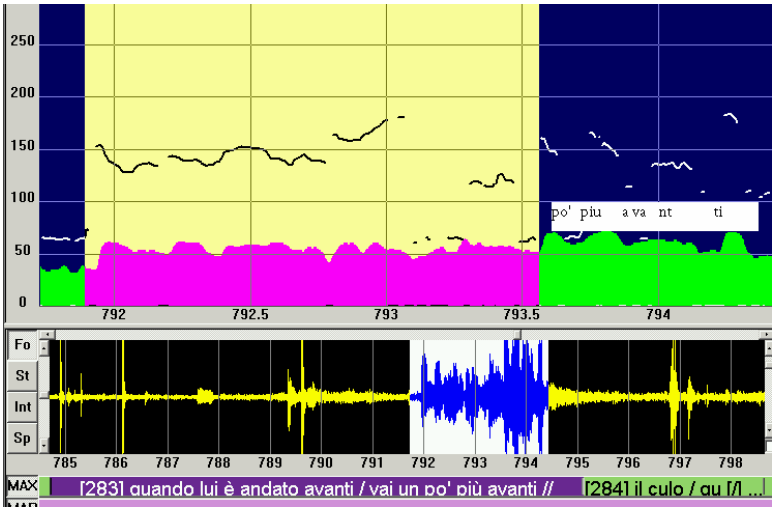

Figura 2

L'esempio è composto da un turno di un solo enunciato articolato in due unità tonali, rispettivamente con funzione di *topic* (in chiaro nel tracciato) e di *comment* (in blu nel tracciato). In particolare si percepisce un cambiamento della spinta emotiva all'interno dell'unità di comment. Mentre l'enunciato è in corso di esecuzione, il guidatore inesperto mostra di non saper misurare la distanza con l'auto davanti, che sta lasciando il distributore, e rischia di toccarla. Il parlante risponde aumentando la *tensione emotiva* e la verbalizzazione viene ad essere connotata emozionalmente in corrispondenza alla risposta emotiva stessa. Il cambiamento avviene dunque all'interno dell'enunciato ed è collocato nella costruzione "vai un po' più avanti", specificamente a partire dall'espressione "po'".

Il cambiamento di tensione emotiva è percepibile in connessione a qualità vocali:

- un salto repentino di F0 sulla sillaba "po'";
- un aumento dell'intensità del segmento interessato rispetto all'attacco dell'unità di comment e specificamente sulle toniche, a cui si attribuisce valore di "concitazione".

In relazione all'ipotesi che l'enunciato sia l'entità di riferimento dell'emozione si deve notare che qui la connotazione emotiva non riguarda olisticamente l'intero enunciato, ma solo la parte appena evidenziata. Quindi l'espressione dell'emozione *si inserisce* nell'enunciato, ed è espressa simultaneamente ad una sola parte di esso.

L'esempio può essere considerato, per così dire, l'eccezione che conferma la regola, in quanto è evidente che il parlante non ha connotato emozionalmente le parti dell'enunciato

riguardate dall'espressione emotiva in quanto parole o specifiche unità di informazione. Ovvero non sono le parole "po' più a vanti" o l'informazione dell'atto direttivo di comment che risultano l'oggetto del contenuto emotivo. In altri termini l'enunciato intero risulta caratterizzato dal cambio di tensione emotiva. Se ne deduce che, dal punto di vista delle qualità acustiche, non è necessario che caratteri olistici riguardino tutto l'enunciato perché questo esprima una emozione. E' interessante notare che, così come abbiamo visto che una connotazione emotiva non altera il valore dei movimenti prosodici che determinano il valore illocutorio, vediamo qui che il cambiamento di spinta emotiva non comporta necessariamente l'interruzione del programma locutivo e illocutivo di un enunciato, che in effetti si conclude correttamente senza cambi di programma.

Questo esempio permette di evidenziare una distinzione nelle modalità di categorizzazione delle emozioni da parte dei parlanti competenti. Questi giudicano, peraltro con una notevole uniformità, che la connotazione emotiva concerne un *cambiamento di tensione* nell'enunciazione, che passa dall'essere non marcata ad essere connotata, appunto, da notevole tensione. Diversamente dall'esempio precedente (rabbia-paura), però, l'attribuzione del termine *tensione* all'emozione in oggetto, non risulta determinata senza la griglia di scelte predefinite, mentre è determinata se ai parlanti competenti sono fornite le alternative fornite dalla griglia. Per cui, sebbene l'emozione risulti riconoscibile, la sua etichetta convenzionale non è accessibile al di fuori di un sistema esplicito di categorizzazione.

Da un punto di vista semantico sembra che i predicati che identificano alcune emozioni siano facilmente accessibili e disponibili ai parlanti, in particolare se corrispondono a predicati nominali di livello base, come rabbia, paura, tristezza (Rosh, 1978). Al contrario alcune scelte lessicali sono meno accessibili, se il loro *label* ha un significato tecnico come in questo caso, o comunque non corrisponde ad un concetto di base del sistema di categorizzazione naturale.

Questa differenza è certamente importante per la realizzazione di schemi di annotazione delle emozioni,[9] schemi che debbono assicurare sia l'applicazione dei label corrispondenti ad espressioni di livello base sia adeguata interpretazione delle emozioni che non hanno invece un predicato di livello base corrispondente nel lessico, come in questo caso *l'aumento di tensione*.

2.3. Emozione di tristezza: stimoli esterni e interni

Abbiamo visto che le sessioni che comportano una presenza di forti stimoli pragmatici registrano anche una frequenza significativa di enunciati marcati emotivamente. L'emergenza dell'emozione, però, può non solo essere provocata da uno stimolo esterno, ma anche dall'elaborazione, da parte del soggetto, del contenuto che è oggetto della verbalizzazione. Ovvero l'emozione insorge nella verbalizzazione in connessione ad una rappresentazione interna al soggetto, derivata dal dialogo o dal flusso del discorso del parlante stesso, che stimola una risposta emotiva.

L'esempio seguente tratto da un "*recit de vie*" dello stesso parlante (ifammn18) esemplifica bene questo tipo di realizzazione emotiva, generata all'interno di un processo di significazione propriamente umano.

---

[9] Si considerino per esempio gli schemi di annotazione di progetti quali DAMSL o VERBMOBIL (Fischer, 1999) e in particolare Crags (2004)

*MAX: ... io ci &ten [///] c' avrei tenuto tanto / che la Martina avesse / stato un pochino anche dentro nella mi' casa / <insomma> / ecco //$
%com: Martina è la figlia del parlante, alla cui nascita è stata comprata la stufa di cui si parla
*SAB: [<] [53]<certo> //$
%com: appena percepibile
*MAX: [54] ma invece / questa cosa non è successa //$
%com: alla nascita di Martina il parlante si è trasferito insieme alla figlia nella casa della sua compagna
Scarica l'audio

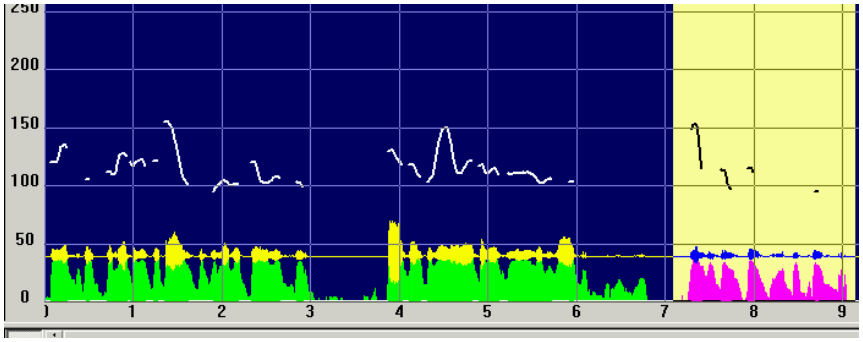

Figura 3

L'enunciato 54, in chiaro nel tracciato corrispondente all'esempio in Figura 3, riporta la circostanza che la figlia del parlante, contrariamente ai suoi desideri, non è stata mai portata a dormire nella sua casa. L'enunciato è considerato univocamente dai parlanti competenti un enunciato assertivo, e più specificamente una conclusione, connotata però da una chiara espressione emotiva di *tristezza*. L'emozione è, come nel primo esempio, riconosciuta univocamente anche al di fuori dello schema di annotazione e si inscrive nell'alveo delle emozioni universali, identificate dal predicato nominale di livello base "tristezza".

E' probabile che un certo numero di tali esempi siano in effetti riscontrabili nei racconti di vita, che è un contesto adatto all'emergenza di ricordi che assumono significati, positivi o negativi, sufficientemente forti da determinare l'insorgenza della spinta emozionale.Va sottolineata però l'episodicità di tali manifestazioni dell'emozione all'interno dell'interazione verbale ordinaria, che risulta quindi, stando al nostro corpus, *povera* dal punto di vista del contenuto emozionale. Il sondaggio porta risultati chiari da questo punto di vista. Nel testo considerato, l'enunciato appena descritto è l'unico caso su 146 enunciati nel quale vi sia una chiara espressione dell'emozione, mentre nel testo ifamdl08 ho riscontrato due enunciati su 214 enunciati rispettivamente connotati da rabbia e da rilassatezza e nessun enunciato marcato emozionalmente nel testo ifammn03, su 531 enunciati.

L'esempio propone una conferma ulteriore all'ipotesi che l'emozione pertenga effettivamente all'atto linguistico. In questo caso siamo in presenza di un enunciato composto da sette parole e che, dal punto di vista della struttura informativa, si distribuiscono una unità di informazione di *Incipit* "ma invece", seguita da una unità di *Comment*, composta a sua volta dalla frase "questa cosa 'un è successa". E' rilevante sottolineare, dunque, che il giudizio relativo alla connotazione emozionale, riguarda in

effetti l'intero enunciato e non parti di esso. Il giudizio "esprime emozione di *tristezza*" è relativo all'enunciato e non pertiene specificamente nessuna delle sue parti, né informative, né locutive, che non sono in sé connotate emozionalmente dalla tristezza.

Allo stesso modo, e questo è particolarmente evidente, la spinta emotiva non si estende aldilà dell'enunciato stesso, che appunto introduce la nota di tristezza prima assente.

Sottolineiamo quindi che l'enunciato è l'unità di riferimento della manifestazione emotiva indipendentemente dalla sua strutturazione informativa e dal numero di parole che lo compongono, e indipendentemente dall'origine dello stimolo che determina l'insorgenza dell'emozione.

L'analisi acustica evidenzia in effetti caratteri propri delle emozioni a "bassa attivazione"(Shrerer, 2003): a) la frequenza media, che si attesta su valori bassi, anche per il locutore maschile in questione (inferiori ai 100 hz); una realizzazione con scarsa intensità, che è stata sempre considerata una caratteristica tipica dell'emozione di tristezza.

Tali notazioni confermano la complementarietà dei tratti caratteristici della tristezza rispetto a quelli osservati nell'emozione di rabbia (ad "alta attivazione") e parallelamente conferma ciò che abbiamo notato nell'esempio precedente, ovvero che i tratti in questione pertengono caratteri olistici (macro-prosodici) del pattern tonale che contiene l'enunciato stesso.

Sarebbe però a nostro avviso erroneo ascrivere, in senso assoluto, l'espressione di tristezza ai caratteri appena individuati. Abbiamo notato infatti che la nota emotiva di tristezza, se è in effetti un tratto dell'enunciato in questione, non può essere apprezzata nell'ascolto dell'enunciato stesso in isolamento. Per esempio, se i parlanti competenti leggono e comprendono il significato del racconto e ascoltano, poi, in isolamento, il solo enunciato 54, questi, dopo aver notato la realizzazione "bassa", *non* gli assegnano una connotazione triste. In questo banale test l'enunciato risulta assolutamente neutro dal punto di vista emotivo. Il lettore può replicare da solo l'effetto, ascoltando la porzione acustica corrispondente all'enunciato in Figura 3', in isolamento nel file ifammn18-sel2.wav.

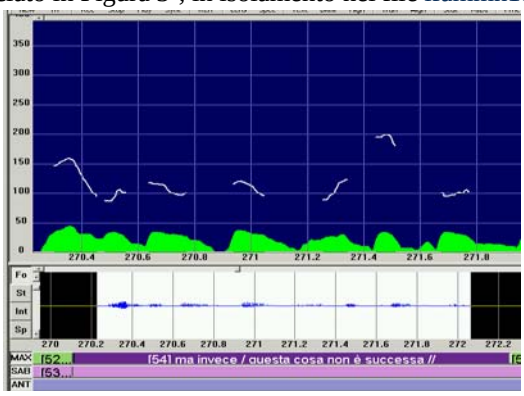

Figura 3'
Scarica l'audio

Se agli stessi informatori, però, tale enunciato viene riproposto all'ascolto insieme al contesto linguistico che lo precede immediatamente, allora tale connotazione emerge con chiarezza.

Possiamo dunque ipotizzare che sia il *processo dinamico* di abbassamento di intensità e di frequenza media, piuttosto che la semplice messa in atto di tali caratteri, a determinare la percezione dell'emozione sottostante.

2.4. Stato emotivo e atto linguistico

L'analisi su corpus conferma la nostra prima ipotesi, ovvero che l'emozione pertenga l'enunciato, ma non ha fino a questo punto contraddetto neppure la seconda ipotesi, ovvero che ad ogni spinta emotiva corrisponda uno e un solo enunciato; infatti nei tre contesti osservati ogni enunciato si realizza puntualmente in connessione ad una *modificazione* dello stato emotivo del soggetto.

Tale richiesta riproduce per la relazione tra emozione ed enunciato la stessa restrizione che vige tra espressione della forza illocutoria ed enunciato, per la quale non si dà che uno stesso atto illocutorio generi due enunciati diversi (Cresti, 2000).

La ricerca su corpus mostra, invece, e in modo teoricamente rilevante, che tale parallelismo non corrisponde ai fatti. Si consideri il turno dialogico seguente, sempre appartenente alla sessione di scuola guida:

*MAX: [42] piano //$ [43] piano //$ [44] piano //$ [45] vai //$ # [46] no //$ [47] ma &e [//] Martina //$ [48] ma che cazzo fai ?$
%sit: Martina accelera invece di frenare
Scarica l'audio

Dal punto di vista illocutorio gli enunciati 46, 47 e 48 sono rispettivamente una *protesta*, un *richiamo* e una *imprecazione* e sono tutti contrassegnati dall'espressione dell'emozione di *rabbia*, universalmente riconoscibile.Coerentemente alla prima ipotesi, l'emozione pertiene ciascun enunciato della serie, ma la seconda ipotesi, ovvero che ogni enunciato è caratterizzato da una diversa spinta emotiva, non è invece soddisfatta. Una corretta descrizione di questo esempio deve orientarsi, piuttosto, verso la constatazione che esiste un unico stato emotivo di rabbia del soggetto sul quale si innestano tre enunciati diversi.

Constatiamo allora, sulla base degli esempi concreti del corpus, che la relazione tra emozione e enunciato è di ordine diverso rispetto alla relazione tra enunciato e forza illocutoria. Non possiamo ipotizzare che la spinta emotiva determini necessariamente uno e un solo enunciato in quanto da un lato è evidente che si danno enunciati anche senza la spinta di una emozione (e sono la stragrande maggioranza), e dall'altro, come si vede qui, che più enunciati possono corrispondere una stessa spinta emotiva, che dovremmo più propriamente chiamare "stato emotivo".

La distinzione di livello tra espressione dell'emozione e espressione della forza deve dunque essere sottolineata anche a livello teorico, se non altro per il diverso rapporto che tali manifestazioni del linguaggio mostrano con il sostrato neuropsicologico. L'azione linguistica (forza) non è generata da una spinta emotiva, ma da una realizzazione affettiva (Cresti, 2000), per cui non deve sorprendere che all'interno di uno stesso stato emotivo possano darsi realizzazioni affettive e atti linguistici diversi.

3. L'ESPRESSIONE DEL PUNTO DI VISTA: UNITÀ DI RIFERIMENTO E CATEGORIZAZIONE

In sintesi abbiamo notato che nel corpus di parlato spontaneo l'espressione dell'emozione nella voce:

- è frequente solo in casi di forte stimolazione esterna, mentre è sporadica nelle conversazioni de plano;
- è fortemente definita e identificabile intersoggettivamente a livello semantico;
- pertiene l'enunciato, anche se non è in rapporto biunivoco con esso (come avviene invece per l'espressione della forza illocutoria);
- se l'emozione insorge all'interno dell'enunciato non comporta l'interruzione della programmazione;
- i suoi indici acustici riguardano caratteristiche olistiche dell'enunciato (salvo il caso particolare che abbiamo visto);
- non modifica l'informazione dei movimenti prosodici che esprimono l'azione linguistica

Il sondaggio evidenzia però, nel parlato spontaneo del corpus, espressioni degli atteggiamenti psicologici dei parlanti, rilevanti alla percezione, alla cui presenza fanno certamente riferimento le notazioni relative ad una "pervasività" dell'informazione emotiva nel parlato spontaneo. Infatti, contrariamente all'espressione vocale delle emozioni, che abbiamo detto è sporadica, tali fenomeni hanno una incidenza quantitativa notevole nel parlato e costituiscono un carattere prominente in tutte le situazioni informali da noi considerate. Ci riferiamo alla notevole quantità di enfatizzazioni espresse dalla voce, e in modo simultaneo anche dal corpo, che comunicano valenze semantiche ricche, e che caratterizzano il parlato spontaneo in situazione naturale.[10]

Notare compiutamente l'incidenza di una varietà di fattori espressivi, ma anche la loro diversità rispetto all'espressione delle emozioni propriamente dette, è in particolare rilevante se si considera la difficoltà di realizzare schemi di annotazione delle emozioni che permettano un accordo sufficiente tra annotatori. Per cui evidenziare i caratteri che distinguono l'espressione di atteggiamenti psicologici dall'espressione delle emozioni vere e proprie è, oltre che interessante teoricamente, una premessa necessaria alla formazione di efficienti sistemi di annotazione.

In questo paragrafo considereremo alcuni contesti linguistici tipici dai quali si percepiscono, attraverso segni di enfasi presenti nella voce, atteggiamenti psicologici del parlante e differenzieremo esplicitamente tali fenomeni dall'espressione delle emozioni vere e proprie.

La partizione di Damasio tra *emozione* e *sentimento*, unita alla possibilità di riferirsi ad una definizione esplicita di *affetto*, aiutano nella distinzione. Non possiamo designare infatti con il termine "contenuto emozionale" informazioni relative al sentimento del parlante (appercezione dello stato interno determinato all'emozione), né le manifestazioni affettive che informano a proposito di tale sentimento. Non possiamo designare con il termine "contenuto emotivo" movimenti affettivi del parlante o espressioni del suo sentimento che si realizzano in modo volontario o semi volontario, come avviene quando enfatizziamo con la voce parti della locuzione.

Sarà dunque interessante notare che alla diversità di livello nel sostrato neuro-psicologico fanno riscontro anche modalità diverse di manifestazione del linguaggio, sia dal punto di vista delle proprietà prosodiche, sia dal punto di vista delle qualità semantiche. In sintesi vedremo che le manifestazioni dello stato psicologico veicolate dall'enfasi:

---

[10] Il riferimento al corpo, ovvero le modificazioni di espressione, postura e i movimenti, simultanei alle enfatizzazioni, è un tratto comune all'espressione delle emozioni che deve essere sottolineato, anche se qui non portiamo dati in proposito.

- sono ad alta frequenza nel parlato spontaneo
- hanno scarsa definizione “semantica”, ovvero è indeterminata la loro qualificazione in una etichetta
- hanno carattere intenzionale
- non caratterizzano l’enunciato, bensì riguardano unità di livello gerarchico inferiore (unità tonale o parola)

3.1. L’espressione del punto di vista nel Comment

Si consideri lo scambio dialogico seguente preso da un testo che riporta un dialogo tra una padrona di casa e la donna della pulizie (ifamdl08):

*ELA: [<] [28]<ecco / lei &che> +$ [29] cioè / preferisce più / come dire / fare le pulizie / o guardare i bambini ?$
*TIZ: [30] senta / guardare i bambini / &eh / impegnan parecchio / io dico la verità //$

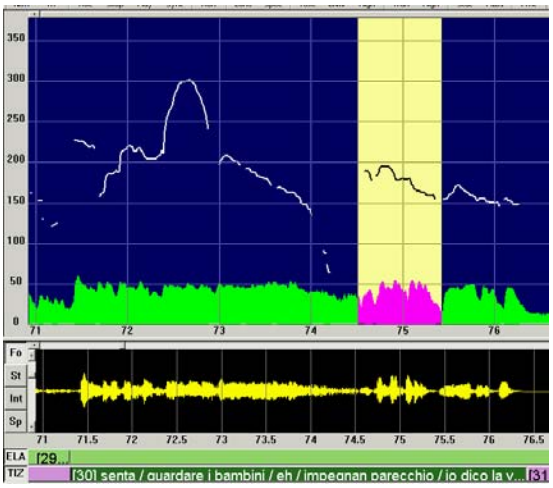

Figura 4

Scarica l’audio

L’enunciato 30 di TIZ, riportato nel tracciato, è un enunciato rappresentativo con forza di risposta, e contiene una sfumatura espressiva percepibile, enfatica, che chiaramente comunica un atteggiamento della parlante rispetto al contenuto espresso. L’enunciato è composto da quattro unità di informazione, rispettivamente con valore di *Incipit*, *Topic*, *Comment*, *Inciso*. I parlanti competenti notano, in particolare, una variazione significativa nelle modalità della locuzione del *comment* “impegnan parecchio” (in chiaro nel tracciato) che esplica, grosso modo, una “valutazione negativa” da parte della parlante, relativa alla pesantezza del lavoro con i bambini.

Osserviamo prima il fenomeno dal punto di vista semantico. Se assumiamo che tale unità di informazione sia marcata espressivamente, vediamo che, diversamente da quanto visto nel paragrafo 2., l’insieme delle risposte da parte degli uditori su cosa il parlante esprima non è univoca, ed è al tempo stesso è incerta in ogni informatore. Coloro che hanno valutato l’enunciato riportano la seguente serie di giudizi sull’atteggiamento psicologico

espresso, giudizi sempre incerti e mai concordanti: “non gli piace il lavoro, esprime *fastidio*”, “il lavoro con i bambini è troppo duro, esprime *stanchezza*”, “i bambini sono cattivi, è *arrabbiata*”, “non è contenta di lavorare con i bambini, è *scontenta*”.

Il giudizio su cosa sia espresso a livello emotivo-affettivo non è definito, ma concorda sul segno negativo che è ascritto al parlante, che, si dice, è implicito nel modo con cui questi ha realizzato la locuzione. Il lettore può replicare la valutazione dall’ascolto.

A nostro parere non è né possibile né significativo riportare ad una categorizzazione univoca il contenuto espressivo veicolato. Infatti l’espressione è semmai caratterizzata dal fatto che le possibilità interpretative sono aperte, e che il contenuto inferito è vago. Ciò che è certo è che il parlante ha voluto esprimere con le modalità enfatiche di locuzione del comment percepibili e discriminabili dall’interlocutore, una sua valutazione o atteggiamento negativo sui fatti da lui riferiti.

Ora, il concetto di “valutazione o atteggiamento negativo”, che è quanto in effetti la voce esprime, non può riguardare l’emozione, almeno all’interno di un quadro neuropsicologico compatibile con quello qui adottato. La valutazione è un processo che coinvolge il giudizio ed è quindi un carattere intenzionale. Vorremmo dire che il parlante esprime nella voce un moto affettivo legato ai significati espressi dalla scelta locutiva o, in altri termini, *esprime il suo punto di vista su ciò che dice.*

Dunque, se questa descrizione (ricalcata sul concetto di *Modus* in Bally) è esatta, non si tratta di una reazione involontaria a uno stimolo interno o esterno, ma piuttosto di una *scelta* nelle modalità di realizzazione vocale, in quanto tale intenzionale e controllata. Identificheremo qui il livello espressivo in questione con il termine di *espressione del punto di vista affettivo del parlante sulla locuzione*.

Questo livello si distingue dall’espressione dell’emozione per la sua intenzionalità, per il fatto che i significati che veicola sono, non solo non universali, ma idiosincratici e di incerta categorizzazione, e si caratterizza inoltre per la sua frequenza. Infatti fenomeni di enfasi ricca di significato occorrono nel parlato con frequenza molto alta, sia in assoluto sia in relazione alla frequenza con cui occorrono, invece, le espressioni dell’emozione. Dal sondaggio effettuato risulta, per esempio, che nel racconto di vita “storia di una stufa” visto prima, nel quale si riscontra un solo enunciato marcato emozionalmente su 146, si contano invece ben 91 contesti che portavano una enfasi e esprimevano informazioni sul punto di vista del parlante sulla locuzione. In questo colloquio con la donna delle pulizie, da una annotazione grossolana, almeno 106 contesti su 214 enunciati sono marcati da un’ enfasi.

Aldilà della distinzione esplicita tra enfasi ed espressione dell’emozione, che certo non è sorprendente, risulta interessante notare che tali manifestazioni espressive hanno caratteri differenziali rispetto all’espressione delle emozioni per il modo con cui si manifestano nella voce.

Se l’unità linguistica pertinente all’espressione delle emozioni è l’enunciato, possiamo chiederci a quale unità linguistica si riferisca *l’espressione del punto di vista sulla locuzione*. L’ipotesi che questa riguardi l’intero enunciato non trova riscontro. Per esempio, in Fig. 4 la parte dell’enunciato precedente al *comment* “senta / guardare i bambini” non condivide tale valutazione negativa. La valutazione negativa non ha per scope l’intero enunciato, ma solo una parte di esso, ovvero l’unità informativa di comment (“impegnan parecchio”). In altri termini, all’interno dello stesso enunciato convivono espressioni di punti di vista diversi del parlante, come tra l’altro è previsto dalla teoria dell’articolazione dell’informazione (Cfr. Cresti, 2000; 2003).

3.1.1 Definire quali specifici caratteri determinino qui la percezione del “punto di vista negativo” non è compito banale. A livello della descrizione dell’unità di informazione notiamo:

- un aumento di velocità di eloquio
- un notevole abbassamento della frequenza media, a fronte di una intensità costante che caratterizzano l’unità di informazione nel suo insieme
- un movimento percettivamente rilevante nel comment che evidenzia i caratteri della risposta, peraltro compatibili anche con il profilo di conclusione (profilo basso discendente) (Cfr. Cresti, Moneglia, Martin, 2003).

Abbiamo fatto pronunciare in laboratorio, alla stessa parlante, un enunciato di risposta di struttura topic - comment nel quale la locuzione del comment fosse identica (“impegna parecchio”), ma priva di sfumatura negativa.[11] Ne è derivata una locuzione equivalente non marcata. Il letture può rendersi conto direttamente del risultato ascoltando l’enunciato. La parte in giallo di Figura 5, corrisponde al suo tracciato di F0 e dal punto di vista delle proprietà acustiche, notiamo che il movimento discendente è tipologicamente identico al corrispondente in Figura 4, ma che la discesa è più repentina. La differenza tra le due locuzioni riguarda infatti esclusivamente l’altezza di F0 dell’attacco, che è di circa 30 hz superiore nell’esempio di laboratorio.

Per verificare l’imputabilità dell’effetto espressivo dell’enunciato del corpus a tale tratto differenziale abbiamo in primo luogo sostituito la locuzione “impegnan parecchio” dell’esempio naturale con la corrispondente di laboratorio, realizzata con stessa tempo e intensità. Ne risulta un enunciato, [tiziana-sel1.wav] il cui tracciato è riportato in Figura 5, grosso modo accettabile, dal quale però, non si evidenziano sfumature espressive.[12]

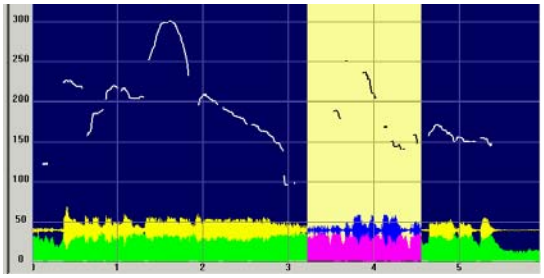

Figura 5

Scarica l’audio Tiziana1
Scarica l’audio tiziana-sel1.wav

In secondo luogo abbiamo modificato sinteticamente l’unità in questione, abbassando l’attacco all’altezza dell’unità incontrata nel corpus (come indicato dal tracciato massiccio in Fig. 6, sopra il tracciato originale), per verificare se in effetti una espressione modale negativa divenisse di nuovo percepibile nella voce.

---

[11] L’enunciato è stato selezionato all’interno di una serie di enunciati prodotti dalla locutrice. Domanda: “quanto impegnano le scale / i piatti / i dolci ?” Risposta: “le scale / i piatti / i dolci ecc. / impegnan parecchio”.

[12] Si percepisce però chiaramente una variazione di programma.

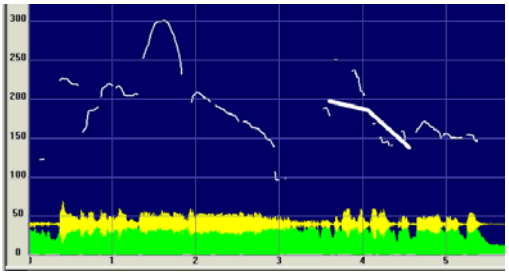

Figura 6
Scarica l'audio

Come il lettore può verificare ascoltando il file Tiziana-sel1-remod, il risultato sembra confermare questa ipotesi. In effetti un abbassamento del livello di attacco del *comment*, e una conseguente diminuzione della F0 media, porta, secondo i parlanti competenti che hanno giudicato l'esperimento, ad un marcamento della scelta lessicale, che viene così a veicolare caratteri espressivi in tutto simili a quelli dell'esempio originale.[13]

3.2. L'espressione del punto di vista nell'inciso

L'espressione di un punto di vista all'interno dell'enunciato è un carattere che riguarda in particolare alcune unità di informazione. Ci riferiamo in particolare alle unità di *inciso* e di *discorso riportato*, che per definizione hanno *sempre* una modalità specifica, diversa da quella dell'enunciato in cui compaiono. Tali unità sono state studiate recentemente nel corpus LABLITA[14] e occorrono con frequenza notevole. In particolare sull'insieme degli enunciati composti da più di una unità di informazione, l'unità informativa di *Inciso* è presente nel 19% (Cfr. Tucci, 2003) e il riportato ha una incidenza, calcolata sulla base dei dati di Giani 2005, intorno al 7%.

Non è sorprendente, dunque, che in presenza di tali unità una *espressione del punto di vista* si riscontri nella voce.

Nel monologo seguente, tratto dal file ifamnn03, il parlante racconta un episodio di un suo viaggio in Africa.

*CLA: ... venne [/] venne la sera / dove eravamo accampati / dieci anni fa / ... stiamo un altro giorno qui / perché ci dice / che c' era un matrimonio //$ [25] in un villaggio +$ [26] sicché / si va / dico / bah / allora stiamo qui / un matrimonio / è un' occasione per vederlo //$ [27] e si va / a vede' questo matrimonio / che noi / s' era già visto un matrimonio / però / insomma era + [28] era [//] non era un gran che / però / non so / c' è stato un qualche cosa / di [/] di / un po' di tensione / no &e [/] non [/] non [/] non è scattata quella molla / quell' interesse //$

Nell'enunciato 27 riportato nel tracciato, l'unità di informazione sottolineata nel turno, ed evidenziata in chiaro in Figura 7, è caratterizzata espressivamente. L'enunciato è un enunciato interrotto composto da una unità di *Incipit*, da un *Topic*, e da un *Inciso*, che precede alcuni frammenti.

---

[13] Nonostante il disturbo introdotto dalla sintesi.
[14] http://lablita.dit.unifi.it

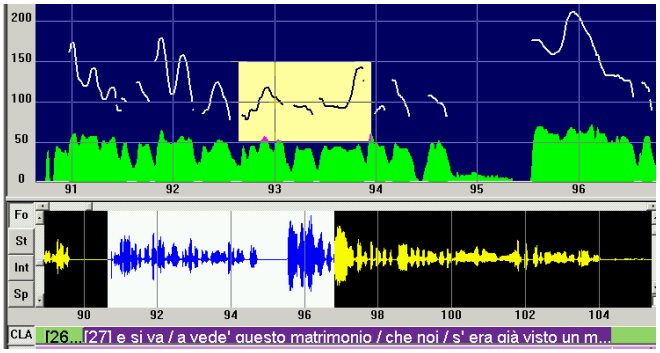

Figura 7
Scarica l'audio

L'unità è classificabile come Inciso in quanto presenta sia i caratteri funzionali che i caratteri prosodici tipici di tale unità. Dal punto di vista funzionale gli Incisi sono *inserti metalinguistici che esplicitano la valutazione modale del soggetto parlante sull'enunciato in corso di produzione* (Tucci, 2003). Dal punto di vista prosodico (Tucci e Firenzuoli, 2002) gli incisi sono realizzati in unità tonali specifiche che mostrano:

- un abbassamento repentino della media di F0 rispetto ai valori delle unità tonali precedenti e seguenti;
- un aumento della velocità di eloquio;
- un andamento di F0, generalmente piatto, opzionalmente caratterizzato da movimenti percettivamente rilevanti.[15]

Possiamo osservare che le proprietà differenziali rispetto all'espressione delle emozioni riscontrate nell'esempio precedente di enfasi nel *comment* si ripetono. A livello semantico i parlanti competenti concordano che il modo con cui è intonata l'unità tonale comunica informazioni sull'atteggiamento del parlante, ovvero che il punto di vista espresso è negativo. L'espressione di tali caratteri è certamente da considerare "intenzionale" quanto la scelta delle parole a cui si lega. I giudizi sullo specifico contenuto negativo oscillano: "per lui il fatto non era interessante, esprime *svalutazione*", "vuol dire che ne ha viste tante, esprime *sufficienza*", "vuol dire che era andato solo per *compiacere*". Anche in questo caso, quindi, la definizione del contenuto espresso dalla voce è scarsa e tutt'altro che universalmente riconoscibile.

Parallelamente a quanto abbiamo visto nell'esempio in 3.2 e contrariamente a tutti i casi di espressione delle emozioni, possiamo rilevare facilmente che la valutazione negativa non

[15] "In alcuni casi è stata rilevata la possibilità che occorresse, in posizione seguente al nucleo tonale, una porzione di coda tonale con movimento ascendente. In tal caso il profilo dell'Inciso viene a caratterizzarsi come una sequenza di due movimenti, il primo (che corrisponde al nucleo) con andamento piatto o discendente [o] o [D], il secondo (che corrisponde alla coda tonale) con andamento ascendente [1]. La coda tonale è inoltre caratterizzata da una riduzione della velocità tipica del nucleo, in corrispondenza del movimento [1], che si situa solitamente in corrispondenza delle ultime due sillabe a destra dell'unità". Tucci e Firenzuoli 2002.

riguarda l'intero enunciato. Non riguarda in particolare l'unità di *topic* ("e si va a vedè questo matrmonio") che è pronunciata in modo neutro, ma riguarda solo il contenuto espresso nell'unità di informazione di inciso ("s'era già visto un matrimonio"). Anche in questo caso, come nel precedente, si deve dunque rilevare che uno stesso enunciato contiene più punti di vista (neutro e negativo).

Se è vero che la locuzione comunica un contenuto valutativo negativo, la percezione di tale contenuto può risultare, in teoria, sia da caratteri dell'esecuzione vocale sia dalle presupposizioni pragmatiche che sottostanno all'interpretazione della locuzione: se un matrimonio tribale era stato già visto in precedenza, allora non era interessante. Ma i presupposti semantico-pragmatici, se possono certamente influenzare l'assegnazione di un punto di vista, non lo determinano necessariamente. Rimane dunque da comprendere come le modalità di esecuzione vocale selezionino, invece, univocamente tale interpretazione.

L'abbassamento della frequenza media dell'unità tonale rispetto a quelle che precedono è il tratto tipico di tutte le unità di informazione di Inciso e tale forma non subisce modificazioni in questo esempio. Notiamo però che l'unità è caratterizzata da un movimento percettivamente rilevante di marcata salita sull'ultima sillaba (non tonica) evidenziata nel dettaglio in Figura 8. Tale tratto viene percepito come una forma di enfasi.

Abbiamo fatto l'ipotesi che la percezione della sfumatura affettiva sia connessa alla collocazione di tale movimento percettivamente rilevante. In particolare si ipotizza che la collocazione del movimento abbia qualità tali da marcare la scelta lessicale operata dal soggetto ("s'era già visto un matrimonio") evidenziandone la significatività e di conseguenza la connotazione negativa.

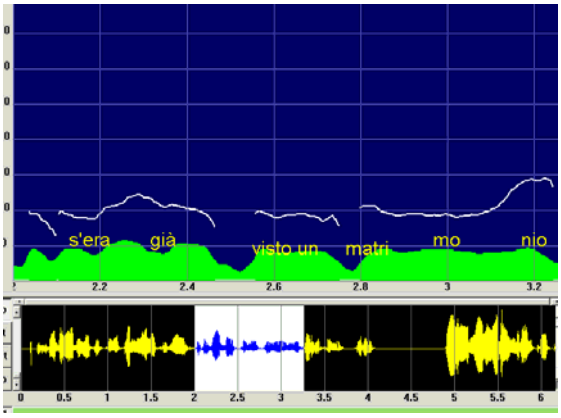

Figura 8

Un argomento a sostegno di questa ipotesi può essere ottenuto attraverso una procedura di normalizzazione in laboratorio e successiva modificazione sintetica che riproduce l'effetto di modalizzazione. Abbiamo realizzato in laboratorio un enunciato identico dal punto di vista della struttura informativa e accentuale, ma con una scelta lessicale fortemente demodalizzata ("impresario" al posto di "matrimonio"), ovvero un inciso in tutto simile ma che non comporta alcuna possibilità di presupposizione pragmatica negativa:

“allora siamo andati a vede’ questo matrimonio / che noi s’era già visto un impresario / però / insomma / era”

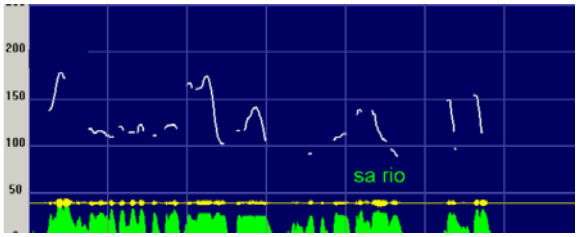

Figura 9

Scarica l’audio

Il lettore può verificare percettivamente la corrispondenza degli andamenti intonativi tra l’esempio naturale e l’esempio di laboratorio, e constatare che quest’ultimo presenta un inciso, parallelo a quello naturale, nel quale però non sono percepibili sfumature affettive o giudizi impliciti.

Come il tracciato in Figura 9 evidenzia, nell’inciso di laboratorio il movimento rilevante di F0 è collocato sulla sillaba tonica di “impresario”, anziché sull’ultima sillaba di “matrimonio”.

A partire da tale descrizione abbiamo una prima convalida descrittiva all’ipotesi che la collocazione del picco di F0 dopo la tonica, e il parallelo abbassamento di F0 sulla tonica, siano una strategia di enfasi che sottolinea la scelta lessicale, e ha come esito percettivo il marcamento di un punto di vista del parlante.

A conferma dell’ipotesi abbiamo poi prodotto due varianti di sintesi: 1) l’unità presa da corpus è stata modificata spostando il picco di F0 sulla tonica, come nell’enunciato di laboratorio (matrimonio-sel-remod); 2) l’enunciato di laboratorio è stato modificato spostando il picco di F0 sulla post-tonica, come nell’esempio marcato (matrimonio-max-remod2).

Nei grafici seguenti la linea massiccia indica ancora il profilo sintetico di F0 che modifica il profilo naturale (sottile). Il lettore può verificare che, se il picco del movimento viene posizionato sulla tonica di “matrimonio” dell’enunciato del corpus, l’effetto “modalità negativa” svanisce.

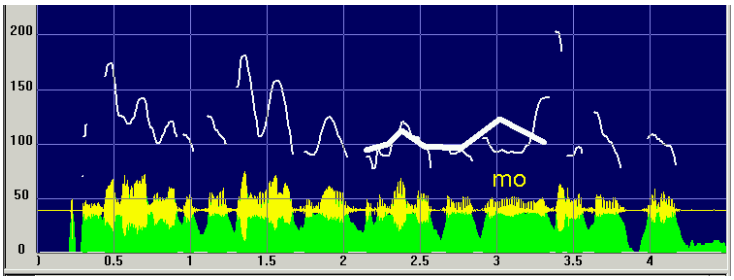

Figura 10

Scarica l’audio

Parallelamente, nell'esempio normalizzato, il picco viene portato sull'ultima sillaba e la tonica è marcata da un abbassamento di F0, come nel tracciato seguente (matrimonio max-remod.wav). L'effetto percettivo che si ottiene è di nuovo un marcamento, questa volta però tutti gli informatori concordano sul fatto che il giudizio sottinteso non è negativo, ma semmai positivo.

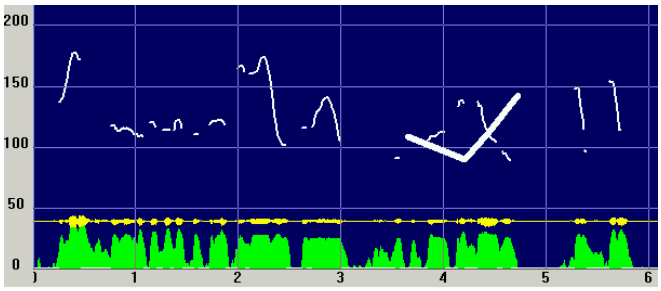

Figura 11

Scarica l'audio

L'enfasi ottenuta attraverso lo spostamento del picco di F0 dalla sua sede normale segnala, dunque, un valore della scelta semantica operata e la modalizza. Il valore positivo dato al contenuto della locuzione è infatti del tutto assente nell'esempio non sintetico.

E' assai rilevante notare, dunque, che l'enfasi non determina l'interpretazione con un suo tratto vocale: la scelta "valore positivo vs. valore negativo" non è funzione della forma dell'enfasi prosodica, che è identica, bensì della scelta lessicale.

In effetti non è sensato dare al fatto di aver visto un impresario un valore negativo, mentre è più probabile che il valore pragmatico sia positivo (magari, si ipotizza, voleva produrre un reportage dell'evento), per cui il marcamento dato dall'enfasi indica ancora un fatto espressivo, ma la determinazione semantica di cosa sia espresso risulta dal giudizio sul valore della locuzione, che l'enfasi sottende, ma non esprime. Forse è per questo che i giudizi sulla semantica di tali fattori espressivi risultano sempre non definiti: la loro semantica non è specificata da tratti dalla voce.

3.3. In conclusione abbiamo verificato che l'espressione del punto di vista affettivo del parlante si distingue dall'espressione dell'emozione per le seguenti proprietà:

- Alta frequenza nel corpus di parlato
- Scarsa definizione semantica
- Controllo volontario del locutore
- Non pertiene l'enunciato, ma l'unità informativa in cui si trova, della quale marca la scelta lessicale, evidenziandone il punto di vista sottostante.

I modi con cui tale informazione è realizzata nel segnale vocale, se da un lato hanno in comune il fatto di riguardare specificamente l'unità di informazione e non l'intero l'enunciato, dall'altro lato sono ascrivibili all' enfasi, e non è possibile dalla base osservativa presentata proporre alcuna generalizzazione. Dobbiamo rilevare però che, così come l'espressione dell'emozione in un enunciato non altera l'informazione prosodica che

ne esprime la valenza illocutoria, allo stesso modo l'enfasi, che esprime il punto di vista affettivo, non altera, per quanto possiamo constatare, le caratteristiche formali e funzionali dell'unità di informazione in cui compare. Entrambi i fenomeni, espressione dell'emozione e espressione del punto di vista, si confermano dunque, da questa analisi, strettamente *paralinguistici.*

Vorrei concludere questo lavoro rimarcando che la pertinenza all'unità di informazione, e non all'enunciato, dell'espressione enfatica del punto di vista convalida due assunti importanti della teoria dell'articolazione dell'informazione: a) la scelta lessicale pertiene l'unità di informazione; b) le unità di informazione contengono un punto di vista. I due assunti sono l'uno funzione dell'altro e l'espressività della voce lo testimonia: dato che la scelta lessicale comporta necessariamente un punto di vista, la voce che marca tale scelta si connette necessariamente all'unità linguistica nella quale la scelta si compie. [16]

BIBLIOGRAFIA

Anolli, L. & Ciceri, R., 1995, *La voce delle emozioni*, Milano, Angeli.

Bally, Ch.**,** 1932, *Linguistique gènèrale et linguistique française*. Berna, Francke Verlag.

Craggs, R., 2004. *Annotating emotion in dialogue – Issues and Approaches*, Department of Computer Science, University of Manchester. http://www.cs.man.ac.uk/~craggsr7/papers/ads04CategoricalAnnotation.pdf

Cresti, E., 2000, *Corpus di italiano parlato*. *Campioni*, Firenze, Accademia della Crusca.

Cresti, E., 2003, *Illocution et modalité dans le comment et le topic*, in A. Scarano (a cura di), *Macrosyntaxe et pragmatique: l'analyse linguistique del'oral*, Roma, Bulzoni, pp.133-182.

Cresti, E., Moneglia, M., Bacelar, F., Sandoval, A.M., Veronis, J., Martin, PH., Choukri, K., Mapelli, V., Falavigna, D.; Cid, A., 2002, *The C-ORAL-ROM Project. New methods for spoken language archives in a multilingual romance corpus*, in C. Rodriguez e C. Suarez Araujo (eds.) *Proceedings of the 3th LREC Conference* Paris, ELRA, vol. 1 pp. 2-10.

Cresti, E.. Bacelar do Nascimento, F., Moreno Sandoval, A. Veronis, J. Martin, Ph. Choukri, K., 2004, *The C-ORAL-ROM CORPUS A multilingual resource of spontaneous speech for romance languages***.** in M.T Lino, M.F. Xavier, F. Ferraira, R. Costa, R. Silva (eds.) *Prococeedings of the 4th LREC Conference*, Paris, ELRA, vol. 2, pp. 575-578.

Cresti, E. & Firenzuoli, V., 1999, *Illocution et profils intonatifs de l'italien*, «Revue Française de Linguistique Appliquée» IV(2), pp. 77-98.

[16] Si noti che le unità di appendice che sono un "eco" del testo, coerentemente alla mancanza di scelta lessicale, non sono previste avere un punto di vista. Significativamente tali unità si prevedono anche mancanti di enfasi.

Cresti, E & Moneglia, M. (eds), 2005, *C-ORAL-ROM Integrated Reference Corpora for Spoken Romance Language*s., Amsterdam-Philadelphia,. Benjamins

Cresti, E., Moneglia, M., Martin, Ph., 2003, *L'intonation des illocutions naturelles répresentatives: analyse et validation perceptive*, in Scarano (a cura di), *Macrosyntaxe et pragmatique: l'analyse linguistique del'oral*, ,Roma, Bulzoni, pp. 243-264;

Crystal, D., 1975, *The English Tone of Voice*. London: Edward Arnold.

Damasio, A., 2000, *Emozione e conoscenza*, Adelphi, Milano

Danieli, M., Garrido, J. M., Moneglia, M.; Panizza, A., Quazza, S., Swerts, M., 2004, *Evaluation of Consensus on the Annotation of Prosodic Breaks in the Romance Corpus of Spontaneous Speech "C-ORAL-ROM*, in M.T Lino, M.F. Xavier, F. Ferraira, R. Costa, R. Silva (eds.) *Prococeeding of the 4th LREC Conference* , vol. 4, Paris, ELRA, pp. 1513-1516

Firenzuoli, V., 2003, *Le Forme Intonative di Valore Illocutivo dell'Italiano Parlato: Analisi Sperimentale di un Corpus di Parlato Spontaneo (LABLITA).* PhD thesis, Università degli Studi di Firenze.

Fischer, K 1999, Annotating emotional language data. Technical report, Verbmobil - report 236.

Giani, D., 2005, *Il discorso diretto riportato nel parlato e nello scritto letterario*. PhD thesis, Università degli studi di Firenze.

't Hart, J., Collier, R., Cohen, A., 1990, *A Perceptual Study on Intonation. An Experimental Approach to Speech Melody*. Cambridge, Cambridge University Press.

Karcevsky, S., 1931, *Sur la phonologie de la phrase* «Travaux du Cercle Linguistique de Prague» IV, pp.188-228.

Magno-Caldognetto, E., 2002,, *I correlati fonetici delle emozioni*, In C. Bazzanella, P. Kobau, (a dura di) , *Passioni, emozioni, affetti*, McGraw-Hill, pp.197-213.

Magno Caldognetto, E., Zmarich, C., Ferrero, F.E., 1998, *Indici acustici macroprosodici dello stato emotivo del parlante*, in *Atti del XXVI convegno Nazionale di Acustica*, pp. 263-268.

Moneglia, M., 1999, *Senso e immagini consce nel lessico verbale italiano*, «Studi di linguistica teorica e applicata», XXVIII, 9, pp.149-172.

Moneglia, M., Fabbri, M., Quazza, S., Panizza, A., Danieli,. M, Garrido, J. M., Swerts, M., 2005, *Evaluation of consensus on the annotation of terminal and non-terminal prosodic breaks in the C-ORAL-ROM corpus* in E. Cresti, M. Moneglia (eds.). *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*, Amsterdam / Philadelphia, Benjamins, pp.257-276.

Rosch, E., 1978, *Principles of categorization* in E. Rosch e F. Lloyd (eds.) *Cognition and categorization*, Hillsdale NJ, LEA, pp.27-48.

Scherer, K.R., 1998, *Facets of Emotions: Recent Researchs*, Hillsdale, Erlbaum;

Scherer, K.R., 2003, *Vocal Communication of emotion: A Review of research paradigms*, «Speech Communication», 40, pp. 227-256.

Tucci, I. & Firenzuoli, V., 2002, *L'unita informativa di inciso: correlati intonativi*. In G. Marotta, N. Nocchi (a cura di). *La coarticolazione : Atti delle XIII giornate di studio del gruppo di fonetica sperimentale (A.I.A.)*, ETS, Pisa.

Tucci, I., 2003, C*aratteristiche morfosintattiche e intonative dell'unità informativa di Inciso*. In F. Albano Leoni, F. Cotugno, M. Pettorino, R., Savy (a cura di). 2003. *Atti del Convegno "Il parlato italiano"*. Napoli, M.D'Auria Editore.