

PARAMETRI PER LO STUDIO DELLE MODIFICAZIONI ARTICOLATORIE DEL PARLATO EMOTIVO

E. Magno Caldognetto, F. Cavicchio, P. Cosi, C. Drioli, G. Tisato
ISTC CNR Sezione di Padova
{magno, cavicchio, cosi, drioli, tisato}@pd.istc.cnr.it

SOMMARIO

L'articolo presenta la quantificazione dei bersagli labiali vocalici e consonantici /a, b, v/ nel parlato realizzato con emozioni di gioia, paura, collera, tristezza, disgusto, sorpresa. I parametri considerati: Apertura Labiale, Larghezza Labiale, Spostamento del Labbro Superiore e Inferiore, Protrusione del Labbro Superiore e Inferiore, Spostamento Verticale dell'Angolo Sinistro e Destro delle labbra, Spostamento Orizzontale dell'Angolo Sinistro e Destro delle labbra, Asimmetria Orizzontale e Verticale sono stati calcolati con il sistema optoelettronico ELITE. Le analisi dimostrano che i valori di tutti i parametri risultano significativamente diversi nel parlato emotivo per tutti i bersagli rispetto alle produzioni non emotive, che alcuni parametri, p.es. Apertura Labiale, pur nella variazione dei dati, mantengono le distinzioni fonetico-fonologiche, mentre altri, p.es. gli spostamenti degli angoli delle labbra e le asimmetrie, sono determinati prevalentemente dalla produzione delle configurazioni espressive.

1. INTRODUZIONE

Lo studio delle caratteristiche articolatorie del parlato emotivo (Magno Caldognetto et al. 2003a, 2003b, 2004a, 2004b) costituisce uno sviluppo non solo interessante, ma anche indispensabile delle tradizionali ricerche sul parlato qualora si voglia analizzarlo nelle condizioni "ecologiche" di interazione faccia-a-faccia. In tali situazioni la comunicazione è multimodale poiché, oltre alle informazioni linguistiche, vengono normalmente trasmesse anche informazioni extralinguistiche, anatomo-fisiologiche e idiosincratice, e paralinguistiche, relative agli atteggiamenti interpersonali, agli stati affettivi e alle emozioni, e tutte queste informazioni vengono inviate sia nella modalità uditivo-vocale quanto in quella visivo-cinesica.

Il parlato emotivo è un caso emblematico di comunicazione multimodale e la sua complessità giustifica la diversità e numerosità delle ricerche che ad esso sono state dedicate: lo studio degli indici acustici che trasmettono le caratteristiche linguistiche segmentali, soprasegmentali, prosodiche e intonazionali, del parlato; le ricerche sui correlati acustici delle emozioni: modificazioni dell' F_0 , della qualità vocalica, della prosodia e dell'intonazione; l'individuazione degli indici visivi, estratti dai movimenti linguistici di labbra, mandibola e lingua, che trasmettono informazione sui visemi vocalici e consonantici; l'analisi degli indici visivi labiali e facciali correlati alle configurazioni tipiche delle emozioni e infine la determinazione delle caratteristiche della voce e della faccia dei parlanti, che veicolano le informazioni extralinguistiche e idiosincratice che costituiscono il sostrato, la base-line acustica e visiva, su cui si (sovra)impongono le realizzazioni delle informazioni linguistiche e paralinguistiche.

Queste tematiche sono state sviluppate da studiosi di diversa formazione: per esempio, i fonetisti (linguisti, fisici acustici, ingegneri delle telecomunicazioni) hanno studiato prevalentemente le caratteristiche acustiche, articolatorie e percettive delle unità fonologiche segmentali e soprasegmentale, gli psicologi hanno analizzato le configurazioni facciali e gli indici acustici vocali che veicolano le emozioni; ancora gli psicologi hanno elaborato le teorie sulla cosiddetta "Lettura Labiale" (lipreading o speechreading), e sulla costituzione su base percettiva dei Visemi, mentre fisiologi e foniatristi hanno studiato le caratteristiche extralinguistiche della voce dipendenti da sesso, età e stato di salute e infine gli ingegneri delle telecomunicazioni, informatici, specialisti di computer graphics e specialisti di IA si sono dedicati alla sintesi e al riconoscimento automatici del parlato uni- e bi-modale.

Solo recentemente si è affermata l'importanza teorica della multimodalità comunicativa e si sono studiate le interazioni e i sinergismi tra le diverse componenti uditive e visive del parlato, anche emotivo, e il loro ruolo nell'evoluzione, nell'uso e nell'apprendimento del linguaggio da un punto di vista neurofisiologico, antropologico e cognitivista.

Lo sviluppo di queste ricerche è stato indubbiamente favorito anche dalle aumentate potenzialità delle strumentazioni a disposizione degli studiosi: oggi è infatti possibile, per esempio, registrare e analizzare contemporaneamente segnali articolatori e acustici tramite sistemi elettromagnetici (p.es. ARTICULOGRAPH), elettropalatografici (prodotti da KAY ELEMETRIC Co. da RION) oppure optoelettronici (ELITE, OPTOTRAK) e utilizzare programmi informatici per la visualizzazione contemporanea di segnali acustici e immagini (p.es. ADOBE) o per la segmentazione e l'etichettatura di segnali acustici e immagini video registrate (p.es. ANVIL).

Uno sviluppo particolarmente interessante delle ricerche sul parlato emotivo, che integra i risultati delle tematiche sopra elencate, è l'implementazione delle Facce Parlanti Espressive, cioè i programmi di sintesi bimodale uditivo-visiva del parlato che si propongono, dal punto di vista fonetico-fonologico, come simulazioni delle relazioni causali tra movimenti articolatori e prodotto acustico, come metodo di indagine dei rapporti tra percezione uditiva dei segnali acustici del parlato ed estrazione di informazioni fonologiche dai movimenti articolatori visibili e come banco di prova delle interazioni tra movimenti articolatori e configurazioni facciali espressive. Proprio per le loro caratteristiche di ecologicità (poiché riproducono le condizioni di interazione faccia a faccia trasmettendo informazione uditiva e visiva), di naturalezza (poiché riproducono la fusione di informazione linguistiche, paralinguistiche ed extralinguistiche) e di robustezza (poiché la comunicazione bimodale assicura la comprensibilità del messaggio in condizioni ambientali che alterano la trasmissione del segnale acustico o in caso di deficit uditivi del Ricevente), le Facce Parlanti hanno grandi possibilità applicative come interfacce uomo-macchina in tutti i sistemi di distribuzione di informazione (notiziari, prenotazioni, didattica, e-learning ecc.) e negli accessi a banche di dati, oltre che nell'industria dell'intrattenimento.

2. LO SCOPO DELLA RICERCA

E' noto che le emozioni (provate) inducono nell'Emittente rilevanti cambiamenti nell'attività del sistema nervoso autonomico che causano variazioni del battito cardiaco, del ritmo respiratorio, della sudorazione, ecc (D'Urso e Trentin 2001, pp.27-70) da cui dipendono nel parlato emotivo modificazioni segmentali e soprasegmentali. Questi effetti sono stati indagati prevalentemente dal punto di vista acustico (per una revisione: Scherer 2002 e 2003), con poche eccezioni riguardanti in particolare la forma dell'orifizio labiale e

le correlate variazioni spettrali nell'happy talk (p.es. Aubergé e Cathiard 2003, Schroeder et al. 1998).

All'interno di questa ampia e complessa problematica ci siamo proposti di individuare le regole che presiedono alla fusione tra gli atteggiamenti labiali, determinati dalle caratteristiche articolatorie "canoniche" dei foni vocalici e consonantici dell'italiano, e le configurazioni labiali che veicolano l'espressione delle emozioni. Le immagini videoregistrate riprodotte in Fig. 1, estratte dal corpus di produzioni emotive analizzato, esemplificano le modificazioni che i bersagli labiali consonantici e vocalici da noi scelti, /a, b, v/, subiscono a seconda delle emozioni (cfr. § 8).

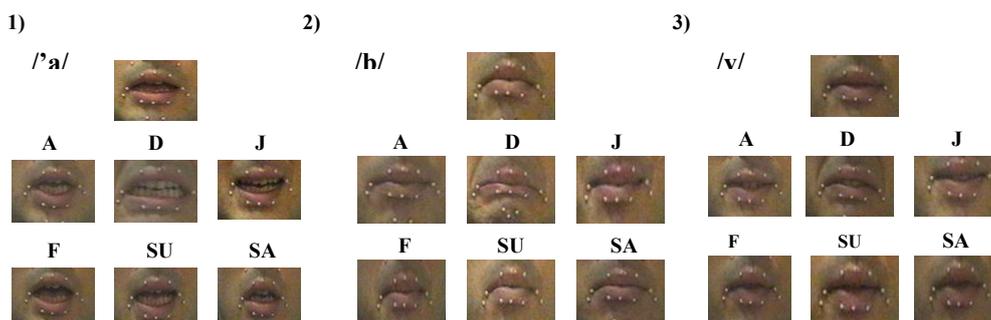


Figure 1.1, 1.2, 1.3. Esempi di configurazioni labiali emotive relative al punto mediano dei bersagli articolatori analizzati /a, b, v/.

Gli interrogativi a cui abbiamo cercato di rispondere si possono riassumere come segue:

- le realizzazioni emotive dei bersagli vocalici e consonantici si differenzino significativamente dalla realizzazione non emotiva e tra loro?
- vi sono parametri labiali i cui valori sono vincolati dalle caratteristiche fonetico-fonologiche dei diversi foni e che quindi risultano meno modificati nel parlato emotivo?
- vi sono, al contrario, parametri che dipendono prevalentemente dalla realizzazione delle configurazioni labiali emotive?

Per rispondere a queste domande si è scelto di analizzare le modificazioni spaziali 3D dei bersagli labiali di /a/, /b/, /v/ nel parlato emotivo, poiché precedenti ricerche avevano descritto, per l'italiano, le caratteristiche 3D dei bersagli labiali delle vocali, toniche e atone (Magno Caldognetto et al. 1995, 1998), di tutte le consonanti coarticolate con /a/ (Cosi e Magno Caldognetto 1996, Magno Caldognetto et al. 1997, 1998) e delle consonanti occlusive bilabiali e costrittive labiodentali coarticolate con /a/, /i/, /u/ (Zmarich e Magno Caldognetto 2003).

Naturalmente nella valutazione dei risultati di questa ricerca-pilota si deve tener conto della specificità delle condizioni sperimentali: l'analisi delle caratteristiche articolatorie del parlato emotivo qui presentata è basata infatti su materiale registrato in laboratorio, come avviene per tutte le ricerche sul parlato che implicano la registrazione di segnali articolatori (cfr. sistemi di risonanza magnetica MRI e fotogrammetrici in Badin et al. 2002; sistemi

1 Le emozioni analizzate, oltre alla produzione non emotiva N, sono Rabbia (Anger), Disgusto (Disgust), Gioia (Joy), Paura (Fear), Sorpresa (Surprise), Tristezza (Sadness).

optoelettronici a marker attivi quali OPTOTRAK in Yehia et al. 2002 o a marker passivi quali ELITE in Magno Caldognetto et al. 1995, 1997, 1998, 2000, 2003a, 2003b, 2004 a), b))².

Comunque i dati raccolti, anche se relativi ad un unico parlante (cfr. § 7), costituiscono secondo noi una interessante base di dati per illustrare le problematiche, discutere la metodologia e proporre le prime regole di interazione/integrazione tra l'informazione linguistica e paralinguistica.

Per quanto riguarda la metodologia, il problema della scelta dei parametri, la valutazione della loro adeguatezza e i criteri di quantificazione dei dati è di fondamentale importanza.

3. LA SCELTA DEI PARAMETRI

3.1 I parametri labiali per le vocali

La letteratura fonetica ha individuato in *spread*, *neutral* e *rounded* le primarie dimensioni significative per la descrizione degli atteggiamenti labiali nelle vocali (per una revisione Ladefoged 2001 pp.407-408, Laver 1994 pp. 278-282.), dato che le caratteristiche di apertura/chiusura delle labbra dipendono principalmente dai movimenti della mandibola, anche se il *rounding*, determinando la riduzione della dimensione labiale orizzontale, determina anch'esso un'espansione della dimensione verticale. Oltre a questi tratti, anche il tratto di protrusione/ritrazione è importante perché non solo può modificare l'area interlabiale, ma anche il volume dell'orifizio labiale. Sebbene vi siano tendenze universali nella cooccorrenza non solo tra le configurazioni labiali e le posizioni della lingua che distinguono le vocali anteriori e posteriori, in quanto di solito le vocali posteriori sono arrotondate e le vocali anteriori appiattite, ma anche tra l'arrotondamento e la protrusione, sono possibili anche delle dissociazioni identificate da studi linguo-specifici (cfr. per una revisione: Farnetani 1999). Per l'italiano le caratteristiche articolatorie labiali³ delle vocali toniche, atone e isolate, sono state quantificate in base ai parametri di Apertura Labiale (LH), Larghezza Labiale (LW), Protrusione Superiore e Inferiore (UP e LL), calcolati su dati optoelettronici registrati e analizzati con il sistema ELITE (Magno Caldognetto et al. 1995, 1998). Il grafico riportato in Fig. 2 presenta i valori medi normalizzati per le tre

² L'estrazione di dati ottenuti dall'elaborazione di immagini televisive, metodo auspicabile perché, non richiedendo l'applicazione di marker o di sensori di alcun tipo, non interferisce con altre attività del parlante, presenta ancora problemi di adeguatezza, in particolare per la valutazione delle caratteristiche tridimensionali dei movimenti e atteggiamenti labiali (cfr. Kroos et al. 2002 per un confronto tra dati ottenuti da sistemi optoelettronici attivi, OPTOTRAK, e da misurazioni estratte da immagini videoregistrate).

³ Le sigle utilizzate nel seguito corrispondono alla terminologia inglese dei parametri: LO=Lip Opening (apertura labiale, indicata anche come LH, altezza labiale), UL=Upper Lip vertical displacement (spostamento verticale del labbro superiore), LL=Lower Lip vertical displacement (spostamento verticale del labbro inferiore), LR= Lip Rounding (arrotondamento, indicato anche con LW=Lip Width), ULP=Upper Lip Protrusion (protrusione del labbro superiore), LLP=Lower Lip Protrusion (protrusione del labbro inferiore), LCX=Left Corner horizontal displacement (spostamento orizzontale dell'angolo sinistro), RCX=Right Corner horizontal displacement (spostamento orizzontale dell'angolo destro), LCY=Left Corner vertical displacement (spostamento verticale dell'angolo sinistro), RCY=Right Corner vertical displacement (spostamento verticale dell'angolo destro), AsymX=Horizontal Asymmetry (asimmetria orizzontale), AsymY= Vertical Asymmetry (asimmetria verticale).

tipologie di vocali esaminate. Le analisi statistiche, riportate negli articoli sopra citati, hanno dimostrato la significatività di tutti i parametri considerati, la correlazione positiva tra protrusione superiore e inferiore ($r=.82$), la correlazione negativa tra larghezza labiale e protrusione superiore ($r=-.81$) e tra larghezza labiale e protrusione inferiore ($r=-.73$), mentre nessuna correlazione è stata individuata tra altezza labiale e larghezza labiale e tra altezza labiale e protrusione superiore e inferiore, risultati che coincidono con i dati ottenuti per l'Inglese (Fromkin 1964, Linker 1982) e il Francese (Abry e Boe 1986).

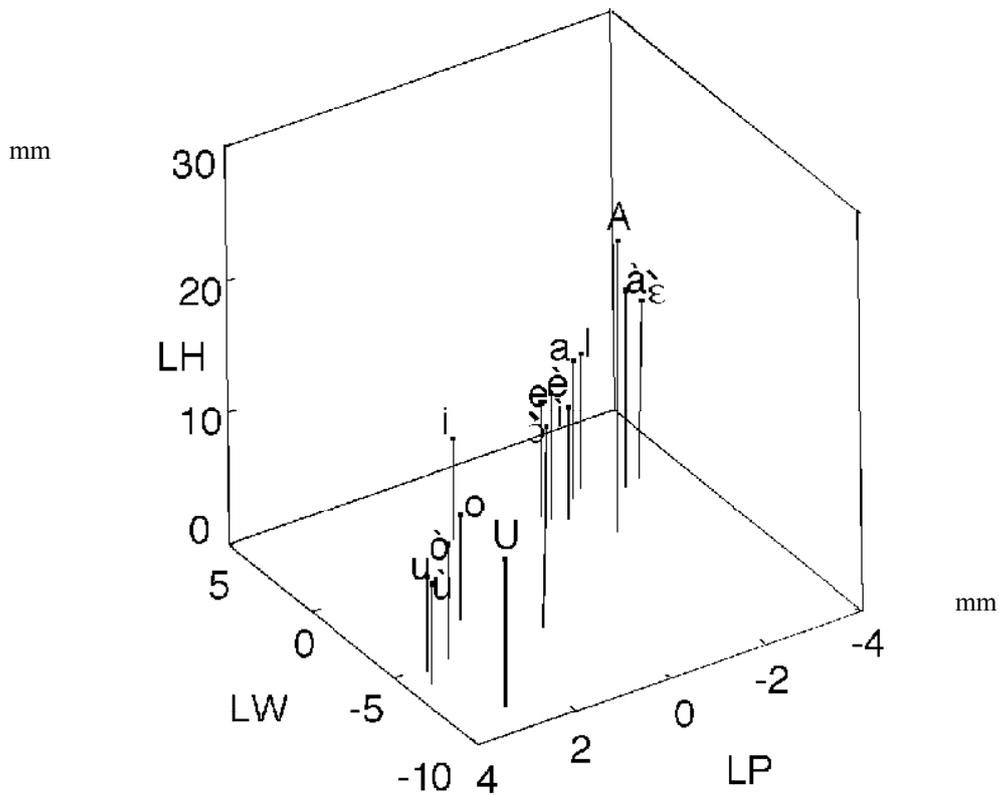


Figura 2. Configurazioni spaziali 3D dei bersagli labiali delle 7 vocali toniche /'i, 'e, 'E, 'a, 'O, 'o, 'u/, delle 5 vocali atone /i, e, a, o, u/ e delle tre vocali cardinali isolate /A, I, U/, basate sui valori normalizzati, espressi in mm, di LH, LW e UP.

3.2 PARAMETRI LABIALI PER LE CONSONANTI

Anche le configurazioni labiali per tutte le consonanti dell'italiano sono state quantificate in base agli stessi parametri utilizzati nelle ricerche sulle vocali e ottenuti con lo stesso dispositivo sperimentale (Magno Caldognetto et al. 1997, 1998).

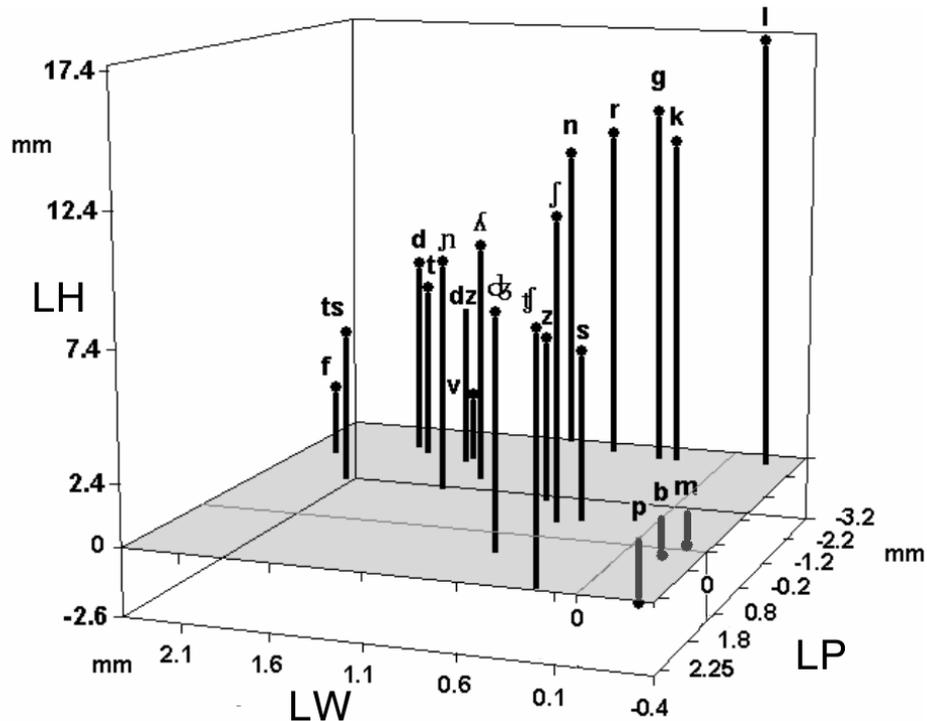


Figura 3. Configurazioni spaziali 3D dei bersagli labiali delle 21 consonanti dell'italiano coarticolate con /a/, basate sui valori normalizzati, espressi in mm, di LH, LW e UP.

Il grafico di fig. 3 presenta la configurazione 3D dei bersagli consonantici in coarticolazione con /a/. Le analisi statistiche (Cluster Analysis) riportate negli articoli sopra citati hanno dimostrato che il parametro più significativo per distinguere tra loro le consonanti è l'Apertura Labiale, che permette di differenziare, 6 gruppi:

- /p, b, m/: -2.535 / -1.273
- /v, f/: 2.652 / 2.844 ;
- /ts, dz, z, s, t, d/: 5.879 / 7.808;
- /tʃ, dʒ, ʎ, ɲ, ɳ, ʎ/: 8.798 / 9.553;

- /ʃ, n, k, r, g/: 11.716 / 14.443;
- /l/: 17.385.

Questi dati, che sono valori medi normalizzati, evidenziano un primo gruppo caratterizzato dalla compressione labiale e costituito dalle consonanti occlusive bilabiali (i valori negativi indicano la riduzione dell'apertura labiale rispetto alla posizione di riposo) e un secondo gruppo caratterizzato da valori di costrizione labiale massima e costituito dalle consonanti costrittive labiodentali. Per queste due classi consonantiche i valori di Apertura Labiale corrispondono alle caratteristiche di modo e luogo di articolazione, dato che per queste consonanti le labbra sono gli articolatori attivi e i loro movimenti sono fonologicamente specificati (Magno Caldognetto et al. 1997, 1998; Zmarich e Magno Caldognetto 2003).

Per tutte le altre classi di consonanti, il cui articolatore primario è la lingua, la configurazione labiale è determinata prevalentemente dal comportamento cooperatorio di lingua-mandibola-labbra, e dagli effetti coarticolatori della vocale contestuale aperta centrale /a/ (Fletcher e Harrington, 1999).

4. LE AZIONI MUSCOLARI SOTTOSTANTI AI PARAMETRI LABIALI FONETICI

E' interessante ricordare quali sono i muscoli coinvolti nella esecuzione delle azioni di rilevanza fonetica da cui dipendono i valori dei parametri labiali citati nel § precedente perché il confronto con la descrizione delle FACS, che sarà riportata nel § 5, evidenzierà casi di attivazione di muscoli implicati nella realizzazione tanto di vocali e consonanti quanto di caratteristiche espressive e casi di muscoli coinvolti quasi esclusivamente nella produzione delle configurazioni labiali emotive.

Secondo Hardcastle (1976, pp. 112-120, ma vedasi anche Farnetani 1999) i muscoli che eseguono le azioni di rilevanza fonetica sono i seguenti:

- la chiusura delle labbra viene realizzata dall'*orbicularis oris* che opera principalmente l'adduzione delle labbra tanto nell'occlusione di [p, b, m] quanto nel contatto del labbro inferiore contro i denti superiori in [f]. In questa azione l'*orbicularis* è assistito dai muscoli che innalzano la mandibola, cioè dal *masseter*, dal *temporalis* e dal *pterygoideus internus*. L'*orbicularis* è implicato anche nell'arrotondamento e nell'arricciamento (*puckering*) delle labbra).
- l'innalzamento del labbro superiore, che caratterizza la produzione della costrizione labiodentale di [f], è determinato dall'azione dello *zygomaticus minor*. Anche il *levator labii superior* opera sinergisticamente al muscolo precedente nella produzione di [f].
- Il *levator labii superior alaeque nasi* solleva specificamente la parte mediana del labbro superiore e contemporaneamente le narici.
- l'abbassamento del labbro inferiore è prodotto dal *depressor labii inferior* che quindi contribuisce al movimento di apertura di [p, b, m].
- l'arrotondamento delle labbra, caratteristico di [u], è realizzato dall'*orbicularis oris*. Nella produzione delle diverse gradazioni di arrotondamento l'*orbicularis* opera in associazione con i muscoli che elevano o abbassano la mandibola.
- la protrusione delle labbra è determinata dall'attività del muscolo *mentalis* che rende protruse le vocali arrotondate e può contribuire alla chiusura delle labbra nelle consonanti. Anche l'*orbicularis*, in particolare con le sue fibre profonde, agisce sinergisticamente con il *mentalis* e con i muscoli che abbassano il labbro inferiore per

realizzare le diverse configurazioni delle vocali arrotondate medio-alte e medio-basse e per realizzare diversi gradi di protrusione.

- la ritrazione degli angoli della bocca è operata da più muscoli: dal *buccinator* che, ritraendo gli angoli della bocca, comprime la labbra contro i denti ed è attivo nella realizzazione delle vocali anteriori [i, e] "appiattite", "stirate" e nell'articolazione di [f], consonante caratterizzata dallo stiramento del labbro inferiore; dallo *zygomaticus maior* che invece stira lateralmente e all'insù gli angoli delle labbra ed è attivo nella produzione delle vocali anteriori appiattite, di [f] e di [s]; infine dal *risorius* che contribuisce alla ritrazione degli angoli delle labbra, come i due precedenti muscoli, nella produzione di [i, e], [s] e [f].
- l'innalzamento degli angoli della bocca è realizzato dal *levator anguli oris* che, innalzando gli angoli delle labbra e il labbro superiore, contribuisce alla fase di chiusura delle occlusive bilabiali [p, b, m] e dallo *zygomaticus maior* che contribuisce alla realizzazione di [f], stirando gli angoli delle labbra verso l'alto e lateralmente.
- l'abbassamento degli angoli della bocca è infine affidato al *depressor anguli oris* e al *platysma*.

5. LE FACS (FACE ACTION CONTROL UNITS)

Dobbiamo principalmente ad Ekman (Ekman, Friesen 1978; Ekman et al. 2002) l'accurata ricerca sulle Action Units che descrivono l'attività muscolare sottostante alle diverse possibili modificazioni espressive della faccia, che possono essere apprese e riconosciute. Le azioni più importanti che determinano modificazioni della forma delle labbra sono quelle che le allungano (*elongate*), le accorciano (*de-elongate*), le assottigliano (*narrow*, con riferimento alla parte rossa delle labbra), le ingrandiscono (*allargano*, *widen*: sempre in riferimento alla parte rossa delle labbra), le appiattiscono (*flatten*), le protrudono (*protruded*), le contraggono (*tighten*), le stirano (*stretch*), le arricciano in dentro (*rolled in*)⁴. Per ciascuna è anche proposta una scala di visibilità che corrisponde a punteggi progressivi di intensità (*slight, marked, pronounced, severe, extreme, maximum*) in cui gli intervalli non sono eguali e le differenze sono comunque molto ridotte, dato che le gamme dei movimenti coinvolti sono limitate spazialmente.

Le FACS coinvolte primariamente nella produzione delle configurazioni labiali emotive si possono suddividere in:

1) unità coinvolte nella azioni di innalzamento e abbassamento⁵ (p.92 e sgg.):

- AU 10: realizza l'innalzamento del labbro superiore (*upper lip raiser*, pp.95-97) grazie all'azione del *levator labii superioris*;
- AU 15: produce l'abbassamento degli angoli delle labbra (*lip corner depressor*, pp.100-102), realizzato dal *depressor anguli oris*: determina la spinta degli angoli verso il basso, quindi contemporaneamente determina la dimensione orizzontale del labbro inferiore, che viene stirato. L'azione può essere graduata in intensità (tale azione comporta creazione di rughe, rigonfiamenti, protuberanze, cambiamenti di forma della pelle sottostante il labbro).

⁴ Anche nel seguito del paragrafo si riportano le etichette proposte dagli Autori, dato che essi volutamente non usano la tradizionale terminologia fonetica articolatoria e intendono sottolineare la funzione espressiva dei movimenti labiali descritti

⁵ Le indicazioni delle pagine si riferiscono ad Ekman et al. 2002.

- AU 16: determina l'abbassamento del labbro inferiore (*lower lip depressor*, pp.113-117) realizzato dal *depressor labii inferiori* che fa abbassare il labbro inferiore, scoprendo i denti (creando anche rughe sulla pelle del mento). Questa azione può essere associata ad abbassamento della mandibola e vi possono essere interazioni tra AU 15 e 17 (pp.118-119).

2) Unità coinvolte nelle azioni labiali orizzontali (p. 145 e sgg.)

- AU 20: crea lo stiramento delle labbra (*lip stretcher*, p.146), realizzato dal *risorius* e dal *platysma* il cui effetto è di stirare lateralmente gli angoli delle labbra;
- AU 14: realizza l'arrotondamento delle labbra (*lip dimpler*, p.148), realizzato dal *buccinator* che determina piccoli avvallamenti orizzontali delle labbra, che stringe e spinge verso l'interno gli angoli delle labbra.

3) Unità coinvolte nelle azioni labiali oblique (p.175 e sgg.)

sono costituite da gruppi di muscoli che spingono verso l'alto ad angolo obliquo la pelle della faccia, in particolare sono importanti per la produzione del sorriso le due seguenti Action Units:

- AU 12: produce l'innalzamento degli angoli delle labbra (*lip corner puller*, pp.178-181): l'azione di questa unità, realizzata dallo *zygomaticus maior*, spinge in su gli angoli delle labbra, che possono essere sia chiuse che aperte. L'innalzamento può essere bi- o uni-laterale.

Sono possibili molte gradazioni di intensità, p. es. l'AU 13, *sharp lip puller* (pp.182-184), determina l'innalzamento massimo degli angoli della labbra.

4) unità coinvolte nelle azioni labiali "orbicolari" (p. 231 sgg.)

Sono 5 unità, realizzate tutte dall'*orbicularis oris*, anche in connessione con altri muscoli, a seconda dell'azione specifica:

- AU 18: determina l'arrotondamento delle labbra tramite l'arricciamento (*lip pucker*);
- AU 22: crea la protrusione delle labbra e quindi un orifizio "ad imbuto" (*lip funneler*);
- AU 23: produce l'appiattimento, l'assottigliamento delle labbra (*lip tightener*);
- AU 24: realizza la compressione delle labbra (*lip presser*);
- AU 28: risucchia le labbra (*lips sucker*).

E' evidente che gran parte delle azioni labiali previste dal FACS coincidono con quelle che realizzano i movimenti articolatori, con l'eccezione, forse, di AU 16 "*lower lip depressor*", di AU 14 "*lip dimpler*" o delle unità coinvolte nelle azioni labiali oblique. Sembra ragionevole quindi ipotizzare che la differenza tra movimenti labiali espressivi e la produzione articolatoria linguistica non sia da individuare nella presenza vs assenza di specifiche unità di azione, ma piuttosto in gradazioni diverse di intensità di attivazione e nelle associazioni di più unità.

6. LE CONFIGURAZIONI LABIALI DELLE EMOZIONI

Sulla base dell'analisi di questi schematici elenchi, si è deciso di integrare l'elenco dei parametri messi a punto per lo studio delle caratteristiche fonetico-fonologiche 3D dei bersagli labiali di /a, b, v/, presentate in molte pubblicazioni precedenti a cui si rimanda per i criteri di quantificazione (Magno Caldognetto et al. 1995, 1997, 1998):

- Apertura Labiale (LO);
- Spostamento Verticale del Labbro Superiore (UL);

- Spostamento Verticale del Labbro Inferiore (LL);
- Arrotondamento Labiale (LR= larghezza Labiale, LW);
- Protrusione del Labbro Superiore (ULP);
- Protrusione del Labbro Inferiore (LLP).

con il calcolo di sei nuovi parametri (2003a, 2003b, 2004 a) e 2004 b)):

- Spostamento Orizzontale dell'Angolo Sinistro (LCX);
- Spostamento Orizzontale dell'Angolo Destro (RCX);
- Asimmetria orizzontale (ASYMX);
- Spostamento Verticale dell'Angolo Sinistro (LCY);
- Spostamento Verticale dell'Angolo Destro (RCY);
- Asimmetria verticale (ASYMY).

7. METODO

Per raccogliere i dati articolatori è stato utilizzato il sistema optoelettronico di analisi del movimento tridimensionale ELITE, che permette anche la registrazione sincrona del segnale acustico (per applicazioni precedenti di questo sistema alla definizione dei visemi dell'italiano su base articolatoria: Magno Caldognetto et al. 1995, 1997, 1998). Questo sistema assicura la massima accuratezza e il minimo disagio per il soggetto in quanto registra gli spostamenti della luce infrarossa riflessa da piccoli (2 mm diametro) marker passivi incollati su differenti punti della faccia e del contorno esterno delle labbra (vedasi Fig. 4).

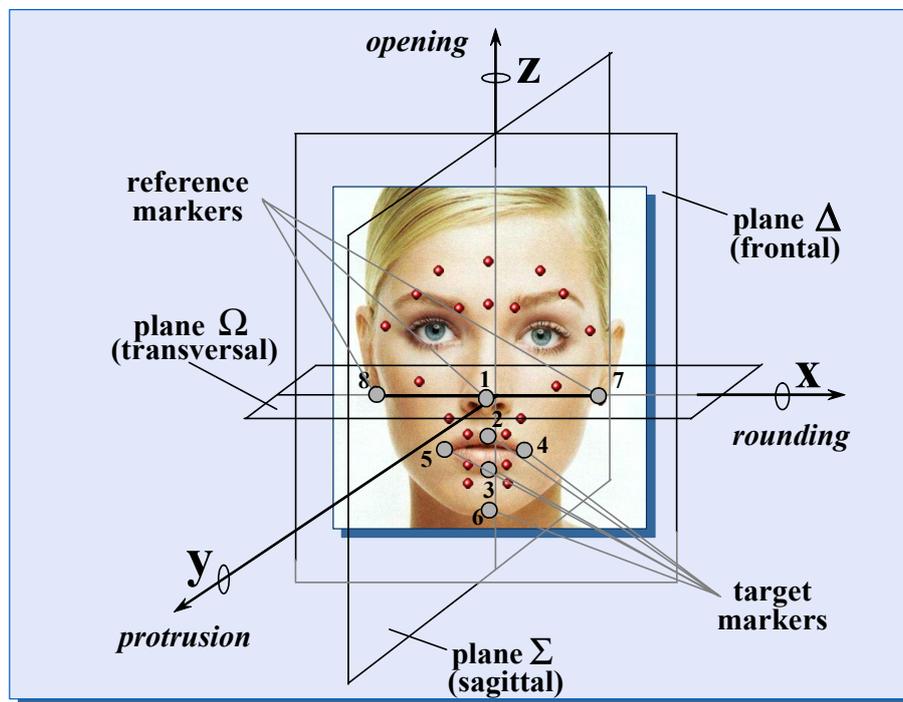


Figura 4. Posizione dei 28 marker e dei piani per la raccolta dei dati articolatori.

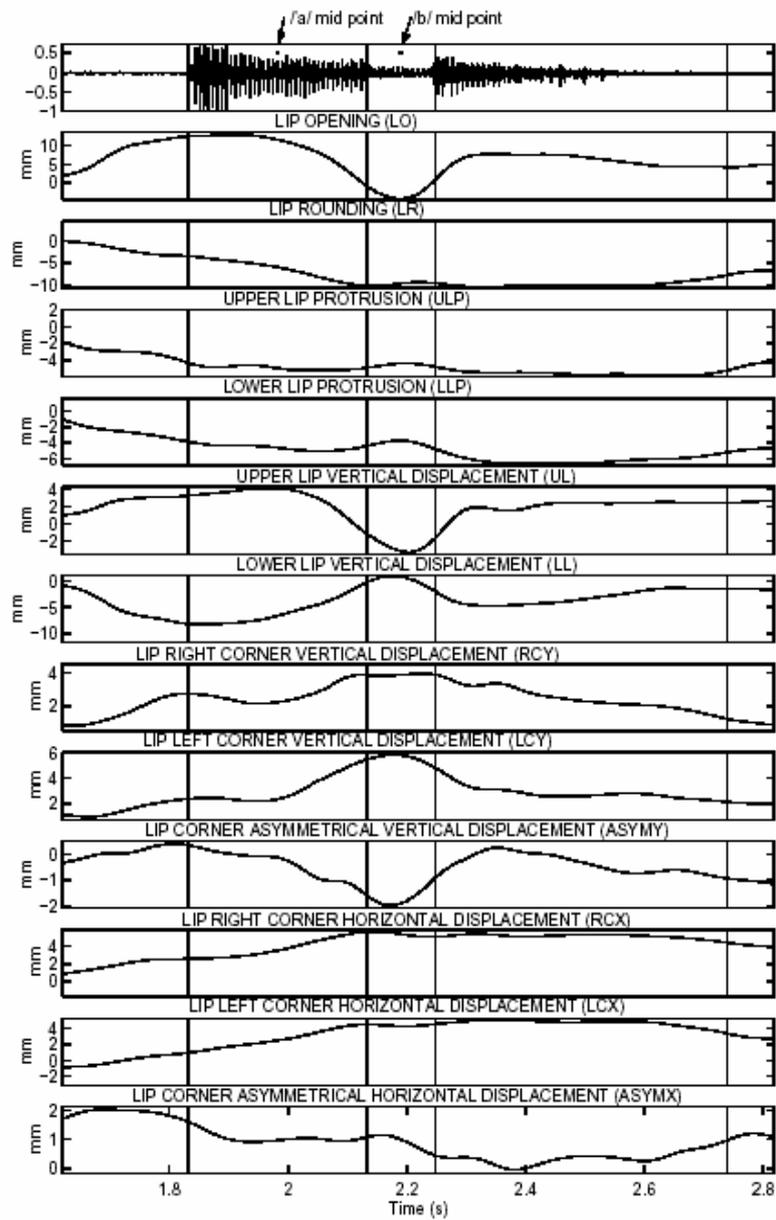


Figura 5. Evoluzione temporale di alcuni parametri cinematici corrispondenti alla sequenza /'aba/ prodotta con emozione di **disgusto**.

8. RISULTATI

Un soggetto con competenze recitative ha pronunciato due strutture fonologiche 'VCV, corrispondenti a due nomi propri femminili, "Aba" /'aba/ e "Ava" /'ava/, basandosi su scenari appropriati, con le sei emozioni primarie: collera, gioia, disgusto, paura, sorpresa, tristezza, oltre alla produzione non emotiva (neutra), corrispondente alla frase dichiarativa. Sono state così ottenute 107 registrazioni, per ciascuna delle quali sono stati calcolati gli andamenti dei 12 parametri sopraelencati per un totale di 2568 registrazioni analizzate.

8.1 Prima normalizzazione

In Tabella 1 sono riportati i valori di tutti i parametri labiali dei tre bersagli /'a, b, v/ nelle produzioni non emotive (N), indispensabili per la valutazione delle variazioni emotive.

I valori di LO (= LH) e di LL variano notevolmente, dato che realizzano le caratteristiche fonetico-fonologiche di massima apertura labiale per /a/, di compressione per /b/ e di massima costrizione per /v/, mentre i valori di LR (=LW) risultano molto simili a quelli della posizione di riposo.

	UL	LL	LO	ULP	LLP	LR	LCX	RCX	ASYMX	LCY	RCY	ASYMY
/a/	-0.04	-12.0	14.0	-1.3	-1.7	1.7	-1.0	-0.7	0.33	3.2	3.4	0.19
/b/	-1.1	-1.3	-0.5	-1.3	0.4	-0.3	0.2	0.1	-0.072	2.1	3.0	0.809
/v/	-0.06	-1.1	1.9	-1.2	-2.3	0.2	-0.08	0.3	0.35	-0.2	-0.3	-0.008

Tabella 1. Quantificazioni parametriche dei bersagli labiali per /'a, b, v/ nella produzione non emotiva.

Gli spostamenti orizzontali degli angoli delle labbra e la relativa asimmetria sono invece minimi, mentre gli spostamenti verticali sono presenti in /a/ (in questo caso determinati dall'apertura della vocale, realizzata con abbassamento della mandibola che determina anche l'abbassamento degli angoli della bocca) e in /b/ (in cui sono connessi alla realizzazione dell'occlusione), ma non nella labiodentale /v/ e la asimmetria è comunque minima in tutti e tre i bersagli.

In Fig.6 a) e 6 b), sulla base dei valori di LO (=LH), LR (=LW) e LLP (cfr. §), vengono riportati i dati delle configurazioni labiali 3D per la vocale tonica e la consonante per ciascuna delle due sequenze /'aba/ e /'ava/ per tutte le produzioni emotive e per la produzione N.

Paura (F) e *collera* (A) determinano i valori più elevati di LO, mentre *disgusto* e *gioia* li riducono drasticamente. Anche i valori di LR vengono modificati: D e J comportano infatti l'appiattimento delle labbra, F e A l'arrotondamento. Le analisi statistiche permettono di affermare che la produzione delle emozioni determina modificazioni significative per tutti i parametri di rilevanza fonetica.

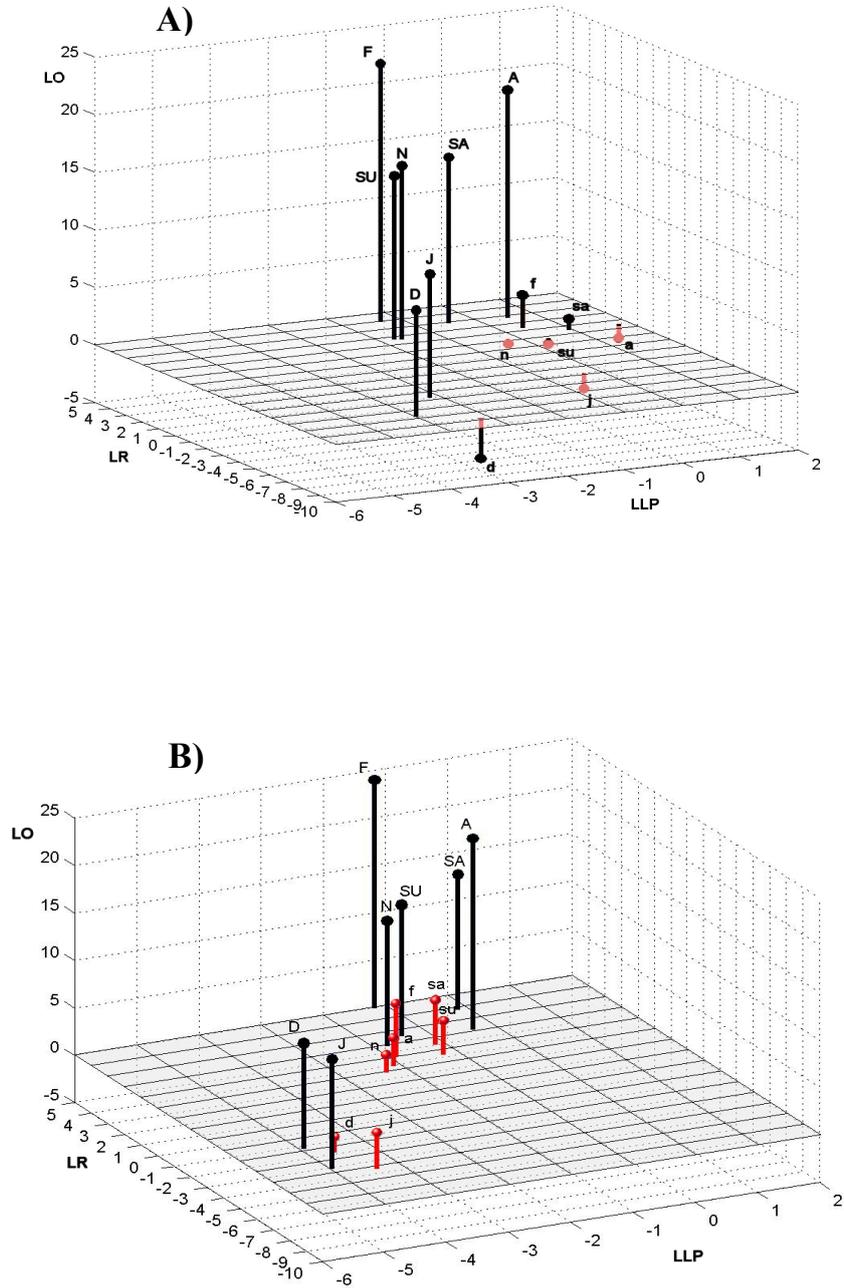


Figura 6 a) e 6 b). Configurazione labiale 3D dei bersagli della vocale tonica e della consonante rispettivamente nelle sequenze /'aba/ e /'ava/ relativa ai valori medi normalizzati di LO, LR e LLP (valori negativi di LR indicano l'appiattimento).

In Fig.7 a) e 7 b), per gli stessi fonemi e per tutte le emozioni, sono riportati i valori delle asimmetrie orizzontali e verticali. In ambedue le sequenze, mentre N non evidenzia nessuna asimmetria (vedasi Tabella 1), le produzioni emotive sembrano essere caratterizzate da questo parametro: in particolare il *disgusto* (D) presenta la più elevata asimmetria verticale e *gioia* (J) quella orizzontale.

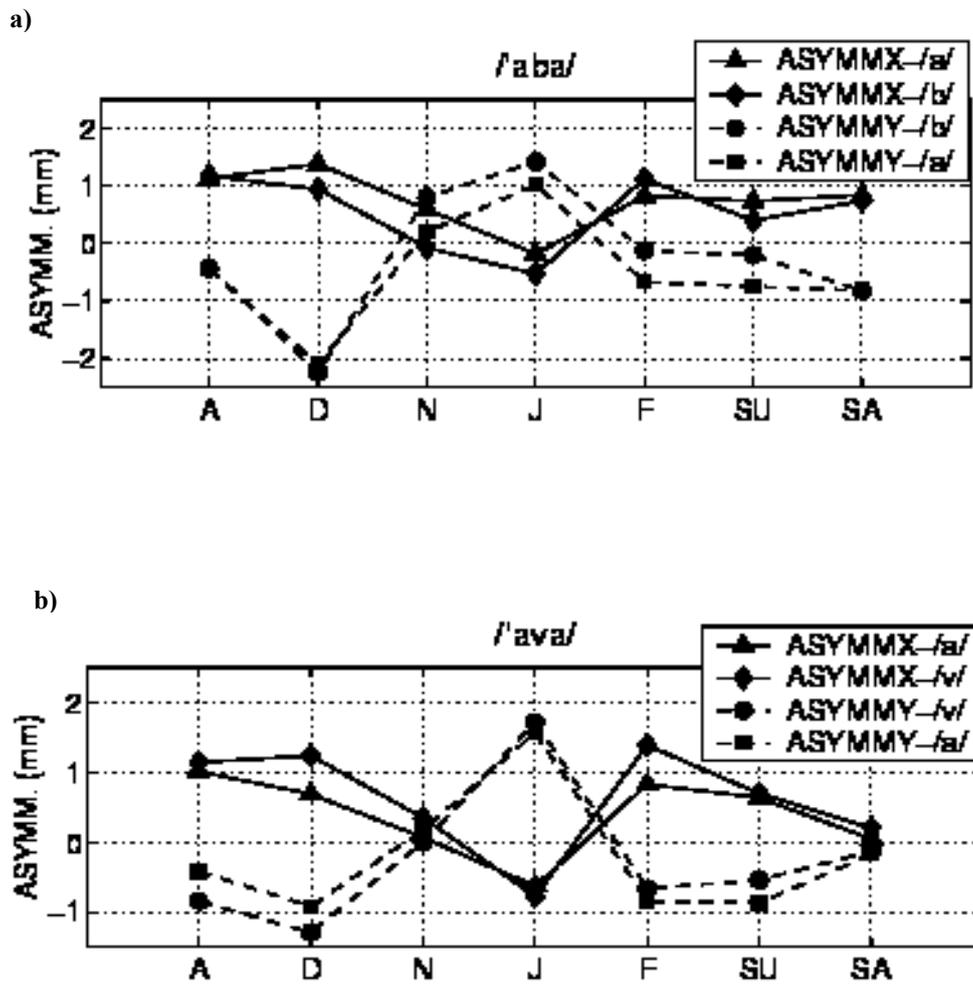


Figura 7 a) e b). Valori medi dei parametri asimmetrie orizzontali e verticali per i bersagli /a/, /b/ e /v/

Per ogni sessione di acquisizione sono stati registrati anche i dati articolatori nella posizione di riposo e i parametri selezionati sono stati normalizzati rispetto a questi valori.

Questa normalizzazione è necessaria in quanto individua l'estensione dei movimenti indipendentemente da forma e dalle dimensioni delle labbra.

Le differenze dei valori rilevati sono determinati dall'effetto delle emozioni sui valori dei parametri di rilevanza fonetica

8.2 Seconda normalizzazione

Dai dati normalizzati rispetto alla posizione di riposo sono stati sottratti successivamente, per ogni parametro delle diverse produzioni emotive, i valori dei parametri relativi alla produzione non emotiva (N), quantificando così la variazione spaziale determinata dalle emozioni.

I risultati di questa seconda normalizzazione sono stati illustrati e discussi in Magno Caldognetto 2003a, 2003b, 2004 a), 2004 b) e 2004 c) e a questi articoli rimandiamo per la loro analisi dettagliata.

Riportiamo nel seguito alcuni cluster di particolare interesse per i problemi che ci siamo posti, cioè per l'identificazione di parametri vincolati dalle caratteristiche articolatorie linguistiche e di parametri che dipendono dalle realizzazioni delle configurazioni emotive.

In Fig. 8 vengono presentati i cluster relativi ai valori del parametro LO per tutte le emozioni nei bersagli /'a, b, v/. Per la vocale centrale aperta, *disgust* e *joy*, con i loro valori fortemente negativi, sono le due emozioni che riducono l'apertura labiale di circa 8 mm, mentre *fear* l'aumenta moderatamente (2.7 mm).

Modificazioni simili si registrano per /b/⁶, mentre i dati relativi al bersaglio /v/ indicano che le produzioni emotive non modificano l'apertura labiale non emotiva, vincolata dalla produzione della costrizione labiodentale "a fessura".

⁶ Si deve ricordare che il rilevamento dell'Apertura Labiale viene eseguita calcolando la distanza tra marker posti sui punti centrali del bordo esterno del labbro superiore e inferiore. Non viene quindi misurato direttamente il grado di contatto tra le labbra e i dati inattesi di Apertura Labiale per /b/ possono essere determinati da particolari configurazioni, "corrugamenti" o "arricciamenti" delle labbra che pure non alterano l'occlusione bilabiale.

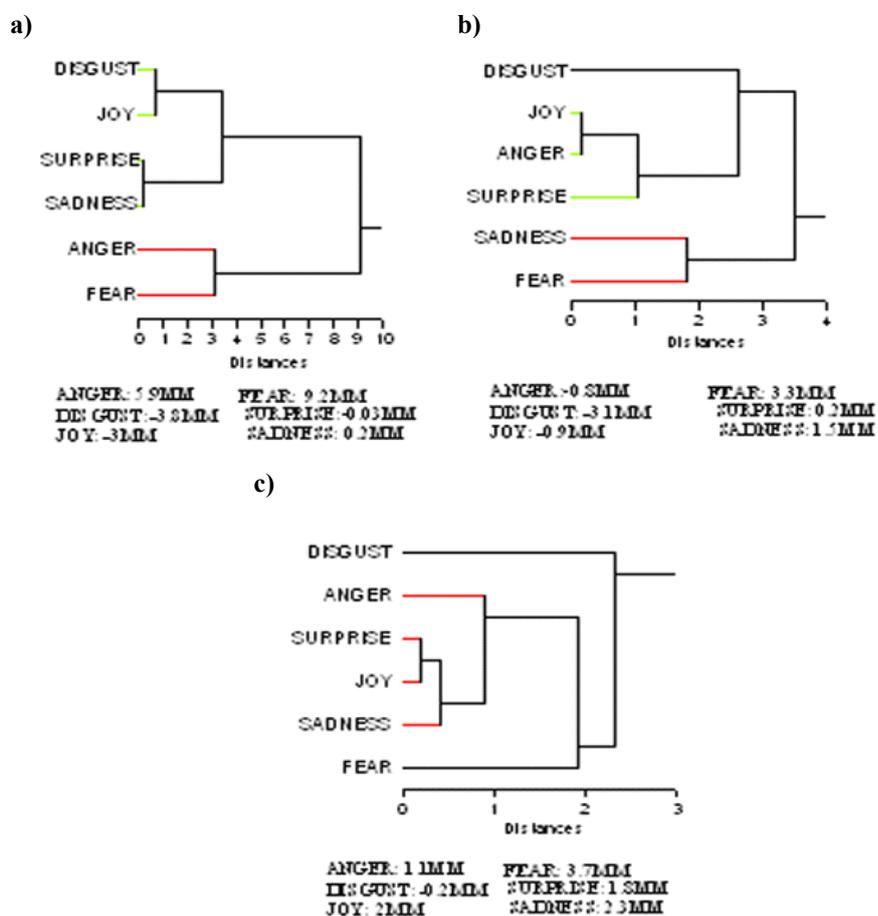


Figura 8. Cluster analysis dei parametri di apertura labiale LO per i bersagli /^ha/, /b/ e /v/.

In fig. 9 sono riportati i raggruppamenti dei valori degli spostamenti verticali dell'angolo sinistro e destro per /^ha/. Lo spostamento più elevato si verifica per *joy* nella vocale /a/⁷, in modo quasi simmetrico, mentre per *disgust* si evidenzia un diverso comportamento dell'angolo destro. In *fear* invece si verifica l'abbassamento di ambedue gli angoli, anche questo quasi simmetrico.

⁷ I valori negativi indicano l'innalzamento, per i criteri di estrazione di questo parametro vedasi Magno Caldognetto et al. 2004

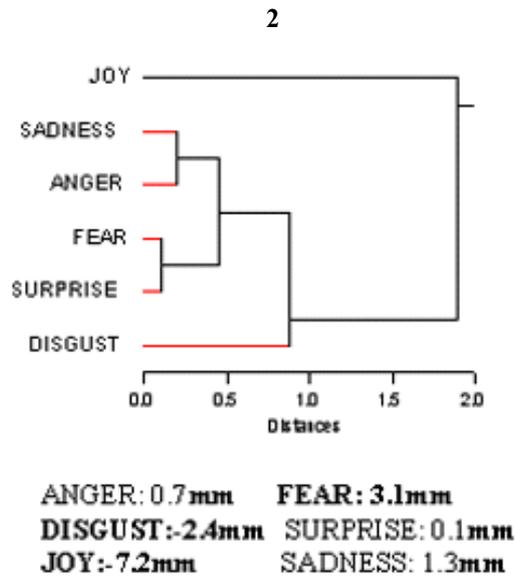
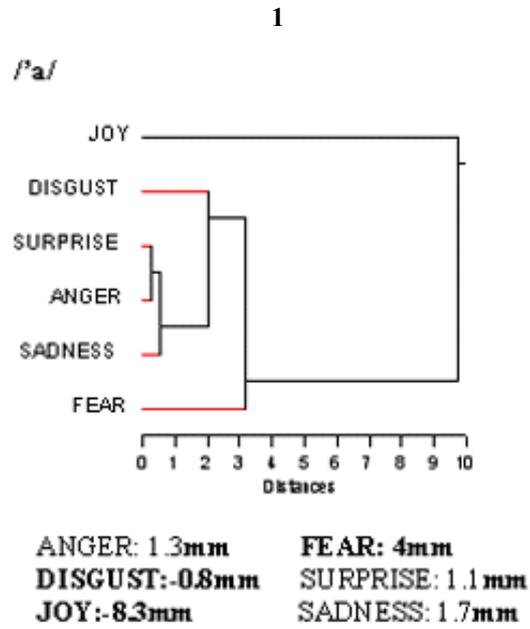
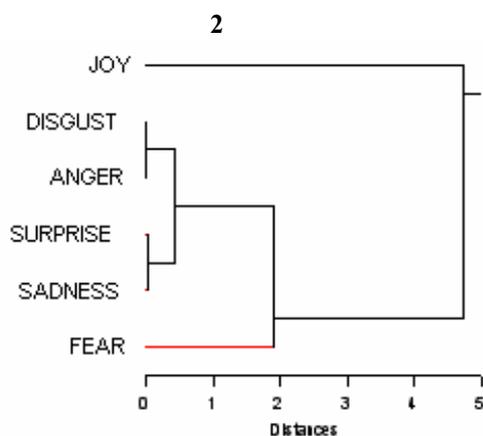


Figura 9.1 e 9.2 Cluster analysis degli spostamenti verticali degli angoli sinistro (1) e destro (2) per la vocale centrale /a/

La stessa tendenza si verifica anche negli spostamenti relativi al bersaglio /b/ (fig. 10.1 e 10.2) sebbene con gamme di valori ridotti; ciò accade anche per /v/ (fig. 11.1 e 11.2), cosicché i parametri che quantificano i movimenti degli angoli e relative simmetrie sono dipendenti soprattutto dalla produzione delle emozioni e non dipendenti dal bersaglio fonetico.



ANGER: 0.5mm	FEAR: 2.3mm
DISGUST: 2.3mm	SURPRISE: 0.6mm
JOY: -5.6mm	SADNESS: 1.3mm



ANGER: -0.8mm	FEAR: 1.4mm
DISGUST: -0.8mm	SURPRISE: -0.4mm
JOY: -5mm	SADNESS: -0.3mm

Figura 10.1 e 10.2. Cluster analysis degli spostamenti verticali degli angoli sinistro (1) e destro (2) per la consonante /b/



Figura 11.1 e 11.2. Cluster analysis degli spostamenti verticali dell'angolo sinistro (1) e destro (2) nella realizzazione della consonante /v/

In conclusione possiamo quindi rispondere positivamente alle domande proposte al paragrafo 2, poiché abbiamo verificato come effettivamente alcune modificazione sono da imputare prevalentemente ai parametri fonetico-fonologici e altri invece prevalentemente alle caratteristiche delle configurazioni emotive.

BIBLIOGRAFIA

- Abry C., Boe L.J., 1986, "Laws" for Lips, «Speech Communication» 5, pp. 97-104.
- Aubergé V., Cathiard M., 2003, *Can we Hear the Prosody of Smile?*, «Speech Communication» 40, pp. 87-98.
- Badin P., Bailly G., Revéret L., Baciú M. and Segebarth C., 2002, *Three-Dimensional Linear Articulatory Modeling of Tongue, Lips and Face, Based on MRI and Video Images*, «Journal of Phonetics» 30, 533-554.
- D'Urso V. & Trentin R., 1998, *Introduzione alla psicologia delle emozioni*, Bari, Editori Laterza,.
- Ekman P. & Friesen W., 1978, *Facial Action Coding System*, Palo Alto (USA), Consulting Psychologist Press Inc.
- Ekman P., Friesen W., Hager J.C., 2002, *New Version of the Facial Action Coding System*, Salt Lake City (Utah), A Human Face Publ., (eBook, Cd).
- Farnetani, E. 1999, *Labial coarticulation*, in *Coarticulation. Theory, Data and Techniques*, W. J. Hardcastle and N. Hewlett, (Editors), Cambridge: Cambridge University Press, pp. 144-163.
- Fletcher J., Harrington J., 1999, *Lip and Jaw Coarticulation*, in W.J. Hardcastle, N. Hewlett (Editors), *Coarticulation. Theory, Data and Techniques*, Cambridge, Cambridge University Press, pp. 164-175.
- Fromkin V., 1964, *Lip Positions in American English Vowels*, «Language and Speech», 7, pp. 217-225.
- Kroos Ch., Kuratate T., Vatikiotis-Bateson E., 2002, *Video-based Face Motion Measurement*, «Journal of Phonetics» 30, pp. 569-590.
- Ladefoged P., 2001, *Vowels and Consonants. An Introduction to the Sounds of Language*, Oxford (UK), Blackwell Publ.
- Laver J., 1994, *Principles of Phonetics*, Cambridge (UK), Cambridge University Press.

Linker W., 1982, *Articulatory and Acoustic Correlates of Labial Activity in Vowels: A Cross-linguistic Survey*, «UCLA Working Papers in Phonetics» 56, pp.1-134.

Magno Caldognetto E., Vagges K., Zmarich C., 1995, *Visible Articulatory Characteristics of the Italian Stressed and Unstressed Vowels*, in *Proceedings of 13th International Congress of Phonetic Sciences, Stockholm*, vol. 1, pp. 366-369.

Magno-Caldognetto, E., Zmarich, C., Cosi, P., Ferrero, F., 1997, *Italian Consonantal visemes: relationships between spatial/temporal Articulatory Characteristics and Coproduced Acoustic Signal*, in C. Benoit and R. Campbell (Editors), *Proceedings of the Workshop on Audio-Visual Speech Processing*, Rhodes, 5-8 September, pp. 117-123.

Magno Caldognetto E., Zmarich C., Cosi P., 1998, *Statistical Definition of Visual Information for Italian Vowels and Consonants*, in *Proceedings of AVSP '98*, D. Burnham, J. Robert-Ribes and E. Vatikiotis-Bateson (Editors), Terrigal (Aus), pp. 135-140.

Magno Caldognetto E., Cosi P., Drioli C., Tisato G., Cavicchio F., 2003a, *Coproduction of Speech and Emotions: Visual and Acoustic Modifications of Some Phonetic Labial Targets in Proceedings of AVSP 03*, September 4-7, St. Jorioz (France), pp. 209-214.

Magno Caldognetto E., Cosi P., Cavicchio F., Drioli C., Tisato G., 2003b, *Modificazioni Articolatorie ed Acustiche Dovute alla Coproduzione di Emozioni e Parlato*, in *Atti delle XIV giornate del GFS Costituzione, Gestione e Restauro di Corpora Vocali*, Viterbo, 4-6 Dicembre, pp. 234-341.

Magno Caldognetto E., Cosi P., Cavicchio F., 2004a, *Modifications of Speech Articulatory Characteristics in the Emotive Speech*, in *Proceedings of Tutorial and Research Workshop Affective Dialogue Systems 2004*, June 14 – 16, Kloster Irsee (Germany), pp. 106-112.

Magno Caldognetto, E.; Cosi, P.; Drioli, C.; Tisato, G.; Cavicchio, F., 2004b, *Modifications of phonetic labial targets in emotive speech: Effects of the coproduction of speech and emotions*, «Speech Communication» 44, pp. 173-185.

Nordstrand, M., Svanfeldt, G., Granstrom, B., House D., 2003, *Measurements of Articulatory Variations and Communicative Signals in Expressive Speech*, in J.L.Schwartz, F.Berthommier, M.A. Cathiard and D. Sodayer (Editors), *Proceedings of AVSP 03*, September 4-7, St. Jorioz (France), pp. 233-237.

Scherer K.R., 1986, *Vocal Affect Expression: A Review and a Model for Future Research*, «Psychological Bulletin» 99, pp. 143-165.

Scherer K.R., 2003, *Vocal Communication of Emotion: A Review of Research Paradigm*, «Speech Communication» 40, pp.227-256.

Schroeder M., Aubergé V., Cathiard M., 1998, *Can we Hear Smile?*, in *Proceedings of ICSLP '98*, Sydney , pp. 559-662.

Yehia H. C., Kuratate T., Vatikiotis-Bateson E., 2002, *Linking Facial Animation, Head Motion and Speech Acoustics*, «Journal of Phonetics», 30, pp.555-568.

Zmarich C., Magno Caldognetto E., 2002, *Effetti coarticulatori 3D sulle caratteristiche spaziotemporali dei movimenti delle labbra nelle consonanti bilabiali e labiodentali*, in *Atti delle XII Giornate del GFS*, Macerata 13-15 dicembre, Roma, Il Calamo, pp. 193-200.

Zmarich C., Magno Caldognetto E., 2003, *Three-dimensional Coarticulatory Effects on Spatial Characteristics of Labial Targets for 'VpV and 'VfV Italian Consonants in Asymmetric Bisyllables*, in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcellona, 3-9 August, pp. 3121-3124.