

MODIFICAZIONI ARTICOLATORIE DOVUTE ALLA COPRODUZIONE DI EMOZIONI E PARLATO

E. Magno Caldognetto, P. Cosi, F. Cavicchio, C. Drioli, G. Tisato
ISTC - CNR "Fonetica e Dialettologia" Sez. di Padova
e-mail: magno @pd.istc.cnr.it

INTRODUZIONE

I dati attualmente disponibili sulle modificazioni delle caratteristiche articolatorie nel parlato emotivo provengono principalmente dall'analisi di caratteristiche acustiche (Banse e Scherer 1996, Scherer 2003, Scherer et al. 2003). In questo articolo viene invece affrontato il problema della loro descrizione in termini di modificazioni dei parametri cinematici articolatori, specificamente labiali, registrati e analizzati tramite il sistema optoelettronico ELITE. Data la tecnica sperimentale adottata, è stato analizzato il parlato emotivo prodotto in laboratorio e non quello spontaneo. Come è stato da più parti sottolineato, la scelta del parlato emotivo prodotto in laboratorio viene di solito preferito per motivi metodologici, poiché in questa condizione vengono controllati il materiale linguistico (vocali, sillabe, non parole, parole, frasi semanticamente non emotive, brani), i parlanti (attori vs soggetti naive), gli scenari (per determinare la produzione dell'emozione desiderata, di solito le emozioni primarie sopra citate, le possibili varianti di una stessa emozione, p. es. collera "fredda" vs collera "calda", sorpresa piacevole vs sorpresa spiacevole, e la loro gradualità) e possono essere predisposte condizioni ottimali di registrazione. Il parlato emotivo così prodotto risulterà ovviamente stereotipato e prototipico (Scherer 1986, 2003), ma attraverso test di percezione è comunque possibile verificare, oltre alla correttezza della produzione, anche la sua adeguatezza. Scopo della ricerca qui presentata è la quantificazione dei movimenti labiali nel parlato emotivo che risultano dall'interazione tra le restrizioni fonetico-fonologiche che regolano le produzioni delle caratteristiche di modo e luogo di articolazione di fonemi vocalici e consonantici e le configurazioni labiali tipiche delle emozioni primarie: collera, gioia, disgusto, paura, sorpresa e tristezza (Ekman e Friesen 1978, Ekman et al. 2002). Specificamente in questo studio sono state analizzate le modificazioni della vocale centrale aperta /a/, dell'occlusiva bilabiale /b/ e della costrittiva labiodentale /v/ (consonanti per le quali le labbra costituiscono l'articolatore primario).

1. METODO

Per raccogliere i dati articolatori è stato utilizzato un sistema optoelettronico di analisi 3D del movimento che permette anche la registrazione sincrona del segnale acustico¹.

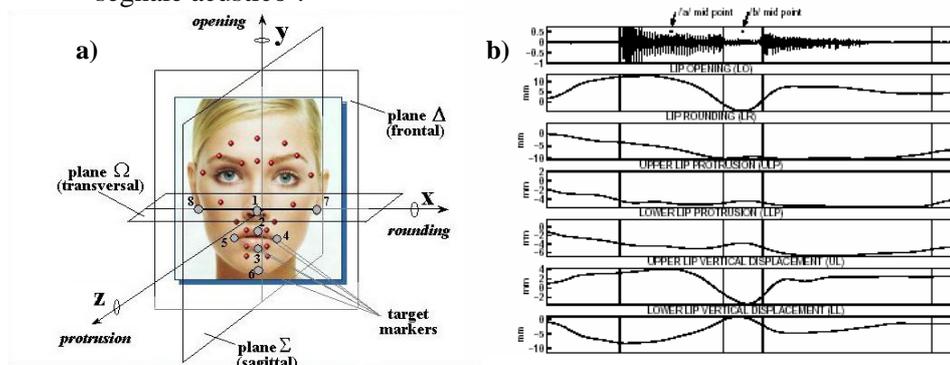


Fig. 1 a) e b) Posizione dei 28 marker e dei piani per la raccolta dei dati articolatori ed evoluzione temporale di alcuni parametri cinematici corrispondenti alla sequenza /aba/ prodotta con emozione di disgusto.

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 e alla posizione di riposo. Sono state così ottenute 107 registrazioni, per ciascuna delle quali sono stati calcolati gli andamenti dei seguenti 12 parametri (Magno Caldognetto et al. 2003): Apertura Labiale (LO)²; Spostamento Verticale del Labbro Superiore (UL); Spostamento Verticale del Labbro Inferiore (LL); Arrotondamento Labiale (LR); Protrusione del Labbro Superiore (ULP); Protrusione del Labbro Inferiore (LLP); Spostamento orizzontale dell’Angolo Sinistro (LCX) e dell’Angolo Destro (RCX); Spostamento Verticale dell’Angolo Sinistro (LCY) e dell’Angolo Destro (RCY); Asimmetria orizzontale (AsymX) e verticale (AsymY). In questo articolo vengono presentati solo i dati relativi ai primi 6 parametri che quantificano le caratteristiche fonetico-fonologiche dell’orifizio labiale. L’analisi è stata condotta sui valori medi che caratterizzano i bersagli di /a/, /b/ e /v/, estratti in corrispondenza dei punti

¹ Per applicazioni precedenti di questo sistema alla definizione dei visemi dell’italiano su base articolatoria cfr. Magno Caldognetto et al. 1998.

² Le sigle utilizzate corrispondono alla terminologia inglese dei parametri: LO= Lip Opening, UL= Upper Lip vertical displacement, LL= Lower Lip vertical displacement, LR= Lip Rounding, ULP= Upper Lip Protrusion, LLP= Lower Lip Protrusion, LCX= Left Corner horizontal displacement, RCX= Right Corner horizontal displacement, LCY= Left Corner vertical displacement, RCY= Right Corner vertical displacement, AsymX= Horizontal Asymmetry, AsymY= Vertical Asymmetry.

mediani dei corrispondenti segnali acustici (Fig.1 b)³. I dati registrati da ELITE, elaborati rispetto ai piani riportati in Fig.1 a)(Magno Caldognetto et al. 1998, 2003, 2004), sono stati sottoposti a due normalizzazioni. La prima ha implicato, per ciascuna registrazione, ciascun parametro e ciascuna emozione, la sottrazione dei valori relativi alla posizione di riposo dai dati originali di ELITE, in modo da individuare l'estensione dei movimenti indipendentemente dalla forma e dimensione delle labbra⁴. Successivamente, tutti i valori dei parametri articolatori, già normalizzati rispetto alla posizione di riposo, sono stati ulteriormente normalizzati rispetto ai dati corrispondenti della produzione *non emotiva* (N), al fine di evidenziare le modificazioni dei parametri dovute esclusivamente alla espressione visiva delle emozioni. Poiché i valori finali dei diversi parametri sono determinati, rispetto all'originario sistema di estrazione dei dati di ELITE, dalle successive sottrazioni dei valori relativi alla posizione di riposo e a N, ciascuna caratterizzata da propri valori positivi o negativi, è necessario elencare il significato fonetico dei valori finali risultanti per i diversi parametri: per LO i valori positivi segnalano un aumento dell'apertura labiale, i valori negativi una sua riduzione rispetto a N; per UL e LL i valori positivi indicano un aumento dello spostamento di UL e/o LL mentre i valori negativi quantificano la riduzione dello spostamento rispetto alla posizione di N; per LR, i valori negativi quantificano lo stiramento delle labbra, i valori positivi l'arrotondamento; per ULP e LLP i valori negativi segnalano l'appiattimento delle labbra, i valori positivi l'aumento della protrusione.

2. ANALISI DEI RISULTATI

Per ciascun parametro, ai dati risultanti per le diverse emozioni è stato applicato l'*Average Linkage Method between Groups*, un metodo statistico che ha permesso di raggruppare insieme, sulla base dei valori medi reali, le emozioni che hanno determinato le modificazioni più simili tra loro. In Figg. 3, 4 e 5 vengono riportati, per ogni parametro, i cluster con i valori medi relativi ad ogni emozione. Questi dati permettono di verificare, per i bersagli /'a/, /b/ e /v/ e per i diversi parametri, effetti simili per emozioni diverse e, all'opposto, effetti diversi di una stessa emozione a seconda della qualità dei bersagli.

2.1 Modificazioni del bersaglio di /'a/

³ Per l'analisi di tutti e 12 i parametri e le loro correlazioni con il segnale acustico, vedasi Magno Caldognetto et al., 2004 e 2004, submitted.

⁴ Per le analisi di significatività dei dati ricavati dalla prima normalizzazione cfr. Magno Caldognetto et al. 2003.

Per quanto riguarda i dati relativi alla vocale /'a/ (mediati tra le vocali toniche di /'aba/ e /'ava/), *Disgusto* e *Gioia* si differenziano dalle altre emozioni per valori negativi di LO (corrispondenti a una riduzione dell'Apertura Labiale rispetto a N), per i valori positivi di UL e di LL (corrispondenti a un aumento dello spostamento rispetto a N), per i valori negativi di LR (corrispondenti allo stiramento delle labbra) e di ULP e di LLP (corrispondenti alla ritrazione della labbra). *Collera* e *Paura* determinano invece una Apertura Labiale più ampia rispetto alle altre emozioni (valori positivi di LO), mentre la *Paura* è caratterizzata dall'arrotondamento massimo (valori positivi di LR).

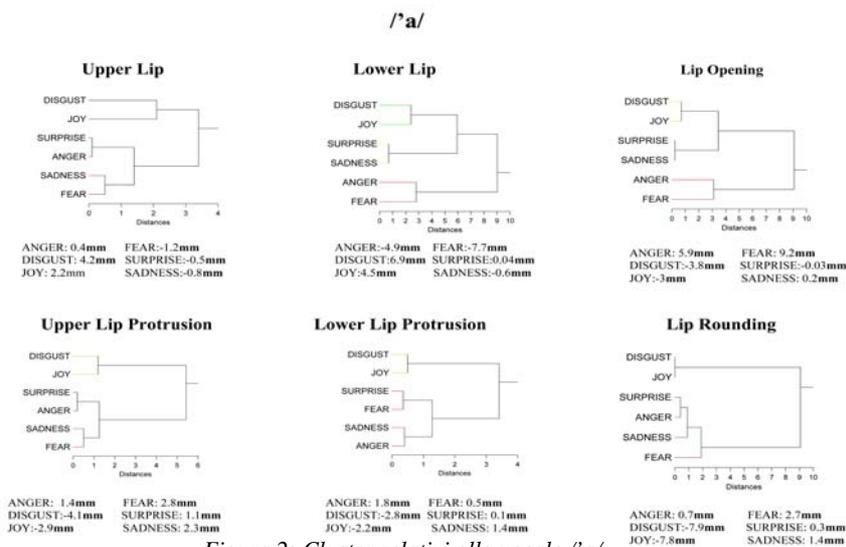


Figura 2: Cluster relativi alla vocale /'a/

2.2 Modificazioni del bersaglio /b/

Per la consonante /b/, il *Disgusto* presenta la più alta riduzione dell'Apertura Labiale (valori negativi di LO), il massimo innalzamento del Labbro Inferiore e del Labbro Superiore rispetto a N (segnalato dai valori positivi di LL e UL), il massimo stiramento labiale (valori negativi di LR), la massima riduzione di ULP e LLP (corrispondente alla ritrazione delle labbra). Anche la *Gioia* è caratterizzata da un innalzamento di UL e LL rispetto al neutro (valori positivi) e dei valori di LR, corrispondenti all'appiattimento delle labbra. La *Paura* determina l'aumento massimo dell'Apertura Labiale (valori positivi di LO), la diminuzione dello spostamento di UL e LL rispetto a N, l'aumento dell'arrotondamento (valori positivi di LR) e della Protrusione Superiore (valori positivi di ULP, valore condiviso con *Tristezza*) e Inferiore (valori positivi di LLP). La *Collera* è caratterizzata da valori massimi positivi di LLP.

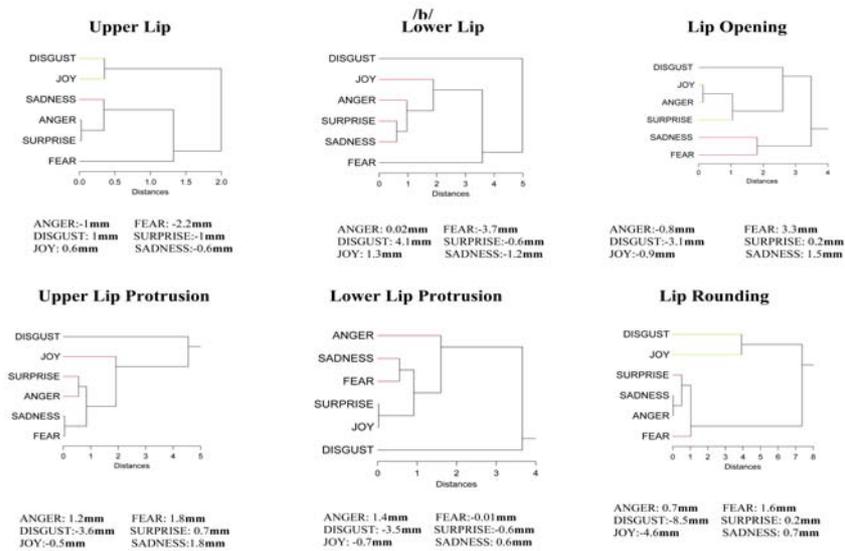


Figura 3: Cluster relativi alla consonante /b/

2.3 Modificazioni del bersaglio /v/

Nella costrittiva labiodentale /v/, il Disgusto presenta la massima riduzione dell'Apertura Labiale (valori negativi di LO), l'innalzamento di UL e LL (valori positivi) rispetto a N, la riduzione della protrusione del labbro Superiore e Inferiore (valori negativi di ULP e LLP). Anche la Gioia è caratterizzata dall'innalzamento di UL e dalla riduzione massima di ULP e LLP (corrispondente alla ritrazione delle labbra). La Paura presenta invece i valori massimi di Apertura Labiale (valori positivi di LO), la riduzione dello spostamento del Labbro Superiore e Inferiore (valori negativi di UL e LL), la riduzione dell'arrotondamento (valori negativi di LR) e l'aumento della protrusione del Labbro Superiore (valori positivi di ULP).

3. CONCLUSIONI

L'analisi in cluster ha evidenziato che le emozioni modificano sempre i dati articolatori della produzione N dei 3 bersagli considerati. In particolare le configurazioni labiali tipiche di *Disgusto* e *Gioia* riducono i valori di LO e di LR (e conseguentemente modificano i dati di UL, LL, ULP, LLP), mentre *Paura*, al contrario, determina l'aumento massimo dei valori di LO e LR. E' interessante notare che queste sistematiche modificazioni presentano però delle diversità nei valori specifici a seconda del bersaglio vocale o consonantico e del parametro considerato: la modificazione, nella sua sovrapposizione ai valori di N, non assume un valore sempre eguale. Vedasi

il caso di LO i cui valori variano a seconda del grado di apertura/chiusura di /a, b, v/. Quindi nel parlato emotivo le configurazioni labiali delle emozioni non possono violare le caratteristiche fonetiche linguistiche e queste costituiscono, a loro volta, un vincolo alla produzione di tali configurazioni.

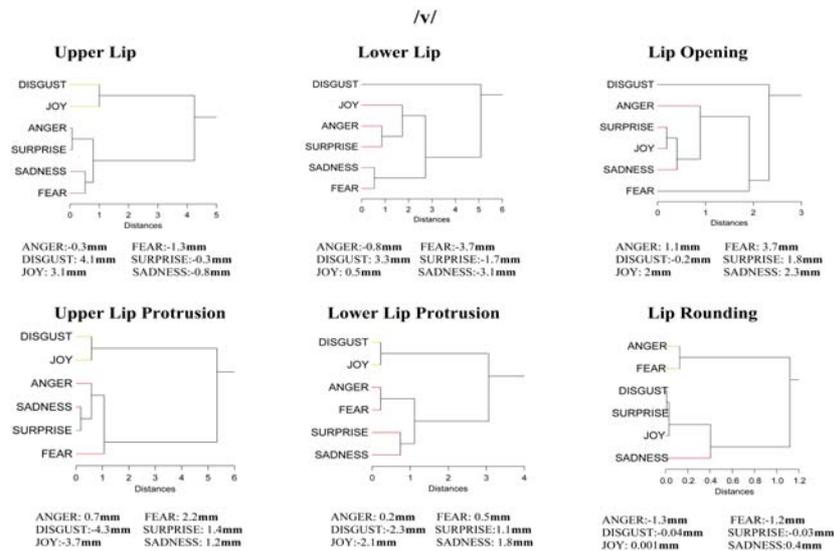


Figura 4: Cluster relativi alla consonante /v/

Bibliografia

- Banse, R. & Scherer, K. R. (1996) Acoustic profiles in vocal emotion expression, *Journal of Personality and Social Psychology*, 70, 614-636.
- Ekman, P. & Friesen, W. V. (1978) *Facial Action Coding System* (Consulting Psychologists Press Inc.), USA: Paolo Alto.
- Ekman, P., Friesen, W. V. & Hager, J. C. (2002) *New version of the Facial Action Coding System (eBook)*, Salt Lake City, USA: A Human Face Publ.
- Magno Caldognetto, E., Zmarich, C. & Cosi, P. (1998) Statistical definition of visual information for Italian vowels and consonants, in *Proceedings of AVSP'98* (Terrigal, Australia), pp. 135-140.
- Magno Caldognetto, E., Cosi, P., Drioli, C., Tisato, G. & Cavicchio, F. (2003) Coproduction of speech and emotion: bimodal audio-visual changes of consonant and vowel labial targets, in *Proceedings of AVSP03*, pp. 209-214, S. Joriz, France.
- Magno Caldognetto E., Cosi P., Drioli C., Tisato G. & Cavicchio F., *Visual and Acoustic Modifications of Phonetic Labial Targets in the Emotive Speech: Effects of the Coproduction of Speech and Emotions*, submitted to *Speech Communication Special Issue AVSP03*.
- Scherer, K. R., Johnstone, T. & Klasmeyer, G. (2003) Vocal expression of emotion, *Handbook of the Affective Sciences* (R. J. Davidson, H. Goldsmith, K. R. Scherer, editors) pp. 433-456. New York: Oxford University Press.
- Scherer, K. R. (1986) Vocal affect expression: a review and a model for future research, *Psychological Bulletin*, 99, 143-165.
- Scherer, K. R. (2003) Vocal communication of emotions: a review of research paradigms, *Speech Communication*, 40, 227-256.