

# EMOVO, DATABASE DI PARLATO EMOTIVO PER L'ITALIANO

Iacopo Iadarola  
Dipartimento di Studi Glottoantropologici e Discipline musicali  
dell'Università di Roma "La Sapienza"  
Fondazione Ugo Bordini  
[iacopo.iadarola@gmail.com](mailto:iacopo.iadarola@gmail.com)

## 1. SOMMARIO

EMOVO è il primo database di parlato emotivo per la lingua italiana, ed è stato realizzato per conto del Dipartimento di Studi Glottoantropologici e Discipline musicali dell'Università di Roma "La Sapienza" e della Fondazione Ugo Bordini, sotto la supervisione del Prof. Paolo Di Giovine e dell'Ing. Andrea Paoloni.

Sono stati convocati 6 attori, 3 maschi e 3 femmine, di provata professionalità, e si sono fatte loro recitare 14 frasi (assertive, interrogative, elenchi) in base ai 6 stati emotivi di base (disgusto, paura, rabbia, gioia, sorpresa, tristezza) più lo stato neutro. Tali emozioni sono le ben note "big six" riscontrabili in molta della letteratura inerente al parlato emotivo, nonché quelle prese in considerazione dal CNR di Padova, che ha condotto i principali studi sul parlato emotivo italiano. Le registrazioni sono state effettuate, con adeguate strumentazioni professionali, presso i laboratori della Fondazione Ugo Bordini, dove è stata garantita agli attori la possibilità di autoindursi lo stato emotivo desiderato in base alle proprie tecniche professionali. Si è optato per una frequenza di campionamento di 48 kHz, 16 bit stereo, formato wav. Ogni singola elicitazione delle frasi è stata etichettata ed archiviata in un database di tipo Access, in modo da consentirne una rapida individuazione. In totale sono stati archiviati 588 record: 6 attori X 14 frasi X 7 stati. Il totale del materiale registrato, escluse le pause, è all'incirca di 10 minuti per attore, quindi un'ora per l'intero database, che verrà distribuito su cd-rom a chi ne farà [richiesta](#). Nel presente articolo è possibile averne qualche saggio (cf. § 2.2.1, § 2.2.3)

Inoltre, EMOVO è stato validato con un test di discriminazione delle emozioni su due frasi ('la casa forte vuole col pane', 'il gatto sta scorrendo nella pera'), condotto, parallelamente e separatamente, dall'autore e dall'Università della Calabria. Entrambi i test hanno avuto un campione di 12 soggetti riconoscenti, i quali hanno dovuto di volta in volta indicare, scegliendo fra due possibili risposte, lo stato emotivo delle frasi ascoltate. Il test ha avuto esito positivo in quanto è risultata un'accuratezza complessiva dei riconoscimenti pari all'80%, senza significative discrepanze fra i risultati dei due test. E' stata valutata inoltre la scala di riconoscibilità delle emozioni, e si è riscontrata una sostanziale concordanza con la scala di riconoscibilità desunta dalla letteratura al riguardo.

## 2. I DATABASE DI PARLATO EMOTIVO E EMOVO

### *2.1 I corpora del parlato emotivo: quadro d'insieme, problematiche e prospettive*

Nel campo degli studi sul parlato emotivo c'è ad oggi urgente bisogno non solo di ulteriori e più approfondite analisi, ma soprattutto di standardizzazioni e convenzioni condivise dalla comunità scientifica al fine di rendere tali analisi comparabili e convergenti.

Per troppo tempo la nostra conoscenza del parlato emotivo è progredita a rilento a causa della mancanza di sforzi congiunti da parte degli studiosi: e questa è stata una conseguenza, più che di problemi logistici e organizzativi, della mancanza di un modello teorico (nella concettualizzazione delle emozioni) e di prassi comuni (soprattutto nella rilevazione dei

parametri acustici). Ora, sia dal punto di vista teorico sia dal punto di vista pratico, siamo lontani dall'auspicabile risultato di potersi riferire a modelli unanimemente accettati. A cavaliere fra i due versanti si trova una delle questioni fondamentali da affrontare per il progresso degli studi del parlato emotivo: la creazione di database linguistici emotivamente connotati. Dopo che nella comunità scientifica si sono effettuati centinaia di esperimenti isolati su materiale linguistico estremamente limitato, e dopo che si è assodato che l'enorme mole dei risultati così ottenuta è lungi dall'essere soddisfacente, si è cominciato a sentire l'urgenza di partire da altre basi: di passare dai *dataset* ai database. Infatti, se uno degli obiettivi della comunità scientifica è la comparazione e la convergenza dei risultati, la prima ovvia cosa da fare è cercare di partire da basi comuni di dati, oltre che da premesse teoriche comuni.

Un caso illuminante in questo senso è la creazione del network europeo HUMAINE, lanciato nel 2004 con l'adesione di 33 partner da 14 paesi europei, con l'obiettivo di sviluppare sistemi capaci di registrare, modellare e influenzare stati emotivi umani:

Tali sistemi potrebbero essere di importanza centrale nelle interfacce future, ma i loro presupposti teorici non sono abbastanza avanzati da renderci sicuri sul loro effettivo potenziale e sulle modalità per il loro sviluppo. Uno dei motivi è che le conoscenze al riguardo sono disperse fra varie discipline. HUMAINE intende raggruppare i leader del settore delle discipline-chiave in un programma finalizzato ad ottenere maggiore integrazione intellettuale.<sup>1</sup>

Il programma prevede l'organizzazione di aree di ricerca con relativi laboratori, nonché la creazione delle infrastrutture necessarie, quali appunto "databases recognising cultural and gender diversity." Sul sito del network si può trovare molto materiale interessante da un punto di vista teorico e metodologico, nonché utili software di analisi, ma un database HUMAINE ancora non è disponibile.

E' ora utile sottoporre a disamina le caratteristiche e le problematiche di cui lo sviluppo di un database deve tener conto. Innanzitutto va definito il numero, il sesso dei parlanti, la lingua, le emozioni, le occorrenze di una data emozione, eventuali configurazioni funzionali/sociali: la diversità dei risultati ottenuti nei vari studi può avere valore euristico soltanto se è ben chiaro l'assetto di tali parametri, dal momento che l'espressione vocale di un'emozione può cambiare notevolmente di parlante in parlante, di cultura in cultura, o ancora in base al sesso o alla situazione sociologica. Per quanto riguarda le emozioni, la loro selezione è fondamentale: se di solito la scelta "di default" nei database è quella di utilizzare le emozioni di base (fra cui le più consuete: paura, rabbia, tristezza, gioia), bisogna anche tener conto del fatto che nella vita di tutti i giorni le emozioni sono fortemente condizionate dalle *display rules* (i condizionamenti socioculturali all'atto espressivo, studiati in particolar modo da Paul Ekman), e da altre profonde influenze culturali (cf. Harré 1986): il risultato è che nella realtà difficilmente si hanno emozioni pure, ben più spesso svariati stati emotivi sfumati, la cui codifica presenta difficoltà ancora maggiori di quante già ne comporti la codifica delle emozioni di base. Un database ideale dovrebbe contemplare entrambe le tipologie di stati emotivi, il che renderebbe la sua portata di gran lunga più estesa dei database e dei *dataset* attualmente esistenti.

Altro elemento qualificante di un database è costituito dal fatto di essere *naturale* o meno. Si apre qui una questione dibattutissima sui *pro* e sui *contro* riguardo al materiale

---

<sup>1</sup> Cf. il sito: [www.emotion-research.net](http://www.emotion-research.net)

raccolto naturalmente rispetto al materiale ottenuto artificialmente. Nel primo caso, ovviamente, i vantaggi sono dati da una maggiore rappresentatività, da parte del materiale analizzato, della reale espressione delle emozioni (ma si tenga presente quanto detto poc'anzi sulla scarsa frequenza di emozioni pure nella vita di tutti i giorni). Gli svantaggi sono di ordine pratico, deontologico e teorico: registrare del parlato emotivo in situazioni reali comporta inevitabilmente che il suono non sia pulito e che possa difficilmente essere confrontato con altro materiale raccolto in modo analogo. Inoltre, la registrazione di parlato emotivo dal vero comporta spinose questioni di *copyright*, e conseguentemente di accessibilità ai dati da parte di altri studiosi. Infine, come etichettare il materiale così ottenuto? Non c'è nessuna garanzia sull'attendibilità della classificazione fornita per lo stato emotivo "catturato", e ciò ovviamente comporta seri problemi di sistemazione e comparabilità dei dati. La comparabilità dei dati così ottenuti è ovviamente inficiata anche dalla estrema diversità delle situazioni da cui sono sorte le emozioni analizzate. Il materiale raccolto finora dal vivo, infatti, proviene dalle fonti più disparate: registrazioni di piloti di aerei prima che questi si schiantassero al suolo, giornalisti che commentano eventi disastrosi o che intervistano persone appena uscite da traumi fisici o psichici, *talkshow* (Johannes et al. 2000, Cowie – Douglas-Cowie 1996, Eldred – Price 1958, Hargreaves et al. 1965, Frolov et al. 1999, Huttar 1968, Kuroda et al. 1979, Niwa 1971, Roessler – Lester 1976, 1979, Simonov – Frolov 1973, Sulc 1977, Utsuki – Okamura 1976, Williams – Stevens 1969, 1972, Zuberbier 1957, Zwirner 1930). Se sulla variabilità di questi dati si può soprassedere all'interno di un singolo *dataset* per un esperimento, il problema certo non è di facile risoluzione nel caso di un intero database. La soluzione di compromesso è spesso quella di limitare la gamma di emozioni studiate, come nel caso del database *Vanderbilt II* (usato in France et al. 2000), che contempla solo la depressione e gli stati emotivi tendenti al suicidio, o del *SUSAS* (Hansen – Bou-Ghazale 1997), incentrato sullo stress. Ugualmente limitato a pochi stati emotivi è ovviamente il *Geneva airport lost luggage study* (Scherer – Ceschi 1997, 2000), in cui sono archiviate le reazioni emotive dei passeggeri nell'aeroporto di Ginevra, ai quali si era fatto credere che i loro bagagli fossero andati perduti, come in una sorta di "*candid camera*", sia pure con fini scientifici.

Le difficoltà su elencate non si pongono, invece, nel caso di materiale elicitato in laboratorio, sebbene anche qui si presentino varie questioni metodologiche di non facile soluzione, *in primis* sui soggetti elicitanti: attori professionisti o non-professionisti? Nel *Berlin database* (Kienast – Sandlmeier 2000, Paeschke – Sandlmeier 2000) le emozioni sono registrate in laboratorio dalla bocca di parlanti non attori, in quanto i professionisti produrrebbero, secondo gli autori, emozioni stereotipate. A ciò si potrebbe obiettare, pirandellianamente, che non solo gli attori, ma qualsiasi persona in pubblico recita, in misura minore o maggiore. Inoltre, a giudicare dai buoni risultati di riconoscimento delle emozioni elicitate da attori professionisti,<sup>2</sup> si può tenere per acquisita, in linea generale, la capacità di un attore di rispecchiare la realtà dell'espressione emotiva, per quanto delle

---

<sup>2</sup> Banse – Scherer 1996, Bortz 1966, Coleman – Williams 1979, Costanzo et al. 1969, Cosmides 1983, Davitz 1964, Fairbanks – Pronovost 1939, Fairbanks – Hoaglin 1941, Fonagy 1978, 1983, Fonagy – Magdics 1963, Green – Cliff 1975, Höffe 1969, Kaiser 1962, Kienast et al. 1999, Klasmeyer – Meier 1999, Klasmeyer – Sandlmeier 1997, Klasmeyer – Sandlmeier 1999, Klasmeyer 1999, Kotlyar – Morozov 1976, Levin – Lord 1975, Paeschke et al. 1999, Scherer et al. 1972a,b, 1973, Sedlacek – Sychra 1963, Sobin – Alpert 1999, Tischer 1993, Van Bezooijen 1984, Wallbott – Scherer 1986, Whiteside 1999, Williams – Stevens 1972.

differenze sicuramente possano permanere. In particolar modo si è convenuto che le emozioni simulate dagli attori ci diano informazioni sulle componenti socioculturali delle espressioni emotive, mentre le ricorrenze naturali di un'emozione ci permettono di indagare con più facilità le sue caratterizzazioni psico-fisiologiche: e va puntualizzato che l'aspetto socioculturale non è certo meno importante di quello psico-fisiologico, dal momento che il le emozioni e il parlato emotivo trovano gran parte della propria ragione d'essere per l'appunto nella funzione comunicativa e di interazione sociale. Altri svantaggi dell'uso di attori in laboratorio sono legati al tipo di materiale elicitato: per lo più testi scritti e monologhi. Ciò ovviamente preclude l'analisi di importanti aspetti dell'espressione emotiva legati all'interrelazione e all'improvvisazione.

Come si può vedere gli argomenti a favore e a sfavore sono numerosi gli uni e gli altri, e la scelta di una soluzione o di un'altra in ultima analisi dipende ancora dalle idiosincrasie dei singoli studiosi. Sembrerebbe auspicabile utilizzare database naturali, ma resta ancora da vedere se le attuali tecnologie ci possono permettere di far fronte al problema in modo efficace. Per compendiare il doppio corno del problema, riporto infine un'affermazione di Scherer 1981: "Se i campioni di parlato usati nelle ricerche di simulazione non sono abbastanza naturali, i campioni costituiti da parlato spontaneo non sono abbastanza emotivi."

Vanno menzionati, ancora, tentativi di mediare fra le due posizioni: sono i casi di materiale "seminaturale", che prevede l'induzione di emozioni in laboratorio. I principali metodi di induzione emotiva sono stati l'uso di droghe (Helfrich et al. 1984), l'assegnazione di compiti particolarmente difficili per generare stress, la presentazione di film o diapositive atti a indurre date emozioni (Alpert et al. 1963, Bachorowski – Owren 1995, Bonner 1943, Brown 1980, Fernandez – Picard 2000, Havrdova – Moravek 1979, Hicks 1979, Karlsson et al. 1998, Markel et al. 1973, Plaikner 1970, Roessler – Lester 1979, Scherer 1977, 1979, Scherer et al. 1985, Skinner 1935, Tolkmitt – Scherer 1986). I problemi che si accompagnano a queste procedure sono in primo luogo metodologici (di fronte a uno stesso stimolo un soggetto può reagire con emozioni diverse, e anche qui sorge il problema di come si possano etichettare inequivocabilmente le emozioni elicitate), e in secondo luogo di efficienza: "these procedures often produce only weak affect", come sottolinea Scherer 2003, p. 232. Un'altra soluzione di tipo "seminaturale" è invece il "bootstrapping": l'uso di materiale naturale per guidare l'elicitazione di materiale recitato. I risultati di questa procedura sono tuttavia a mio parere inversamente proporzionali alla bravura di un attore: nel senso che un vero attore professionista dovrebbe essere in grado da sé di fare del "bootstrapping" mentale prima della propria *performance*, non così nel caso di attori meno competenti, che si affidano a moduli emotivi stereotipati. Ma sulla questione degli attori ritorneremo in seguito.

Infine, di un database è importantissima l'*accessibilità*: sua vera e propria ragione d'essere, se si considera che l'istanza di creare database nasce proprio da un bisogno sempre più urgente di condivisione dei saperi. L'accessibilità comporta senz'altro l'uso di formati simili per i descrittori e per il materiale raccolto (ad esempio, la questione se archiviare i file audio in formato WAV o MPEG: quest'ultimo è indubbiamente un formato più versatile per l'archiviazione e la trasmissione di dati, ma comporta minori informazioni sul segnale). Un altro problema legato all'accessibilità è quello già accennato del *copyright*: il materiale raccolto dal vivo abbisogna del consenso da parte dei soggetti, ma anche degli enti televisivi di cui si sono eventualmente registrate le trasmissioni: questi ultimi possono dare il nulla osta per un singolo esperimento, ma possono negarlo per una libera diffusione del database (è il caso dei *Belfast* e *Reading-Leeds database*).

## 2.2 Il database EMOVO

Un *corpus* emotivo per l'italiano ancora non è stato realizzato, e anche dal punto di vista dei *dataset* il materiale raccolto non è molto. Ricordo in particolare i seguenti studi:

- AA. VV. 2008  
Svariati casi di studio contenuti negli Atti del 1° convegno nazionale GSCP (gruppo di studio della comunicazione parlata), a cura di Emanuela Magno Caldognetto, Federica Cavicchio e Piero Cosi. Segnalo in particolar modo Moneglia 2008, che contiene un sondaggio sulle manifestazioni acustiche delle emozioni nel *corpus* di parlato spontaneo C-ORAL-ROM (Cresti – Moneglia 2005)
- Anolli – Ciceri 1992  
1 frase pronunciata da 2 parlanti non attori. Stati emotivi: collera, gioia, paura, tristezza, disprezzo, tenerezza.
- De Dominicis 1992  
2 frasi, una sintetizzata ed una naturale, pronunciate da 1 attore. Più che stati emotivi sono stati presi in esame atti linguistici, con finalità di analisi del profilo intonativo: richiesta, provocazione, ordine, concessione di facoltatività, informazione, valutazione, rivelazione, domanda sì/no.
- Kori – Magno Caldognetto 1986  
3 vocali pronunciate da 1 attore. Stati emotivi: collera, disgusto, gioia, paura, sorpresa, tristezza.
- Kori – Magno Caldognetto 1990  
1 vocale sintetizzata. Stati emotivi: collera, disgusto, gioia, paura, sorpresa, tristezza.
- Kori – Magno Caldognetto 2003  
2 frasi, una in giapponese e una in italiano, per 2 attori, un giapponese e un italiano. Stati emotivi: collera, disgusto, gioia, paura, sorpresa, tristezza.
- Magno Caldognetto et al. 1998  
1 parola pronunciata da 3 parlanti non attori. Stati emotivi: collera, disgusto, gioia, paura, sorpresa, tristezza.
- Magno Caldognetto – Ferrero 1996  
1 brano e 1 frase pronunciati da 1 attore. Stati emotivi: collera, disgusto, paura, tristezza.

Constatata pertanto la mancanza, per il parlato emotivo, di un *corpus* di dati più esteso, non ho ritenuto inutile la creazione del database di parlato emotivo EMOVO, pur non potendo rispondere a tutte le istanze presentate nello scorso paragrafo, dati i limiti in termini di tempo e risorse a disposizione. Ho tentato comunque di gettare le basi in questa direzione, nella quale ulteriori sviluppi saranno sicuramente necessari, tanto in termini di analisi di dati quanto di raccolta e sistemazione degli stessi.

Descriverò ora la struttura di EMOVO:

### 2.2.1 Scelta delle emozioni

Come già anticipato nel primo capitolo, si è deciso di focalizzare l'attenzione sulle “*big six*”. Pertanto le *performance* emotive espresse nel database ineriscono ai seguenti stati (cliccando sui link su potrà sentire l'emozione corrispondente):

- [disgusto](#)
- [gioia](#)
- [paura](#)
- [rabbia](#)
- [sorpresa](#)
- [tristezza](#)
- (stato emotivo neutro)

Sono queste le emozioni che hanno ricevuto più attenzione dalla comunità scientifica italiana e pertanto si è ritenuto opportuno mantenersi entro questo assortimento di emozioni di base piuttosto che analizzarne altre, le quali non avrebbero potuto beneficiare di termini di paragone derivanti dai precedenti studi. Per evitare interpretazioni troppo diverse di queste emozioni, si sono date apposite spiegazioni almeno per le emozioni di più problematica definizione. In particolar modo, si è spiegato ad ogni attore che la rabbia da esprimere è quella identificata da Scherer come *calda*. Mentre, per quel che attiene al disgusto, si è sottolineato come si intendesse un disgusto fisico più che morale<sup>3</sup>. Per la sorpresa ho invitato gli attori a mantenere questo stato emotivo per l'intera durata della frase, e di non usare a tal fine una tonia interrogativa. Per le altre emozioni non ho ritenuto necessario fornire ulteriori spiegazioni. Infine, è stato richiesto agli attori di non esprimere indicatori emotivi espliciti quali risate, singhiozzi ecc., che avrebbero falsato i test di riconoscimento.

### 2.2.2 Attori

Sebbene convenga con quanti sostengono la necessità di analisi su materiale naturale, ho ritenuto opportuno, data l'esigua mole di studi per la lingua italiana, di rafforzare le nostre conoscenze innanzitutto su un *corpus* di materiale recitato, il quale sarà utile se non altro come termine di paragone per eventuali studi sul parlato emotivo naturale. Abbiamo inoltre già sottolineato come si sia convenuto da parte di molti studiosi sul fatto che gli attori possono darci informazioni probanti sulle componenti socioculturali delle emozioni, mentre il materiale non recitato ce ne segnalerebbe gli aspetti psico-fisiologici. Infatti, anche in altri paesi dove sono stati elaborati molti più studi, come in Germania, l'uso di attori per l'elicitazione delle emozioni a tutt'oggi non è affatto disdegnato.

A differenza di alcuni studiosi tedeschi, si è ritenuto non irrilevante l'impiego di attori *professionisti*. Come già accennato nel primo paragrafo, c'è chi sostiene che gli attori professionisti produrrebbero emozioni stereotipate, a differenza dei parlanti non professionisti cui si è richiesto di esprimere una data emozione.<sup>4</sup> Non si capisce su quali basi si sostenga ciò, e ritengo invece la situazione esattamente ribaltata. E' proprio la ragion

---

<sup>3</sup> E' stato riscontrato come negli studi relativi a questo stato emotivo i valori della media del pitch dipendano dalle modalità di elicitazione dell'emozione: là dove il disgusto è stato indotto tramite materiale audiovisivo, si è rilevata una media del pitch maggiore rispetto al parlato non emotivo; laddove è stato simulato, una media minore. A mio parere, ciò rivela come questa etichetta si presti ad essere rianalizzata in due stati emotivi abbastanza distinti: un disgusto "fisico", caratterizzato da una media maggiore del pitch e da una maggiore attivazione; e un disgusto "morale", culturale, etico o che dir si voglia: questo è caratterizzato da una media minore del pitch, suscitando un *arousal* minore (eventualmente attenuato da *display rules*), in quanto si tratta ovviamente di un'emozione meno fisica.

<sup>4</sup> Cf. ad esempio Burkhardt et al. 2005.

d'essere dell'attore quella di riprodurre le emozioni della vita reale, e una delle prime cose che un attore impara è di non essere impostato, ma "naturale". In generale, si ricorre invece agli stereotipi quando non si ha perizia performativa, esattamente come nel caso dei disegni dei bambini, o appunto in quello di parlanti non professionisti cui si richiede di "fare la rabbia" o altre emozioni: da alcune prove preliminari in laboratorio, con parlanti non attori, ho potuto verificare ciò ottenendo *performance* pessime, che sono migliorate solo nel momento in cui si è chiesto ai parlanti di essere più naturali, cioè impartendo loro un primo ovvio rudimento di recitazione.

Pertanto, ho scelto 6 attori professionisti, 3 maschi e 3 femmine, e mi son premurato di verificare i loro trascorsi professionali, in particolar modo la loro conoscenza del famoso metodo Stanislavski, consistente nell'autoindursi una data emozione ricordando le situazioni della propria vita in cui si è provata intensamente la medesima emozione. La loro età media è di 27 anni: non mi sono preoccupato di scegliere attori di età molto diverse in quanto allo stadio attuale degli studi rimane ancora indeterminata e del tutto negletta la correlazione fra età ed espressione delle emozioni. Ho dunque preferito uniformarmi agli standard di ricerca, piuttosto che disperdere energie in nuove direzioni.

Gli attori sono stati classificati come [M1](#), [M2](#), [M3](#), [F1](#), [F2](#), [F3](#). Ecco le loro referenze professionali:

**M1:**

Francesco Maiorca. 30 anni, palermitano. Formazione presso: laboratorio teatrale a cura Bibi Bianca; Accademia di formazione per attori di musical "MASTMASTER" a cura di Sofia Amendolea; Seminario di Giancarlo Sepe "La Musica in Movimento" presso il Teatro La Comunità; Seminario di educazione vocale e canto a cura di Lena Biolcati e Patrizia Troiani; Seminario di uso della voce "La Vita Bestia" a cura di Filippo Timi; Laboratorio teatrale "Masculi e Fimmini" a cura di Davide Enia; Laboratorio teatrale "Il contatto" a cura di Marco Maltauro. Partecipazione come attore in vari spettacoli teatrali e cortometraggi, nonché del film "Ho ammazzato Berlusconi" di Daniele Giometti e Gianluca Rossi, attualmente in lavorazione.

**M2:**

Simone Arcese. 27 anni, romano. Conseguimento dei titoli di attore rilasciati dalla Scuola di recitazione "RIBALTE", di Enzo Garinei, e dal Centro Sperimentale di Cinematografia. Partecipazione a svariati film (di registi quali Michele Soavi, Daniela Stroppa, Pupi Avati; ruolo di protagonista in *Saremo film* di Ludovica Marineo) e sceneggiati televisivi ("Incantesimo", "Distretto di polizia").

**M3:**

Filippo Sandon. 30 anni, padovano. Ha studiato con Marina De Luca, attrice e regista padovana, per 2 anni. Con lei ha iniziato a lavorare in teatro, allestendo due spettacoli teatrali. Ha proseguito gli studi a Roma con Beatrice Bracco e Michael Margotta dell'Actor's centre. Ha lavorato in film per cinema e televisione con registi come Luchetti, Soavi, Vanzina e Perelli, anche in parti di coprotagonista.

**F1:**

Simona Augelli. 28 anni, gaetana. Diplomata presso la Scuola di Teatro "Auroville" diretta da Augusto Zucchi; laboratori teatrali con: Tullio Solenghi, Giuliana De Sio, Massimo

Dapporto, Paolo Giuranna. Svariate rappresentazioni teatrali e partecipazione a spot televisivi e alla fiction "Un posto al sole".

**F2:**

Maria Cristina Papeo. 23 anni, foggiana. Laureata in DAMS nel curriculum "regista e coordinatore di gruppi di teatro"; formatasi come attrice con un Master Class in clown-attore tenuto da Vladimir Olshansky (artista ospite del Cirque du Soleil) e con due laboratori teatrali tenuti rispettivamente da Paolo Alessandri e Pino Casolaro. Attrice in due rappresentazioni teatrali.

**F3:**

Chiara Fiorelli. 25 anni, romana. Qualifiche di "attore" e "aiuto regista teatrale" conseguite presso la Scuola Internazionale di Teatro "Circo a Vapore" di Roma, riconosciuta dalla Regione Lazio. Esperienze lavorative: otto messe in scena, presso i teatri "Euclide", "Anfitrione" e "In Portico" di Roma, nell'ambito della didattica triennale della scuola di teatro; tournée estive di teatro di strada in numerose piazze italiane; partecipazione a vari festival tra cui Caorle, Aosta, Carpineto Romano e la Notte Bianca di Roma; interprete di *Trincee* allo "Spazio Uno" di Roma; partecipazione come attrice a vari cortometraggi.

2.2.3 *Materiale linguistico*

Riporto qui l'elenco delle frasi da elicitare:

- 1 [Gli operai si alzano presto.](#)
- 2 [I vigili sono muniti di pistola.](#)
- 3 [La cascata fa molto rumore.](#)
- 4 [L'autunno prossimo Tony partirà per la Spagna: nella prima metà di ottobre.](#)
- 5 [Ora prendo la felpa di là ed esco per fare una passeggiata.](#)
- 6 [Un attimo dopo s'è incamminato...ed è inciampato.](#)
- 7 [Vorrei il numero telefonico del Signor Piatti.](#)
- 8 [La casa forte vuole col pane.](#)
- 9 [La forza trova il passo e l'aglio rosso.](#)
- 10 [Il gatto sta scorrendo nella pera](#)
- 11 [Insalata pastasciutta coscia d'agnello limoncello.](#)
- 12 [Uno quarantatré dieci mille cinquantasette venti.](#)
- 13 [Sabato sera cosa farà?](#)
- 14 [Porti con te quella cosa?](#)

Ho ritenuto funzionale ai fini proposti l'uso di frasi isolate. Nella maggior parte dei database si è optato per una scelta simile e l'unica questione che si potrebbe sollevare è quella relativa alla maggiore o minore efficacia, ai fini dell'elicitazione e dell'analisi di un'emozione, della lettura di brani anziché frasi. A mio parere l'impiego di brani da colorare emotivamente nella *performance* presenta numerose questioni metodologiche: se da una parte è vero che l'impiego di un brano sufficientemente lungo permette meglio di una frase di cogliere lo sviluppo di un'emozione nel tempo, dall'altra ciò complica di molto la già spinosa questione dell'etichettamento dell'emozione<sup>5</sup>. In un brano si può cambiare facilmente lo stato emotivo: si può passare dal neutro all'emotivo e viceversa (e di una

---

<sup>5</sup> Per cui si rimanda a D'Urso – Cavicchio – Magno Caldognetto 2008.



stessa emozione si possono rendere stati più o meno intensi, in modo del tutto “*speaker-specific*”) se non passare addirittura ad un’altra emozione: come si classificherà tutto ciò? Nello studio di Anolli – Ciceri 1992 si è risolto il problema isolando una frase al centro del brano e limitando la funzione e l’impiego del brano ad una semplice preparazione dell’antecedente emotivo del parlante. Tutto ciò mi è sembrato superfluo nel caso di attori professionisti, la cui preparazione interiore supplisce di per sé alla contestualizzazione emotiva della frase.

Per quanto riguarda il contenuto delle frasi, ai fini di non falsare i test di riconoscimento dell’emozione, una condizione necessaria è che il valore semantico di tale contenuto sia emotivamente neutro. A tal fine è possibile utilizzare frasi che non si riferiscano specificatamente a una data espressione emotiva, e frasi “*nonsense*”. Se le prime possono porre qualche problema all’attore per la loro collocazione nella giusta situazione emotiva,<sup>6</sup> le seconde comportano il rischio di essere recitate in una maniera stereotipata in quanto l’attore può non riuscire a “sentirle” come naturali.<sup>7</sup> Consapevoli di queste problematiche, ho deciso di usare entrambe le categorie (frasi 1-7, 13, 14 semanticamente neutre, frasi 8-10 *nonsense*), aggiungendo due frasi che si situano in una posizione intermedia rispetto alle frasi nonsense e a quelle semanticamente non emotive: un elenco di numeri e un elenco di alimenti (frasi 11,12).

Delle frasi semanticamente neutre ho inoltre formulato delle versioni brevi (1-3) e lunghe (4-7).

Ai fini delle analisi spettroacustiche sono state soddisfatte le seguenti condizioni basilari:<sup>8</sup>

- presenza nelle frasi di tutti i fonemi della lingua italiana
- presenza in ogni frase di un giusto bilanciamento fra consonanti sorde e sonore
- assenza di consonanti occlusive iniziali (in quanto in tale posizione non è possibile misurare la durata della fase di occlusione).

Per quanto riguarda eventuali analisi prosodiche si è tenuto conto dei quattro tipi di tonia descritti da Canepari 1985. La *tonia conclusiva* è analizzabile nel database nelle frasi 1-5 e 7-13. Di norma la tonia conclusiva viene impiegata quando il parlante comunica di aver terminato il suo enunciato con un significato semplicemente denotativo: è ovvio pertanto che la sua ricorrenza può cambiare in base allo stato emotivo in questione. Con la *tonia sospensiva* invece il parlante avverte il suo ascoltatore che sta seguendo un’informazione importante, avvalendosi di una pausa per fini di *suspence*: è quanto è stato richiesto di fare agli attori per la frase 6. La *tonia divisiva* è la tonia non marcata, la cui funzione è la semplice suddivisione dell’enunciato in diversi segmenti fono-sintattici, senza particolari sfumature semantiche: agli attori è stato spiegato che è questo il significato dei due punti nella frase 4. Infine la *tonia interrogativa* è rappresentata dalle frasi interrogativa parziale 14, che si distingue dalla frase 13, interrogativa totale (o a risposta sì/no, che è invece caratterizzata di norma da una tonia conclusiva). Questa distinzione è importante perché, come rilevano Scherer et al. 1984, il profilo intonativo di una domanda può

---

<sup>6</sup> Frasi emotivamente neutre in assoluto in realtà non esistono, ed è del tutto dipendente dalle idiosincrasie e dai vissuti personali dei singoli interpreti il fatto che una frase “neutra” possa addirsi ad un’emozione piuttosto che ad un’altra.

<sup>7</sup> Cf. ad esempio Scherer 1981.

<sup>8</sup> Colgo qui l’occasione per ringraziare la Dottoressa Magno Caldognetto per i preziosi suggerimenti datimi al riguardo.

cambiare in base a questi due casi.<sup>9</sup> Come specificato per la tonia conclusiva, anche per gli altri tipi di tonia la loro effettiva realizzazione, secondo i criteri descrittivi di Canepari o di altri studiosi, può cambiare notevolmente di emozione in emozione: in alcune emozioni date tonie possono essere annullate o modificate a fini espressivi più specifici, e secondo quali modalità ciò avvenga sarà precipuamente il risultato delle analisi prosodiche che si potranno effettuare sul materiale offerto da questo database.

#### 2.2.4 Registrazione e archiviazione dei dati

La *performance* degli attori è stata registrata in una stanza dei laboratori della Fondazione Ugo Bordoni a Roma, in orari che permettessero adeguate condizioni di quiete acustica. Sono stati impiegati due microfoni professionali marca SHURE modello SM58LC ed un registratore digitale MARANTZ modello PMD670. Le registrazioni sono state eseguite con una frequenza di campionamento di 48 kHz, 16 bit stereo, formato wav. Gli attori hanno avuto la possibilità di muoversi liberamente e ciò ha ovviamente inficiato i dati relativi all'intensità assoluta del segnale, dipendente dalla lontananza della bocca dal microfono. In altri casi si è invece dovuto provvedere all'abbassamento manuale del volume di registrazione, onde evitare la saturazione del segnale nei casi di emozioni che generano livelli di energia elevati, quali la rabbia.

Si è consentito agli attori di avere tutto il tempo necessario per prepararsi all'autoinduzione dell'emozione, concedendo loro le pause richieste e consentendo la ripetizione delle frasi la cui esecuzione non li convinceva.

Ogni sessione di registrazione è durata all'incirca un'ora. Successivamente le frasi registrate sono state isolate, una per *file*, ed archiviate nel database in 6 cartelle diverse corrispondenti ai 6 attori. Ogni *file* è stato etichettato nel seguente modo:

##### ***neu-m1-b1.wav***

Nella prima parte del nome è indicato lo stato emotivo (**neutro**, **disgusto**, **gioia**, **paura**, **rabbia**, **sorpresa**, **tristezza**).

Nella seconda parte del nome è identificato l'attore/attrice (**m1**, **m2**, **m3**, **f1**, **f2**, **f3**) che recita la frase.

Nella terza parte del nome è indicato il tipo di frase:

le frasi 1-3 sono state classificate rispettivamente. b1, b2, b3 (**brevi**)

le frasi 4-7 sono state classificate risp. l1, l2, l3, l4 (**lunghe**)

le frasi 8-12 sono state classificate risp. n1, n2, n3, n4, n5 (**nonsense**)

le frasi 13,14 sono state classificate risp. d1, d2 (**domande**).

Infine, tutte le frasi sono state inserite in un database elaborato col programma Access, grazie al quale è possibile ordinare i *record* a piacere in base all'emozione, al testo della frase, al tipo di frase o all'attore/attrice, nonché procedere all'ascolto della frase cliccando su un apposito *link*.

In totale sono stati archiviati 588 *record*: per ognuno dei 6 attori sono state registrate 98 frasi, corrispondenti alle 14 frasi pronunciate nei 6 stati emotivi più lo stato neutro.

Il totale del materiale registrato, escluse le pause, è all'incirca di 10 minuti per attore, quindi un'ora per l'intero database. EMOVO è attualmente disponibile su un singolo cd-rom: sarà distribuito a che ne farà [richiesta](#).

---

<sup>9</sup> In particolare rilevano come uno contorno del pitch discendente venga percepito come neutro in una "WH-question", come aggressivo in una "Y/N question."

### 3. VALIDAZIONE DEL DATABASE EMOVO

#### 3.1 Il decoding del parlato emotivo

La capacità umana di riconoscere le emozioni in base a meri stimoli vocali ha interessato psicologi e psichiatri sin dall'inizio del secolo passato, in special modo fra gli anni '50 e '80. Gli studi al riguardo si sono configurati per lo più in questo modo: si è richiesto a dei soggetti ascoltatori di indicare lo stato emotivo di stimoli vocali fatti loro ascoltare (nella maggior parte dei casi, elicitati da attori), facendo loro scegliere fra una lista più o meno lunga di alternative; successivamente sono state calcolate le percentuali di stimoli correttamente interpretati ed è stato confrontato tale risultato con le percentuali di risposte che ci si sarebbe aspettati se il soggetto avesse deciso a caso. In alcuni studi sono state inoltre elaborate delle "matrici di confusione", per inquadrare meglio i tipi di errori e di false attribuzioni cui sono andati incontro i soggetti. Secondo Scherer – Johnstone – Klasmeyer 2003, a tirare le somme di tutti questi studi (all'incirca una trentina) si può quantificare la percentuale di accuratezza umana nel riconoscere le emozioni, in base a soli stimoli vocali, intorno al 60%: una percentuale cinque volte maggiore di quanto ci si aspetterebbe per effetto del caso. E' un livello di accuratezza notevole, se si tiene in considerazione che in molti di questi studi gli stimoli vocali erano assai brevi e che gli stati emotivi con cui li si caratterizzava erano rari o poco studiati (come l'orgoglio o la gelosia), se paragonati alle più comuni emozioni di base. Anche altri studiosi riportano valori analoghi di percentuale di accuratezza: 65% per Van Bezooijen 1984; 56% per Scherer – Banse – Wallbott – Goldbeck 1991.

Questi valori sono inferiori, in media, di un 15% rispetto alla capacità umana di riconoscere emozioni in base alla mimica facciale (cf. Ekman 1994): in realtà, la differenza fra riconoscimento a partire dalla voce e riconoscimento dalla mimica facciale cambia alquanto di emozione in emozione, come si può vedere nella seguente tabella (fonti: Ekman 1994 per la prima riga, Scherer – Banse – Wallbott 2001 per la seconda):

|                        | neutro | rabbia | paura | gioia | tristezza | disgusto | sorpresa | media      |
|------------------------|--------|--------|-------|-------|-----------|----------|----------|------------|
| <b>mimica facciale</b> | N.P.   | 78%    | 77%   | 95%   | 79%       | 80%      | 88%      | <b>78%</b> |
| <b>voce</b>            | 74%    | 77%    | 61%   | 57%   | 71%       | 31%      | N.P.     | <b>62%</b> |

A falsare i valori medi sono soprattutto i valori relativi a gioia e disgusto, fin troppo riconoscibili nel volto di una persona. Per quanto riguarda il riconoscimento vocale possiamo rilevare come vengano identificate meglio la tristezza e la rabbia, seguite dalla paura. E' interessante notare come questi primati di riconoscimento siano da inquadrare per Johnstone – Scherer 2000 nel contesto dell'evoluzione darwinianamente intesa: essa avrebbe selezionato e favorito, per la rabbia e la paura, sviluppi nell'espressione vocale, in modo che gli antenati dell'uomo potessero avvertirsi e minacciarsi in modo esplicito, anche a lunghe distanze. Avrebbe invece selezionato e favorito sviluppi biologici relativi alla mimica facciale per emozioni quali disgusto e gioia, di modo che fra compagni, nel senso etimologico del termine, ci si potesse capire al volo. Ciò presuppone, naturalmente, che l'espressione facciale sia in qualche modo indipendente dalla volontà dell'individuo, cosa nel del tutto ovvia.

Per quanto riguarda invece le attribuzioni erronee, dalle matrici di confusione dei vari studi si possono evincere dati importanti sul processo di *decoding* del soggetto, e di riflesso, sulla natura del parlato emotivo. Dallo studio di Banse – Scherer 1996 è emerso, ad esempio, che le confusioni non avvenivano casualmente fra tutte le emozioni, ma tra

famiglie di emozioni: segno che non si trattava di veri e propri errori, ma di approssimazioni (dallo studio delle quali emergono rapporti di contiguità e prossimità fra le emozioni, secondo il modello dimensionale di cui si è parlato nel primo capitolo: resterebbe da indagare più in dettaglio quali dimensioni siano maggiormente coinvolte nel generare confusione). Per citare degli esempi, dallo studio sopra ricordato è emerso che il disgusto è stato confuso solo con emozioni negative, che la rabbia calda è stata confusa quasi esclusivamente con la rabbia fredda e la riprovazione, che l'interessamento è stato confuso con l'orgoglio e la gioia molto più frequentemente che con le altre emozioni. Da altri studi riporto ulteriori correlazioni: tristezza confusa con amore; vergogna confusa con tristezza; disprezzo confuso con collera o disgusto; gioia e sorpresa reciprocamente confuse (Fairbanks – Pronovost 1939, Davitz – Davitz 1959, Davitz 1964, Dimitrovsky 1964, Albas et al. 1976, Van Bezooijen 1984). Tuttavia, per quanto si possano trovare correlazioni sistematiche nelle matrici di confusione all'interno dei singoli studi, scarseggiano correlazioni fra le matrici di confusioni dei vari studi. Ciò è dovuto senz'altro ai seguenti motivi:

- diverso assortimento e/o etichettamento delle emozioni impiegate
- diversità nelle impostazioni dei test, in particolar modo riguardo al tipo di domande
- diverse moli di dati raccolte, in particolar modo riguardo al numero dei soggetti e alla quantità del materiale sottoposto ai test, se vocali, frasi, brani o altro.

### 3.2 Descrizione del test di validazione

Gli studi citati nel paragrafo precedente sono stati spesso criticati per aver usato un *set* ristretto e limitato di emozioni da far scegliere ai soggetti ascoltatori, caratterizzandosi pertanto come test di discriminazione più che test di riconoscimento. Nel nostro caso, invece, il problema non si pone, essendo il fine principale del test non il riconoscimento né la discriminazione, ma la validazione del database EMOVO, ovvero garantire la bontà delle *performance* emotive degli attori coinvolti. Tale test è stato somministrato a due gruppi di soggetti ascoltatori: i risultati verranno presentati separatamente e congiuntamente.

#### 3.2.1 Campione

- Primo gruppo:  
12 soggetti<sup>10</sup>  
- Secondo gruppo:  
12 soggetti<sup>11</sup>

---

<sup>10</sup> Il primo gruppo è costituito da 6 maschi e 6 femmine, di età compresa fra i 21 e i 32 anni. Riporto qui i loro nomi non solo a mo' di ringraziamento, ma anche per indicarne il sesso, la cui influenza sulla performance riconoscitiva verrà presa in considerazione nella discussione dei risultati. **Soggetto 1:** Stefano V. Talone, 24 anni. **Soggetto 2:** Nicola Villa, 23 anni. **Soggetto 3:** Elisabetta Lopalco, 24 anni. **Soggetto 4:** Erica Pulvirenti, 22 anni. **Soggetto 5:** Beatrice Bertini, 29 anni. **Soggetto 6:** Michelangelo Iadarola, 26 anni. **Soggetto 7:** Marina Sforza, 22 anni. **Soggetto 8:** Valerio Friello, 21 anni. **Soggetto 9:** Chiara Carmosino, 21 anni. **Soggetto 10:** Luigina Ruffolo, 23 anni. **Soggetto 11:** Valerio Pescatore, 32 anni. **Soggetto 12:** Rio Olmedo, 22 anni.

<sup>11</sup> Il secondo gruppo è costituito da studenti della Facoltà di Lettere e Filosofia dell'Università della Calabria, cui è stato somministrato il mio test grazie alla gentile

### 3.2.2 Procedura

Ai soggetti è stata fornita oralmente una sintetica descrizione dei fini del test. Si è successivamente somministrato loro un pre-test per acclimatarli all'esperimento e al software impiegato a tal fine. Il test preliminare è consistito nell'ascolto di 6 frasi, una per ogni stato emotivo, una per attore: di frase in frase il soggetto doveva scegliere fra una coppia di risposte possibili (ex: rabbia/gioia, paura/tristezza).

In seguito si è somministrato al soggetto ascoltatore il test di validazione completo, così costituito:

Per ogni attore, fra le 14 tipologie di frasi del database, sono state scelte due frasi *nonsense*:

- la casa forte vuole col pane
- il gatto sta scorrendo nella pera

ciò per ridurre al minimo l'eventualità che il contenuto semantico delle frasi potesse indirizzare o sviare l'attribuzione emotiva degli ascoltatori. Di frasi ne sono state scelte due in modo che, nel caso di due domande con diverse emozioni da riconoscere ma con una medesima coppia di risposte, si potesse usare ora una frase, ora l'altra. Ovviamente, di tali tipologie di frasi sono state selezionate all'interno del database tutte le coloriture emotive, per poterle presentare al soggetto ascoltatore pronunciate nei sette stati emotivi: neutro, disgusto, gioia, paura, rabbia, sorpresa, tristezza.

Quindi la frase di un attore, per ogni stato emotivo, è stata presentata al soggetto ascoltatore sei volte, fornendo ogni volta una coppia diversa di risposte alternative: una risposta giusta per lo stato emotivo in questione, una risposta sbagliata relativa a un altro stato emotivo scelto fra gli altri sei.

Riporto, a mo' di esemplificazione, la parte iniziale della strutturazione del test:

#### *Domanda 1*

in quale stato emotivo pensi sia pronunciata la domanda? (frase da ascoltare nello stato emotivo **neutro**).

- neutro
- disgusto

#### *Domanda 2*

in quale stato emotivo pensi sia pronunciata la domanda? (frase da ascoltare nello stato emotivo **neutro**).

- neutro
- gioia

[...]

#### *Domanda 6*

in quale stato emotivo pensi sia pronunciata la domanda? (frase da ascoltare nello stato emotivo **neutro**).

- neutro
- tristezza

#### *Domanda 7*

in quale stato emotivo pensi sia pronunciata la domanda? (frase da ascoltare nello stato emotivo di **disgusto**).

- disgusto
- neutro

---

collaborazione del Prof. Luciano Romito e del Dott. Vincenzo Galatà, su invito dell'Ing. Andrea Paoloni della Fondazione Ugo Bordoni.

(In quest'ultima domanda, essendosi presentata la stessa coppia di risposte della domanda 1, oltre allo stato emotivo della frase da sentire è stato cambiato anche il testo, come spiegato prima)

E così via, per un totale di 42 domande (sei domande per sette stati emotivi). L'ordine delle domande, così come quello delle risposte, è stato opportunamente reso casuale tramite un'apposita funzione del software di *test-making*. Lo stesso software ha fornito l'interfaccia da presentare al soggetto ascoltatore, contenente il testo della domanda, le due caselle delle risposte da segnare, un pulsante per sentire il file audio abbinato alla domanda, un pulsante per procedere alla domanda successiva. E' stato concesso al soggetto di poter riascoltare la frase della domanda tutte le volte che ritenesse necessario. Non è stato concesso al soggetto di poter controllare le risposte precedenti. Orientativamente, tutti i soggetti hanno impiegato dieci minuti di tempo per rispondere alla batteria di 42 domande.

Ovviamente, la batteria delle 42 domande così strutturate, è stata approntata per ogni attore e per ogni attrice. A ogni soggetto ascoltatore sono state somministrate due batterie di domande, una per un attore maschio e una per un'attrice, secondo questo schema.

Batteria relativa a M1

Somministrata ai soggetti 1 2 3 4 di entrambi i gruppi di ascolto

Batteria relativa a F1

Somministrata ai soggetti 1 2 3 4 id.

Batteria relativa a M2

Somministrata ai soggetti 5 6 7 8 id.

Batteria relativa a F2

Somministrata ai soggetti 5 6 7 8 id.

Batteria relativa a M3

Somministrata ai soggetti 9 10 11 12 id.

Batteria relativa a F3

Somministrata ai soggetti 9 10 11 12 id.

### 3.3 *Presentazione dei risultati del test*

Mi sono premurato di calcolare l'accuratezza del riconoscimento correggendo i dati percentuali delle risposte secondo questa formula:

$$\text{accuratezza} = [(\text{risposte corrette} - \text{risposte sbagliate}) / \text{numero delle domande}] \cdot 100$$

Vale a dire si è sottratto il numero delle risposte sbagliate da quello delle risposte corrette, e si è divisa la differenza per il numero totale delle domande.

Con tale formula, che si è applicata in tutti i casi, si è voluto eliminare l'effetto della casualità nelle risposte date dai soggetti ascoltatori. I risultati sono stati qui approssimati per chiarezza espositiva.

### 3.3.1 Risultati relativi al primo gruppo

#### 3.3.1.1 Risultati relativi agli attori

| <b>RIQUADRO 1</b>             |                   |                    |                |                           |
|-------------------------------|-------------------|--------------------|----------------|---------------------------|
| <b>BATTERIA RELATIVA A M1</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 1                             | 37                | 5                  | 42             | 76%                       |
| 2                             | 40                | 2                  | 42             | 90%                       |
| 3                             | 36                | 6                  | 42             | 71%                       |
| 4                             | 40                | 2                  | 42             | 90%                       |
|                               |                   |                    |                | <b>accuratezza per M1</b> |
| 1,2,3,4                       | 153               | 15                 | 168            | <b>82%</b>                |
| <b>BATTERIA RELATIVA A F1</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 1                             | 36                | 6                  | 42             | 71%                       |
| 2                             | 39                | 3                  | 42             | 86%                       |
| 3                             | 37                | 5                  | 42             | 76%                       |
| 4                             | 36                | 6                  | 42             | 71%                       |
|                               |                   |                    |                | <b>accuratezza per F1</b> |
| 1,2,3,4                       | 148               | 20                 | 168            | <b>76%</b>                |
| <b>BATTERIA RELATIVA A M2</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 5                             | 40                | 2                  | 42             | 90%                       |
| 6                             | 37                | 5                  | 42             | 76%                       |
| 7                             | 37                | 5                  | 42             | 76%                       |
| 8                             | 38                | 4                  | 42             | 81%                       |
|                               |                   |                    |                | <b>accuratezza per M2</b> |
| 5,6,7,8                       | 152               | 16                 | 168            | <b>81%</b>                |
| <b>BATTERIA RELATIVA A F2</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 5                             | 41                | 1                  | 42             | 95%                       |
| 6                             | 38                | 4                  | 42             | 81%                       |
| 7                             | 40                | 2                  | 42             | 90%                       |
| 8                             | 40                | 2                  | 42             | 90%                       |
|                               |                   |                    |                | <b>accuratezza</b>        |

|                         |                   |                    |                |                          |
|-------------------------|-------------------|--------------------|----------------|--------------------------|
|                         |                   |                    |                | per F2                   |
| 5,6,7,8                 | 159               | 9                  | 168            | 89%                      |
| BATTERIA RELATIVA A M3  |                   |                    |                |                          |
| soggetto ascoltatore    | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto |
| 9                       | 37                | 5                  | 42             | 76%                      |
| 10                      | 40                | 2                  | 42             | 90%                      |
| 11                      | 37                | 5                  | 42             | 76%                      |
| 12                      | 36                | 6                  | 42             | 71%                      |
|                         |                   |                    |                | accuratezza per M3       |
| 9,10,11,12              | 150               | 18                 | 168            | 79%                      |
| BATTERIA RELATIVA A F3  |                   |                    |                |                          |
| soggetto ascoltatore    | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto |
| 9                       | 40                | 2                  | 42             | 90%                      |
| 10                      | 38                | 4                  | 42             | 81%                      |
| 11                      | 40                | 2                  | 42             | 90%                      |
| 12                      | 41                | 1                  | 42             | 95%                      |
|                         |                   |                    |                | accuratezza per F3       |
| 9,10,11,12              | 159               | 9                  | 168            | 89%                      |
| ACCURATEZZA RELATIVA A: |                   |                    |                |                          |
| ATTRICI                 | risposte corrette | risposte sbagliate | numero domande |                          |
| 85%                     | 466               | 38                 | 504            |                          |
| ATTORI                  |                   |                    |                |                          |
| 81%                     | 455               | 49                 | 504            |                          |
| COMPLESSIVA             |                   |                    |                |                          |
| 83%                     | 921               | 87                 | 1008           |                          |



3.3.1.2 Risultati relativi ai soggetti ascoltatori del primo gruppo:

| <b>RIQUADRO 2</b>            |                        |                           |                       |                    |
|------------------------------|------------------------|---------------------------|-----------------------|--------------------|
| <b>soggetto</b>              | <b>risposte giuste</b> | <b>risposte sbagliate</b> | <b>numero domande</b> | <b>Accuratezza</b> |
| 1                            | 73                     | 11                        | 84                    | 74%                |
| 2                            | 79                     | 5                         | 84                    | 88%                |
| <b>3</b>                     | <b>73</b>              | <b>11</b>                 | <b>84</b>             | <b>74%</b>         |
| <b>4</b>                     | <b>76</b>              | <b>8</b>                  | <b>84</b>             | <b>81%</b>         |
| <b>5</b>                     | <b>81</b>              | <b>3</b>                  | <b>84</b>             | <b>93%</b>         |
| 6                            | 75                     | 9                         | 84                    | 79%                |
| 7                            | 77                     | 7                         | 84                    | 83%                |
| 8                            | 78                     | 6                         | 84                    | 86%                |
| 9                            | 77                     | 7                         | 84                    | 83%                |
| <b>10</b>                    | <b>78</b>              | <b>6</b>                  | <b>84</b>             | <b>86%</b>         |
| 11                           | 77                     | 7                         | 84                    | 83%                |
| 12                           | 77                     | 7                         | 84                    | 83%                |
| soggetti maschi              | 459                    | 45                        | 504                   | 82%                |
| <b>soggetti femmine</b>      | <b>462</b>             | <b>42</b>                 | <b>504</b>            | <b>83%</b>         |
| <b>risultati complessivi</b> | <b>921</b>             | <b>87</b>                 | <b>1008</b>           | <b>83%</b>         |

(in neretto sono evidenziati i soggetti di sesso femminile)

3.3.1.3 Risultati relativi alle emozioni del primo gruppo

Tale tabella va letta nel seguente modo:

- emozioni nelle ordinate confuse con emozioni nelle ascisse nel X% dei casi
- quando l'emozione dei due assi è la stessa il valore nella casella indica l'accuratezza di riconoscimento di quell'emozione per tutti i soggetti del gruppo

| <b>EMOZIONE ELICITATA</b> | <b>RIQUADRO 3</b> | <b>EMOZIONE PERCEPITA</b> |                 |              |              |               |                 |                  |
|---------------------------|-------------------|---------------------------|-----------------|--------------|--------------|---------------|-----------------|------------------|
|                           |                   | <b>Neutro</b>             | <b>Disgusto</b> | <b>Gioia</b> | <b>Paura</b> | <b>Rabbia</b> | <b>Sorpresa</b> | <b>Tristezza</b> |
|                           | <b>Neutro</b>     | <b>96%</b>                | 1,3%            | 0%           | 0%           | 1,3%          | 0%              | 1,3%             |
|                           | <b>Disgusto</b>   | 3%                        | <b>76%</b>      | 3%           | 6%           | 6%            | 2%              | 6%               |
|                           | <b>Gioia</b>      | 1%                        | 3%              | <b>64%</b>   | 6%           | 8%            | 15%             | 3%               |
|                           | <b>Paura</b>      | 1%                        | 10%             | 0%           | <b>72%</b>   | 4%            | 6%              | 7%               |
|                           | <b>Rabbia</b>     | 0%                        | 3%              | 0%           | 3%           | <b>93%</b>    | 0%              | 1%               |
|                           | <b>Sorpresa</b>   | 1,5%                      | 3%              | 3%           | 1,5%         | 0%            | <b>83%</b>      | 8%               |
|                           | <b>Tristezza</b>  | 1,3%                      | 0%              | 1,3%         | 1,3%         | 0%            | 0%              | <b>96%</b>       |

### 3.3.2 Risultati relativi al secondo gruppo

#### 3.3.2.1 Risultati relativi agli attori

| <b>RIQUADRO 4</b>             |                   |                    |                |                           |
|-------------------------------|-------------------|--------------------|----------------|---------------------------|
| <b>BATTERIA RELATIVA A M1</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 1                             | 35                | 7                  | 42             | 67%                       |
| 2                             | 37                | 5                  | 42             | 76%                       |
| 3                             | 39                | 3                  | 42             | 86%                       |
| 4                             | 41                | 1                  | 42             | 95%                       |
|                               |                   |                    |                | <b>accuratezza per M1</b> |
| 1,2,3,4                       | 152               | 16                 | 168            | <b>81%</b>                |
| <b>BATTERIA RELATIVA A F1</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 1                             | 37                | 5                  | 42             | 76%                       |
| 2                             | 35                | 7                  | 42             | 67%                       |
| 3                             | 37                | 5                  | 42             | 76%                       |
| 4                             | 41                | 1                  | 42             | 95%                       |
|                               |                   |                    |                | <b>accuratezza per F1</b> |
| 1,2,3,4                       | 150               | 18                 | 168            | <b>79%</b>                |
| <b>BATTERIA RELATIVA A M2</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 5                             | 39                | 3                  | 42             | 86%                       |
| 6                             | 34                | 8                  | 42             | 62%                       |
| 7                             | 39                | 3                  | 42             | 86%                       |
| 8                             | 36                | 6                  | 42             | 71%                       |
|                               |                   |                    |                | <b>accuratezza per M2</b> |
| 5,6,7,8                       | 148               | 20                 | 168            | <b>76%</b>                |
| <b>BATTERIA RELATIVA A F2</b> |                   |                    |                |                           |
| soggetto ascoltatore          | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 5                             | 38                | 4                  | 42             | 81%                       |
| 6                             | 31                | 11                 | 42             | 48%                       |
| 7                             | 40                | 2                  | 42             | 90%                       |
| 8                             | 41                | 1                  | 42             | 95%                       |
|                               |                   |                    |                | <b>accuratezza per F2</b> |
| 5,6,7,8                       | 150               | 18                 | 168            | <b>79%</b>                |
| <b>BATTERIA RELATIVA A M3</b> |                   |                    |                |                           |

| soggetto ascoltatore           | risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
|--------------------------------|-------------------|--------------------|----------------|---------------------------|
| 9                              | 35                | 7                  | 42             | 67%                       |
| 10                             | 39                | 3                  | 42             | 86%                       |
| 11                             | 36                | 6                  | 42             | 71%                       |
| 12                             | 38                | 4                  | 42             | 81%                       |
|                                |                   |                    |                | <b>accuratezza per M3</b> |
| 9,10,11,12                     | 148               | 20                 | 168            | <b>76%</b>                |
| <b>BATTERIA RELATIVA A F3</b>  |                   |                    |                |                           |
| soggetto ascoltatore           | Risposte corrette | risposte sbagliate | numero domande | accuratezza per soggetto  |
| 9                              | 36                | 6                  | 42             | 71%                       |
| 10                             | 39                | 3                  | 42             | 86%                       |
| 11                             | 35                | 7                  | 42             | 67%                       |
| 12                             | 40                | 2                  | 42             | 90%                       |
|                                |                   |                    |                | <b>accuratezza per F3</b> |
| 9,10,11,12                     | 150               | 18                 | 168            | <b>79%</b>                |
| <b>ACCURATEZZA RELATIVA A:</b> |                   |                    |                |                           |
| <b>ATTRICI</b>                 | risposte corrette | risposte sbagliate | numero domande |                           |
| <b>79%</b>                     | 450               | 54                 | 504            |                           |
| <b>ATTORI</b>                  |                   |                    |                |                           |
| <b>78%</b>                     | 448               | 56                 | 504            |                           |
| <b>COMPLESSIVA</b>             |                   |                    |                |                           |
| <b>78%</b>                     | 898               | 110                | 1008           |                           |

### 3.3.2.2 Risultati relativi ai soggetti ascoltatori del secondo gruppo

| <b>RIQUADRO 5</b>            |                          |                           |                       |                    |
|------------------------------|--------------------------|---------------------------|-----------------------|--------------------|
| <b>soggetto</b>              | <b>risposte corrette</b> | <b>risposte sbagliate</b> | <b>numero domande</b> | <b>Accuratezza</b> |
| 1                            | 72                       | 12                        | 84                    | 71%                |
| 2                            | 72                       | 12                        | 84                    | 71%                |
| 3                            | 76                       | 8                         | 84                    | 81%                |
| 4                            | 82                       | 2                         | 84                    | 95%                |
| 5                            | 77                       | 7                         | 84                    | 83%                |
| 6                            | 65                       | 19                        | 84                    | 55%                |
| 7                            | 79                       | 5                         | 84                    | 88%                |
| 8                            | 77                       | 7                         | 84                    | 83%                |
| 9                            | 71                       | 13                        | 84                    | 69%                |
| 10                           | 78                       | 6                         | 84                    | 86%                |
| 11                           | 71                       | 13                        | 84                    | 69%                |
| 12                           | 78                       | 6                         | 84                    | 86%                |
| <b>risultati complessivi</b> | <b>898</b>               | <b>110</b>                | <b>1008</b>           | <b>78%</b>         |

### 3.3.2.3 Risultati relativi alle emozioni del secondo gruppo

Tale tabella va letta nel seguente modo:

- emozioni nelle ordinate confuse con emozioni nelle ascisse nel X% dei casi
- quando l'emozione dei due assi è la stessa il valore nella casella indica l'accuratezza di riconoscimento di quell'emozione per tutti i soggetti del gruppo

| <b>EMOZIONE ELICITATA</b> | <b>RIQUADRO 6</b> | <b>EMOZIONE PERCEPITA</b> |                 |              |              |               |                 |                  |
|---------------------------|-------------------|---------------------------|-----------------|--------------|--------------|---------------|-----------------|------------------|
|                           |                   | <b>Neutro</b>             | <b>Disgusto</b> | <b>Gioia</b> | <b>Paura</b> | <b>Rabbia</b> | <b>Sorpresa</b> | <b>Tristezza</b> |
|                           | <b>Neutro</b>     | <b>90%</b>                | 0%              | 0%           | 0%           | 7%            | 0%              | 3%               |
|                           | <b>Disgusto</b>   | 3%                        | <b>60%</b>      | 1%           | 5,5%         | 14%           | 11%             | 5,5%             |
|                           | <b>Gioia</b>      | 3%                        | 5,5%            | <b>67%</b>   | 8%           | 5,5%          | 5,5%            | 5,5%             |
|                           | <b>Paura</b>      | 3%                        | 4%              | 4%           | <b>74%</b>   | 1%            | 1%              | 13%              |
|                           | <b>Rabbia</b>     | 1,3%                      | 0%              | 1,3%         | 3%           | <b>90%</b>    | 3%              | 1,3%             |
|                           | <b>Sorpresa</b>   | 1%                        | 3%              | 6%           | 1%           | 1%            | <b>78%</b>      | 10%              |
|                           | <b>Tristezza</b>  | 3%                        | 4%              | 0%           | 4%           | 0%            | 0%              | <b>89%</b>       |

### 3.3.3 Risultati relativi a tutto il campione

#### 3.3.3.1 Risultati relativi agli attori

| <b>RIQUADRO 7</b>              |                   |                    |                |                           |
|--------------------------------|-------------------|--------------------|----------------|---------------------------|
| <b>BATTERIA RELATIVA A M1</b>  |                   |                    |                |                           |
| soggetti ascoltatori           | risposte corrette | Risposte sbagliate | numero domande | accuratezza <sup>12</sup> |
| 1,2,3,4 <sup>13</sup>          | 305               | 31                 | 336            | <b>82%</b>                |
| <b>BATTERIA RELATIVA A F1</b>  |                   |                    |                |                           |
| soggetti ascoltatori           |                   |                    |                |                           |
| 1,2,3,4                        | 298               | 38                 | 336            | <b>77%</b>                |
| <b>BATTERIA RELATIVA A M2</b>  |                   |                    |                |                           |
| soggetti ascoltatori           |                   |                    |                |                           |
| 5,6,7,8                        | 300               | 36                 | 336            | <b>79%</b>                |
| <b>BATTERIA RELATIVA A F2</b>  |                   |                    |                |                           |
| soggetti ascoltatori           |                   |                    |                |                           |
| 5,6,7,8                        | 309               | 27                 | 336            | <b>84%</b>                |
| <b>BATTERIA RELATIVA A M3</b>  |                   |                    |                |                           |
| soggetti ascoltatori           |                   |                    |                |                           |
| 9,10,11,12                     | 298               | 38                 | 336            | <b>77%</b>                |
| <b>BATTERIA RELATIVA A F3</b>  |                   |                    |                |                           |
| soggetti ascoltatori           |                   |                    |                |                           |
| 9,10,11,12                     | 309               | 27                 | 336            | <b>84%</b>                |
| <b>ACCURATEZZA RELATIVA A:</b> |                   |                    |                |                           |
| ATTRICI                        | risposte corrette | risposte sbagliate | numero domande |                           |
| <b>82%</b>                     | 916               | 92                 | 1008           |                           |
| ATTORI                         |                   |                    |                |                           |
| <b>79%</b>                     | 903               | 105                | 1008           |                           |
| <b>COMPLESSIVA</b>             |                   |                    |                |                           |
| <b>80%</b>                     | 1819              | 197                | 2016           |                           |

<sup>12</sup> Relativa all'attore/attrice per tutti i soggetti di entrambi i campioni

<sup>13</sup> Qui ovviamente ci si riferisce, come nelle celle sottostanti, ai soggetti 1,2,3,4 di entrambi i campioni

### 3.3.3.2 Risultati relativi ai soggetti ascoltatori

Come si evince dal riquadro precedente, il campione complessivo dei 24 soggetti ascoltatori ha dimostrato un'accuratezza dell'**80%**.

### 3.3.3.3 Risultati relativi alle emozioni

Tale tabella va letta nel seguente modo:

- emozioni nelle ordinate confuse con emozioni nelle ascisse nel X% dei casi
- quando l'emozione dei due assi è la stessa il valore nella casella indica l'accuratezza di riconoscimento di quell'emozione per tutti i soggetti del campione

| EMOZIONE ELICITATA | RIQUADRO 8 | EMOZIONE PERCEPITA |          |       |       |        |          |           |
|--------------------|------------|--------------------|----------|-------|-------|--------|----------|-----------|
|                    |            | Neutro             | Disgusto | Gioia | Paura | Rabbia | Sorpresa | Tristezza |
|                    | Neutro     | 93%                | 1%       | 0%    | 0%    | 4%     | 0%       | 2%        |
|                    | Disgusto   | 3%                 | 67%      | 2%    | 6%    | 10%    | 6%       | 6%        |
|                    | Gioia      | 2%                 | 4,3%     | 65%   | 7%    | 7%     | 10,3%    | 4,3%      |
|                    | Paura      | 2%                 | 7%       | 2%    | 74%   | 3%     | 3%       | 9%        |
|                    | Rabbia     | 1%                 | 1%       | 1%    | 3%    | 92%    | 1%       | 1%        |
|                    | Sorpresa   | 1%                 | 3%       | 4%    | 1%    | 1%     | 81%      | 9%        |
|                    | Tristezza  | 2%                 | 2%       | 1%    | 3%    | 0%     | 0%       | 92%       |

## 4. CONCLUSIONI

### 4.1 Discussione dei risultati del test

Si comincerà col prendere in considerazione i dati relativi alla riconoscibilità delle *performance* degli attori: per il primo gruppo l'accuratezza complessiva è dell'**83%**, per il secondo del **78%**, per ambedue dell'**80%**. Entrambi i gruppi di ascoltatori hanno riconosciuto meglio le attrici (**82%**) che gli attori (**79%**): per un così lieve scarto, e tenuto conto dell'esiguità del campione di ascolto, sarebbe azzardato fare generalizzazioni sulla migliore capacità delle attrici di esprimere emozioni.<sup>14</sup> Tenuto invece presente che l'obiettivo del test è precipuamente quello di garantire la bontà delle *performance* degli attori coinvolti nella creazione del database EMOVO, ci si può ritenere più che soddisfatti: un'accuratezza di riconoscimento dell'**80%** è di molto al di sopra dei valori medi desunti dalla letteratura sul riconoscimento delle emozioni, come riportati in § 3.1, che si collocano intorno al 62%: tuttavia non si dovrebbe dare troppa importanza a questo scarto, dal

<sup>14</sup> Per il Prof. Paolo Di Giovine, su comunicazione personale, la migliore riconoscibilità delle emozioni espresse dalle attrici potrebbe esser dovuta, *ceteris paribus*, al fatto che la voce femminile ricopre un gamma di frequenze maggiore rispetto a quella maschile. Ciò potrebbe senz'altro darsi, considerato il ruolo preminente che i livelli di pitch giocano nell'espressione dell'emozione, ma necessiterebbe di una verifica sperimentale che purtroppo esula dal presente lavoro.

momento che quest'ultimo valore si riferisce a test di riconoscimento senza dubbio più difficili del test sopra presentato, con risposte aperte o con molte più alternative di scelta.

Prendiamo ora invece in considerazione le *performance* dei singoli attori, partendo dai risultati di tutto il campione di ascolto: come è visibile dal riquadro 7, gli attori più riconoscibili risultano essere stati F2 e F3 (**84%**), a seguire M1 (**82%**), M2 (**79%**), e infine F1 e M3 (**77%**). Quel che è più importante rilevare è però il fatto che fra gli attori/attrici non ci sono scarti significativi che possano pregiudicare l'impiego, per eventuali analisi, del materiale elicitato da un attore piuttosto che da un altro: la riconoscibilità è sufficientemente garantita per tutti gli attori e le attrici.

Un'osservazione analoga va fatta per l'abilità dei soggetti ascoltatori: a parte il soggetto 6 del secondo gruppo di ascolto, non si sono rilevate significative devianze dall'indice di accuratezza dell'80%, relativo alla riconoscibilità delle emozioni da parte di tutti i soggetti. Un'altra differenza è quella relativa alle percentuali di riconoscimento, nel primo gruppo di ascolto, dei soggetti di sesso femminile rispetto a quelli di sesso maschile (cf. riquadro 2): uno scarto così lieve su un campione così esiguo può bastare, se non altro, a farci riflettere in rapporto ai casi in cui si è registrata una maggiore abilità di riconoscimento da parte delle donne (Dusenberry – Knower 1939), e, per converso, ai casi in cui si è registrata una sostanziale omogeneità fra i sessi nelle abilità riconositive (Hall 1978, Van Bezooijen 1984, Anolli – Ciceri 1992).

Restano da considerare ora i risultati pertinenti alle emozioni. Come si può vedere dal riquadro 8, nella scala di riconoscibilità delle emozioni c'è una sostanziale concordanza degli indici di riconoscibilità del campione del nostro test con gli indici di riconoscibilità delle emozioni desunti dalla letteratura (cf. § 3.1)<sup>15</sup>:

|                                   |              |                 |                 |             |                |                |
|-----------------------------------|--------------|-----------------|-----------------|-------------|----------------|----------------|
| <b>Emozioni nel nostro test</b>   | Neutro (93%) | Tristezza (92%) | Rabbia (91%)    | Paura (73%) | Disgusto (67%) | Gioia (65%)    |
| <b>Emozioni nella letteratura</b> | Rabbia (77%) | Neutro (74%)    | Tristezza (71%) | Paura (61%) | Gioia (57%)    | Disgusto (31%) |

Se non si dà troppa importanza ai singoli livelli, che differiscono fra loro solo per pochi punti percentuale, si nota come la concordanza sia molto significativa: il fattore di correlazione fra le due serie di dati è di **0.87**. Per quanto riguarda, invece, il livello inferiore dei valori della letteratura rispetto ai nostri, ciò è del tutto normale. Ricordiamo quanto affermato poc'anzi: essendo il fine del test la validazione della qualità del materiale registrato del database EMOVO, e non la valutazione della capacità umana di riconoscere le emozioni, si è progettato il test in modo meno complesso che nel secondo caso, con un conseguente livello di difficoltà minore.

Ma quel che è più importante notare è come siano nettamente più riconoscibili la rabbia e la tristezza rispetto alla gioia e al disgusto, la paura avendo una riconoscibilità media. Ciò concorda perfettamente con quanto affermano Scherer – Johnstone – Klasmeyer 2003, p. 444, in una loro sintesi teorica sul parlato emotivo: “Sadness and anger are generally best

<sup>15</sup> La sorpresa, di cui non dispongo una media degli indici di riconoscibilità degli studi in cui è stata presa in considerazione, è risultata riconoscibile a un livello medio-alto (81%), fra la paura e la rabbia/tristezza

recognized vocally, followed by fear. Joy has rather mixed accuracy percentages in different studies, possibly due to differences with respect to quiet happiness versus elated joy being portrayed by the actors.” Senz’altro è quel che è avvenuto nel nostro caso: in italiano il termine “gioia” si presta molto di più ad essere inteso come contentezza esaltata che come serena contentezza: sono memore del fatto che a tre attori su sei, di quelli da me convocati, la prima idea che è venuta in mente per autoindursi lo stato di gioia è stato “l’aver vinto la lotteria” (di fronte a tale associazione io non mi espressi per evitare di condizionare la *performance*). Il risultato è quindi stato una gioia presumibilmente più riconoscibile rispetto alla gioia nelle sue rimanenti accezioni. Un discorso analogo va fatto per il disgusto, il cui basso indice di riconoscibilità desunto dalla letteratura è dovuto indubbiamente all’ambiguità semantica su cui mi sono soffermato nella nota 3, per cui nei vari studi sono stati persino rilevati due profili acustici del disgusto molto diversi, l’uno con una media del pitch maggiore al parlato non emotivo, l’altro con una media del pitch minore. Essendomi invece premurato di chiarire bene agli attori che si richiedeva il disgusto “fisico” piuttosto che quello “morale”, ne è conseguita una resa senza dubbio più chiaramente riconoscibile.

Infine, qualche considerazione sulle matrici di confusione. Come si vede dal riquadro 8, le confusioni fra due emozioni che si sono verificate maggiormente (stabilendo una soglia d’interesse del 5% dei casi) sono:

- 10,3 % gioia confusa con sorpresa
- 10% disgusto con rabbia
- 9% sorpresa con tristezza
- 9% paura con tristezza
- 9% gioia con paura
- 7% gioia con rabbia
- 7% paura con disgusto
- 6% disgusto con paura
- 6% disgusto con sorpresa
- 6% disgusto con tristezza

Queste confusioni sono tutte asimmetriche, a parte quella fra disgusto e paura: una confusione simmetrica fra due emozioni attesta che esse sono molto più vicine fra loro che non emozioni confuse asimmetricamente, e, in effetti, non è difficile scorgere una forte analogia fra il disgusto e la paura, emozioni volte entrambe all’avversare e/o evitare lo stimolo che ha generato l’emozione.

Per quanto riguarda le altre confusioni, nella letteratura sono state date due distinte interpretazioni che ne rendano conto: per Van Bezooijen 1984 le emozioni si confondono in base al livello di attivazione, per cui le emozioni ad alto livello di attivazione si confondono con altre emozioni ad alto livello, e così anche per le emozioni a basso livello. Per Dawes – Kramer 1966 invece le confusioni dovrebbero imputarsi alle contiguità semantiche fra le emozioni, ad es. la rabbia col disprezzo, o, come scritto poc’anzi, la paura col disgusto. Per le confusioni che sono risultate dal nostro test, la prima spiegazione potrebbe rendere conto della gioia confusa con la paura, e di disgusto confuso con sorpresa; la seconda spiegazione di sorpresa con tristezza, paura con tristezza, disgusto con tristezza (se accettiamo le



contiguità semantiche del cerchio di Ekman<sup>16</sup>). Gli altri casi potrebbero invece essere spiegati in entrambe le direzioni.

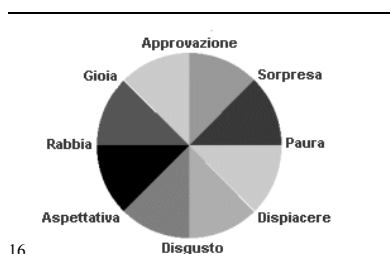
Un'altra questione, alquanto negletta nella letteratura, è quella relativa alla capacità di un'emozione non di essere confusa, ma di confondere. Per renderci conto di questo fenomeno, riporto qui una tabella col conteggio di tutte le volte che uno stato emotivo è stato erroneamente attribuito allo stato emotivo ascoltato, dato non direttamente evincibile dal riquadro 8 (dove, colonna per colonna, è invece possibile vedere le percentuali dei casi di erronea attribuzione di un'emozione sbagliata a quella ascoltata):

| neutro | disgusto | gioia | Paura | rabbia | Sorpresa | tristezza |
|--------|----------|-------|-------|--------|----------|-----------|
| 16     | 26       | 14    | 28    | 35     | 31       | 47        |

E' molto interessante notare come l'emozione che è stata attribuita erroneamente più volte, la tristezza, è stata anche una delle emozioni più riconoscibili (riconosciuta con un'accuratezza del 92%! ). Non stupisce invece che lo stato emotivo neutro, riconoscibile con un'accuratezza del 93%, sia stato attribuito erroneamente molte volte di meno. Caso diametralmente analogo alla tristezza è invece quello della gioia, emozione fra le più difficilmente riconoscibili (accuratezza del 65%), eppure attribuita erroneamente pochissime volte. Tali correlazioni sembrano avallare il pensiero di Leopardi: è proprio come se l'orecchio umano non sapesse bene cos'è la gioia, anche se capirebbe bene cosa non lo è. Parimenti, saprebbe bene cosa è la tristezza, eppure, non riuscirebbe a capire cosa non lo è.

Detto questo, non posso spingermi in ulteriori speculazioni sulle confusioni del nostro test, dal momento che questo non è stato progettato per tal fine (bisognerebbe disporre di un campione più ampio, e con domande diversamente impostate) e dal momento che non sono state fatte analisi elettroacustiche delle frasi, indispensabili, fra le altre cose, per stabilire con certezza il livello di attivazione delle ricorrenze delle emozioni (in particolar modo quelle relative alla sorpresa e al disgusto): accertato il livello di attivazione e i valori degli altri parametri acustici, si potrà avere una migliore comprensione dei motivi delle confusione delle emozioni, individuando, ad esempio, valori simili di parametri che soggiacciono alle emozioni confuse fra di loro nei test di riconoscimento.

Fatte queste precisazioni, possiamo ritenere le *performance* degli attori, il riconoscimento dei soggetti e la resa delle emozioni pienamente soddisfacenti: la validazione del database di parlato emotivo EMOVO ha avuto esito positivo.



16

#### 4.2 Possibili impieghi di EMOVO

Il database EMOVO può prestarsi ai seguenti impieghi:

- **Analisi spettroacustiche:** assodato, grazie al test di validazione, che le emozioni che abbiamo registrato sono soddisfacentemente rappresentative delle emozioni percepite, il materiale del *corpus* può essere sottoposto a numerose ulteriori misurazioni, come analisi macroprosodiche di pitch, intensità e durata, analisi microprosodiche di detti indici e delle formanti vocaliche, analisi relative alla *voice quality*, analisi combinate, in vario modo, di tutti o alcuni di questi parametri. Ma più che gli indici macroprosodici, che presumibilmente confermeranno i risultati degli altri esperimenti, sarà interessante misurare gli andamenti del profilo intonativo in base alle diverse strutture semantiche e morfosintattiche offerte dal database, territorio ancora manchevole di generalizzazioni, soprattutto per l'italiano. In particolar modo, sarà interessante osservare il co-variare del contorno del pitch in relazione da una parte con la tonia conclusiva, sospensiva, interrogativa o divisiva prevista dalla struttura delle frasi del database, dall'altra con il variare dell'emozione. Infine, le analisi dei profili acustici delle emozioni del database EMOVO potrebbero servire da modello per la messa a punto di regole prosodiche per la sintesi del parlato emotivo italiano, regole che finora sono state inferite a partire da materiale relativo ad altre lingue o, per l'italiano, da analisi su *dataset* quantitativamente limitati.
- **Analisi di decoding:** sul materiale registrato possono effettuarsi numerosi test di riconoscimento, con finalità e procedure diverse rispetto al test presentato nel terzo capitolo: *in primis* test di discriminazione di emozioni con campioni più ampi e risposte con più alternative, al fine di corroborare e correggere le osservazioni fatte nello scorso paragrafo. Altri test potrebbero invece usare tutte le frasi offerte dal database, al fine di capire quali tipologie fraseologiche inibiscano o aiutino il riconoscimento di ciascuna emozione. Inoltre, tramite l'impiego di appositi programmi, come Melodyne, il materiale registrato potrebbe essere manipolato o "mascherato" in base ai più svariati parametri, dalla velocità d'eloquio al profilo intonativo, e si potrebbe studiare, tramite appositi test di riconoscimento, come queste modificazioni influiscano sulla riconoscibilità delle emozioni, e in che misura. Infine, così come le analisi sul materiale di EMOVO potrebbero fornire profili acustici per la sintesi del parlato emotivo, allo stesso modo i test di riconoscimento sul nostro materiale potrebbero fornire indispensabili modelli di riferimento per la messa a punto di sistemi di ASR (automated speech recognition) del parlato emotivo.
- **Analisi interculturali:** i dati estrapolati da esperimenti su EMOVO potrebbero essere messi a confronto con quelli relativi alle emozioni nelle altre lingue, per gettare luce su questioni cruciali circa gli aspetti culturali vs. universali della codifica e decodifica delle emozioni.
- **Comparazioni fra materiale recitato e materiale dal vivo:** come ricordato nel § 2.1, se è auspicabile lo studio delle emozioni con materiale genuino non recitato, tuttavia in fase preliminare è necessario l'impiego di materiale elicitato in laboratorio. Quest'ultimo infatti consente di eliminare eccessive variabili di ricerca capaci di soffocare i primi avanzamenti della conoscenza in questo campo. In molti casi, in una frase registrata dal vivo non ci si potrà mai concentrare in modo ottimale sulla misurazione di un parametro, perché difficilmente si potranno scevrare le misurazioni del dato parametro dalle influenze e dalle covarianze che

con esso hanno i numerosissimi altri parametri che la “realtà” offre, molti dei quali difficilmente misurabili. E comunque sia, potrebbe darsi che lo scarto fra materiale recitato e materiale naturale sia veramente minimo: per stabilire ciò, sarebbero utili test progettati esclusivamente per mettere alla prova le due tipologie di materiale in modo sistematico: “this would seem to be one of the highest priorities in the field”, come afferma Klaus Scherer<sup>17</sup>

## RINGRAZIAMENTI

Un sentito ringraziamento va al Prof. Paolo Di Giovine dell’Università di Roma La Sapienza e all’Ing. Andrea Paoloni della Fondazione Ugo Bordoni di Roma, senza il supporto dei quali la realizzazione di EMOVO non sarebbe stata possibile. Si ringraziano inoltre il Prof. Luciano Romito e il Dott. Vincenzo Galatà per la loro collaborazione alla validazione del database.

## BIBLIOGRAFIA

- AA. VV. (2008), *Comunicazione parlata e manifestazione delle emozioni, atti del I convegno GSCP, Padova, 29 novembre-1 dicembre 2004* (a cura di Emanuela Magno Caldognetto, Federica Cavicchio e Piero Così), Liguori, disponibile in e-book.
- Alpert M., Kurtzberg R. L. & Friedhoff A. J. (1963), *Transient voice changes associated with emotional stimuli*, in “Archives of general psychiatry” 8, pp. 362-365,
- Anolli L. & Ciceri R. (1992), *La voce delle emozioni*, Franco Angeli, Milano.
- Bachorowski J. A. & Owren M. J. (1995), *Vocal expression of emotion: acoustic properties of speech are associated with emotional intensity and context*, in “Psychological science” 6, pp. 219-224.
- Banse R. & Scherer K.R (1996), *Acoustic profiles in vocal emotion expression*, in “Journal of personality and social psychology” 70, pp. 614-636.
- Bortz J. (1966), *Physikalisch-akustische Korrelate der vokalen Kommunikation*, in “Arbeiten aus dem Psychologischen Institut der Universität Hamburg” 9.
- Brown B. L. (1980), *The detection of emotion in voice qualities*, in Giles H., Robinson W. P. & Smith P. M. (a cura di), *Language: social perspectives*, Pergamon press, Oxford.
- Burkhardt F., Paeschke A., Rolfes M., Sendlmeier W. & Weiss B. (2005), *A database of german emotional speech*, in Interspeech 2005, Lisbona, pp. 1517- 1520, pp. 187-200.
- Canepari L. (1985), *L’intonazione*, Liguori, Napoli.
- Cauldwell R. (2000), *Where did the anger go? The role of context in interpreting emotion in speech*, in *ISCA Workshop on speech and emotion*, settembre 2000, Belfast, pp. 127-131.
- Coleman R. F. & Williams R. (1979), *Identification of emotional states using perceptual and acoustic analyses*, in Lawrence V. & Weinberg B. (a cura di), *Transcript of the eighth symposium: care of the professional voice (part I)*, The voice foundation, New York.
- Cresti E. & Moneglia M. (2005) (a cura di), *C-ORAL-ROM Integrated reference corpora for spoken romance languages*, Benjamins, Amsterdam-Philadelphia.

---

<sup>17</sup> Scherer 2003, p. 247.

- Davitz J. R. (1964), *Auditory correlates of vocal expressions of emotional feeling*, in Davitz J. R. (a cura di), *The communication of emotional meaning*, McGraw-Hill, New York.
- Davitz J. R. & Davitz L. J. (1959), *The communication of feelings by content-free speech*, in "Journal of communication" 9, pp. 6-13.
- Dawes R. M. & Kramer E. (1966), *A proximity analysis of vocally expressed emotion*, in "Perceptual and motor skills" 22, pp. 571-574.
- De Dominicis A. (1992), *Intonazione e contesto. Uno studio su alcuni aspetti del discorso in contesto e delle sue manifestazioni intonative*, Edizioni dell'Orso, Alessandria.
- Dimitrovsky L. (1964), *The ability to identify the emotional meaning of vocal expressions at successive age levels*, in Davitz J. R. (a cura di), *The communication of emotional meaning*, McGraw-Hill, New York.
- D'Urso V., Cavicchio F. & Magno Caldognetto E. (2008), *Le etichette lessicali nelle ricerche sperimentali sulle emozioni: problemi teorici e metodologici*, In Magno Caldognetto E., Cavicchio F. & Così P. (a cura di), *Comunicazione parlata e manifestazione delle emozioni, atti del I convegno GSCP, Padova, 29 novembre-1 dicembre 2004*, Liguori, disponibile in e-book.
- Douglas-Cowie E., Campbell N., Cowie R. & Roach P. (2003), *Emotional speech: towards a new generation of databases*, in "Speech communication" 40, pp. 33-60.
- Dusenberry D. & Knower F. H. (1939), *Experimental studies of the symbolism of action and voice II- a study of the specificity of meaning in abstract tonal symbols*, in "Quarterly journal of speech" 25, pp. 67-75.
- Ekman P. (1994), *Strong evidence for universals in facial expression: a reply to Russel's mistaken critique*, in "Psychological bulletin" 115, pp. 268-287.
- Eldred S. H. & Price D. B. (1958), *A linguistic evaluation of feeling states in psychotherapy*, in "Psychiatry" 21, pp. 115-121.
- Fairbanks G. & Hoaglin L. W. (1941), *An experimental study of the durational characteristics of the voice during the expression of emotion*, in "Science" 221, pp. 1208-1210.
- Fairbanks G. & Pronovost W. (1939), *An experimental study of the pitch characteristics of the voice during the expression of emotion*, in "Speech monographs" 8, pp. 85-90.
- Fónagy I. (1978), *A new method of investigating the perception of prosodic features*, in "Language and speech" 21, pp. 34-49.
- Fónagy I. (1983), *Emotions, voice and music*, in Sundberg J. (a cura di), *Research aspects on singing*, Payot, Parigi.
- Fónagy I. (1987), *Semantic diversity in intonation*, in Proceedings of International congress of phonetic sciences 1987, Tallinn, pp. 468-471.
- Fónagy I. (1989), *La vive voix*, Payot, Parigi.
- Fónagy I. & Magdics K. (1963), *Emotional patterns in intonation and music*, in "Zeischrift für phonetische. Sprachwissenschaft und Kommunikationforschung" 16, pp. 293-326.

- Green R. S. & Cliff N. (1975), *Multidimensional comparison of structures of vocally and facially expressed emotion*, in "Journal of perceptive psychophysiology" 17, pp. 429-438.
- Hargreaves W. A., Starkweather J. A., & Blacker K. H. (1965), *Voice quality in depression*, in "Journal of abnormal psychology" 70, pp. 218-220.
- Havrdova Z. & Moravek M. (1979), *Changes of the voice expression during suggestively influenced states of experiencing*, in "Activitas nervosa superior" 21, pp. 33-35.
- Helfrich H., Standke R. & Scherer K. R. (1984), *Vocal indicators of psychoactive drug effects*, in "Speech communication" 3, pp. 245-252.
- Johnstone T. & Scherer K. R. (2000), *Vocal communication of emotion*, in Lewis M. – Haviland J. (a cura di), *Appraisal processes in emotion: theory, methods, research*, Oxford university press, New York, pp. 271-284.
- Kienast M. & Sendlmeier W. F. (2000), *Acoustical analysis of spectral and temporal changes in emotional speech*, in Sendlmeier W. (a cura di), *Speech and signals*, Frankfurt am M., pp. 157-168.
- Kienast M., Paeschke A. & Sendlmeier W. F. (1999), *Articulatory reduction in emotional speech*, in Eurospeech 1999, Budapest, pp. 117-120.
- Klasmeyer G. (1999), *Akustische Korrelate des stimmlich emotionalen Ausdrucks in der Lautsprache*, in Wodarz H. W., Heike G., Janota P. & Mangold M. (a cura di), "Forum phoneticum" 67, pp. 1-238.
- Klasmeyer G. & Meier T. (1999), *Rhythm in Emotional Speech*, in DAGA-Tag der ASA-Tagung 99 in Berlin, CD-Rom zur ASA-EAA-DEGA Gemeinschaftstagung Berlin 99.
- Klasmeyer G. & Sendlmeier W.F. (1997), *The classification of different phonation types in emotional and neutral speech*, in "Forensic Linguistics" 4, pp. 104-124.
- Klasmeyer G. & Sendlmeier W.F. (1999), *Voice and emotional states*, in Kent R. - Ball M. (a cura di), *Voice quality measurement*, Singular publishing group, San Diego, pp. 339-359.
- Kori S. & Magno Caldognetto E. (1986), *Caratterizzazioni elettroacustiche delle emozioni in italiano*, in "Rivista italiana di acustica" 10, pp.33-39.
- Kori S. & Magno Caldognetto E. (1990), *Indici acustici e percettivi delle emozioni*, in *Scritti in onore di Lucio Croatto*, Microprint, Padova.
- Kori S. & Magno Caldognetto E. (2003), *La caratterizzazione fonetica delle emozioni: primi dati da uno studio cross-linguistico italiano-giapponese*, in Così P., Magno Caldognetto E. & Zamboni A. (a cura di), *Voce-canto-parlato. Studi in onore di Franco Ferrero*, Unipress, Padova, pp. 187-200.
- Kotlyar G.M. & Morozov V.P. (1976), *Acoustical correlates of the emotional content of vocalized speech*, in "Journal of acoustics of the academy sciences of the USSR" 22, pp. 208-211.
- Kuroda I., Fujiwara O., Okamura N. & Utsuki N. (1976), *Method for determinino pilot stress through analysis of voice communication*, in "Aviation, space, and environmental medicine" 47, pp. 528-533.

- Ladd D. R., Scherer K. R. & Silverman K. E. A. (1986), *An integrated approach to studying intonation and attitude*, in Johns-Lewis C. (a cura di), *Intonation in discourse*, College-Hill Press, San Diego, pp. 125-138.
- Ladd D. R., Silverman K. E. A., Tolkmitt F., Bergman G. & Scherer K. R. (1985), *Evidence for the independent function of intonation contour type, voice quality and f0 range in signaling speaker affect*, in "Journal of the acoustical society of America" 78, pp. 435-444.
- Levin H. & Lord W. (1975), *Speech pitch frequency as an emotional state indicator*, in "Institute of electrical and electronics engineers transactions on system, man and cybernetics" 5, pp. 259-273.
- Magno Caldognetto E. & Ferrero F. E. (1996), *Macro e micro variazioni prosodiche dipendenti dalle scelte paralinguistiche del parlante*, in Fedi F. & Paoloni A. (a cura di), *Atti delle VI giornate di studio del GFS, "Caratterizzazione del parlatore"*, Roma 23-24 novembre 1995, pp. 95-107.
- Markel N. N., Bein M. F., & Phillis J. A. (1973), *The relationship between words and tone-of-voice*, in "Language and speech" 16, pp. 15-21.
- Moneglia M. (2008), *Le unità di riferimento per l'espressione delle emozioni: dati dal corpus C-ORAL-ROM*, in Magno Caldognetto E., Cavicchio F. & Cosi P. (a cura di), *Comunicazione parlata e manifestazione delle emozioni, atti del I convegno GSCP, Padova, 29 novembre-1 dicembre 2004*, Liguori, disponibile in e-book.
- Niwa S. (1971), *Changes of voice characteristics in urgent situations*, in "Reports of the aeromedical laboratory, Japan air self-defence force" 11, pp. 246-251.
- Paeschke A., Kienast M. & Sandlmeier W. F. (1999), *f0 contours in emotional speech*, in *Proceedings of International congress of phonetic sciences 1999*, San Francisco, pp. 929-933.
- Paeschke A. & Sendlmeier W. F. (2000), *Prosodic characteristics of emotional speech: Measurements of fundamental frequency movements*, in *SpeechEmotion 2000*, pp. 75-80.
- Plaikner D. (1970), *Die Veränderungen der menschlichen Stimme unter dem Einfluss psychischer Belastung*, tesi di dottorato, Innsbruck Universität.
- Poggi I. & Magno Caldognetto E. (2003), *Il parlato emotivo. Aspetti cognitivi, linguistici e fonetici*, in *Atti del Convegno Il parlato italiano*, D'Auria, Napoli, cd-rom.
- Polzin T. S. & Waibel A. (2000), *Emotion-sensitive human-computer interfaces*, in *ISCA workshop on speech and emotion*, Settembre 2000, Belfast, pp. 201-206.
- Roessler R. & Lester J. W. (1976), *Voice predicts affect during psychotherapy*, in "Journal of nervous and mental disease" 163, pp. 166-176.
- Scherer K. R. (1977), *The effect of stress on the fundamental frequency of the voice*, in "Journal of acoustical society of America" 62, pp. 25-26.
- Scherer K. R. (1979), *Nonlinguistic vocal indicators of emotion and psychopathology*, in Izard C. E. (a cura di), *Emotion and personality in psychopathology*, Plenum Press, New York.

- Scherer K. R. (1984), *On the nature and function of emotion: a component process approach*, in Scherer K.R. & Ekman P. (a cura di), *Approaches to emotion*, Erlbaum, Hillsdale.
- Scherer K. R., Koivumaki J. & Rosenthal R. (1972a), *Minimal cues in the vocal communication of affect: judging emotions from content-masked speech*, in "Journal of psycholinguistic research" 1, pp. 269-285.
- Scherer K. R., Koivumaki J. & Rosenthal R. (1972b), *Mediating interpersonal expectancies via vocal cues: different speech intensity as a means of social influence*, in "European journal of social psychology" 2, pp. 163-176.
- Scherer K. R., London H. & Wolf J. (1973), *The voice of confidence: paralinguistic cues and audience evaluation*, in "Journal of research in personality" 7, pp. 31-44.
- Scherer K. R., Johnstone T. & Klasmeyer G. (2003), *Vocal expression of emotion*, in Davidson D. J., Goldsmith H. & Scherer K. R. (a cura di), *Handbook of the Affective Sciences*, Oxford university press, New York and Oxford, pp. 433-456.
- Scherer K. R., Banse R., Wallbott H. G. & Goldbeck T. (1991), *Vocal cues in emotion encoding and decoding*, in "Motivation and emotion" 15, pp. 123-148.
- Sedlacek K. & Sychra A. (1963), *Die Melodie als Faktor des emotionellen Ausdrucks*, in "Folia phoniatica" 15, pp. 89-98.
- Skinner E. R. (1935), *A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness*, in "Speech monographs" 2, pp. 81-137.
- Sobin C. & Alpert M. (1999), *Emotion in speech: the acoustic attributes of fear, anger, sadness, and joy*, in "Journal of psycholinguistic research" 28, pp. 347-365.
- Sokolov J. N. (1963), *Perception and the conditioned reflex*, Pergamon press, Oxford.
- Sulc J. (1977), *To the problem of emotional changes in the human voice*, in "Activitas nervosa superior" 19, pp. 215-216.
- Van Bezooijen R. (1984), *The characteristics and recognizability of vocal expression of emotion*, Foris, Dordrecht.
- Wallbott H. G. & Scherer K. R. (1986), *Cues and channels in emotion recognition*, in "Journal of personality and social psychology" 51, pp. 690-699.
- Williams C. E. & Stevens K. N. (1972), *Emotion and speech: some acoustical correlates*, in "Journal of the acoustical society of America" 52, pp. 1238-1250.
- Zuberbier E. (1957), *Zur Schreib- und Sprechmotorik der Depressiven*, in "Zeitschrift für Psychotherapie und medizinische Psychologie" 7, pp. 239-249.
- Zwirner E. (1930), *Beitrag zur Sprache des Depressiven*, in "Phonometrie" 3, pp. 171-187.