

---

UNIVERSITÀ  
DELLA CALABRIA



Università della Calabria

Dipartimento di Scienze Politiche e Sociali

Dipartimento di Lingue e Scienze dell'Educazione

**Corso di Dottorato di Ricerca in  
"Politica, Cultura e Sviluppo"**

**XIX ciclo**

TESI DI DOTTORATO

*La comunicazione delle emozioni in una lingua seconda:  
un'indagine acustico-uditiva*

Settore Scientifico Disciplinare L-LIN/01

---

**Coordinatore**

Ch.mo Prof. Alberto Ventura

**Dottoranda**

dott. Emanuela Paone

**Supervisore  
scientifico**

Ch.mo Prof. Luciano Romito

# Indice

<b>Indice delle figure .....</b>	<b>6</b>
<b>Indice delle tabelle .....</b>	<b>9</b>
<b>Abstract .....</b>	<b>6</b>
<b>Introduzione .....</b>	<b>9</b>

## **Capitolo 1- Il fenomeno emotivo e le sue componenti**

<b>1.1. Introduzione .....</b>	<b>14</b>
<b>1.2 La componente biologica.....</b>	<b>17</b>
1.2.1 Emozioni come meccanismi di risposta .....	17
1.2.2 Le emozioni primarie.....	20
<b>1.3 La componente neuropsicologica .....</b>	<b>22</b>
1.3.1 Il fulcro dell'esperienza emotiva: il cervello.....	22
1.3.2 Il sistema limbico.....	24
<b>1.4 La componente cognitiva .....</b>	<b>24</b>
1.4.1 Emozioni come valutazioni .....	26
1.4.2 Emozioni come scopi.....	27
1.4.3 Il modello componenziale delle emozioni .....	28
1.4.4 Teorie dimensionali .....	29
<b>1.5 La componente culturale.....</b>	<b>30</b>
1.5.1 Una doppia eredità .....	31
1.5.2 L'acquisizione del linguaggio .....	33
1.5.3 I modelli culturali .....	34
1.5.4 Categorizzazione delle emozioni e lessico emotivo .....	35

1.5.5 Il ruolo della cultura nella manifestazione delle emozioni .....	38
1.5.6 Stili comunicativi.....	39
<b>1.6 La componente espressiva.....</b>	<b>41</b>
1.6.1 L'espressione facciale delle emozioni .....	41
1.6.2 I gesti .....	42
1.6.3 La relazione tra emozione e voce .....	43
1.6.4 Le informazioni trasmesse dalla voce.....	45
1.6.5 La comunicazione emotiva .....	47
<b>1.7 Conclusioni .....</b>	<b>49</b>

## **Capitolo 2- Il parlato emotivo**

<b>2.1 Introduzione .....</b>	<b>51</b>
<b>2.2. Per una definizione dei segni vocali non verbali.....</b>	<b>52</b>
2.2.1 fenomeni prosodici .....	54
2.2.2 Le funzioni dell'intonazione .....	55
2.2.3 Tre codici biologici.....	57
<b>2.3 Il quadro teorico .....</b>	<b>60</b>
2.3.1 L'interazione tra il piano linguistico e quello paralinguistico.....	60
2.3.1.1 Approcci descrittivi e sperimentali .....	62
2.3.2 Approcci discreti e dimensionali .....	63
2.3.3 Il modello componenziale.....	65
2.3.4 Un modello teorico per la comunicazione vocale delle emozioni .....	66
<b>2.4 Lo scenario metodologico.....</b>	<b>69</b>
2.4.1 Raccolta e selezione del materiale emotivo: alcune problematiche .....	69
a) Produzioni emotive naturali.....	70
b) Induzione sperimentale.....	71
c) Produzioni simulate, scenari emotigeni, etichette lessicali .....	73

2.4.2 Il processo di decodifica .....	76
2.4.2.1 I test percettivi .....	76
<b>2.5 I profili acustici delle emozioni primarie.....</b>	<b>77</b>
<b>2.6 Tendenze comuni nel riconoscimento delle emozioni a livello intralinguistico .....</b>	<b>79</b>
2.6.1 Il ruolo dei tratti paralinguistici nel riconoscimento delle emozioni .....	80
<b>2.7 Il parlato emotivo in una prospettiva transculturale .....</b>	<b>81</b>
2.7.1 Universalità e specificità culturale: verso un approccio interazionista .	81
2.7.2 Stato dell'arte degli studi interculturali sul parlato emotivo .....	84
2.7.2.1 Le ricerche sull'encoding .....	85
2.7.2.2 Le ricerche sul decoding .....	89
<b>2.8 Conclusioni .....</b>	<b>95</b>

## **Capitolo 3-La comunicazione delle emozioni in una lingua non nativa**

<b>3.1 introduzione .....</b>	<b>97</b>
<b>3.2 La dimensione emotiva nel processo di acquisizione di una L2 .....</b>	<b>97</b>
3.2.1 La competenza comunicativa.....	97
3.2.2 Intelligenza emotiva e competenza emotiva in L2 .....	100
3.2.3 La comunicazione emotiva in una lingua non nativa: alcune problematiche .....	103
<b>3.3. L'acquisizione delle strutture prosodiche in una L2 .....</b>	<b>105</b>
3.3.1 La prosodia dell'interlingua .....	105
<b>3.4 Il parlato emotivo in L2: stato dell'arte .....</b>	<b>108</b>
3.4.1 Gli studi in italiano L2 .....	114
<b>3.5 Conclusioni .....</b>	<b>117</b>

## Capitolo 4- Obiettivi e metodi della ricerca

<b>4.1 Oggetto della ricerca.....</b>	<b>119</b>
<b>4.2 Obiettivi della ricerca .....</b>	<b>120</b>
4.2.1 Il processo di encoding .....	121
4.2.2 Il processo di decoding .....	122
<b>4.3 L'impianto metodologico.....</b>	<b>122</b>
4.3.1 La codifica delle emozioni: strumenti e procedure.....	122
4.3.1.1 Partecipanti .....	122
4.3.1.2 Le emozioni oggetto di indagine .....	123
4.3.1.3 I dialoghi induttori .....	124
La struttura dei dialoghi.....	125
4.3.1.4 La frase bersaglio.....	127
4.3.1.5 Strumenti e procedure di registrazione .....	128
4.3.1.6 La selezione degli stimoli .....	130
4.3.1.7 L'analisi acustica: parametri indagati.....	131
4.3.1.8 Le analisi statistiche.....	132
4.3.2 La decodifica delle emozioni: strumenti e procedure.....	133
4.3.2.1 Il test uditivo degli apprendenti russi.....	134
4.3.2.2 Il test uditivo degli italiani .....	137
<b>4.4. Conclusioni .....</b>	<b>138</b>

## Capitolo 5- Analisi e discussione dei dati

<b>5.1 Introduzione .....</b>	<b>139</b>
<b>5.2 <i>Encoding</i>: risultati .....</b>	<b>139</b>
5.2.1 La variazione dei parametri all'interno di ogni corpus .....	139
5.2.2 Le produzioni emotive nell'italiano nativo (Ita_L1).....	141
5.2.3 Le produzioni emotive nell'italiano non nativo (Ita_L2).....	145

5.2.4 Le produzioni emotive in russo L1 (Ru_L1) .....	149
5.2.5 Il confronto tra i gruppi.....	152
5.2.5.1 Ita L1- Ru L1 .....	153
5.2.5.2 Ita L1- Ita L2.....	156
5.2.5.3 Ita L2-Ru L1 .....	162
5.2.6 Discussione dei dati relativi al processo di encoding .....	164
La variazione dei parametri nelle produzioni emotive .....	166
Le produzioni emotive in italiano L2 .....	167
<b>5.3 Risultati relativi al processo di <i>decoding</i> .....</b>	<b>168</b>
5.3.1 Risultati del test uditivo sottoposto agli apprendenti russi (U_Ita_L2) .....	169
5.3.1.1 Analisi dei riconoscimenti corretti.....	169
5.3.1.2 Analisi dei riconoscimenti erronei.....	174
5.3.2 Risultati del test uditivo sottoposto agli italiani nativi .....	175
5.3.2.1. Riconoscimenti corretti in base alla tipologia di enunciato.....	176
5.3.2.2 Riconoscimenti globali in funzione della lingua dei locutori .....	177
5.3.2.3 Riconoscimenti corretti per tipologia di enunciato in Ita L1, Ita L2 e Ru L1 .....	179
a)Ita L1-Ita L2.....	183
b) Ita L1-Ru L1 .....	183
c) Ita L2-Ru L1 .....	184
5.3.2.4 Differenze di riconoscimento in funzione dei locutori russi .....	184
5.3.2.5 Analisi dei riconoscimenti erronei.....	187
5.3.3 Discussione dei dati relativi al processo di decoding .....	188
<b>Conclusioni .....</b>	<b>195</b>
<b>Bibliografia.....</b>	<b>197</b>
<b>Appendice .....</b>	<b>217</b>

## Indice delle figure

Fig.1-1. Modello della lente di Brunswik (1956) applicato alla comunicazione vocale delle emozioni. ....	69
Fig. 5-1. Confronto ita l1- ru l1: durata e vda, valori normalizzati.....	153
Fig. 5-2. Parametri frequenziali (f0 media, em).....	154
Fig. 5-3. Intensità .....	154
Fig. 5-4 Confronto Ita L1- Ru L1: durata tot.; durata vt, e VdA .....	156
Fig. 5-5. Valori frequenziali ( <i>onset</i> e <i>offset</i> ): confronto ita l1-ita l2.....	157
Fig. 5-6. Valori frequenziali (f0 media; EM): confronto Ita L1-Ita L2.....	157
Fig. 5-6. Valori relativi all'intensità: confronto ita l1-ita l2 .....	158
Fig. 5-7 Parametri temporali (durata e VdA): confronto ita l2-ru l1.....	162
Fig. 5-8 Parametri frequenziali (f0 media; em): confronto ita l2-ru l1 .....	162
Fig. 5-9 Parametri relativi all'intensità: confronto ita l2-ru l1 .....	163
Fig. 5-10. Percentuali decodifiche totalizzate dagli apprendenti nelle due lingue.....	171
Fig. 5-11. Percentuali decodifiche globali totalizzate dagli apprendenti in base alla tipologia di enunciato .....	171
Fig. 5-12. percentuali decodifiche totalizzate dagli apprendenti nelle due lingue in base alla tipologia di enunciato .....	171
Fig.5-13. Media % decodifiche globali totalizzate dagli uditori italiani in base alla tipologia di enunciato .....	176
Fig.5-14. Media % decodifiche corrette per lingua (italiano L1, italiano L2 e russo L1).....	178
Fig.5-15. Media % decodifiche corrette per tipologia di enunciato in Ita L1, Ru L1 e Ita L2 .....	182
Fig.5-16. Media % riconoscimenti totalizzati dalle produzioni dei tre apprendenti (FR1, MR1, MR2) .....	184
Fig.5-17. Media % riconoscimenti totalizzati dalle produzioni dei tre apprendenti (FR1, MR1, MR2) in base alla tipologia di enunciato .....	186

## Indice delle tabelle

Tab. 4-1. Prospetto partecipanti .....	123
Tab. 4-2. Termini emotivi in italiano e in russo.....	124
Tab. 4-3. Tipologie di eloquio e relativi <i>encoder</i> .....	132
Tab. 4-4. Pagina iniziale del test rivolto agli apprendenti.....	134
Tab. 4-5 Opzioni di risposta del test uditivo .....	135
Tab. 4-6. Stimoli somministrati per tipologia di enunciato .....	136
Tab. 4-7. Informazioni linguistiche relative agli encoder in lingua russa.....	137
Tab. 5-1 Prospetto dei parametri osservati.....	140
all'interno di ciascun corpus in funzione della tipologia di enunciato .....	140
Tab. 5-2. Risultati test anova in relazione al fattore emozione all'interno dei tre corpora.....	141
Tab. 5-3 Risultati t test ita l1: confronto Neutro-Emozione.....	142
Tab. 5-4. Risultati t test ita l2: confronto neutro-emozione .....	145
Tab. 5-5. Risultati t test: confronto neutro-emozione .....	149
Tab. 5-6. Prospetto dei valori normalizzati all'interno delle produzioni emotive in italiano L1 (Ita L1), italiano L2 (Ita L2) e russo L1 (Ru L1).....	152
Tab. 5-7. Risultati t test relativi alla VdA: confronto Ita L1-Ru L1 .....	156
Tab. 5-8. Risultati t test (collera): confronto Ita L1-Ita L2 .....	159
Tab. 5-9. Risultati t test (tristezza): confronto Ita L1-Ita L2.....	160
Tab. 5-10. Risultati t test (gioia): confronto Ita L1-Ita L2 .....	161
Tab. 5-11. Risultati t test (tristezza, gioia, neutro): confronto Ita L1-Ita L2.....	164
Tab. 5-12. Statistica decodifiche corrette totalizzate dagli apprendenti nelle due lingue.....	169
Tab. 5-13. Risultati t test: confronto decodifiche corrette in russo e in italiano in relazione alla gioia e alla collera. ....	172
Tab. 5-14. Risultati t test: confronto decodifiche corrette in base alla tipologia di enunciato (in russo ).....	173
Tab. 5-15. Risultati t test: confronto decodifiche corrette in base alla tipologia di enunciato (in italiano.....	173
Tab. 5-16 Matrice errori in Ita L1 e Ru L1 .....	174
Tab. 5-17. Statistica decodifiche corrette (globali) e relative percentuali .....	176

Tab. 5-17. Risultati t test: confronto medie riconoscimenti corretti in base alla tipologia di enunciato. ....	177
Tab. 5-19. Statistica decodifiche corrette in funzione della lingua dei locutori.....	178
Tab. 5-20. Risultati T TEST: confronto Ita L1-Ita L2; Ita L1-Ru L1 .....	179
Tab. 5-21. Media % decodifiche corrette per tipologia di enunciato in Ita L1, Ru L1 E Ita L2 .....	179
Tab. 5-22. Risultati t test Ita L1: confronto medie riconoscimenti per enunciato.....	180
Tab. 5-23. Risultati t test Ita L2: confronto medie riconoscimenti per enunciato.....	180
Tab. 5-24. Risultati t test Ru L1: confronto medie riconoscimenti per enunciato .....	181
Tab. 5-25. Risultati t test ita l1-ita l2: confronto medie riconoscimenti (tristezza; gioia) .....	183
Tab. 5-26. Risultati t test Ita L1-Ru L1: confronto medie riconoscimenti (gioia; neutro) .....	183
Tab. 5-27 Medie % decodifiche corretti per tipologia di enunciato.....	185
Tab. 5-28. T-test: confronto medie riconoscimenti MR1-MR2;MR2-FR1 .....	187
Tab. 5-29. Matrice errori in Ita L1, Ru L1 E Ita L2 .....	187

# Abstract

---

La presente ricerca si è posta l'obiettivo di verificare a quali difficoltà andassero incontro apprendenti di italiano di origine russa con un livello di competenza intermedio (B1/B2) in contesti comunicativi che implicassero il ricorso alla prosodia emotiva della lingua target per veicolare e interpretare significati emotivi.

Il disegno sperimentale si è articolato in due fasi: la prima relativa alla raccolta e all'analisi del materiale emotivo (*encoding*), la seconda relativa alla verifica uditiva di quest'ultimo (*decoding*). In relazione al processo di codifica delle emozioni, l'obiettivo generale è stato quello di analizzare e confrontare le produzioni emotive realizzate da due gruppi linguistici aventi una differente L1 (l'italiano e il russo). Nel dettaglio, le produzioni prese in esame sono riconducibili a tre tipologie: a) parlato emotivo in italiano L1, prodotto cioè da parlanti nativi; b) parlato emotivo in italiano L2, prodotto da apprendenti di origine russa, c) parlato emotivo in russo L1, prodotto dai medesimi apprendenti. In secondo luogo, l'obiettivo è stato quello di verificare se e in che misura la L1 degli apprendenti influisse sul parlato emotivo in italiano L2 (ipotesi del *transfer* prosodico).

In relazione al processo di decodifica, l'indagine uditiva è stata necessaria al fine di verificare l'efficacia comunicativa delle produzioni analizzate. A questo proposito sono stati elaborati e implementati due test uditivi rivolti rispettivamente ad un gruppo di 10 apprendenti russi e ad un gruppo di 33 parlanti nativi di italiano. L'obiettivo del test somministrato agli apprendenti era quello di verificare in primo luogo le abilità di decodifica di questi ultimi rispetto alle produzioni emotive realizzate dagli italiani nativi e dai parlanti russi. Per quanto concerne la verifica uditiva rivolta agli uditori italiani, il primo obiettivo è stato quello di stabilire se e in che misura le produzioni emotive realizzate dagli apprendenti in italiano L2 trovassero corrispondenza nei giudizi affettivi espressi dal campione di nativi. In secondo luogo, la ricerca si è proposta di testare la capacità degli italiani nel decodificare le produzioni emotive realizzate in italiano L1 e in russo L1.

In relazione alle emozioni considerate, l'indagine si è focalizzata su tre stati emotivi: collera, tristezza e gioia. In una prospettiva dimensionale, si tratta di emozioni caratterizzate da livelli di attivazione, valutazione e potere differenti in grado di incidere

sulla variazione dei segnali vocali. Il materiale emotivo è stato elicitato mediante dei dialoghi induttori, tradotti in entrambe le lingue. I dialoghi riproducevano il contesto comunicativo di una telefonata informale tra due amici legati da un rapporto di familiarità e confidenza. Per cogliere le variazioni vocali rispetto all'eloquio normale dei partecipanti, è stato elaborato un dialogo "neutro", ovvero privo di reazioni emotive importanti, in cui prevale invece il carattere informativo della telefonata, che si svolge tra un tecnico e un cliente. L'intenzione comunicativa è stata associata ad una frase bersaglio, la stessa per ogni dialogo. Complessivamente, sono state sottoposte ad analisi acustica 60 produzioni emotive e 20 neutre. I parametri indagati sono stati suddivisi in tre macro-categorie: ritmico-temporali, intonativi e relativi all'ampiezza.

Le analisi condotte sui tre corpora hanno restituito risultati per certi versi congruenti con l'ipotesi che l'espressione vocale delle emozioni si fondi su aspetti universali nonché su aspetti specificamente linguistici e culturali. Le produzioni analizzate hanno mostrato una tendenza simile nella modulazione delle risorse prosodiche, tuttavia, sembrerebbe che il fattore "emozione" non abbia influito allo stesso modo sulla variazione dei parametri all'interno di ciascun corpus. Nell'italiano nativo, tutte le produzioni emotive subiscono variazioni significative in termini di durata e di velocità di articolazione, mentre sul piano frequenziale, è l'escursione melodica il parametro maggiormente soggetto a variazioni significative. Anche l'intensità varia sensibilmente in tutte le produzioni emotive. Nelle produzioni realizzate dagli apprendenti in italiano e in russo, i parametri temporali e frequenziali non subiscono variazioni significative, soltanto l'intensità dimostra una certa variabilità.

In relazione al parlato emotivo degli apprendenti, i risultati ottenuti lasciano supporre che gli studenti siano riusciti a modulare le risorse prosodiche della L2, tuttavia questa elasticità non si è manifestata in egual misura in tutte le emozioni, per cui in alcuni contesti emotivi la prosodia emotiva della L1 viene utilizzata per sopperire a queste difficoltà. Ad esempio, le produzioni degli italiani si caratterizzavano per una maggiore estensione tonale, variabilità di  $f_0$  e intensità, nonché per una velocità di articolazione più elevata. D'altro canto, le produzioni in lingua russa si caratterizzavano per una minore variabilità, soprattutto in termini di estensione tonale, per cui i segnali acustici impiegati dagli apprendenti russi possono essere stati percepiti come poco salienti sul piano comunicativo o aver dirottato il giudizio degli italiani verso altre emozioni. Infatti, i risultati dei test uditivi hanno contribuito a chiarire in che misura le produzioni emotive oggetto di indagine venissero decodificate dai due gruppi di uditori,

aventi una diversa L1. In primo luogo, le percentuali di decodifiche corrette ottenute in entrambi i test superano la soglia della pura casualità. Questo significa che le tre emozioni considerate (e l'eloquio neutro) sono state discriminate a prescindere dalla lingua degli *encoder*. Questo dato è in linea con gli studi precedenti che hanno confermato la capacità degli esseri umani nel decodificare il parlato emotivo anche in una lingua diversa dalla propria (cfr. Scherer, Banse, Wallbott 2001; Scherer *et al.* 2003; Thompson, Balkwill 2006; Galatà, Romito 2010). Ciò non toglie che, osservando le percentuali ottenute dalle due lingue, non sia possibile cogliere dati altrettanto significativi e coerenti con l'ipotesi che vi siano delle peculiarità linguo-specifiche in grado di spiegare le basse percentuali di riconoscimento e l'elevato numero di riconoscimenti erronei associati ad alcune emozioni.

In relazione alle abilità di decodifica degli apprendenti, anche in questo caso il loro livello di competenza (B2) e di conoscenza della lingua e della cultura target ha favorito il riconoscimento di alcune emozioni, come la tristezza e in misura minore la gioia. Tuttavia, gli stimoli presentati nella L1 degli studenti sono stati identificati con percentuali significativamente più elevate. Inoltre, l'alta variabilità dei riconoscimenti erronei lascia supporre che il processo di decodifica abbia destato maggiori difficoltà in italiano rispetto alla loro L1.

Considerando i risultati del test uditivo sottoposto al gruppo di italiani, emergono alcune tendenze simili in relazione ai riconoscimenti corretti ed erronei. Tuttavia, anche in questo caso, l'accuratezza nel riconoscimento è significativamente inferiore se si considerano le percentuali totalizzate da ciascuna emozione in Ita L1, Ita L2 e Ru L1.

I risultati ottenuti, pur presentando dei limiti, dati principalmente dal numero poco ampio di partecipanti e dall'impiego di materiale emotivo simulato/indotto, vanno forse interpretati in una prospettiva interazionista, che concilia la dicotomia universalità vs. specificità culturale, promuovendo una visione dinamica della comunicazione vocale, sensibile cioè agli aspetti creativi del linguaggio nonché ai vincoli sociali e culturali.

In quest'ottica, l'acquisizione di una lingua seconda implica l'accesso a queste peculiarità linguo-specifiche. Per questa ragione, sembrerebbe opportuno riservare nelle pratiche didattiche uno spazio alla comunicazione emotiva e in particolare alla prosodia emotiva della lingua target (cfr. De Marco, Paone, 2016), nonché alle altre componenti non verbali che entrano in gioco nel corso delle interazioni quotidiane con i nativi.

# Introduzione

---

La linea di congiunzione tra il fenomeno emotivo e il contesto linguistico e culturale rappresenta il punto di partenza di questo lavoro, ovvero la dimensione emotiva nel processo di acquisizione di una lingua seconda, nelle sue modalità di espressione e percezione.

Sono tre i concetti intorno a cui ruota la ricerca: il primo è il concetto di *cultura*, intesa come mediazione tra l'uomo e la realtà. Essa incrementa e altera i processi comunicativi, filtrando la percezione che un uomo ha del mondo (cfr. Hall 1968; Tomasello 1999; Anolli 2006). Il secondo concetto è quello di *emozione*, intesa come reazione psicofisica ad eventi esterni che compromettono o consentono il raggiungimento degli obiettivi (Poggi 2001). Tali reazioni hanno indubbiamente un carattere biologico e dunque universalmente condiviso (Darwin 1872/1982), tuttavia, esse sono fortemente influenzate anche da norme culturali (cfr. *display rules*, Ekman 1982) e da processi cognitivi individuali. Uno degli obiettivi della ricerca è valutare in che misura la cultura influenzi un aspetto importante della dimensione emotiva, ovvero l'aspetto vocale. Difatti, la *voce* (terzo concetto su cui si fonda il progetto) è un canale non verbale potente, in grado di trasmettere informazioni importanti sul parlante, molte delle quali sono trasmesse in modo inconsapevole. Queste caratteristiche hanno una base biologica, tuttavia i modelli culturali e le valutazioni individuali possono incidere sulla variazione di tali aspetti (cfr. Zinken, Knoll, Panksepp 2008). In quest'ottica, il fenomeno emotivo sarà considerato non in base alla oramai *vexata quaestio* natura vs. cultura, ma attraverso una prospettiva *dialettica*, in cui la cultura "è solo una delle molte facce che la natura può assumere" (Tomasello, 2005, p. 249).

Alla luce di quanto esposto, si ipotizza che comunicare in una lingua straniera non significhi soltanto accedere alle sue strutture linguistiche, ma che implichi anche la comprensione di alcuni aspetti non verbali, quali i gesti, le espressioni del viso, la prossemica, il tono della voce ecc. Maggiore è la distanza tra la lingua d'origine (L1) e la lingua target (L2), maggiori sembrano essere le difficoltà sperimentate (cfr. Dewaele 2005); soprattutto in un contesto emotivo, saper esprimere le proprie emozioni ed interpretare quelle altrui diventa essenziale ai fini di una comunicazione efficace (cfr. competenza emotiva, Saarni, 1999).

Il primo capitolo introduce il fenomeno emotivo e le sue componenti, partendo dalla componente biologica e neuropsicologica, per poi circoscrivere gradualmente il livello di analisi alla sua dimensione cognitiva e culturale. Questo percorso, dall'esterno (l'eredità biologica) verso l'interno (l'eredità culturale), tenterà di offrire una prospettiva multidimensionale delle dinamiche che determinano l'esperienza emotiva sul piano fisiologico e cognitivo e che ne vincolano la manifestazione o l'inibizione sul piano sociale.

Il secondo capitolo introduce i principali nodi teorici che fanno capo allo studio della comunicazione vocale delle emozioni, e in generale al rapporto tra il piano linguistico e quello paralinguistico. In seconda battuta, si focalizza sugli approcci (discreti, componenziali e dimensionali) maturati in questo ambito di ricerca a partire dalle teorie sviluppatesi in ambito psicologico e cognitivo, e sui risultati empirici riportati da questo tipo di ricerche. Inoltre, viene delineato lo scenario metodologico che caratterizza lo studio sul parlato emotivo. Il capitolo si conclude con una ricognizione degli studi condotti in una prospettiva interlinguistica e transculturale, che ha lo scopo di evidenziare gli aspetti linguistici, culturali e sociali che regolano la comunicazione delle emozioni sul piano vocale.

Il terzo capitolo affronta gli aspetti più problematici legati alla dimensione emotiva in una lingua seconda (L2) partendo dal concetto di competenza emotiva proposto da Saarni (1999) e concludendo con le osservazioni e i risultati empirici delle indagini che hanno esplorato la relazione tra emozioni e lingua seconda. Ampio spazio è riservato inoltre alla ricognizione degli studi sul parlato emotivo in L2 e agli aspetti più controversi evidenziati dalle indagini in questione.

Il quarto capitolo illustra il disegno sperimentale e gli obiettivi della presente ricerca, nonché le tecniche e le procedure impiegate nelle due fasi che compongono l'indagine, ovvero il processo di codifica e quello di decodifica delle emozioni in un contesto di acquisizione di una L2.

Il quinto capitolo è dedicato all'analisi e alla discussione dei dati, e propone una possibile interpretazione dei risultati sulla base del Modello della Lente di Brunswik (1956) rielaborato da Scherer (1978; 1982) e adattato alla comunicazione vocale delle emozioni.

# Capitolo 1

---

## Il fenomeno emotivo e le sue componenti<sup>1</sup>

### 1.1. Introduzione

Un racconto di Isaac Asimov, *Notturmo*, narra di un pianeta, Lagash, situato in una galassia lontana e illuminato da sei soli. I suoi abitanti non hanno mai sperimentato l'oscurità della notte, né sanno cosa siano le stelle. Alcuni astronomi, barricati nella loro torre d'avorio, tentano di risolvere il mistero che avvolge la storia del pianeta, la cui civilizzazione sembra avere un carattere ciclico: puntualmente, all'apice della loro cultura, si verifica un evento che porta alla distruzione di ogni forma di civiltà. Il Libro delle Rivelazioni, una raccolta di miti e credenze religiose (l'unica testimonianza scritta di questo popolo in relazione a questo evento catastrofico), narra che

Ogni duemilacinquantanni Lagash entrava in un'immensa caverna, così che tutti i soli scomparivano, e che su tutto il mondo calava l'oscurità totale! Poi apparivano cose chiamate Stelle, che derubavano gli uomini della loro anima e li lasciavano simili a bruti privi di razionalità, al punto da distruggere la civiltà che essi stessi avevano costruito (Asimov, 1973/2015, p. 50).

La paura e il panico, suscitati dalla notte, scatenano negli abitanti del pianeta forme di follia collettiva, che degenerano nella distruzione delle città e di ogni artefatto culturale, allo scopo di riavere la luce attraverso il fuoco e le fiamme. Non avendo mai sperimentato il buio, non hanno sviluppato tecnologie basate sull'illuminazione elettrica (in una prospettiva darwiniana, potremmo dire che si sono adattati all'ambiente circostante). Il terrore dell'oscurità e il misticismo creato dai culti religiosi ostacolano il lavoro degli scienziati, che tentano di divulgare, forse troppo tardi, la spiegazione scientifica di questo fenomeno, ovvero l'esistenza di "luna", mai osservata in cielo, che renderebbe plausibile l'eclissi totale e la conseguente oscurità. Per salvaguardare la civiltà e la cultura, decidono di costruire un rifugio, illuminato con torce rudimentali, in

---

<sup>1</sup> Alcune sezioni dei capitoli 1, 2, 3 relative agli approcci teorici e allo stato dell'arte sono state già pubblicate nel volume scritto insieme a Anna De Marco (cfr. De Marco, Paone, 2016) che raccoglie una serie di tecniche e strumenti glottodidattici finalizzati allo sviluppo di una competenza emotiva nella lingua target.

cui accogliere poche persone e conservare la documentazione necessaria per contribuire alla ricostruzione della civiltà.

L'unico psicologo presente nel racconto annuncia con parole profetiche ciò che accadrà alle loro menti di fronte ad un'esperienza che non hanno mai provato:

Il suo cervello non è stato costruito per una concezione del genere, proprio come non è in grado di concepire l'infinito o l'eternità. Lei può soltanto parlarne. Una frazione della realtà è sufficiente a sconvolgerla, e quando questa realtà subentrerà in tutta la sua concretezza, il suo cervello si troverà di fronte a un fenomeno che esula dai limiti della sua comprensione. Lei impazzirà in modo completo e permanente! [...] E un altro paio di millenni di lotta penosa sfumeranno nel nulla. (p. 58).

Infatti, quando l'ultimo dei sei soli sta per sparire nel cielo, gettando il pianeta nell'oscurità, il terrore pervade gli stessi astronomi, i quali, pur essendo consapevoli della causalità degli eventi, sono invasi dalla medesima follia collettiva che anima gli abitanti delle città (rivoltatisi contro gli scienziati e decisi a distruggere l'Osservatorio).

L'incapacità di comprendere e assimilare una realtà sconvolgente è forse dovuta al fatto che questa realtà non è stata codificata e categorizzata dal sistema culturale del piccolo mondo che, ignaro di cosa sia la notte, aspetta il giorno dell'oscurità con un terrore mistico, smorzato da un ostinato scetticismo che rifiuta una verità tanto spaventosa. Anzi, le uniche forme di trasmissione culturale si fondano su miti oscuri che non fanno che alimentare il terrore dell'oscurità. Dal canto loro, gli scienziati tentano di arginare questa reazione collettiva, preservando la conoscenza, tuttavia le loro menti restano paralizzate di fronte all'ignoto, all'oblio della notte e alla comparsa delle stelle. La razionalità soccombe di fronte ad un evento che schiaccia ogni consapevolezza, la paura si tramuta in panico e il panico in follia, forse perché non esiste un pericolo da cui fuggire, se non la sensazione claustrofobica di oppressione suscitata dall'assenza di luce:

Theremon si rialzò barcollante, con la gola chiusa che gli impediva il respiro e tutti i muscoli contratti da un'intensità di terrore e di panico al di là di ogni sopportazione. Stava per impazzire e lo sapeva, e in un angolo perduto del suo cervello un ultimo frammento di lucidità stava dibattendosi, per arginare l'ondata irrefrenabile di terrore. Era orribile diventare pazzo e sapere di impazzire: sapere che, tra un istante, sarebbe stato lì, fisicamente, e tuttavia la sua vera essenza si sarebbe spenta, sarebbe stata inghiottita da una nera follia (ivi, p. 84).

Questo racconto mette in luce almeno due aspetti cruciali: da un lato, la reazione collettiva di panico e follia che sconvolge anche le menti più razionali di fronte ad un evento che va al di là della propria esperienza quotidiana e che rappresenta una delle paure ancestrali di un popolo, e dall'altro, l'annientamento del sapere scientifico, degli

artefatti culturali, e della civiltà stessa, conseguente a questo episodio di follia. Questi due elementi, considerati nell'insieme, suggeriscono l'immagine di un popolo, la cui storia sembra vincolata ad una certa linea di sviluppo, "spalmata" nell'arco di duemila anni, al di là della quale non esiste possibilità di progresso (almeno fintanto che la follia collettiva non avrà lasciato spazio a forme di collaborazione e organizzazione sociale, in grado di conservare e trasmettere il sapere da una generazione all'altra). Questo fenomeno blocca lo sviluppo culturale e scientifico del pianeta, poiché viene meno la trasmissione di saperi e artefatti culturali sui quali un popolo costruisce la propria civiltà (una questione che verrà approfondita più avanti, cfr. PAR. 1.5).

Com'è possibile che un evento tanto banale, come il sopraggiungere delle tenebre, susciti una reazione emotiva di tali proporzioni? La mente razionale non riesce ad arginare un terrore che nasce dall'incapacità di comprendere una realtà nuova, che non ha avuto né il tempo, né il modo di essere cognitivamente e culturalmente assimilata. Sono in gioco diversi fattori: da un lato, l'aspetto puramente biologico, rintracciabile nella reazione automatica ad un evento sconosciuto, dall'altro, una componente psicologica, cognitiva, che tenta invano di valutare lo stimolo ed elaborare una reazione adeguata, in base alle esperienze precedenti e a paradigmi comportamentali culturalmente condivisi. Tuttavia, ciò che manca, in questo caso, è l'esperienza stessa dell'oscurità: gli abitanti di Lagash hanno una concezione della realtà *diversa* da quella degli abitanti della Terra e, pertanto, anche la loro reazione sarà diversa, determinata da una paura primitiva che preclude l'accesso alla comprensione e alla categorizzazione di un fenomeno tanto temuto. Allo stesso modo, le recenti scoperte degli scienziati non restano che sterili informazioni di fronte all'entità dell'evento osservato, né possono essere accettate, in un tempo così breve, dalla gente comune. Il racconto si conclude infatti con la discesa delle tenebre e l'apparizione delle stelle, che assistono indifferenti alle conseguenze rovinose di un'emozione intensa e feroce come la paura.

Che cos'è, dunque, un'emozione? Definire un fenomeno complesso come quello emotivo implica, in un certo senso, la necessità di adottare una prospettiva *dinamica*, in grado di cogliere le molteplici dimensioni di un fenomeno che, da un lato, si intreccia con la storia evolutiva dell'"animale uomo" e, dall'altro, trova una sua peculiare declinazione all'interno delle coordinate culturali, sociali e linguistiche in cui si realizza. In quest'ottica, il fenomeno emotivo sarà considerato non in base alla oramai *vexata quaestio* natura vs. cultura, ma attraverso una prospettiva *dialettica*, in cui la

cultura "è solo una delle molte facce che la natura può assumere" (Tomasello, 2005, p. 249).

L'acceso dibattito che riguarda l'antitesi natura/cultura affonda le sue radici in ambito filosofico, nell' Europa nel XVIII secolo, prima che il pensiero e le osservazioni di Darwin gettassero luce sui processi biologici legati all'evoluzione della specie. Come sottolinea Tomasello, la prospettiva evoluzionista avrebbe dovuto rendere obsoleto il dibattito (ivi, p. 69), in quanto essa fornisce una spiegazione circa i meccanismi di selezione naturale e di adattamento all'ambiente circostante che determinano certi risultati piuttosto che altri. L'obiettivo, dunque, "non è decidere se una struttura sia o no "innata", ma piuttosto determinare i processi all'opera nel suo sviluppo" (ivi, p. 71), tenendo in considerazione le influenze esterne nonché i processi di trasmissione culturale.

Abbandonando dicotomie quanto mai limitative, tanto in un senso quanto nell'altro, il fenomeno emotivo verrà dunque considerato nella sua componente biologica e neuropsicologica, per poi circoscrivere gradualmente il livello di analisi alla sua dimensione cognitiva e culturale. Questo percorso, dall'esterno (l'eredità biologica) verso l'interno (l'eredità culturale), tenterà di offrire una prospettiva multidimensionale delle dinamiche che determinano l'esperienza emotiva sul piano fisiologico e cognitivo e che ne vincolano la manifestazione o l'inibizione sul piano sociale.

## **1.2 La componente biologica**

Nella specie umana, come in quella animale, gli stati emotivi sono il risultato di un insieme di reazioni specifiche in grado di fornire una risposta di adattamento all'ambiente circostante. A questi meccanismi di risposta automatici corrisponde l'attivazione di risposte fisiologiche preposte alla sopravvivenza e alla salvaguardia della specie. I progenitori dell'uomo, in costante bilico tra la lotta e la difesa dall'ambiente circostante, hanno sviluppato questi meccanismi di risposta per fronteggiare pericoli o fuggire da potenziali minacce, garantendo la sopravvivenza della specie anche in condizioni estreme.

### *1.2.1 Emozioni come meccanismi di risposta*

Fu Darwin il primo a intuire la funzione sostanzialmente adattiva di alcuni comportamenti espressivi (legati a emozioni intense) in relazione al processo di

selezione naturale. Già nel 1859, con *L'origine delle specie*, aveva messo in luce i tre principi che regolano il processo evolutivo, sconvolgendo il senso comune dell'epoca, che considerava immutabili le specie animali e vegetali. Il primo principio, la *sovraabbondanza*, consente la riproduzione di una prole più numerosa rispetto a quella che arriverà fino all'età adulta; il secondo, la *variazione*; garantisce la trasmissione di tratti e aspetti variabili di generazione in generazione; il terzo principio, la *selezione*, fa sì che i discendenti le cui variazioni si adattano meglio all'ambiente circostante abbiano più probabilità di sopravvivere, riprodursi e trasmettere a loro volta queste caratteristiche (cfr. Oatley, 2007, p. 40).

Partendo da questi principi, Darwin si chiese se la specie umana condividesse con quella animale non solo meccanismi biologici, ma anche tendenze e comportamenti sul piano psicologico. Nella sua celebre monografia, *L'espressione delle emozioni negli uomini e negli animali* (1872/1982) racconta la sua visita al parco zoologico di Londra, durante la quale sperimentò su se stesso una reazione di paura. Avvicinatosi ad una lastra di vetro dietro cui si trovava una vipera africana, decise a non ritrarsi nel caso in cui il rettile avesse tentato di attaccarlo, constatò con grande stupore, che, nonostante la volontà e la ragione, il suo corpo aveva reagito al movimento della vipera, allontanandosi istintivamente. Come scrive Oatley: "Darwin proponeva che ciò che aveva governato il suo comportamento era stato "installato" nel suo cervello molto tempo prima: un dispositivo cerebrale che lo faceva indietreggiare all'attacco del serpente" (Oatley, 2007, p. 42).

Darwin intuisce il ruolo di primo piano svolto da questo meccanismo di risposta automatico all'interno del processo di evoluzione e selezione naturale. L'attenzione del naturalista è rivolta in modo particolare ai movimenti espressivi associati a emozioni intense (come la rabbia o la paura), rintracciabili anche nei primati non umani. Con il termine *espressione*, si riferisce ad un comportamento o a un cambiamento corporeo visibile determinato da un'emozione (cfr. Oatley, 2007). Secondo Darwin, in origine, tali movimenti svolgevano una funzione differente da quella espressiva, ovvero di adattamento e risposta agli stimoli esterni: sottrarsi a una minaccia, alleviare qualche sofferenza o appagare un desiderio; il processo evolutivo ha fatto sì che i meccanismi in questione, al pari di quelli biologici, sviluppassero in seguito funzioni specifiche, implicate nella manifestazione di determinati fenomeni emotivi. Nel caso del pianto, ad esempio, "abbiamo acquisito l'abitudine di contrarre i muscoli che circondano gli occhi [...] dai nostri progenitori i quali contraevano questi muscoli come reazione alla

spiacevole sensazione che avevano sperimentato [...] quando lanciavano forti urli" (Darwin 1872/1982, trad. it., p. 409). La collera doveva inizialmente essere espressa "con gesti di minaccia, con movimenti frenetici, con l'arrossamento della pelle e con lo sguardo torvo" (ivi, p. 416), ovvero attraverso una serie di automatismi biologici che prefiguravano un attacco.

Anche il riso, secondo Darwin, doveva essere praticato dai progenitori dell'uomo già prima che assumessero una forma completamente umana. Darwin suffraga la validità di tale deduzione grazie all'osservazione di alcune specie di scimmie, che "quando sono contente, emettono suoni ripetuti, molto simili a quelli della nostra risata, spesso accompagnati da movimenti vibratorii delle mascelle o delle labbra, dallo spostamento [...] degli angoli della bocca [...] e perfino dal brillare degli occhi"(ivi, p. 415).

In sostanza, i movimenti che accompagnavano emozioni intense erano compiuti originariamente per uno scopo distinto da quello espressivo. Ad esempio, l'emissione delle lacrime era dovuta probabilmente alla contrazione convulsa delle palpebre e al gonfiore dei globi oculari conseguenti allo sforzo di urlare. Nel caso del disgusto, ipotizza invece che i progenitori dell'uomo avessero la capacità di rigurgitare in modo volontario il cibo potenzialmente nocivo. Con lo sviluppo del linguaggio, il rigurgito volontario perse la sua funzione primaria: comunicare ai propri simili la pericolosità di un alimento. Oggi tale capacità è venuta meno, tuttavia si manifesta ancora nell'uomo come azione involontaria- la nausea- scaturita da un pensiero, un sapore o un odore ripugnanti.

Un caso particolare è costituito dalla paura e dalle sue diverse espressioni (dal sussulto della sorpresa al terrore più intenso). I principi dell'abitudine, dell'associazione e dell'eredità, afferma Darwin, offrono una spiegazione di alcuni segni di queste emozioni: l'apertura totale degli occhi e della bocca e il sollevamento delle sopracciglia consentivano di osservare attentamente l'ambiente circostante e di ascoltare eventuali suoni. Inoltre, l'abitudine umana a tutelare la propria sopravvivenza ha spinto i nostri progenitori ad affrontare i nemici con la lotta o a fuggire impulsivamente. Tali pratiche, secondo Darwin, innescavano intense reazioni fisiologiche (accelerazione del battito cardiaco e della respirazione, dilatazione delle narici) che avevano come risultato una forte prostrazione fisica (pallore, sudorazione, tremore). Per tale ragione, ancora oggi, per abitudine ed eredità, si manifestano gli stessi effetti a livello fisiologico.

### 1.2.2 Le emozioni primarie

La prospettiva evuzionista di Darwin suggerisce dunque che le alterazioni fisiologiche associate ad uno stato emotivo siano il risultato di un lungo processo di adattamento all'ambiente circostante, nel corso del quale questi meccanismi di risposta hanno garantito la sopravvivenza della specie. Questa prospettiva è alla base di un approccio teorico che, cavalcando l'onda della filogenesi, postula un numero discreto di emozioni primarie, considerate come unità discrete afferenti a meccanismi biologici innati, e dunque universali, dalle quali avrebbero origine le altre emozioni, in questo caso secondarie e non necessariamente condivise sul piano interculturale. C'è da dire che Darwin non opera una distinzione tra emozioni primarie<sup>2</sup> e secondarie, ma si limita piuttosto a descrivere i comportamenti espressivi che ha rinvenuto nelle varie specie animali e nell'uomo.

In base a queste teorie, il carattere innato di tali emozioni sarebbe comprovato dalla persistenza di meccanismi simili non solo nell'uomo, ma anche in altre specie animali (soprattutto nei primati). La loro funzione, come sosteneva Darwin, sarebbe essenzialmente adattiva. Plutchik (1980) ad esempio, suggerisce un insieme di comportamenti adattivi che accomuna tutti gli organismi viventi in relazione agli stimoli esterni ed interni (in termini di risposte efficaci di adattamento). Nel complesso, queste risposte comportamentali innate sono rintracciabili non solo negli esseri viventi più evoluti, ma anche nelle forme più elementari di vita animale e vegetale. Ad esempio, l'*incorporazione* è legata alla funzione fisiologica dell'ingestione dei cibi e all'accettazione di stimoli esterni positivi all'interno dell'organismo; il *rifiuto*, ovvero l'espulsione, di ciò che è stato già introdotto nell'organismo, in quanto nocivo per l'organismo stesso; la *protezione*, ovvero l'insieme dei comportamenti atti ad allontanare situazioni di minaccia e di dolore per l'organismo; la *distruzione*, ovvero l'abbattimento di qualsiasi barriera che ostacoli la soddisfazione dei propri bisogni; la *riproduzione*, che controlla le tendenze di avvicinamento e contatto legate al rapporto sessuale; la *reintegrazione*, che include l'insieme di reazioni conseguenti alla perdita di qualcosa posseduto in precedenza o al quale si era legati affettivamente; l'*orientamento* ovvero la reazione ad uno stimolo nuovo o inusuale (la sorpresa); e infine

---

<sup>2</sup> La distinzione tra emozioni primarie e secondarie affonda le sue radici nella tradizione filosofica. Già Descartes, ne *Le passioni dell'anima* (1642/2003) opera una prima classificazione di quelle che definisce emozioni fondamentali (l'Odio, la Meraviglia, ecc.), annoverando in questa categoria anche stati emotivi e sentimenti che in seguito sono stati esclusi, come l'odio.

*l'esplorazione*, ovvero l'insieme dei comportamenti necessari alla perlustrazione e alla ricognizione di un determinato ambiente.

Dal punto di vista evolutivistico, tali comportamenti sarebbero alla base di un sistema di controllo che gli organismi attivano in presenza di eventi esterni che rischiano di comprometterne la sopravvivenza. Le emozioni primarie sarebbero dunque regolate da meccanismi innati e caratterizzate da un insieme di costituenti motivazionali e fenomenologici propri (D'Urso, Trentin, 1998).

Non c'è accordo sul numero delle emozioni primarie, in quanto i diversi modelli proposti comprendono variabilmente dalle cinque alle nove emozioni di base. Tomkins (1962) teorizza l'esistenza di otto emozioni fondamentali: interesse, sorpresa, gioia, angoscia, paura, vergogna, disgusto e collera, alle quali corrisponderebbero determinate risposte fisiologiche e specifiche espressioni facciali. In particolare, l'espressione facciale di ogni emozione sarebbe codificata geneticamente nel sistema nervoso autonomo (SNC), e avrebbe dunque una base innata. Gli impulsi e le spinte motivazionali sarebbero inoltre una manifestazione dei bisogni dell'organismo, accentuati successivamente dalle emozioni.

La teoria delle emozioni differenziali di Izard (1977), in accordo con Tomkins sulla natura innata degli stati emotivi di base, si fonda su cinque assunti fondamentali. Il primo sostiene che ogni emozione sia costituita da tre componenti: un substrato neurale, un *pattern* espressivo specifico (rintracciabile soprattutto nelle espressioni facciali), e un determinato sentimento (positivo o negativo). Il secondo afferma che ogni emozione abbia delle proprietà motivazionali necessarie alla sopravvivenza dell'individuo e della specie, e che sia per natura adattiva. Il terzo puntualizza la discretezza delle emozioni: ogni stato emotivo determina esperienze interiori diverse e porterà a comportamenti differenti. Il quarto sostiene che le emozioni interagiscono tra di loro (attivando, amplificando o riducendosi a vicenda). In quest'ottica, uno stato emotivo può essere il risultato di un complesso aggregato di emozioni individuali. Infine, in base al quinto presupposto, ogni emozione interagisce influenzando altri processi fisiologici importanti, come l'omeostasi, la percezione, le risposte motorie, gli elementi cognitivi ecc. Sebbene, di norma, i diversi processi interagiscano tra di loro, quello emotivo, secondo Izard, mantiene una propria autonomia rispetto a quelli cognitivi (il cui ruolo non è preponderante nelle emozioni).

La teoria in questione prevede l'esistenza di nove emozioni di base: gioia, sorpresa, interesse, dolore, rabbia, disgusto, disprezzo, vergogna e paura. Esse, in

quanto unità discrete, avrebbero una corrispondenza con determinate espressioni facciali, in grado di costituire precisi indizi sullo stato emotivo in atto.

La mancanza di un accordo sul numero di emozioni primarie sembrerebbe suggerire che vi siano ancora degli elementi da chiarire in relazione a ciò che è considerato universalmente condiviso o soggetto a variabili culturali. Come già sostenuto in precedenza, una simile dicotomia appare riduttiva, soprattutto in considerazione del fatto che, di per sé, il contesto culturale rappresenta l'ecosistema all'interno del quale si sviluppa la specie umana (Tomasello, 2005, Anolli, 2006).

### **1.3 La componente neuropsicologica**

La prospettiva di Darwin chiarisce la funzione delle emozioni in chiave evolutiva. Per spiegare le dinamiche che determinano l'esperienza emotiva è necessario chiamare in causa un'ulteriore componente del fenomeno emotivo: i meccanismi neurofisiologici e neuropsicologici che innescano le risposte dell'organismo.

#### *1.3.1 Il fulcro dell'esperienza emotiva: il cervello*

La complessità del cervello umano, in base all'ipotesi suggerita da Paul MacLean (1990), è il risultato dell'evoluzione di tre tipi di cervello:

- il *cervello rettiliano* (il più antico), sede dei comportamenti primitivi, legati alla caccia e alla difesa del territorio;

- il *cervello paleomammaliano* (ereditato dai mammiferi antichi), destinato alla gestione dei bisogni legati all'autoconservazione e alla conservazione della specie; rispetto al cervello dei rettili, quello dei mammiferi si differenzia per tre aspetti: "le madri dei mammiferi accudiscono la prole che è dipendente da loro; i mammiferi fanno segni vocali gli uni agli altri e si dedicano a giochi improvvisati, [...] vivono in un universo sociale ben diverso da quanto può essere osservato nei rettili" (Oatley, 2007, p. 89);

- il *cervello neomammaliano*, dal quale dipendono i processi cognitivi più evoluti (MacLean, 1990), rintracciabile nella parte del cervello definita neocorteccia (o corteccia), che comprende oltre l'80% del volume totale, comunemente ritenuta la sede del pensiero. Infatti, come spiega Oatley (2007, p. 90): "la corteccia prefrontale, è associata a quel tipo di pensiero [...], noto come progettazione. È anche implicata nella creazione di qualsiasi tipo di strumento ed è l'organo del linguaggio".

A intuire il ruolo centrale del cervello nella percezione dell'esperienza emotiva furono due fisiologi statunitensi, Walters Bradford Cannon (1927) e Philip Bard (1934). Inizialmente, la teoria dello psicologo William James (1884) aveva indirizzato le ricerche verso gli organi periferici del corpo (in particolare i visceri)<sup>3</sup>. Secondo James, infatti, il corpo costituiva una sorta di cassa di risonanza in grado di registrare i minimi cambiamenti al livello dell'inconscio.

Le ricerche di Cannon e Bard si focalizzarono invece sull'attivazione delle regioni cerebrali; con i loro esperimenti dimostrarono che, separando chirurgicamente il cervello dagli organi periferici, le reazioni emotive suscitate dalla paura continuavano a verificarsi. Non erano i visceri a determinare il fenomeno emotivo, bensì il talamo, una struttura del sistema nervoso centrale. In linea con l'ipotesi evoluzionista di Darwin, Cannon sosteneva che tale regione dell'encefalo si fosse sviluppata e potenziata al fine di garantire una risposta immediata agli stimoli esterni, mediante l'attivazione dei muscoli e dei visceri.

Oggi sappiamo che le risposte fisiologiche dell'organismo sono essenzialmente legate all'attivazione di tre sistemi distinti: il sistema nervoso centrale (SNC), il sistema nervoso autonomo (SNA) e il sistema endocrino. Il sistema nervoso centrale è sede dell'amigdala e dell'ipotalamo, cruciali nell'attivazione degli stati emotivi (cfr. D'Urso, Trentin, 1998). Il sistema nervoso periferico, composto da gruppi di nervi, funge da collegamento tra il SNC e l'organismo, ed è sede del SNA (al quale si deve una serie di risposte autonome che concernono la salivazione, l'apparato cardiovascolare, quello gastroenterico ecc.). Il sistema nervoso autonomo è composto dal sistema simpatico e da quello parasimpatico. Il primo è finalizzato essenzialmente alla produzione di energia, mentre il secondo spinge nella direzione opposta, ovvero alla sua conservazione. Ad esempio, i sintomi legati al sistema simpatico si rifletteranno nella respirazione, nel tono muscolare, nel battito cardiaco e così via. Il sistema parasimpatico, invece, agirà sulla contrazione dello stomaco, sul controllo del pianto

---

<sup>3</sup> Secondo la sua ipotesi (formulata parallelamente al fisiologo Peter Lange, cfr. Lange 1885) la percezione di uno stimolo potenzialmente pericoloso (come la presenza di un orso), innesca alcune reazioni fisiologiche che preparano l'organismo alla fuga; è la percezione di questi cambiamenti corporei a determinare l'emozione: pertanto, non fuggiamo perché abbiamo paura, ma abbiamo paura perché fuggiamo (ivi, pp. 189-190).

ecc.<sup>4</sup> Il sistema endocrino, infine, funge da mediatore tra SNC e SNA, convogliando le informazioni che interessano le funzioni endocrine al SNC.

### *1.3.2 Il sistema limbico*

Le ricerche attuali hanno confermato il ruolo centrale di tre strutture cerebrali: l'ipotalamo, l'ippocampo (struttura localizzata nella parte mediana di entrambi i lobi temporali, particolarmente importante nella formazione della memoria a lungo termine) e l'amigdala (responsabile della memoria emozionale di ciascun individuo e dell'elaborazione di reazioni di paura e aggressività (Ledoux, 2000; 2003). Questo complesso sistema di strutture è definito sistema limbico ed è situato al di sotto della corteccia celebrale (costituisce lo strato del cervello che MacLean ha definito "paleomammaliano").

L'amigdala svolge funzioni di notevole importanza in circostanze in cui risulta necessario adottare una reazione di fuga o difesa dai pericoli, offrendo una rapida valutazione dello stimolo. Si tratta di una valutazione superficiale, fondata su un'osservazione globale e approssimativa, che tuttavia è in grado di garantire la sopravvivenza di un individuo, innescando una reazione di paura (necessaria per fuggire o attaccare). La corteccia, invece, interviene per stabilire se si tratti di un pericolo reale o immaginato, offrendo dunque una valutazione più precisa, ma impiegando un tempo maggiore, durante il quale l'amigdala allerta l'organismo, preparandolo all'azione.

## **1.4 La componente cognitiva**

Se in tempi remoti le emozioni hanno salvaguardato la specie umana, attraverso risposte fisiologiche automatiche, in tempi più recenti, soprattutto in virtù dell'evoluzione del linguaggio, delle scienze e delle tecniche, i pericoli da cui fuggire o da affrontare sono diventati altri, per cui si suppone che al di là delle strutture neurofisiologiche interessate, vi siano dei processi mentali in grado di offrire una valutazione dello stimolo e di innescare una reazione adeguata agli scopi.

---

<sup>4</sup> Come spiegano D'Urso e Trentin: "In condizioni di rilassamento e di tono edonico blandamente positivo si ha una dominanza del sistema parasimpatico mentre l'insieme delle risposte attivate dal simpatico ha un significato difensivo in quanto predispongono l'organismo a fronteggiare i pericoli ambientali con tutte le energie disponibili" (D'Urso, Trentin, 1998, pp. 28-29).

A partire dagli anni '60 del secolo scorso, il contributo della psicologia cognitiva ha messo in luce alcuni meccanismi legati all'esperienza emotiva, tralasciati dagli approcci teorici evuzionistici, che, tendenzialmente circoscrivono l'influenza degli aspetti cognitivi nella manifestazione di uno stato emotivo. Al contrario, le teorie cognitive pongono in primo piano i processi di elaborazione e valutazione mentale nell'esperienza emotiva.

In quest'ottica, gli stati emotivi non sarebbero soltanto il frutto di risposte innate agli stimoli esterni, ma sarebbero determinati da processi cognitivi, a loro volta condizionati dall'esperienza individuale, dall'apprendimento e dalla corrispondenza con paradigmi comportamentali socialmente condivisi (Galati, 1993).

Ad esempio l'emozione, nella teoria cognitivo-attivazionale di Schachter e Singer (1962), si configura come il risultato dell'interazione fra due costituenti di natura diversa: il primo, legato all'aspetto fisiologico, consiste nell'attivazione dell'organismo (*arousal*); il secondo, puramente psicologico, corrisponde alla percezione di questo stato di attivazione e alla sua interpretazione in relazione ad uno stimolo emotigeno. Inoltre, la prospettiva di Schachter e Singer attribuisce all'interpretazione cognitiva il compito di definire la natura dell'emozione stimolata nella fase di *arousal*. Mentre l'attivazione fisiologica reagisce a contatto con un dato evento emotigeno, l'aspetto cognitivo interviene ad etichettare il tipo di emozione stimolata. L'attivazione fisiologica, dunque, non basterebbe a produrre un fenomeno emotivo, ma è necessario che tale fenomeno sia inserito in una categoria emozionale specifica, affinché venga opportunamente decodificato. In sostanza, l'emozione scaturirebbe dall'*arousal* e da due diversi processi cognitivi, il primo sarebbe implicato nella percezione e nel riconoscimento dello stimolo, il secondo determinerebbe una connessione tra l'atto cognitivo in questione e l'*arousal* (Anolli, Ciceri, 1992).

Le argomentazioni di Schachter e Singer partono dal presupposto che le reazioni fisiologiche legate ad un'emozione (accelerazione del battito cardiaco, sudorazione, tensione muscolare) informano il cervello della presenza di uno stato di attivazione (*arousal*). Tuttavia, poiché tali reazioni sono simili in emozioni diverse, non siamo in grado di identificare il tipo di stato emotivo che stiamo provando. Soltanto sulla base delle informazioni legate al contesto fisico e sociale e alla conoscenza del tipo di emozione che solitamente si verifica in contesti simili, siamo in grado di etichettare l'emozione provata. Inoltre, sempre in base alle ipotesi formulate dai due studiosi,

l'emozione implicherebbe l'attivazione fisiologica, qualora essa non fosse innescata, l'organismo non potrebbe sperimentare alcuna emozione.

#### 1.4.1 Emozioni come valutazioni

Già nel IV secolo a.C., Aristotele aveva suggerito che l'insorgere di un'emozione fosse il risultato di una valutazione individuale, in grado di modificare soggettivamente l'interpretazione delle circostanze e di innescare un'azione, sulla base di credenze e convincimenti personali<sup>5</sup>.

Nell'assetto teorico di Aristotele, l'emozione si configura dunque come il risultato di una credenza e si traduce in una tendenza all'azione. Questo modello, teorizzato nel IV secolo a.C., ha trovato un riscontro concreto e una sua continuità negli approcci nati in seno alla psicologia delle emozioni e alla psicologia cognitiva. E' in questo contesto che trova la sua definizione il concetto di *appraisal*, introdotto da Magda Arnold nel 1960, inteso come una "valutazione precognitiva, non riflessa" di un evento (Poggi, 2008, p. 17), rispetto al quale ciascun individuo elabora una valutazione positiva o negativa. Tale valutazione è legata al probabile raggiungimento o alla compromissione di uno scopo (cfr. PAR. 1.4.2).

Arnold (1960) suggerisce che il fenomeno emotivo sia il risultato di due processi: la percezione e la valutazione (*appraisal*, appunto). Mentre per percezione si intende l'effettivo riconoscimento di un evento emotigeno, indipendente dal soggetto, l'*appraisal*, che segue immediatamente la percezione, si configura come la valutazione - positiva o negativa- di tale evento in relazione al soggetto (cfr. Anolli, Ciceri, 1992). L'emozione costituisce quindi la diretta conseguenza dell'*appraisal*.

Le teorie dell'*appraisal* tentano di comprendere e identificare i contenuti mentali che determinano la percezione di un'esperienza in termini emotivi. In quest'ottica, la natura (positiva o negativa) di uno stimolo non scaturisce l'emozione, ma è la valutazione personale dell'individuo a condizionare l'emergenza o meno di uno stato emotivo. Come scrive Frijda (2007, p. 93):

---

<sup>5</sup> Ad esempio, Aristotele suggerisce che l'ira derivi dalla credenza di aver subito un'offesa, un oltraggio immeritato, che genera sofferenza e il desiderio di infliggere il medesimo dolore attraverso la vendetta; il timore invece è determinato dall'imminenza di un pericolo in grado di suscitare dolore e sofferenza. Anche in questo caso, la disposizione d'animo del soggetto ha un peso specifico rilevante nella manifestazione di tale emozione: si ha paura quando si ritiene di poter subire conseguenze rovinose per la propria persona; viceversa, il timore viene meno quando si crede di essere in grande prosperità (Aristotele, 1996, p. 173) o quando si ritiene di aver già subito ogni male e si guarda con distacco al futuro (ad esempio, nel caso di un condannato a morte).

Emotions can be distinct by how the events that elicit them are appraised. They may be seen as representing different 'core relational themes' or different 'goal juncture contingencies'. Sadness typically consists of experiencing a personal loss, anger of experiencing offence or frustration, and fear of perceived serious threat.

In sostanza, la valutazione positiva o negativa di un evento innesca una reazione adeguata agli scopi. Arnold definiva tale proprietà delle emozioni "tendenza all'azione". Per Frijda (1986) il fulcro di un'emozione si ritrova in una condizione mentale di "preparazione all'azione" che implica anche una concreta attivazione fisiologica dell'organismo.

#### *1.4.2 Emozioni come scopi*

Come qualsiasi atto comunicativo, anche le emozioni sottendono a scopi specifici, in relazione alle probabilità di raggiungimento o compromissione degli obiettivi. Tale dicotomia si esprime attraverso lo stato emotivo che il soggetto sperimenterà: l'emozione percepita sarà positiva nel caso in cui gli scopi prefissati vengano realizzati, viceversa lo stato emotivo sarà negativo in caso di mancato raggiungimento degli obiettivi.

Le emozioni, dunque, possono essere classificate in base al tipo di scopo e alla valenza personale e sociale che assumono. Ad esempio, emozioni come odio e amore gestiscono i rapporti con le altre persone in base al tipo di scopo (aggressione/cooperazione) e hanno quindi una dimensione sociale (come stima e ammirazione). Al contrario, sono essenzialmente emozioni legate al concetto di sé (imbarazzo, la vergogna, l'orgoglio ecc. Ciascuna di esse esprimerà l'effettiva realizzazione di sé o la compromissione dell'immagine all'interno del contesto sociale.

Altri tipi di emozioni dipendono invece dall'autovalutazione del soggetto (Poggi, Magno Caldognetto, 2004), ovvero dal grado di controllo che si ritiene di possedere: ad esempio, la rabbia è generata da un sentimento di dominanza, mentre la paura rispecchia una condizione di sottomissione.

In linea generale, gli stati emotivi esprimono, consciamente o inconsciamente, degli scopi predeterminati, i quali stimolano un *arousal* fisiologico e predispongono l'organismo ad una reazione specifica (in seguito ad una valutazione cognitiva del contesto), consentendo quindi la realizzazione dello scopo. La paura, ad esempio,

essendo un'emozione attivante, predispone l'organismo alla fuga e alla difesa, mentre la rabbia (anch'essa attivante) prepara all'attacco.

Al contrario, la tristezza o il dolore sono invece delle emozioni che tendono ad inibire; in particolare la prima ha una funzione di "reintegrazione" (Anolli, Ciceri, 1992), ovvero proietta il soggetto in una dimensione riflessiva, necessaria al superamento dello stato di sconforto.

#### *1.4.3 Il modello componenziale delle emozioni*

Negli anni '80 del secolo scorso, Klaus Scherer (cfr. Scherer, 1984; 2009) ha elaborato un modello componenziale delle emozioni (CPM, *Component Process Model*). Contrariamente a quanto sostenuto dalle teorie discrete delle emozioni, il CPM non ammette l'esistenza di un ristretto e limitato numero di emozioni innate, poiché, come spiega lo stesso Scherer "the emotion process is considered as a continuously fluctuating pattern of change in several organismic subsystems" (Scherer, 2009, p.1320).

Il CPM prevede la presenza di cinque componenti che racchiudono, in uno schema ben definito, la complessità del processo (cfr. Galati, 1993): una *componente emotiva*, che assicura l'elaborazione dell'informazione, ricavata dall'evento emotigeno; una *componente neurofisiologica*, che verifica l'attivazione del Sistema Nervoso Autonomo, del Sistema Nervoso Centrale e del Sistema Endocrino nel fenomeno emotivo (cfr. PAR. 1.2); una *componente motivazionale*, che gestisce gli scopi e i bisogni legati alla risposta emotiva; una *componente espressivo- motoria*, abilitata alla manifestazione non verbale delle emozioni; una *componente soggettiva*, che coordina la fase di monitoraggio e riflessione del fenomeno emotivo.

La prospettiva componenziale di Scherer vede come punto focale il concetto di *appraisal*. L'emozione è intesa come una reazione ad eventi emotigeni che predispone l'azione e valuta le differenti alternative. Ad esempio, un'azione caratterizzata da un alto contenuto emotivo come l'aggressione, afferma Scherer, viene predisposta da emozioni quali la rabbia o la paura, tuttavia la loro esecuzione non è determinata soltanto dallo stato emotivo in questione ma da una molteplicità di fattori.

Il modello di Scherer suggerisce che ciascun evento venga valutato dal soggetto in base ad una serie di criteri: questi tengono conto sia dei bisogni fisiologici dell'organismo sia di quelli legati al contesto socio-culturale. Pertanto, alla base del

processo emozionale si instaura il rapporto tra *arousal* (attivazione fisiologica) e *appraisal* (valutazione cognitiva).

Secondo Scherer, il processo di valutazione "does not necessarily require a complex cognitive calculus but often occurs in an automatic, unconscious and effortless fashion"(Scherer, 2009, p. 1309). Lo stimolo esterno, dunque, viene processato attraverso una serie di controlli gestiti gerarchicamente (dai livelli più bassi e semplici a quelli più elevati e complessi) che Scherer definisce *Stimulus Evaluation Checks* (Controlli Valutativi dello Stimolo). Ad esempio, il primo controllo valuta il grado di novità di uno stimolo (*Novelty Check*) e innesca una risposta automatica dell'organismo, spesso necessaria per la sopravvivenza di un individuo. Il secondo controllo (anch'esso primordiale) garantisce una valutazione dello stimolo in base al grado di piacevolezza/spiacevolezza intrinseca (*Intrinsic Pleasantness Check*). Gli altri controlli consentono di valutare se lo stimolo in questione sia in linea con il raggiungimento di uno scopo o se al contrario rappresenti un ostacolo alla realizzazione degli obiettivi (*Goal/need Significance Check*); misurano le potenzialità dell'individuo e le probabilità che questi ha di gestire e affrontare gli effetti di eventi emotigeni specifici (*Coping Potential Check*); valutano infine il rapporto tra lo stimolo, le norme sociali e il concetto di sé, misurandone la compatibilità (*Norm/self-Compatibility Check*).

In sintesi, il modello proposto da Scherer sottolinea le diverse sfumature qualitative che costituiscono uno stato emotivo. Ogni risposta emozionale richiede un numero variabile di controlli, in base al grado di complessità dell'emozione stessa. Un'ulteriore specificazione è individuabile, inoltre, all'interno delle stesse categorie emozionali, in base al tipo di scopo e al grado di potere legato all'evento. Pertanto, non esisterà un solo tipo di rabbia o di tristezza, ma sarà possibile riconoscere le diverse sfaccettature di ciascuna emozione in funzione delle suddette variabili.

#### 1.4.4 Teorie dimensionali

La prospettiva dimensionale, sviluppatasi già a partire dagli anni '50, non risulta totalmente incompatibile con le teorie componenziali. Come spiega Galati (1993, p.17):

[...] l'esperienza emozionale si organizza non sulla base di componenti elementari o di emozioni primarie ma a partire da strutture di organizzazione della risposta molto generali, quali la tendenza all'avvicinamento o all'allontanamento, la valutazione della piacevolezza o spiacevolezza degli eventi che causano l'emozione. Perciò l'individuo imparerebbe a organizzare e a mettere in atto nel corso del suo sviluppo tali modalità generali di risposta. Le singole emozioni risulterebbero tutte secondarie rispetto alle

dimensioni di organizzazione della risposta, da ritenersi invece primarie e innate.

Già Wilhelm Wundt<sup>6</sup> (1897/1907) aveva elaborato una teoria fondata sulla natura dimensionale degli stati emotivi. Nella prospettiva proposta da Wundt, le emozioni sono concepite come dei fenomeni variabili, sensibili a tre differenti dimensioni, organizzate in base ad una scala bipolare di sensazioni (D'Urso, Trentin, 1998): la qualità edonica (piacevole/spiacevole), l'attività (tensione/rilassamento), l'eccitazione (eccitazione/calma)

Ciascun sentimento (*feeling*), sostiene Wundt, può incanalare, in base al contesto, tutte e tre le dimensioni o soltanto due, fino ad essere rappresentata da una sola dimensione. Tuttavia, le tesi di Wundt, nonostante facciano riferimento a discipline scientifiche, non sono il frutto di una ricerca empirica, in quanto riportano delle semplici supposizioni.

Uno dei primi a condurre degli esperimenti in questa direzione fu Harold Schlosberg, all'interno della ricerca sul riconoscimento delle espressioni facciali. I risultati dei suoi studi suggerivano che la percezione di stimoli visivi fosse organizzata intorno a due dimensioni, a partire, cioè, dal grado di piacevolezza o spiacevolezza dello stimolo e dal grado di attenzione o rifiuto verso l'esterno (Schlosberg, 1952).

Le ipotesi di Schlosberg hanno sostanzialmente trovato conferma in ricerche successive, che hanno arricchito lo studio delle dimensioni emotive, estendendolo al campo dell'organizzazione semantica del lessico emotivo (cfr. PAR. 1.5.4)

## **1.5 La componente culturale**

Alla luce di quanto esposto/discusso nei paragrafi precedenti, circa il peso della componente biologica e di quella cognitiva nell'esperienza emozionale, è opportuno riconsiderare questi aspetti all'interno delle variabili culturali, linguistiche e sociali in cui si realizzano.

---

<sup>6</sup> Wundt, fondatore del primo laboratorio di psicologia sperimentale a Lipsia, impostò la propria ricerca seguendo il metodo introspettivo: questo si basava sulla stimolazione percettiva del paziente, chiamato a descrivere ciò che gli accadeva intorno (in modo simultaneo), esprimendo introspettivamente le sensazioni che la propria psiche percepiva. La tecnica introspettiva divenne subito molto popolare nell'ambito scientifico, poiché offriva un punto di vista "interno" al paziente, una sorta di finestra sul suo stato mentale.

In effetti, se consideriamo le emozioni come meccanismi di adattamento all'ambiente circostante, è ragionevole supporre che in condizioni ambientali, culturali e sociali differenti questi meccanismi riflettano, in un certo senso, le scelte e i condizionamenti che i membri di un sistema culturale hanno operato per adattarsi a quel particolare contesto. Ciò naturalmente non esclude, ma anzi moltiplica, la variabilità di queste "scelte culturali" che, pur riflettendo la continuità filogenetica delle emozioni, determina processi di categorizzazione dell'esperienza emozionale diversificati, soggetti a loro volta a norme che ne regolano la manifestazione o l'inibizione.

### 1.5.1 Una doppia eredità

La cultura, come suggerisce Anolli, è un'entità *invisibile*, che esiste, allo stesso tempo, dentro e fuori dalle menti (Anolli, 2006, p. 9). Essa si traduce nelle credenze, nei simboli, nei significati e nelle pratiche comunicative che ogni sistema culturale ha trasmesso ai propri membri attraverso diverse forme di apprendimento, nonché negli artefatti materiali, nelle espressioni artistiche ecc., che sono il risultato di altrettante scelte culturali.

In che modo si realizza la trasmissione di questi saperi e artefatti? Tomasello (2005) postula l'esistenza di una doppia eredità: una biologica, l'altra culturale. Ad entrambe è legato lo sviluppo cognitivo della specie umana, la quale si differenzia da quella dei primati non umani per la sua capacità, biologicamente ereditata, che le consente di adattarsi e vivere secondo cultura (ivi, p. 73). Questa eredità biologica segue quella che Tomasello ha definito la *linea di sviluppo individuale*, ovvero "tutto ciò che l'organismo conosce e apprende da sé", senza la mediazione delle altre persone o di artefatti culturali. Al contrario, l'eredità culturale dipende in buona misura da tutto ciò che un organismo apprende attraverso "il tentativo di vedere il mondo dal punto di vista degli altri" (ivi, p. 72); in base alla teoria proposta da Tomasello, la capacità di comprendere i conspecifici come esseri intenzionali al pari di sé è all'origine di quelle forme di apprendimento culturale (nonché per imitazione) che consentono la trasmissione di saperi e artefatti di generazione in generazione, contribuendo all'evoluzione della specie. L'eredità culturale si fonda infatti su due processi gemelli: la sociogenesi, che consente la creazione di artefatti e di pratiche culturali, e l'apprendimento culturale, che determina la trasmissione di queste creazioni da una generazione all'altra.

Già dalle prime fasi di sviluppo, i bambini si impegnano in protoconversazioni durante le quali lo scambio comunicativo si fonda sulla condivisione dell'attenzione e si realizza attraverso lo sguardo, il contatto fisico, le vocalizzazioni (cfr. Tomasello, 2005, p. 84). Queste interazioni, basate sull'alternanza dei turni, costituiscono il contesto ideale in cui esprimere e condividere emozioni. I bambini ripetono per imitazione i movimenti corporei degli adulti (come tirare fuori la lingua o aprire la bocca). Questa tendenza ad adeguarsi agli stati emotivi degli adulti, rappresenta secondo Stern (1985) un processo di sintonizzazione affettiva nonché un processo di identificazione con i conspecifici.

È tuttavia intorno ai nove mesi che inizia ad emergere una comprensione più profonda degli altri in quanto esseri intenzionali al pari di sé. Come spiega Tomasello, “i bambini cominciano a guardare in modo flessibile e non occasionale là dove gli adulti stanno guardando (seguendo la direzione dello sguardo), a impegnarsi con loro in episodi interazione sociale mediata da un oggetto (impegno congiunto), a usare gli adulti come punti di riferimento sociale e ad agire sugli oggetti allo stesso modo degli adulti (apprendimento imitativo)” (Tomasello, 2005, pp. 83-84). La sintonizzazione dell'attenzione con quella degli adulti in relazione alle entità esterne costituisce il primo passo per comprendere, attraverso un processo di analogia con il sé, che gli altri, in quanto agenti intenzionali, intervengono su queste entità per raggiungere determinati scopi. Una volta consolidata tale cognizione, i bambini iniziano a comportarsi e ad agire sugli oggetti ricorrendo alle strategie degli adulti, a comunicare attraverso i gesti e a sfruttare queste conoscenze per sviluppare ulteriori strategie.

La cultura può essere intesa dunque come un ambiente sociale all'interno del quale si realizza lo sviluppo della specie umana. I membri di una stessa cultura condividono non solo usanze e tradizioni culinarie, stili di vita, abitudini quotidiane, ma anche un certo bagaglio di esperienze in grado di influenzare lo sviluppo cognitivo dei nuovi membri. Il processo attraverso il quale queste pratiche sociali vengono condivise si fonda sulla focalizzazione dell'attenzione: il bambino apprende per *emulazione* il modo in cui un adulto manipola un oggetto, scoprendone caratteristiche e funzioni; per *imitazione* apprende invece le azioni intenzionali degli adulti (Tomasello, *ivi*, p. 105). Se l'attenzione può essere considerata una sorta di percezione intenzionale, vi è dunque alla base una selezione delle possibili prospettive attraverso le quali è possibile osservare la realtà. Gli adulti guidano i bambini verso la comprensione del mondo focalizzandone l'attenzione verso quegli aspetti che il contesto culturale in cui vivono

ritiene funzionali alla salvaguardia/tutela del proprio ambiente sociale. Il processo evolutivo passa anche attraverso la trasmissione di queste pratiche e di questi artefatti culturali, in grado di garantire, attraverso un processo di selezione, la continua evoluzione di comportamenti e strumenti utili alla conservazione della specie.

### *1.5.2 L'acquisizione del linguaggio*

Nelle prime fasi di sviluppo, l'acquisizione del linguaggio ha un ruolo strutturante, in quanto essa non solo implica l'esposizione degli infanti a informazioni culturalmente rilevanti, ma "implica anche la condivisione sociale- la strutturazione culturale- dei modi in cui i bambini prestano attenzione e danno veste concettuale ai vari aspetti del mondo" (Tomasello, 2005, p. 198), attraverso processi di categorizzazione e di assunzione di prospettiva

Le categorie linguistiche consentono ai bambini di interpretare la realtà attraverso l'assunzione di prospettive differenti e di ricorrere efficacemente ai simboli linguistici per fini comunicativi. Anche in questo caso, vi è alla base una selezione delle possibili prospettive attraverso le quali esprimere l'intenzione comunicativa rispetto alle esigenze dell'ascoltatore e, in questo senso, le peculiarità della comunicazione linguistica umana riflettono l'ampio panorama di scelte che ciascun sistema culturale ha a disposizione per descrivere la realtà. "I simboli linguistici sono artefatti simbolici particolarmente preziosi per lo sviluppo del bambino perché esprimono i modi in cui le generazioni che si sono succedute nel tempo in un gruppo sociale hanno trovato utile categorizzare e interpretare il mondo a fini comunicativi" (Tomasello, 2005, p. 27). In un certo senso, si tratta di una forma di selezione, pari a quella che la natura opera, scegliendo tra le variazioni genetiche, quelle che meglio si adattano all'ambiente circostante (cfr. PAR. 1.1). Attraverso i simboli linguistici, il bambino è capace di adottare in modo simultaneo prospettive differenti in relazione ad un medesimo stimolo percettivo (Tomasello, *ibid.*).

Inoltre, come spiega Anolli (2006, p. 73), "l'acquisizione del linguaggio articola ed espande la mente culturale del bambino, che diventa in grado di affrontare situazioni cognitive, emotive e sociali sempre più complesse", giungendo verso i quattro anni di età a formulare delle rappresentazioni mentali non solo di oggetti ed eventi, ma anche

delle rappresentazioni mentali altrui (come prevede la cosiddetta Teoria della Mente<sup>7</sup>, ToM, cfr. Premack, Woodruff 1978), attraverso un processo di identificazione con l'altro.

Partendo da questa considerazione, è logico supporre che se già dalle prime fasi di sviluppo i nuovi membri di una cultura apprendono a orientarsi e a comunicare con il mondo esterno ricorrendo a pratiche culturali e attraverso la natura prospettica del linguaggio, questo graduale processo di adattamento all'ambiente sociale investa anche la sfera emozionale, sia per quel che riguarda la categorizzazione dell'esperienza emotiva sia per quel che concerne i vincoli che ne regolano la manifestazione e lo stile comunicativo da adottare.

### 1.5.3 I modelli culturali

Per spiegare la funzione della cultura all'interno della specie umana, è necessario considerare la cultura non soltanto come un insieme di pratiche, comportamenti, innovazioni e strumenti materiali, ma anche come il modo in cui queste realtà sono organizzate (Anolli, 2006, p. 84). I modelli culturali offrono agli esseri umani una rappresentazione appropriata della realtà, attraverso una chiave di lettura delle esperienze vissute, e orientano le azioni in base al contesto. Costituiscono una sorta di mappa mentale per la comprensione di situazioni ed eventi, ma sono dettati anche da modelli socialmente condivisi che indicano come agire e comportarsi in determinate circostanze.

Essere membri di una cultura significa infatti "condividere gli schemi mentali, i modelli comunicativi, le categorie interpretative e concettuali sottese alle conoscenze *proposizionali* (quelle che riguardano il «che cosa si conosce» e a quelle *procedurali* (quelle riguardanti il «come» si conosce). Ogni cultura si configura, pertanto, come una *costellazione di credenze*, adeguatamente estesa per affrontare i vari aspetti della realtà fisica e sociale" (Anolli, 2006, p. 96). Alcune credenze (definite *centrali* o *primarie*) sono considerate essenziali per la conservazione della propria identità culturale, per il raggiungimento di uno *status* di prestigio nella comunità culturale di riferimento, per cui la loro violazione è comunemente condannata e punita. Altre credenze, definite *secondarie* o *periferiche*, sono invece più suscettibili a modificazioni individuali, per

---

<sup>7</sup> La Teoria della Mente è intesa come "la capacità di «leggere» la mente degli altri (*mindreading*), nonché di interpretare, spiegare e prevedere le loro azioni, attribuendo ad essi stati e processi mentali, quali desideri, modelli interpretativi, credenze e intenzioni" (Anolli, 2006, pp. 73-74).

cui la loro violazione ha un margine di accettabilità e non ha ripercussioni rilevanti nel sistema culturale (cfr. *ibid.*).

Questo complesso sistema di credenze investe anche la sfera emozionale e il lessico emotivo elaborato da ogni sistema culturale per interpretare e indicare situazioni ed eventi comunicativi.

#### *1.5.4 Categorizzazione delle emozioni e lessico emotivo*

I modelli culturali sono in grado di definire la prospettiva emotiva con cui interpretare e comprendere una determinata situazione comunicativa. Come suggerisce Anolli (2006, p. 102):

Non vi sono stimoli che producano in modo costante e universale una specifica emozione. Persone provenienti da culture differenti, avendo diverse valutazioni di un analogo evento, proveranno anche emozioni differenti.

La teoria cognitivo-attivazionale Schacter e Singer (1962, cfr. PAR. 1.4) è stata spesso chiamata in causa per spiegare il ruolo della cultura nel processo di categorizzazione delle emozioni (cfr. Russell, 1991). In effetti, se l'esperienza emotiva di un individuo dipende in buona misura dall'interpretazione cognitiva degli eventi, ovvero da una loro collocazione in una specifica categoria emozionale, il processo di categorizzazione di un'emozione può giocare un ruolo determinante nei processi cognitivi successivi: "If emotion words vary with culture then persons from different cultures might encode, respond to, and remember emotions in correspondingly different ways" (Russell, 1991, p. 417).

Ogni cultura assume una determinata focalizzazione su certi eventi della realtà, scartandone altri. Questa forma di focalizzazione è stata spesso definita *focalità emotiva*, intesa come "disposizione culturale a rispondere in modo elettivo con certe condotte emotive a fronte di specifici eventi" (Anolli, 2006, p. 103). Secondo Levy (1984), la sensibilità emotiva di ogni sistema culturale comporta l'*ipercognitivizzazione* o l'*ipocognitivizzazione* di alcuni fenomeni strettamente legati alla socialità. Ad esempio, un'emozione è ipercognitivizzata quando, oltre ad essere privilegiata all'interno di un contesto culturale, possiede un bagaglio di termini, espressioni linguistiche ed extralinguistiche abbastanza esteso da consentire la manifestazione delle più piccole sfumature di uno stato emotivo, in base al contesto sociale. Per contro, vi sono emozioni ipocognitivizzate che mancano degli stessi mezzi linguistici ed extralinguistici per essere manifestate (ad esempio, il tahitiano non ha un termine o un

concetto per indicare la tristezza, che viene designata piuttosto come una condizione fisica di malattia, indotta da uno spirito maligno, cfr. Anolli, 2006).

Ogni sistema culturale spinge, in un certo senso, verso l'esaltazione di valori, credenze, emozioni utili per i propri meccanismi interni di auto-conservazione. Ad esempio, nella cultura cinese, la vergogna rappresenta un forte strumento di controllo sociale (esistono diversi termini per designarla, in base al contesto di riferimento), mentre l'ansia non trova un corrispettivo lessicale (*ibid.*). Questo processo di focalizzazione determina lo *stile emotivo* di una cultura, poiché orienta la selezione di alcune esperienze emotive e l'esclusione di altre, l'adozione di una certa condotta piuttosto che un'altra.

Seguendo la tradizione darwiniana, alcuni approcci considerano la comunicazione delle emozioni- sul piano dell'espressione e del riconoscimento- parte di un'eredità biologica. Tuttavia, per classificare le emozioni di base, la maggior parte degli psicologi ha fatto appello alle categorie lessicali della lingua inglese (*anger, disgust, joy, fear*), considerandole universali. Ciò non riflette il principio di relatività linguistica (cfr. Whorf, 1956), in base al quale le categorie linguistiche non sono universali, ma sarebbero il prodotto di variabili linguistiche e culturali.

Le ricerche compiute in ambito lessicale (che rientrano all'interno del filone delle *lexical theories*) sottolineano il ruolo preponderante del linguaggio nell'interpretazione degli stimoli esterni e delle reazioni fisiologiche in atto. In altre parole, l'essenza di un'emozione risiederebbe nel modo in cui il soggetto etichetta l'esperienza emozionale; in base a questa prospettiva, sostenere l'universalità di un'emozione implicherebbe la presenza di un'etichetta lessicale corrispondente in tutte le lingue naturali.

La variabilità di ogni sistema culturale ha portato ad una diversa distribuzione e quantità di termini emotivi: "l'estensione del lessico emotivo, infatti, consente di esprimere e discriminare i propri stati effettivi" (Anolli, 2006, p. 101). Lingue come l'inglese o l'olandese hanno migliaia di termini emotivi, mentre ad esempio, il lessico emotivo cinese contiene circa 750 parole, e quello di comunità non alfabetizzate come gli Ifaluk (in Micronesia) contiene poco meno di una sessantina di termini (Russell, 1991). Tuttavia, "il fatto che un'emozione non sia lessicalizzata non vuol dire che non sia provata. L'assenza di una parola non dimostra l'assenza di un'emozione, ma la presenza di una parola significa certamente la presenza di una certa esperienza emotiva" (Anolli, 2006, p.101).

Il termine "emozione" non è universale, ma assume dei connotati specificamente culturali, in base alle prospettive adottate per descrivere uno stato emotivo<sup>8</sup>. Alcune culture possiedono inoltre termini specifici per indicare emozioni non condivise da altri popoli, come la parola giapponese *amae*, che indica lo stato d'animo generato da una dipendenza totale in una relazione affettiva, o il termine *fago*, in uso presso gli Ifaluk, che indica il sentimento che si prova quando muore una persona, quando è malata o parte per un viaggio, ma anche la sensazione piacevole che si avverte in presenza di una persona degna di ammirazione (Russell, 1991). Gli Ifaluk possiedono inoltre due termini diversi (*metagu* e *rus*) per indicare la paura, a seconda che il pericolo sia situato nel presente o nel futuro, così come gli eschimesi distinguono la paura per un danno fisico dal timore verso un danno di natura sociale.

L'organizzazione semantica del lessico emotivo è stata oggetto di indagine (soprattutto da parte degli approcci dimensionali, cfr. PAR. 1.4.4). Due sono i principali metodi di ricerca: il differenziale semantico e lo *scaling* multidimensionale. Il primo approccio, proposto per la prima volta da Osgood (cfr. Osgood, 1952; Osgood, Suci, Tannenbaum, 1957) si basa sull'analisi delle correlazioni tra campi semantici e varie scale di misura. Attraverso un test, ciascun soggetto è chiamato a valutare 20 concetti appartenenti alla propria lingua associandoli a 50 coppie di aggettivi, ad esempio buio-luminoso, pesante-leggero ecc.

L'analisi dimostra che l'agglomerazione degli elementi si deve principalmente a tre fattori (cfr. Galati, 1993): la valutazione (legata alla dimensione *pleasantness/unpleasantness*), l'attività (in riferimento alla dimensione fisiologica; la potenza (ovvero al controllo esercitato dal soggetto nel contesto specifico).

Esperimenti successivi condotti da Russell (cfr. Russell, 1980) e altri ricercatori hanno avallato questa teoria, confermando principalmente l'esistenza delle prime due dimensioni, ovvero valutazione e attività.

Il secondo metodo, legato all'analisi del lessico emozionale, si definisce *scaling multidimensionale* poiché si richiede a ciascun soggetto di valutare le somiglianze all'interno di un campione di termini dal contenuto emotivo. I risultati offrono una spiegazione sull'organizzazione generale dei significati, misurando la vicinanza o la

---

<sup>8</sup> Ad esempio, in italiano, francese e spagnolo, i termini emozione, *émotion*, *emocion*, si riferiscono ad un episodio emotivo specifico, improvviso, e particolarmente forte, e il loro significato è meno ampio rispetto a quello dell'inglese *emotion*. Vi è inoltre una differenza fra il termine inglese *feeling*, che designa anche una condizione fisica, e il termine italiano *sentimento*, che invece si riferisce a uno stato d'animo a livello cognitivo (così come il francese *séñtiment*, o *sentimiento* in spagnolo, cfr. Anolli, 2006).

lontananza spaziale tra le parole utilizzate. Ciò consente, inoltre, di estrapolare gli eventuali fattori di agglomerazione (cfr. Galati, 1993).

In linea generale, le ricerche compiute in ambito lessicale si sono focalizzate per lo più sulla lingua inglese, sebbene alcune abbiano incluso nella loro analisi altre lingue. Dai risultati ottenuti, osserva Galati, si evince che "non tutte le lingue, soprattutto se molto lontane tra loro, posseggono lo stesso bagaglio terminologico-concettuale per esprimere le emozioni"(ivi, p. 187).

#### *1.5.5 Il ruolo della cultura nella manifestazione delle emozioni*

Come già sostenuto da Darwin, i movimenti facciali esulano, in una certa misura, dalla volontà dell'individuo, in virtù del loro carattere ereditario. Tuttavia, in alcune condizioni gli esseri umani sono in grado di esercitare un controllo più profondo - e soprattutto cosciente- sui movimenti facciali, modificandoli o adattandoli al contesto comunicativo in cui si trovano.

A tal proposito, alcuni autori (cfr. Buck; 1984; 2007) postulano l'esistenza di una comunicazione spontanea e di una comunicazione simbolica. La comunicazione spontanea è determinata biologicamente e si fonda su contenuti non verbali e non falsificabili (Buck, 2007, p. 93).

Al contrario, la comunicazione simbolica (volontaria) "is learned and culturally patterned, its elements are symbols which bear an arbitrary relationship to the referent [...] and its content consists of falsifiable statements or propositions" (ibid.).

In linea di massima, questi due livelli di comunicazione coesistono all'interno degli scambi comunicativi, sebbene il loro peso vari in funzione del contesto. Ad esempio, in situazioni formali, la comunicazione simbolica prevale su quella spontanea, viceversa in circostanze legate all'intimità, si impone quella spontanea, lasciando spazio all'aspetto emozionale.

In altre circostanze, la manifestazione di un'emozione può essere deliberata e cosciente, e può anche non trovare corrispondenza con l'effettiva dimensione interiore di un individuo. Questo tipo di comunicazione è stata definita pseudospontanea (Buck, VanLear, 2002), poiché "while it is voluntary from the sender's point of view, it uses the display mechanism that can stimulate preattunements in the receiver, so if the receiver is taken in it is as if it were a veridical display" (Buck, 2007, p. 93).

In quest'ottica, l'espressione delle emozioni può essere veicolata da una comunicazione simbolica, spontanea o pseudo-spontanea. Il contesto- sociale e

culturale- gioca un ruolo determinante in relazione al tipo di emozione provata e alla sua intensità. Le circostanze esterne in grado di stimolare una reazione emotiva possono variare in funzione delle norme culturali e dei comportamenti socialmente approvati. Ne consegue che: a) non tutte le emozioni possono essere apertamente manifestate; b) alcuni contesti richiedono l'esibizione di un determinato stato emotivo conforme al tipo di circostanza in cui ci troviamo; c) il background culturale influisce sull'applicazione dei principi esposti nei punti a) e b).

Gli studi svolti da Ekman e Friesen (1969; 1976; 1978)- basati sul riconoscimento delle espressioni facciali da parte di valutatori di diversa origine ed estrazione sociale hanno evidenziato il peso della componente culturale all'interno del fenomeno emotivo, postulando l'esistenza di alcune norme - definite *display rules* - in grado di influenzare la manifestazione di un'emozione. In quest'ottica, le espressioni facciali delle emozioni di base (cfr. PAR. 1.6.1), pur rimanendo universali, sarebbero soggette ad una serie di regole espressive culturalmente e socialmente determinate. Sono regole che misurano l'intensità, la valenza, la valutazione sociale e il destinatario dell'emozione in questione e ne determinano l'esibizione o la inibiscono: "A display rule specifies who can show what emotion to whom, when" (Ekman, 1982, p. 62).

Le emozioni primarie si distinguerebbero, secondo Ekman (1992), per una serie di caratteristiche, tra cui l'attivazione di una risposta fisiologica specifica, la presenza di una valutazione automatica dello stimolo (*automatic appraisal*), l'esistenza di antecedenti universali, la rapidità con la quale si attivano e la breve durata della loro manifestazione. Ekman fa riferimento alle emozioni di base tenendo conto delle variabili che ne influenzano l'intensità, l'esibizione o l'inibizione; per tale ragione, ogni emozione primaria costituirebbe una "famiglia" i cui membri rappresentano delle variazioni di uno stesso stato emotivo. Ad esempio, la famiglia della collera comprenderà la rabbia, l'irritazione ecc. (che avranno caratteristiche simili nella conformazione delle strutture muscolari).

#### *1.5.6 Stili comunicativi*

La cultura agisce anche sul piano comunicativo, determinando modalità differenti non solo nella manifestazione delle emozioni (o nella loro inibizione), ma anche nella gestione degli stati emotivi sul piano sociale. Ad esempio, nelle culture orientali, diversamente da quelle occidentali, grande importanza viene conferita ai titoli onorifici, e al ricorso di espressioni che indicano referenza e rispetto dell'autorità. La

violazione di queste norme comporta una seria violazione dei vincoli e delle gerarchie sociali.

Anche il peso della comunicazione verbale e non verbale tende a variare da cultura a cultura. Come già aveva osservato l'antropologo americano Hall (1966/1968), sin da piccoli, in virtù del contesto culturale, i nuovi membri di un sistema culturale sono abituati a cogliere un certo tipo di informazioni dall'ambiente circostante e a scartarne altre. Questo tipo di "selezione" inconscia e automatica abbraccia anche il piano della comunicazione, in quanto "quando parliamo formuliamo solo una parte del messaggio: il resto è riempito dall'ascoltatore. Molto di ciò che non è detto, viene capito al volo, anche se quel che rimane tacito varia da cultura a cultura" (Hall, 1966/1968, p. 139). In particolare, Hall mette in evidenza le modalità differenti attraverso cui un sistema culturale costruisce gli scambi comunicativi, in base al grado di importanza data al *contesto*, ovvero alle informazioni globali che racchiudono un evento e che determinano il significato dell'evento stesso. Secondo Hall, il grado di importanza del contesto all'interno della comunicazione varia da cultura a cultura, differenziando le culture ad alto contesto comunicativo (*high-context cultures*) da quelle a basso contesto comunicativo (*low-context cultures*).

"A high context (HC) communication or message – afferma l'antropologo – is one in which most of the information is either in the physical context or internalized in the person, while very little is in the coded, explicit, transmitted part of the message" (Hall, 1976, p.79). Nelle culture HC, buona parte del significato all'interno di una conversazione è veicolato in modo non verbale, attraverso i gesti e la mimica facciale che arricchiscono il contenuto verbale di elementi di significato e sopperiscono in alcuni casi, alla mancanza di chiarezza trasmessa dalle parole. La conversazione, infatti, si sviluppa maggiormente sul piano del non detto, del sottinteso. I membri di questo tipo di cultura sono in sintonia con il proprio ambiente e pertanto riescono ad esprimere e ad interpretare le emozioni anche in modo non verbale (cfr. Samovar, Porter, McDaniel, 2007). Ad esempio, una cultura di tipo collettivista come quella asiatica, si caratterizza per una serie di modalità comunicative che enfatizzano la gerarchia dei ruoli piuttosto che lasciare spazio all'espressione dell'individuo e alla propria emotività.

Di contro, "a low context (LC) communication is just the opposite; i.e. the mass of the information is vested in the explicit code" (Hall, 1976, p. 79). Le culture a basso contesto comunicativo presentano generalmente una grande diversità all'interno delle comunità; ciò comporta un'attenzione maggiore verso i dettagli in grado di fornire

informazioni precise sul contesto comunicativo. Il messaggio verbale viene espresso in modo esplicito, e poco viene affidato ai segnali non verbali o al contesto fisico.

Naturalmente, ogni cultura presenta degli aspetti legati ad entrambi i tipi di comunicazione (ad alto e basso contesto comunicativo). Allo stesso modo, vi sono *culture della distanza* e *culture della vicinanza* (cfr. Hall, 1966, 1976; Anolli, 2006). Ad esempio, i membri di culture come quella nordeuropea e indiana assumono, durante uno scambio comunicativo, una grande distanza interpersonale rispetto all'interlocutore, evitando il contatto visivo (che mantiene un'angolazione obliqua) e riducendo il contatto corporeo. Per contro, in culture come quelle latine, arabe, sudamericane, uno scambio comunicativo si fonda sulla compartecipazione spaziale, visiva, che ammette spesso il contatto fisico tra gli interlocutori.

Alla luce di quanto esposto nei paragrafi precedenti, risulta evidente quanto sia difficile separare la natura umana dalla nicchia ecologica in cui si sviluppa. La cultura opera influenzando e intervenendo su processi biologici, sociali, cognitivi e linguistici, creando le basi per lo sviluppo di un popolo. L'insieme di questi processi presenta degli elementi simili e parzialmente condivisi sul piano interculturale, sebbene essi tendano ad assumere connotati specificamente culturali e strettamente legati al contesto sociale di riferimento.

## **1.6 La componente espressiva**

La fenomenologia delle emozioni si caratterizza per le sue molteplici modalità di espressione, che interessano aspetti diversi della comunicazione non verbale (il volto, i gesti, la postura del corpo, la voce).

### *1.6.1 L'espressione facciale delle emozioni*

Ekman e Friesen (1976) hanno largamente contribuito all'individuazione dei movimenti facciali implicati nell'espressione delle emozioni, attraverso l'elaborazione di un sofisticato sistema di analisi (*Facial Action Coding System*). Questo si basa sulla codifica di ogni movimento muscolare, definito *unità di azione facciale*, non necessariamente associato ad un'emozione specifica. L'osservazione avviene mediante l'applicazione di elettrodi sui muscoli del viso, che consentono di determinare in che misura ciascun muscolo contribuisca all'alterazione del volto. Le aree interessate sono

rintracciabili nella parte superiore del volto (fronte, sopracciglia e occhi) e in quella inferiore (naso, bocca, mento e guance).

I movimenti facciali vengono dunque interpretati sulla base dell'attivazione di un muscolo o di un gruppo di muscoli. Il sistema permette inoltre di individuare l'intensità espressiva- minima, media o massima- che caratterizza un'espressione emotiva.

Nella maggior parte dei casi, le emozioni facciali sono il risultato della compartecipazione di diversi gruppi muscolari; raramente un'espressione può essere veicolata da una sola unità. Ciò permette di determinare le diverse sfumature di significato all'interno di uno stesso stato emozionale; ad esempio, il sollevamento delle palpebre sarebbe discriminante nella distinzione tra la gioia moderata e l'esultanza; la costrizione delle palpebre contribuirebbe a differenziare la collera fredda (più contenuta) dalla collera calda (Galati, 1993).

Per individuare quali segnali del viso fossero portatori di significato emotivo, Boucher e Ekman (1975) hanno somministrato ad un campione di valutatori una serie di fotografie che riproducevano soltanto una parte del volto (fronte, occhi o bocca), chiedendo di valutare l'efficacia di ciascuna parte nella trasmissione delle emozioni considerate. Dai risultati, è emerso che la fronte è un segnale decisivo nell'espressione della sorpresa, gli occhi lo sono per la tristezza e la paura, mentre la bocca lo è per il disgusto e la felicità.

### *1.6.2 I gesti*

Ekman e Friesen hanno incluso nelle loro ricerche lo studio dei comportamenti non verbali, proponendo una classificazione sulla base della loro funzione. Ad esempio, gli emblemi sono gesti culturalmente codificati, che assumono un valore specifico in virtù del contesto culturale di riferimento. Naturalmente, in alcuni casi, è possibile che una cultura adotti l'emblema di un'altra, o che ne comprenda comunque il significato (Ekman, 2004); i gesti di illustrazione (movimenti deitici riferiti ad un oggetto, o riproduzione di un'azione corporea ecc.) sono solitamente utilizzati per enfatizzare il discorso e contribuiscono a rendere più fluida l'esposizione di un concetto o di un pensiero; alcuni gesti, come annuire con il capo, sorridere, sollevare le sopracciglia, accostarsi all'interlocutore ecc. costituiscono invece un segnale di attenzione e di accordo nei confronti dell'interlocutore. Altri gesti, al contrario, non hanno necessariamente uno scopo comunicativo, ma sono funzionali al superamento di una

condizione di disagio o imbarazzo, come i gesti di manipolazione, legati al contatto di una parte del corpo (ad esempio, le mani) con un'altra (il viso) o con un oggetto.

### 1.6.3 La relazione tra emozione e voce

Sia nella specie umana che in quella animale, la voce costituisce uno strumento potente a fini comunicativi, come già aveva evidenziato Darwin. Anche l'espressione vocale delle emozioni è implicata, secondo Darwin, nel processo di evoluzione. Egli, infatti, ricollegava l'abitudine di emettere "suoni musicali" ad atavici rituali di corteggiamento, accompagnati dall'espressione di emozioni intense quali l'amore, la rivalità e il trionfo: " [...] i progenitori dell'uomo probabilmente emettevano suoni musicali già prima di aver acquisito la capacità di articolare la parola; e di conseguenza, quando la voce viene usata sotto la spinta di una qualche forte emozione, essa tende ad assumere- per il principio dell'associazione- un carattere musicale"(Darwin, 1872/1982, p. 186).

Come spiegano D'Urso e Trentin: "la relazione tra voce ed emozione è basata sull'assunzione che le reazioni fisiologiche tipiche di uno stato emotivo, modificando il respiro, la fonazione e l'articolazione dei suoni, producano delle variazioni apprezzabili negli indici acustici rilevabili nella produzione del discorso"(D'Urso, Trentin, 1998, p.58)

La voce umana, tra le sue molteplici potenzialità, è in grado di veicolare significati e sfumature differenti attraverso la propria modulazione. Tali modificazioni prosodiche, che rappresentano delle entità soprasegmentali, si fondono con le caratteristiche segmentali proprie di ogni lingua.

Poggi e Magno Caldognetto (2004), hanno classificato le risorse segmentali che regolano la natura linguistica del parlato emotivo, distinguendole in risorse lessicali, sintattiche, morfologiche e fonologiche. Le risorse lessicali, ossia verbi, aggettivi, avverbi, sostantivi ed interiezioni, veicolano il contenuto semantico e pragmatico e esplicitano il riferimento a stati d'animo specifici (*mi sento in colpa, sono triste, arrabbiato* ecc.). In particolare, le interiezioni (*toh, maledizione, accidenti, uffa* ecc.) veicolano il messaggio emotivo senza spiegarne la causa, che deve essere necessariamente inferita dal contesto; sul piano sintattico, il parlato emotivo si caratterizza anche per un'alterazione della struttura e dell'organizzazione degli enunciati, un esempio è la dislocazione a sinistra, il cui impiego implica la scelta di enfatizzare un determinato elemento del discorso; sul piano morfologico, la struttura delle parole può

subire un'ulteriore trasformazione nella produzione emotiva. Ad esempio, è molto comune l'uso di diminutivi e vezzeggiativi (in relazione ad emozioni positive, come la gioia) e di dispregiativi (nel caso di emozioni negative, come la rabbia).

Sul piano fonologico, l'articolazione dei suoni risulta alterata dal vissuto soggettivo sperimentato in un determinato contesto. Tali alterazioni possono essere attribuite, ad esempio, al parlare sorridendo o ad altri aspetti salienti della comunicazione non-verbale. In particolare, il rapporto tra i processi articolatori dei singoli suoni e i fenomeni emotivi è stato oggetto di studio da parte della linguistica e della psicologia (cfr. Trojan F., 1962; Fonagy, 1989; Laver, 1980; Scherer, 1984).

Le ricerche portate avanti da Trojan (1948;1962) hanno suggerito l'idea che la componente fonica della comunicazione, pur appartenendo al sistema più evoluto, presenti degli aspetti legati al sistema di comunicazione più antico, basato sull'attivazione di risposte fisiologiche attraverso le quali siamo in grado di veicolare stati d'animo, intenzioni e atteggiamenti individuali (Fónagy, 1972; Balconi, 2000). Le caratteristiche articolatorie osservate da Trojan hanno messo in evidenza alcuni tratti distintivi (come la pressione espiratoria e la tensione muscolare, l'apertura fauce, l'emissione di fiato, la nasalità ecc.) dei correlati articolatori di determinate emozioni, influenzando l'impostazione degli studi successivi (cfr. Trojan, 1962; Fónagy, 1972; Laver, 1980; Scherer, 1986).

L'analisi di Fonagy (1983) sulla correlazione tra movimenti articolatori e espressioni emotive suggerisce che l'articolazione delle occlusive, implicando uno sforzo muscolare intenso per trattenere il flusso d'aria, seguito dal repentino rilascio, possieda "caratteri isomorfici con i vissuti di violenza, collera, aggressività e ostilità. Pertanto la voce della collera è strozzata, quasi che il meccanismo articolatorio del suono configurasse una rappresentazione mimetica dell'atto aggressivo"(Balconi, 2002, p.14).

Sempre in relazione alla tensione muscolare e all'articolazione dura, il fremito o il tremore, associati ad atteggiamenti aggressivi, sono restituiti sul piano fonico soprattutto dalla vibrante /r/, che ricorre spesso nel lessico della collera, riflettendo questo vissuto soggettivo. Viceversa, gli stati emotivi di delusione sarebbero caratterizzati da un generale rilassamento della lingua e del velo.

Per ciò che riguarda l'emissione di fiato, essa può indicare uno stato di profonda eccitazione psichica, dovuta anche alla difficoltà nel gestire l'emozione e alla necessità di proiettarla all'esterno, ed è dunque associata a sentimenti di rifiuto e insofferenza

(Balconi, 2002). Sul piano fonico, questa associazione si rintraccia soprattutto nelle fricative (/v/ e /f/), presenti in alcune onomatopee (come *uffa*).

Alla base di queste osservazioni, vi sono due prospettive teoriche contrapposte che partono da un punto di vista comune: l'associazione tra suono e senso sarebbe dovuta a una precedente congruenza tra il piano dell'espressione e quello del contenuto, ovvero una congruenza isomorfica (in cui cioè segno e significato appartengono alla stessa categoria). Secondo la prospettiva della psicologia della Gestalt, di impronta innatista, queste affinità sarebbero di natura strutturale, dunque il legame tra suono e significato non sarebbe appreso, ma deriverebbe all'esperienza percettiva suscitata dallo stimolo acustico in sé (Balconi, 2000). Al contrario, la prospettiva adottata nell'ambito delle teorie dell'apprendimento (Osgood, 1980; Cytowic, 1989) suggerisce che l'associazione tra le caratteristiche fonetiche e lo stimolo percepito sia il risultato dell'apprendimento di elementi ricorrenti in contesti sensoriali differenti, sulla base dei quali un individuo costruisce dei nessi associativi stabili (cfr. Balconi, 2000; 2002).

#### *1.6.4 Le informazioni trasmesse dalla voce*

Giles, Scherer e Taylor (1979) hanno identificato una serie di segnali extralinguistici strettamente connessi con le informazioni trasmesse dalla voce. Questi indizi, definiti *speech markers*, operano su due livelli e ricoprono funzioni differenti. Le informazioni trasmesse sul primo livello rivelano indizi sull'identità del parlante all'interno di categorie biologiche (come il genere e l'età), geografiche e sociali (come la provenienza e la condizione sociale) e psicologiche.

Molte di queste informazioni vengono trasmesse e decodificate in modo automatico, senza che vi sia la volontà di comunicarle. Ad esempio, il genere e l'età di un individuo sono facilmente desumibili dall'aspetto fisico; tuttavia, la voce rivela indizi sull'identità del parlante in relazione alla sua provenienza geografica, permettendo una categorizzazione del soggetto all'interno delle gerarchie e dei ruoli sociali. Come fanno notare gli autori "Level 1 speech markers then serve the general function of maintaining the social function by identifying and recognizing members who occupy various roles and hierarchical position within it" (ivi, p. 343).

A un secondo livello, questi marcatori assumono una certa salienza nelle interazioni sociali, in qualità di indicatori sintomatici di attitudini e stati psicologici (ivi, p. 345). Nel corso di uno scambio comunicativo, entrano in gioco meccanismi delicati che toccano da vicino l'intimità e la vita privata di un individuo; il ricorso a questi

segnali codificati garantisce una valida scorciatoia per veicolare e interpretare indizi relativi allo stato psicologico del parlante, senza che vi sia la necessità di esplicitare informazioni potenzialmente auto-compromettenti (*ibid.*).

La tendenza a evitare questioni eccessivamente personali è probabilmente legata alla necessità di difendere la propria identità e di salvaguardare l'immagine di sé. Questi meccanismi rientrano nel concetto di "faccia", proposto da Goffman, inteso come "il valore sociale positivo che una persona rivendica per se stessa mediante la linea che gli altri riterranno che egli abbia assunto durante un contatto particolare" (Goffman, 1971/1981, p. 7). In altre parole, il valore della "faccia" non corrisponde all'idea che un individuo ha di se stesso, ma riflette piuttosto la considerazione che gli altri dovrebbero avere di quest'ultimo, in relazione alle qualità positive rivendicate. Come spiega De Marco (2010, p. 115): "La rivendicazione dell'immagine di sé è qualcosa che gli altri devono riconoscere, sia implicitamente che esplicitamente, e cioè riconoscere che l'individuo possenga le qualità che rivendica come proprie». Ne consegue che "il problema della perdita dell'immagine sorge quando le attese rispetto al suo riconoscimento vengono eluse" (ivi, p. 116). Ciò implica una valutazione negativa della persona e rappresenta una minaccia per l'immagine di sé.

In sostanza, "the subtle use of speech markers indicative for these psychological states may be the mechanism whereby these emotionally charged ideas are brought inferentially to the attention of another more appropriately" (Giles, Scherer, Taylor, 1979, p. 6). Naturalmente, il ricorso a questi indicatori varia in funzione del contesto sociale e culturale e dipende in buona misura dal tipo di rapporto che lega gli interlocutori.

I marcatori di secondo livello si inseriscono nella dimensione pragmatica della lingua, rivelando la tendenza dei parlanti a creare condizioni favorevoli alla buona riuscita dello scambio comunicativo, adattandosi al tipo di interlocutore e condividendo, eventualmente, il ricorso agli stessi segnali. Ad esempio, "one of the ways of doing so is to make one's speech more likely the person being addressed. [...] These 'convergences' [...] occur in such a manner that the sender become linguistically more similar to his receiver" (ivi, p. 369). In altre parole, l'adozione di *pattern* comuni all'interlocutore facilita le affinità sul piano cognitivo, promuovendo la comprensione e dunque la riuscita dell'interazione.

Nella comunicazione parlata, i marcatori possono contribuire tanto al successo dell'interazione, quanto alla rottura delle aspettative formulate in precedenza dagli

interlocutori. Gumperz (1992) classifica come spie di contestualizzazione "any verbal and non verbal signs that helps speakers hint at, or clarify, and listeners to make such inferences"(ivi, p. 229). Questi segnali includono aspetti prosodici (tra cui l'intonazione), e paralinguistici (velocità del parlato, vocalizzazioni, risa ecc.), ma anche stilistici e lessicali (il ricorso ad un particolare stile comunicativo e l'adozione di espressioni o modi di dire ad esso conformi). La relazione che lega questi segnali al contesto in cui vengono prodotti è molto forte e si basa dunque su un sottile gioco di inferenze e di rimandi alle conoscenze condivise dai partecipanti. In un certo senso, le spie di contestualizzazione "serve to highlight, foreground or make salient certain phonological o lexical strings vis-à-vis other similar units, that is, they function relationally and cannot be assigned context-independent, stable, core lexical meanings" (ivi, p. 230). Il sistema di attese viene eluso quando questi segnali creano una discrepanza tra gli elementi del contesto; ad esempio, un'intonazione sarcastica, una pausa significativa, o la prominenza data a una parola rispetto alle altre, possono determinare, attraverso un processo di tipo inferenziale, un cambiamento di prospettiva rispetto a quanto precedentemente affermato o supposto. Ciò naturalmente influenza il modo in cui percepiamo l'atteggiamento degli altri (sicuro o insicuro, cortese o maleducato, scostante o socievole).

#### *1.6.5 La comunicazione emotiva*

Concludiamo questo capitolo, proponendo una chiave di lettura dei fenomeni espressivi visti nella loro dimensione pragmatica, tenendo conto del contesto in cui si realizzano e degli scopi sottesi alla loro manifestazione. Questa prospettiva abbraccia globalmente la visione della lingua considerata nel suo contesto d'uso, in cui il successo (o il fallimento) di uno scambio comunicativo è determinato da una combinazione di fattori legati alla situazione comunicativa in atto e alla necessità dei parlanti di raggiungere determinati scopi.

Quando parliamo di scopi, è chiaro il riferimento ad un preciso atto di volontà esercitato da un parlante nel corso di uno scambio comunicativo, che si traduce nell'impiego consapevole di risorse non verbali in grado di veicolare atteggiamenti e stati d'animo. Anche in questo caso, non si tratta di una situazione ideale, in cui la buona riuscita dell'interazione è garantita dalla capacità dei parlanti di codificare e decodificare perfettamente i rispettivi comportamenti verbali e non verbali.

L'idea di un tipo di comunicazione fondata sull'impiego strategico di risorse linguistiche ed extralinguistiche al fine di influenzare le reazioni e i comportamenti altrui sulla base di informazioni di natura affettiva fu elaborata da Marty nel 1908 (Marty, 1908). Il filosofo svizzero distingue due tipi di comunicazione: la *comunicazione emotiva*, in cui l'obiettivo del parlante riveste un ruolo strategico all'interno dello scambio comunicativo, e la *comunicazione emozionale*, in cui la spontaneità delle informazioni affettive trasmesse è maggiormente legata ad una necessità psicofisica del parlante che alla volontà di modificare le reazioni e il comportamento altrui.

La comunicazione emozionale non si fonda necessariamente sull'impiego di strategie linguistiche ed extra-linguistiche ma ricorre agli stessi canali (verbale e non verbale) nella manifestazione di reazioni involontarie o automatiche ad uno stimolo esterno. Un esempio è costituito dalle improvvise manifestazioni di pianto, urla, risa spesso accompagnate da esclamazioni (Ah!, Eh!): la trasmissione di queste informazioni non è legata alla costruzione e all'espressione dell'immagine di sé quanto piuttosto al bisogno primario, catartico, di comunicare la propria reazione ad un evento esterno.

Come ricorda Bonelli (2013), questo tipo di comunicazione, così intesa, è rintracciabile in porzioni limitate del discorso e non si estende a intere enunciazioni, in quanto le esigenze comunicative del parlante seguono in un certo senso coordinate deittiche, legate a un momento preciso dell'enunciazione, per cui la reazione di sorpresa, di disgusto, di paura, assume un carattere intrinsecamente liberatorio per il soggetto esperiente (e non è dunque destinata all'interlocutore).

La comunicazione emotiva, al contrario, viene elaborata a partire dal punto di vista del destinatario, tenendo conto cioè delle possibili reazioni a determinati stimoli di natura affettiva (anche soltanto un sorriso, uno sguardo, un sospiro ecc.). Ciò presuppone un uso consapevole di mezzi e strategie capaci di indirizzare l'attenzione dell'altro verso una precisa direzione, in modo da influenzarne il comportamento (cfr. anche Caffi, Janney, 1994; Caffi, 2001; Bonelli, 2015). Queste strategie sono reciprocamente adottate da ciascun partecipante per costruire e negoziare significati, stimolando la condivisione di emozioni e stati d'animo. Il potere evocativo di questi espedienti è innescato dagli indizi di natura non verbale e paraverbale di cui si compone essenzialmente la comunicazione emotiva.

I membri di una cultura ricorrono a questi marcatori linguistici ed extra-linguistici per evocare significati e suggestioni e per rispondere efficacemente a quelli

altrui, sulla base di segnali in gran parte codificati culturalmente. Su questa cooperazione si fondano essenzialmente le dinamiche della comunicazione emotiva. Uno scambio comunicativo efficace si basa inoltre sulla coordinazione e la sintonia di elementi verbali e non verbali e sulla condivisione di conoscenze da parte degli interlocutori.

## **1.7 Conclusioni**

Alla luce delle questioni sollevate in questo capitolo, risulta evidente la natura complessa e variabile dei fenomeni emotivi. Le ricerche compiute in ambiti diversi, da quello antropologico, a quello psicologico, linguistico e sociale, hanno tentato, a più riprese, di sostenere l'universalità o la specificità culturale di questa componente essenziale della specie umana. Il discorso, come già anticipato all'inizio di questo capitolo, va al di là della semplice antitesi natura/cultura o innato/appreso. Questa dicotomia limita le molteplici prospettive attraverso le quali è possibile osservare il fenomeno emotivo, disconoscendo la profonda relazione tra geni e ambiente. È la cultura stessa a creare una molteplicità di prospettive in relazione all'esperienza emotiva e a determinare, fra le varie scelte possibili, quelle che maggiormente si adattano al suo funzionamento e alla conservazione della sua identità. Poiché si tratta di un processo in continuo sviluppo, strettamente connesso con l'evoluzione della specie umana, appare forse riduttivo cercare di separare l'uno dall'altra; sembra invece più auspicabile il tentativo di osservare il modo in cui processi biologici e variabili culturali intervengano nell'esperienza percettiva, fornendo all'uomo una chiave di lettura (sebbene parziale e limitata) della realtà.



# Capitolo 2

---

## Il parlato emotivo

### 2.1 Introduzione

Lo studio della relazione tra emozione e voce è relativamente recente nel panorama scientifico attuale, che ha a lungo privilegiato altri ambiti di ricerca, tra cui quello delle espressioni facciali, trascurando, in certa misura, il peso della componente vocale nella comunicazione delle emozioni. Questo "ritardo" forse spiega l'assenza di un modello teorico dominante in grado di conciliare posizioni antitetiche e di una metodologia accreditata e condivisa, capace di garantire un certo controllo non solo sul materiale oggetto di analisi, ma anche sulle condizioni in cui viene ottenuto, sulle procedure di elicitazione e sui parametri indagati.

Le ricerche condotte negli ultimi anni (cfr. Albas, McCluskey, Albas 1976; Van Bezooijen 1984; Scherer, Banse, Wallbott 2001; Chen, Gussenhoven, Rietveld 2004; Thompson, Balkwill 2006; Barrett, Bryant 2008; Goudbeek; Broersma 2010) hanno adottato un approccio interculturale, volto ad indagare gli aspetti universali e gli aspetti specificamente culturali che caratterizzano la comunicazione vocale delle emozioni nella specie umana. Il percorso intrapreso da queste ricerche si riflette sostanzialmente in due profili di indagine che permettono di osservare il parlato emotivo da due differenti prospettive. Il primo approccio (*speaker-centered*) privilegia il piano della produzione (*encoding*) nel tentativo di analizzare la struttura acustica del parlato emotivo e le variazioni prosodiche che veicolano emozioni, stati d'animo e atteggiamenti. Il secondo approccio si focalizza invece sul processo di decodifica (o *decoding*) del parlato emotivo, privilegiando la prospettiva dell'ascoltatore (*listener-centered*). L'intento è quello di osservare i meccanismi legati alla percezione e al riconoscimento delle emozioni a partire dagli indizi vocali e dalle modulazioni prosodiche del parlato. Alcuni approcci sono orientati verso un'analisi di tipo globale, che include entrambi i processi di codifica e decodifica, allo scopo di indagare la comunicazione vocale nella sua complessità; tuttavia, sono poche le ricerche che ad oggi hanno optato per un'indagine acustico-percettiva. L'adozione dell'una o dell'altra

prospettiva (o di entrambe) incide naturalmente sull'impostazione metodologica delle indagini e sugli obiettivi della ricerca.

In questo capitolo, saranno introdotti i principali nodi teorici (PAR. 2.3) che fanno capo allo studio delle emozioni vocali, e in generale al rapporto tra il piano linguistico e quello paralinguistico. In seconda battuta, saranno presi in esame gli approcci (discreti, componenziali e dimensionali) maturati in questo ambito di ricerca a partire dalle teorie sviluppatesi in ambito psicologico e cognitivo, e i risultati empirici riportati da questo tipo di ricerche. Successivamente, ampio spazio sarà dato allo scenario metodologico (cfr. PAR. 2.4) che caratterizza lo studio sul parlato emotivo, sia sul piano della codifica sia su quello della decodifica. Uno degli aspetti maggiormente discussi riguarda la natura stessa delle emozioni indagate. Alcuni studi hanno impiegato produzioni emotive spontanee o semi-spontanee, coinvolgendo parlanti *naïf* all'interno di un contesto naturale. Altri studi, invece, hanno optato per un parlato emotivo simulato, avvalendosi di attori o parlanti ingenui. Questo passaggio metterà in luce i punti di forza e i limiti di ciascun approccio. Infine, il parlato emotivo verrà considerato in una prospettiva interlinguistica e transculturale (cfr. PAR. 2.7) che, attraverso una ricognizione dello stato dell'arte, ha lo scopo di evidenziare gli aspetti linguistici, culturali e sociali che regolano la comunicazione delle emozioni sul piano vocale.

## **2.2. Per una definizione dei segni vocali non verbali**

Le informazioni trasmesse dai segni vocali non verbali contribuiscono a veicolare, in modo involontario o consapevole, indicazioni rilevanti sulle caratteristiche intrinseche del locutore (età, genere, provenienza geografica, condizioni di salute, ecc.) e indizi importanti circa le sue intenzioni comunicative, lo stato d'animo in cui si trova, ecc. (cfr. PAR. 1.6.4). Queste informazioni sono legate ad aspetti quali la qualità della voce, le variazioni di tono, la durata, l'intensità, la presenza di altri elementi come il riso o il pianto. L'insieme di questi indizi vocali rientra nel novero dei fenomeni paralinguistici.

Una prima classificazione di questi fenomeni fu elaborata da Hill (1952), che incluse nella definizione di paralinguaggio anche altri indizi non verbali, veicolati dai movimenti del corpo. Come ricordano Anolli e Ciceri (1992), fu Trager (1958) a proporre una definizione più precisa e puntuale, includendo soltanto i segni vocali non

verbali e distinguendo le caratteristiche extralinguistiche, legate cioè al tipo di voce, dalle caratteristiche paralinguistiche, in cui includeva le qualità vocali, il tempo, la durata, il controllo della tonalità, ecc.

La classificazione elaborata da Crystal (1969) qualche anno più tardi, distingueva invece: a) i segnali verbali o segmentali; b) i segnali prosodici e paralinguistici (o soprasegmentali); c) i segnali non linguistici.

Con segnali prosodici si riferiva alle variazioni di tono, di volume, di durata (intonazione e accento) che hanno una funzione contrastiva, per cui sono da considerarsi parte integrante del codice linguistico (Anolli, Ciceri, *ivi*, p. 117). Le caratteristiche paralinguistiche, definite effetti vocali, fanno riferimento a fenomeni percepiti come aventi tono, volume, durata, ma che sono per lo più il risultato di meccanismi fisiologici derivanti dall'attività della faringe o delle cavità nasali e orali (*ibid.*), come il bisbiglio, il falsetto, ecc.

Le categorie proposte da Crystal sembrano tuttavia poco soddisfacenti poiché "se i tratti paralinguistici differiscono da quelli prosodici in quanto non sono altrettanto strettamente integrati con la struttura grammaticale dell'eloquio, ciò che potrebbe essere considerato prosodico in una data lingua potrebbe essere trattato come paralinguistico in un'altra, in forza della diversa funzione linguistica che in essa svolge" (*ivi*, p. 119).

In seguito, la classificazione dei fenomeni vocali è stata elaborata a partire da due criteri di analisi: uno di natura funzionale, l'altro di natura fisica e acustica. Il primo criterio consente una classificazione dei segnali in base ad una scala di linguisticità: la componente verbale risulta la più linguistica (nel senso che è necessaria al codice linguistico); i tratti prosodici vengono considerati più linguistici rispetto a quelli paralinguistici; questi ultimi invece sono più linguistici rispetto ai fenomeni non linguistici (come il tipo di voce o le vocalizzazioni).

Il secondo criterio consente di distinguere i fenomeni vocali in due categorie (*ivi*, p. 119; cfr. anche Laver, Trudgill, 1979; Van Lancker, 1981; Bertinetto, 1987): la prima, quella dei *segnali extralinguistici*<sup>9</sup>, fa riferimento alle caratteristiche intrinseche

---

<sup>9</sup> All'interno delle caratteristiche extralinguistiche è possibile distinguere inoltre: le caratteristiche organiche, "che dipendono esclusivamente dalle configurazioni anatomiche e dalle dimensioni degli organi degli apparati vocali individuali dei locutori" (Anolli, Ciceri, 1992, p. 125) e che permettono di inferire informazioni sul genere, l'età, la corporatura, lo stato di salute di un parlante, e le caratteristiche fonetiche, ovvero "il risultato del modo con cui un individuo abitualmente usa il proprio apparato fonatorio per parlare" (*ibid.*) e sono quindi attribuibili ad un controllo volontario dei muscoli coinvolti nell'articolazione e nell'emissione di suoni.

dell'apparato fonatorio del locutore, dal quale dipendono le qualità vocali e le caratteristiche personali della voce.

I *tratti paralinguistici* (la seconda categoria) sono invece il risultato dell'interazione di parametri quali il tempo, l'altezza del tono, l'intensità e delle qualità vocali. In questa definizione, i termini paralinguistico e prosodico non sono considerati separatamente, ma il secondo viene incluso nel primo (come spiega Sorianello 2006, per molto tempo i fenomeni prosodici non hanno goduto di uno statuto linguistico, cfr. PAR. 2.2.1). I *riflessi vocali* sono invece considerati esterni al codice linguistico, in quanto sono legati a meccanismi fisiologici (starnutare, tossire, sbadigliare, ecc.); possono avere una funzione informativa in relazione alle condizioni fisiche del locutore (se prodotti in modo non intenzionale) o comunicativa se sono introdotti all'interno dell'eloquio in modo intenzionale, in sostituzione delle parole.

Le *vocalizzazioni* sono invece "quell'insieme di suoni vocalizzati ("mhm", "oh", "ah", "ehm", ecc.) che sono in realtà *items* lessicali simili alle interiezioni. Essi vengono caricati e coloriti di sensi diversi mediante le caratteristiche paralinguistiche" (Anolli, Ciceri, 1992, p. 123). Infine, si definiscono *caratterizzatori vocali* azioni come il ridere, il piangere, il singhiozzare, ecc. Anche in questo caso, essi possono costituire una manifestazione dello stato psico-fisico del locutore o possono essere usati in modo intenzionale e simbolico (cfr. comunicazione simbolica, PAR. 1.5.6).

### 2.2.1 fenomeni prosodici

Lo studio dei fenomeni prosodici (come l'accento<sup>10</sup>, il tono<sup>11</sup>, l'intonazione, il ritmo<sup>12</sup>), sebbene sia stato a lungo trascurato dalla ricerca linguistica, ha assunto, soprattutto negli ultimi decenni, le proporzioni di un cambiamento epocale all'interno del panorama scientifico attuale. La svolta è rappresentata, in un certo senso, dalla nascita della fonologia autosegmentale e metrica. Il modello inaugurato da Goldsmith

---

<sup>10</sup> Il dominio dell'accento è la sillaba. "Grazie alla sua funzione culminativa, l'accento pone in rilievo una determinata sillaba all'interno della parola" (Soriano, 2006, p. 20). In lingue come l'italiano, l'accento è in grado di modificare il significato di una parola, in base alla sua posizione (pensiamo ad esempio a *perdono*; *perdòno*, *perdonò*, cfr. Bertinetto; 1981; Nespor, 1981; Soriano, 2006).

<sup>11</sup> Il tono è una caratteristica delle lingue tonali, come il cinese. Le diverse variazioni di altezza sul piano melodico determinano significati differenti sul piano lessicale.

<sup>12</sup> Il ritmo è dato dall'alternanza di sillabe accentate (elementi forti) e sillabe non accentate (elementi deboli), lunghe o brevi, di toni alti e di toni bassi. Il modo in cui questi elementi si alternano tra loro determina l'organizzazione ritmica di una lingua sulla base di schemi variabili (cfr. Pike 1945; Abercrombie 1967); tuttavia, ogni sistema linguistico adotta essenzialmente il principio universale di alternanza ritmica, ovvero l'alternanza di elementi forti e deboli. Come spiega Soriano, il ritmo è dato già a livello fisiologico dall'alternanza tra inspirazione ed espirazione.

(1976) abbandona la linearità del modello segmentale in luogo di una rappresentazione articolata su più livelli paralleli e autonomi (che rendono conto, ad esempio, per i toni di un livello tonale, e per la sillaba di un livello sillabico, cfr. Soriano, 2006, p. 15). I successivi sviluppi (cfr. Pierrehumbert, 1980) hanno portato alla formazione di un nuovo ambito di ricerca, ovvero la Fonologia dell'Intonazione (cfr. Ladd, 1996) e allo sviluppo della cosiddetta *Laboratory Phonology* (cfr. Pierrehumbert, 1980).

Il ruolo centrale della prosodia e in particolar modo dell'intonazione all'interno della comunicazione viene rivendicato a partire dall'idea che i tratti prosodici "partecipino direttamente alla dinamica comunicativa in modo distintivo e contrastivo" (Soriano, 2006, p. 16). Per molto tempo infatti, gli elementi prosodici, per il loro carattere non discreto, sono stati relegati ad una funzione meramente espressiva. Come ricorda Di Russo (2011, p. 37), Trubeckoj (1939) è stato il primo a considerare la sillaba, l'accento, il tono e l'intonazione dal punto di vista prosodico, mentre Bolinger (1986) pur focalizzandosi sull'intonazione, prende in esame anche altri fenomeni, quali la sillaba, l'accento, e il ritmo. In ambito italiano, Bertinetto e Magno Caldognetto si sono incentrati in particolare sul ritmo e l'intonazione (cfr. Bertinetto, Magno Caldognetto, 1993), distinguendo le funzioni di quest'ultima in *extralinguistiche* (legate alla dimensione fisiologica del locutore), *paralinguistiche* (deputate cioè alla trasmissione di informazioni pertinenti attitudini, stati d'animo ed emozioni), e *linguistiche* (legate all'organizzazione sintattica e alla modalizzazione). Soriano (2006) ha proposto una descrizione dei fenomeni prosodici in relazione ai maggiori correlati acustici del parlato, ovvero la frequenza fondamentale (cfr. PAR. 2.2.2), la durata<sup>13</sup>, e l'intensità<sup>14</sup>.

### 2.2.2 Le funzioni dell'intonazione

L'intonazione, elemento comune a tutte le lingue del mondo, può essere descritta come l'alternarsi di movimenti ascendenti e discendenti sul piano melodico, che assume

---

<sup>13</sup> La durata si riferisce all'estensione di un suono o di un enunciato dal punto di vista temporale. L'unità di misura di riferimento è in secondi o in millisecondi (ms). In alcune lingue, la durata temporale di un suono può rappresentare un tratto distintivo e determinare quindi differenze di significato. La durata di un fonema è suscettibile a diverse variabili, come la velocità di eloquio del parlante, l'estensione dell'enunciato. Ovviamente, le caratteristiche intrinseche della persona determinano la gestione temporale della stringa fonetica.

<sup>14</sup> L'intensità di un suono fa riferimento all'ampiezza delle onde sonore che lo compongono. L'unità di misura è il decibel (dB), ossia il rapporto tra l'intensità stimata e un valore di intensità prestabilito. In genere, il valore relativo all'intensità è soggetto a diverse variabili: la quantità di energia articolatoria impiegata durante la fonazione, la pressione dell'aria proveniente dai polmoni, il livello di tensione delle corde vocali (cfr. Soriano, 2006).

un significato pragmaticamente codificato all'interno di una stessa comunità di parlanti. Essa assolve già nei primi mesi di vita "una funzione comunicativa essenziale e vicaria, configurandosi quale forza illocutiva primaria" (Soriano, 2006, p. 39).

Le funzioni dell'intonazione sono molteplici. In primo luogo, essa ha una funzione distintiva sul piano linguistico, e consiste nella segmentazione del parlato in unità intonative<sup>15</sup>, ovvero unità discrete costituite da un gruppo di parole coerente e coeso sul piano testuale (ivi, p. 27). L'intonazione svolge inoltre una funzione sintattica, in quanto esplicita la forza illocutiva di un enunciato, differenziando un enunciato assertivo da uno interrogativo (funzione modale) e "partecipa all'organizzazione tematica del parlato e alla sua scansione in unità minime di informazione" (ivi, p. 62), assolvendo una funzione semantico-pragmatica, in grado cioè di veicolare significati specifici e sfumature importanti sul piano comunicativo (ad esempio, l'introduzione di un'informazione nuova rispetto a quelle già date sarà segnalata da una certa salienza melodica).

Molte lingue ricorrono a strutture intonative che ricoprono funzioni simili, segnalando significati affini sul piano comunicativo. Ad esempio, l'opposizione semantica tra contorno discendente (associato a un significato conclusivo, tipico degli enunciati assertivi) e contorno ascendente (legato a un significato di incertezza e sospensione, tipico degli enunciati interrogativi) costituisce uno degli universali prosodici, sebbene vi siano sistemi linguistici che ricorrono a questa opposizione per segnalare significati esattamente opposti. Come spiega Soriano "la ricerca dei tratti intonativi universali non è priva di alcuni aspetti contraddittori. [...] Molti comportamenti intonativi, più che universali, sono da intendersi come atteggiamenti dominanti non marcati e come tali suscettibili di diversificarsi lungo un continuum di possibilità fonetiche" (ivi, p. 41).

Il correlato acustico dell'intonazione è la frequenza fondamentale<sup>16</sup> o  $f_0$ . Per frequenza fondamentale si intende il numero di vibrazioni glottidali (ossia di

---

<sup>15</sup> L'unità tonale è costituita da un accento intonativo nucleare, da ulteriori accenti ed è caratterizzata da una marca di confine (Soriano, 2006, p. 27).

<sup>16</sup> A livello acustico, la frequenza può essere misurata attraverso l'applicazione del teorema di Fourier (Fast Fourier Transform) grazie al quale le onde sonore complesse possono essere scomposte in una serie di onde semplici, dette armoniche. Queste sono dei multipli interi della prima armonica, definita appunto frequenza fondamentale.

oscillazioni complete) compiute dalle pliche vocali in una unità di tempo. L'unità di misura è l'Hertz (Hz). Tale parametro si rileva solo nei suoni sonori<sup>17</sup>.

Sul piano percettivo, il valore della frequenza fondamentale è responsabile delle differenze di altezza dei suoni (*pitch*): ad una  $f_0$  alta si deve la sensazione uditiva di suoni acuti, mentre una  $f_0$  bassa sarà associata percettivamente a suoni gravi<sup>18</sup>. Poiché la frequenza di vibrazione di un corpo, in questo caso delle pliche vocali, dipende dalla sua conformazione fisica, dalla sua grandezza ecc., nelle donne e nei bambini si registrano cicli di chiusura e apertura molto rapidi (per l'estensione ridotta della laringe). Ciò determina un valore alto della frequenza fondamentale, che si traduce sul piano percettivo, in un tono più acuto. Viceversa, negli uomini la frequenza di vibrazione è notevolmente ridotta, da qui la sensazione uditiva di gravità sonora.

### 2.2.3 Tre codici biologici

Come spiegano Thompson e Balkwill (2006, p. 407), "evolutionary theorists have proposed that prelinguistic humans relied on prosody to communicate their intentions (e.g., mating, defense, aggression), and that this feature of speech has survived because of its evolutionary adaptive properties (Brown 2000)".

Ohala (1983; 1984) ha suggerito che le variazioni prosodiche determinate dalla frequenza fondamentale nella comunicazione umana abbiano una base etologica. Partendo dalle osservazioni di Morton (1977) sul comportamento animale, Ohala sostiene che alcune specie (mammiferi e uccelli) sfruttino la correlazione tra la dimensione della laringe e la frequenza di vibrazione dei muscoli vocali per comunicare

---

<sup>17</sup> Come spiega Sorianello (2006, p. 33): "A livello fisiologico, le vibrazioni delle pliche vocali si devono a un principio aerodinamico. L'aria proveniente dai polmoni sale passando attraverso la trachea e poi attraverso la glottide. Poiché il passaggio a livello glottidale è piuttosto stretto, la pressione glottidale aumenta, mentre l'aria fuoriesce più velocemente. Di conseguenza, le pliche vocali, inizialmente separate, in virtù della loro natura fortemente elastica si accostano, generando fasi alterne di chiusura e di apertura. La velocità con cui le pliche vocali compiono questo ciclo di abduzione/adduzione determina il valore della frequenza fondamentale".

<sup>18</sup> Tra il piano acustico e quello percettivo non esiste un rapporto lineare, in quanto la sensibilità del sistema uditivo umano non è uniforme per tutte le frequenze udibili; l'orecchio umano non è in grado di discriminare due suoni simili sulla base di differenze minime in termini di frequenza e intensità, per cui in presenza di due suoni differenti, ma vicini per intensità e frequenza, essi sembreranno uguali. Se consideriamo due suoni, rispettivamente di 100 e 200 Hz, la differenza percepita dal nostro orecchio sarà identica a quella percepita tra due suoni con una frequenza raddoppiata (ovvero, tra i 200 e i 400 Hz). In altre parole, il sistema uditivo umano avverte sempre lo stesso scarto a livello di altezza nel momento in cui la frequenza viene raddoppiata. Questo intervallo è detto ottava e rappresenta dunque la distanza tra una frequenza e il suo valore doppio (Sorianello, 2006). La mancanza di isomorfismo tra proprietà acustiche e sensazioni uditive ha portato all'adozione di scale psicoacustiche in grado di misurare la distanza a livello percettivo; tra quelle che consentono una valutazione percettiva del *pitch*, vi è ad esempio la scala in semitoni (ST).

relazioni di dominanza o sottomissione in situazioni di ostilità. Come spiega Gussenhoven (2002), questa correlazione è data dal fatto che un suono grave suggerisce che le dimensioni dell'animale che ha emesso la vocalizzazione siano maggiori. Per cui, suoni a bassa frequenza sono associati a comportamenti aggressivi, mentre suoni ad alta frequenza, quindi acuti (come il guaito del cane), comunicano un atteggiamento remissivo. La tendenza dei profili intonativi a seguire quello che viene definito il *codice universale della frequenza* (Ohala, 1984) sembra avere dunque delle basi "biologiche".

Gussenhoven (2002), riprendendo la definizione di Ohala, suggerisce altri due codici per spiegare il significato universale associato a molte variazioni prosodiche: il *codice dello sforzo* e il *codice della produzione*. Durante il processo di produzione, un incremento dell'energia articolatoria si riflette sul segnale vocale, determinando una maggiore precisione nell'articolazione dei suoni, ma anche un'estensione tonale più ampia (codice dello sforzo). Per estensione tonale (o *pitch range*) si intende la differenza tra il valore massimo e il valore minimo di  $f_0$  rilevabili nel profilo intonativo di un enunciato. Secondo l'autore, "speakers exploit this fact by using *pitch range* to signal meanings that can be derived from this effect of the expenditure of effort. A frequent interpretation is that the speaker is being forceful because he believes the contents of his message are important, an informational meaning" (ivi, p. 4). Il codice della produzione si basa invece sulle dinamiche respiratorie che portano un parlante ad adoperare maggiore energia all'inizio di un enunciato rispetto alla parte finale, determinando sul piano frequenziale, valori di  $f_0$  più alti all'inizio e più bassi in corrispondenza della fine dell'enunciato. Questa differenza trova una sua spiegazione nei meccanismi di espirazione: "at the beginning of the exhalation phase, subglottal air pressure will be higher than towards its end. A natural consequence of the fall-off in energy is a gradual drop in intensity, and a weak, gradual lowering of the fundamental frequency, known as 'declination'"(ivi, p. 6)<sup>19</sup>.

Secondo l'autore, gli esseri umani fanno ricorso alle modulazioni del *pitch* per veicolare significati linguistici e paralinguistici sfruttando queste tre condizioni fisiologiche legate alla produzione del segnale vocale. Tuttavia, non è necessario che le

---

<sup>19</sup> La declinazione intonativa, considerata uno degli universali prosodici, è generalmente presente negli enunciati assertivi e fa riferimento all'unità prosodica nella sua globalità. Essa si traduce nel "progressivo abbassamento della curva di  $F_0$  nel corso dell'emissione di una porzione di parlato che corrisponde ad un gruppo respiratorio o breath-group (cioè della fase espiratoria che accompagna la produzione orale)" (De Dominicis, 2010, p. 52). Come spiega Soriano: "l'enunciato assertivo è considerato compiuto solo quando la frequenza fondamentale raggiunge un livello frequenziale adeguatamente basso, mentre il mancato raggiungimento del valore minimo conferisce all'enunciato un'impressione percettiva di sospensione e di continuazione" (2006, p. 44).

condizioni si verifichino realmente perché questi significati siano veicolati, "it is enough to create the effect" (ivi, p. 2). La comunicazione realizzata attraverso il codice dello sforzo, ad esempio, non implica necessariamente un maggiore impiego di energia da parte del parlante, il quale può semplicemente ricorrere ad un'estensione tonale più ampia (ovvero all'effetto fisiologico) per segnalare significati importanti sul piano comunicativo (*ibid.*). Questi codici non sono dunque automatici, ma possono essere utilizzati in modo volontario per comunicare informazioni relative al messaggio (*informational meanings*) o al parlante (*affective meanings*).

Ad esempio, il codice della frequenza può essere impiegato per segnalare un certo grado di certezza (*pitch* basso) o incertezza (*pitch* alto) sul piano informativo. Le interpretazioni affettive generalmente associate a questo codice sono molteplici. Come già evidenziato da Ohala, i significati più ovvi sono legati a dimensioni quali: dominanza, controllo, sicurezza vs. sottomissione, vulnerabilità, insicurezza. Queste dimensioni comprendono altri significati, come affabilità (*friendliness*) e cortesia (*politeness*), associati ad un *pitch* alto; un *pitch* basso invece suggerisce aggressività ma anche senso di protezione (*protectiveness*)<sup>20</sup>. Sul piano informativo il codice dello sforzo conferisce enfasi al messaggio, consentendo al parlante di focalizzare l'attenzione dell'interlocutore verso un elemento del discorso ritenuto importante, ricorrendo ad un'estensione tonale maggiore. Un'estensione tonale ristretta può essere utilizzata invece per segnalare una negazione o la ritrattazione di un'informazione data in precedenza.

. Le interpretazioni affettive legate a questo codice comprendono emozioni come la sorpresa, ma anche altri stati d'animo come l'agitazione, entrambi veicolati attraverso un'estensione maggiore del *pitch*; il *pitch range* risulta infatti sensibile al livello di attivazione determinato da uno stato emotivo. Ad esempio, la collera e la sorpresa sono di norma caratterizzate da un *pitch range* ampio, diversamente dalla tristezza (o dalla noia) in cui lo spazio melodico si assesta nelle fasce frequenziali più basse, determinando un'estensione tonale ridotta (cfr. Pittam, Scherer, 1993; Banse, Scherer, 1996). Meno comunemente, un'estensione tonale più ampia viene interpretata anche come un segnale di cortesia (cfr. Chen, Gussenhoven, Rietveld, 2004).

Il codice della produzione viene principalmente impiegato sul piano informativo per veicolare significati legati al contenuto del messaggio, sfruttando gli "effetti" delle dinamiche respiratorie. Ad esempio, l'introduzione di un'informazione nuova viene

---

<sup>20</sup> Queste interpretazioni si basano sui giudizi affettivi riportati negli esperimenti percettivi in relazione alle scale semantiche considerate (cfr. Uldall, 1964).

segnalata attraverso un innalzamento del *pitch* in corrispondenza dell'inizio di una frase, mentre un *pitch* basso segnala la continuazione del *topic*. Il decremento del *pitch* verso la fine di un enunciato esprime un significato conclusivo, viceversa, un *pitch* alto a fine enunciato trasmette un'impressione percettiva di continuazione (cfr. Chen, Gussenhoven, Rietveld, 2004).

L'insieme di queste interpretazioni offre delle indicazioni sulle funzioni paralinguistiche del *pitch*, tuttavia resta da chiarire in che misura questi segnali veicolino significati universalmente condivisi. Alla luce dei risultati riportati da recenti ricerche (cfr. Chen, Gussenhoven, Rietveld, 2004; Chen, 2005; Chen, Fon, 2011) è stata avanzata l'ipotesi che la percezione degli andamenti intonativi paralinguistici sia determinata non solo da codici biologici universali, ma anche da fattori linguo-specifici che regolano l'implementazione di questi codici (cfr. PAR. 2.7.2.2). A supporto di questa teoria, gli studi interlinguistici sulla percezione delle interrogative (cfr. Gussenhoven, Chen 2000; Makarova 2000) hanno evidenziato una sensibilità linguo-specifica<sup>21</sup> all'altezza tonale, riconoscendo che al di là dei codici biologici, "language background is an active component in listeners' interpretation strategies, in addition to the universal, nonlinguistic strategy" (Gussenhoven, Chen, 2000, p. 4).

## 2.3 Il quadro teorico

### 2.3.1 L'interazione tra il piano linguistico e quello paralinguistico

Uno dei principali nodi teorici che è al centro del dibattito sulla codifica del parlato emotivo riguarda l'interazione tra le componenti verbali e quelle paraverbali. Due si configurano come i modelli di rappresentazione dominanti: il modello dei *canali paralleli* (detto anche della covarianza) e il modello del *canale unico*, noto anche come modello delle configurazioni (cfr. Scherer, Ladd, Silverman, 1984, Ladd *et al.*, 1985). La teoria dei canali paralleli considera i segni vocali non verbali come un sistema semiotico parallelo e indipendente rispetto al contenuto linguistico. Ciascun canale

---

<sup>21</sup> A questo proposito, Ladd (1981) aveva formulato due ipotesi, una esplicitamente universalista (*Strong Universalist Hypothesis*), in base alla quale un movimento ascendente viene interpretato universalmente come una domanda, mentre un movimento discendente segnala un'affermazione; e l'altra linguo-specifica (*Nuclear Tone Hypothesis*), in base alla quale la correlazione tra la forma e la funzione dei contorni intonativi sarebbe influenzata da fattori linguistici e sarebbe dunque arbitraria. Alcuni studi (cfr. Jarman, Cruttenden 1976; Gordon 1999, Gussenhoven, 2000) hanno portato prove a sostegno della seconda ipotesi, evidenziando la tendenza di alcuni sistemi linguistici come l'inglese di Belfast o il chickasaw (una lingua nativa americana) a ricorrere ad un'intonazione ascendente per le assertive e una discendente per le interrogative.

(verbale e non verbale) veicolerebbe un solo tipo di informazione e sarebbe di per sé autonomo: soltanto attraverso la loro successiva comunione l'enunciato emotivo acquisterebbe il suo significato intrinseco. In quest'ottica, "un eventuale ascoltatore è in grado di riconoscere l'emozione provata (o simulata) dal locutore, anche quando il contenuto verbale della frase viene eliminato o privato di qualsiasi valore semantico" (Anolli, Ciceri, 1992, p. 210).

La teoria del canale unico si fonda invece sull'ipotesi che l'espressione vocale delle emozioni sia dovuta ad un'unica codifica delle caratteristiche linguistiche e di quelle paralinguistiche (cfr. Fonagy, 1987; 1989) in quanto la valenza emotiva e comunicativa di un enunciato sarebbe espressa solo attraverso la pianificazione condivisa dei due livelli. In base a questo approccio, gli elementi soprasegmentali non costituirebbero degli indici indipendenti, ma varierebbero in base all'enunciato a cui sono associati. In altre parole, segnali soprasegmentali identici potrebbero essere interpretati in modo differente in base al contesto frasale in cui sono realizzati (cfr. Scherer, Ladd, Silverman, 1984).

Come ricorda Di Russo (2011, p. 27): "La teoria del canale unico [...] non rifiuta né la distinzione tra componenti linguistiche e paralinguistiche del senso né quella tra livello segmentale e soprasegmentale, che anzi ammette, ma suppone che l'elaborazione delle sfumature paralinguistiche, in via privilegiata soprasegmentali, influisca sull'elaborazione delle caratteristiche linguistiche del senso, lessicali e sintattiche come pure morfologiche e fonologiche".

La dicotomia che oppone il piano linguistico a quello paralinguistico sembra non rendere conto della complessità dell'interazione linguistica e dell'intreccio di risorse lessicali, morfologiche, sintattiche e prosodiche che riflette le funzioni della lingua (da quella referenziale a quella espressiva) e specifica le intenzioni comunicative del parlante. Per cogliere il suggerimento di Di Russo (ivi, p. 32):

Occorre [...] un modello capace di rendere conto non solo delle tracce sintattiche, lessicali e morfologiche ma anche di quelle prosodiche e fonetiche alle quali la voce dà corpo, e che sappia cogliere il gioco di opposizioni, negazioni e relazioni che le regola, in funzione delle intenzioni e delle esigenze comunicative, che presiedono alla significazione e si mescolano nel senso.

### 2.3.1.1 *Approcci descrittivi e sperimentali*<sup>22</sup>

I due modelli, elaborati nel tentativo di fornire una lettura dei processi di codifica del parlato emotivo, si differenziano, oltre che sul piano teorico, anche su quello metodologico. Il modello del canale unico si fonda su un approccio descrittivo (o linguistico) e si basa sull'analisi di parametri acustici e sulla verifica percettiva del parlato emotivo (affidata a valutatori esterni). L'ipotesi di base è che l'intonazione implichi una serie di categorie distintive (analoghe alle opposizioni fonologiche o alle categorie grammaticali) e che questo si rifletta, in termini statistici, sui giudizi affettivi espressi dai valutatori, basati cioè su "configurazioni di categorie variabili" (Scherer Ladd, Silverman., 1984, p. 1347) e non su variabili scalari. Per tale ragione, i parametri acustici non sono considerati come indici riconducibili a specifici stati emotivi, come rabbia o tristezza, ma sono inseriti in categorie fonologiche (intonazione discendente/ascendente) che possono di volta in volta veicolare un significato emotivo diverso in base al contesto in cui è realizzata la frase. In altri termini, "the specific message conveyed by an utterance is an interpretation, an active inference by the listener, based on the total configuration of the linguistic choices of the context" (ivi, p. 1347; cfr. anche Bolinger, 1984).

Il modello dei canali paralleli procede invece attraverso un approccio sperimentale ed estende la sua analisi all'intonazione, all'intensità, alla qualità della voce<sup>23</sup>, al ritmo, inserendo questi indici acustici in un modello statistico basato sulla covarianza di variabili scalari in relazione al tipo e al grado di intensità dello stato emotivo prodotto dal parlante. Come spiegano Anolli e Ciceri (1992, p. 211): "ad esempio, se vengono identificati i parametri acustici dell'espressione vocale della gioia, saranno presenti in grado più elevato nelle frasi che esprimono uno stato più intenso di gioia, e in grado minore in quelle che ne esprimono di meno".

Sempre sul piano teorico, una questione altrettanto dibattuta, concerne infatti la discretezza o la scalarità degli andamenti intonativi paralinguistici (Poggi, Magno Caldognetto, 2004). Il primo approccio interpreta le risorse soprasegmentali (frequenza fondamentale, intensità e durata) come un sistema di opposizioni, mentre il secondo

---

<sup>22</sup> Questa distinzione è stata operata adottando la classificazione proposta da Anolli e Ciceri (1992, pp. 210-11) in merito agli studi sperimentali sull'espressione vocale delle emozioni.

<sup>23</sup> Uno dei parametri fondamentali per determinare la qualità della voce e la dimensione psicofisica del parlante è l'osservazione della fonazione amodale delle corde vocali. Difatti queste possono assumere delle conformazioni che si distaccano dalle normali posizioni (cfr. Romito, 2000) definendo specifiche alterazioni della voce.

inserisce le loro variazioni in un sistema scalare, che varia sia all'interno del medesimo stato emotivo, sia in relazione a stati emotivi differenti.

La questione che si pone è indubbiamente complessa e ampiamente dibattuta: da un lato, si postula l'esistenza di uno specifico *pattern* intonativo per ogni emozione, in grado di veicolare precisi significati emotivi, in modo indipendente rispetto alla struttura linguistica; dall'altro, in virtù del carattere contraddittorio dei risultati sperimentali ottenuti, è stata messa in dubbio la natura discreta (e dunque categoriale) degli aspetti vocali delle emozioni, tra cui  $f_0$ . Ad esempio, Pakosz (1983) suggerisce che l'intonazione convogli solo informazioni relative al grado di attivazione dello stato emotivo; in quest'ottica, risulterebbe necessaria la compartecipazione di altri elementi del contesto (e di altri canali non verbali, tra cui le espressioni facciali) per disambiguare e definire le emozioni in categorie specifiche.

Sul versante opposto, una tradizione che emerge dall'approccio linguistico allo studio dei contorni intonativi (ovvero il succedersi di movimenti ascendenti e discendenti dovuti alle variazioni della frequenza fondamentale o  $f_0$ ), nell'espressione vocale delle emozioni, ha rivendicato l'esistenza di specifici modelli intonativi per le emozioni (cfr. Fónagy, Magdics, 1963; Halliday, 1970; O'Connor, Arnold, 1973). Tuttavia, i risultati ottenuti sono stati oggetto di discussione e di critiche (cfr. Frick, 1985), e la questione rimane sostanzialmente ancora aperta (cfr. Bänziger, Scherer, 2005).

Alla luce delle più recenti ricerche (basate anche sulla sintesi e la manipolazione del segnale vocale), l'ipotesi più solida suggerisce che il contorno intonativo contribuisca a veicolare informazioni di natura emotiva soltanto in combinazione con altri parametri (tra cui l'escursione tonale, l'intensità, la durata). I risultati riportati da Bänziger e Scherer (2005) mostrano che il livello medio di  $f_0$  e la gamma dei contorni di  $f_0$  variano fortemente in funzione del grado di attivazione delle emozioni, sebbene non ci siano prove sufficienti all'individuazione di conformazioni specifiche del contorno in relazione agli stati emotivi considerati.

### *2.3.2 Approcci discreti e dimensionali*

Come precedentemente illustrato nel CAP. 1, diverse teorie in ambito psicologico hanno tentato di fornire una definizione di emozione. I fautori delle teorie discrete postulano l'esistenza di un numero limitato di emozioni di base (Tomkins 1962; Izard 1977). Questo approccio, che considera l'emozione come una risposta fisiologica

automatica con uno specifico *pattern* espressivo (rintracciabile nella modificazione dei muscoli facciali e nelle variazioni a livello vocale), ha fortemente influenzato la ricerca sull'espressione vocale delle emozioni. Sono molti gli studi che hanno tentato di identificare i profili acustici di un certo numero di emozioni primarie (cfr. per una ricognizione, cfr. Juslin, Laukka, 2003). In generale, i risultati empirici riportati dalle ricerche che hanno adottato questo approccio mostrano una certa congruenza con i modelli dell'*appraisal* (in particolare, il modello componenziale di Scherer, cfr. PAR. 2.3.4), per cui si rimanda al PAR. 2.8 la ricognizione di questi profili.

Negli ultimi anni, all'interno della ricerca sul parlato emotivo, si è consolidato un approccio di tipo dimensionale, sulla scia delle teorie dimensionali elaborate in ambito psicologico (cfr. PAR. 1.4.3). Come spiegano Poggi e Magno Caldognetto, "la condivisione di indici acustici tra le diverse emozioni è stata considerata indicativa della condivisione delle dimensioni psicologiche costitutive delle emozioni, in particolare di quelle di attivazione, valutazione e potere" (Poggi, Magno Caldognetto, 2004, p. 16).

In base alla rassegna degli studi proposta da Juslin e Scherer (2008), l'attivazione risulta la dimensione più studiata, e i dati sono abbastanza coerenti. La dimensione di attivazione caratterizza emozioni come sorpresa, collera e paura rispetto a stati emotivi come la tristezza e il disgusto. Essa, infatti, è strettamente connessa all'importanza dello scopo, pertanto emozioni ad alta attivazione come paura, collera ecc. presentano in genere dei correlati acustici ben definiti: valori alti di  $f_0$ ; grande variabilità di  $f_0$ ; velocità di eloquio elevata, pause brevi, incremento dell'intensità e dell'energia ad alta frequenza (Davitz 1964; Huttar 1968; Uldall 1960; Scherer, Oshinsky 1977; Pittam *et al.* 1990; Pereira 2000; Schroder *et al.* 2001). Le emozioni caratterizzate da una bassa attivazione presentano specularmente dei valori opposti.

La dimensione di valutazione discerne stati emotivi di natura positiva (gioia, speranza ecc.) da stati emotivi prevalentemente negativi (collera, paura). Tali emozioni sono legate rispettivamente al raggiungimento o alla compromissione di uno scopo. I risultati relativi alla dimensione di valutazione sono molto più inconsistenti. Alcuni studi hanno riportato per la valenza positiva i seguenti profili acustici: media di  $f_0$  bassa, grande variabilità di  $f_0$ , velocità d'eloquio sostenuta, valori bassi di intensità (Scherer 1974; Scherer, Oshinsky 1977; Schroder *et al.* 2001; Uldall 1960). La collera (valenza negativa) si caratterizza invece per: incremento dell'intensità, abbassamenti veloci di  $f_0$  e pause più lunghe (Poggi, Magno Caldognetto, *ibid.*).

La dimensione di potenza riflette la condizione di dominanza o sottomissione legata al raggiungimento di uno scopo, ossia il grado di controllo che ciascun individuo esercita nel perseguire i propri fini. Emozioni legate alla dominanza, come la rabbia, si distinguono, ad esempio, per un'intensità maggiore rispetto alla paura (legata ad una dimensione di sottomissione). Tuttavia, la dimensione di potenza è stata meno studiata, per cui i risultati non offrono un quadro coerente. In relazione alle emozioni ad alta potenza, la media di  $f_0$  non sembra essere un indice affidabile, in quanto in alcuni casi è risultata piuttosto elevata, in altri più bassa; anche la velocità d'eloquio si è caratterizzata per una grande variabilità. I valori relativi all'intensità sono in genere molto alti e abbastanza variabili; si evidenzia inoltre una concentrazione di energia ad alta frequenza (cfr. Uldall 1960; Scherer 1974; Scherer, Oshinsky 1977; Schroder *et al.* 2001).

### 2.3.3 Il modello componenziale

Le ricerche condotte da Scherer ed altri studiosi hanno fornito un contributo fondamentale alle ricerche sulla codifica acustica delle emozioni, sia in ambito intralinguistico sia in ambito interlinguistico. L'applicazione del modello componenziale ai dati ottenuti ha inoltre permesso di osservare i fenomeni emotivi in quanto sistemi complessi, in grado di attivare, sia sul piano fisiologico che su quello cognitivo, una serie di meccanismi di autoregolazione che consentono di valutare lo stimolo emotigeno e di innescare una reazione compatibile con le norme sociali e l'immagine di sé (cfr. CAP. 1).

Secondo Banse e Scherer (1996), le alterazioni fisiologiche innescate dalle emozioni coinvolgono la respirazione, la fonazione, l'articolazione e determinano le variazioni del segnale vocale, avvalorando "the existence of the phylogenetic continuity in the acoustic patterns of vocal affect expression" (ivi, p. 615).

In base al modello componenziale di Scherer (1986), ogni fenomeno emozionale è interpretato in base ad una serie di valutazioni cognitive dello stimolo percepito. Tali controlli, definiti *stimulus evaluation check* (cfr. CAP. 1, PAR. 1.4.2.3), si attivano secondo un ordine ben preciso, misurando di volta in volta la novità dello stimolo (*novelty check*), la piacevolezza intrinseca (*intrinsic pleasantness check*), la rilevanza in relazione allo scopo o al bisogno (*goal/need significance check*), il potenziale di adattamento (*coping potential check*) e la compatibilità dello stimolo con le norme sociali e l'immagine di sé (*norm/self-compatibility check*).

Pertanto, secondo Scherer, le emozioni vanno valutate in base agli antecedenti situazionali sopraelencati; questi infatti determinano dei mutamenti psicofisici riscontrabili anche a livello acustico. Ad esempio, la valutazione del grado di novità dello stimolo ha come effetto una risposta di orientamento, necessaria per catturare ogni tipo di informazione sul nuovo stimolo. Ciò si ripercuote sull'espressione vocale, che viene momentaneamente interrotta, mentre il ritmo respiratorio sale (Anolli, Ciceri, 1992).

Successivamente, in base alla valutazione della piacevolezza/spiacevolezza dello stimolo, si avranno ulteriori cambiamenti a livello vocale. Ad esempio, in caso di stimoli poco piacevoli, la reazione dell'organismo sarà rappresentata da una costrizione muscolare relativa alla faringe. Pertanto, la riduzione generale del canale produrrà un tipo di voce che Scherer definisce stretta, caratterizzata da un'ampiezza ristretta di banda formantica (Anolli, Ciceri, 1992).

Nel caso in cui lo stimolo risulti piacevole invece, l'organismo tende a rilassarsi e tale rilassamento interessa anche il tratto faringeo. La voce prodotta in seguito a tali circostanze viene definita ampia.

#### 2.3.4 *Un modello teorico per la comunicazione vocale delle emozioni*

Gli studi sull'espressione vocale delle emozioni hanno dimostrato che i singoli segnali vocali non sono perfettamente correlati con gli stati emotivi, e che ci sono intercorrelazioni tra i vari indizi vocali (Banse, Scherer 1996; Juslin, Laukka 2001) che in parte riflettono il processo di produzione orale (Juslin, Scherer, 2008, p. 83). Per esempio, un aumento della pressione subglottidale determina un incremento non solo dell'intensità, ma, in una certa misura, anche della frequenza fondamentale. Questi vincoli, legati al processo di produzione orale, potrebbero rendere conto dei risultati discordanti ottenuti da alcune ricerche (*ibid.*).

Le caratteristiche del processo di comunicazione vocale possono essere quantitativamente descritte utilizzando la variante del *modello della lente*<sup>24</sup> di Brunswik (1956), rielaborata da Scherer (1978; 1982). Questo modello percettivo è stato adottato per spiegare il modo in cui gli esseri umani selezionano le informazioni provenienti dall'ambiente circostante, riuscendo ad adattarsi e a sopravvivere in un sistema

---

<sup>24</sup> Originariamente, il modello della lente è stato inteso come un modello di percezione visiva, tuttavia è stato successivamente adottato in diversi ambiti per studiare, ad esempio, la percezione sociale (Brunswik 1956; Scherer 1978), o la percezione della musica (Juslin 2000).

ecologico eterogeneo. Il modello si fa carico di spiegare anche l'intero processo comunicativo (cfr. FIG. 1), che comprende tre fasi: l'*espressione* (codifica o *encoding*) di un'emozione dal punto di vista del mittente, la *trasmissione* del suono, e l'*impressione* (decodifica o *decoding*) da parte del ricevente, intesa come il risultato di un'inferenza rispetto all'emozione espressa (Juslin, Scherer, 2008, p. 83). Riprendendo la terminologia di Brunswik, Scherer afferma che i segnali vocali emessi dal parlante costituiscono degli indicatori distali dell'emozione (*distal cues*), non ancora percepiti dal ricevente. In seguito, diventano degli indicatori prossimali (*proximal cues*), e sono convogliati nel processo percettivo del ricevente che li elabora a partire da categorie soggettive legate alla discriminazione della voce e delle sue modificazioni (Juslin, Scherer, *ivi*, p. 83).

La differenza tra i due indicatori è utile per comprendere la complessità dell'intero processo che, in alcuni casi, non si conclude con la decodifica corretta dell'emozione espressa dal parlante. Secondo Scherer (1993), la comunicazione vocale umana non è determinata soltanto da fattori biologici, che ne confermano la continuità filogenetica, ma è, in una certa misura, soggetta a vincoli culturali e sociali.

Da un lato, fattori interni (legati a processi fisiologici) producono delle modificazioni di stato all'interno dell'organismo che si ripercuotono sugli organi vocali e dunque "i risultanti parametri acustici sono quasi esclusivamente determinati dalla natura e dalla forza dei meccanismi fisiologici" (*ivi*, p. 58), dall'altro, fattori esterni (convenzioni sociali e regole di esibizione) "influenzano la produzione di segni richiesti in situazioni sociali, specificando pattern acustici da realizzare" (*ivi*, p. 57). Scherer definisce *push effect* i fattori interni (determinati da meccanismi biologici, come il tono muscolare) che "spingono" la vocalizzazione in una certa direzione (determinando un incremento della frequenza fondamentale). I fattori esterni, definiti *pull effect* (come le aspettative di un ascoltatore) attirano la vocalizzazione affettiva in un'altra direzione, determinando un maggiore controllo sulla produzione vocale (*ibid.*).

In questo senso, un parlante è potenzialmente in grado di alterare volontariamente il proprio comportamento espressivo (in seguito ad una serie di valutazioni dello stimolo, cfr. *stimulous evaluation checks*, PAR. 2.3.3) e di veicolare informazioni false sul proprio stato emotivo, ricorrendo a una serie di indicatori distali che non riflettono l'emozione esperita realmente. Infatti, il modello della lente sottolinea

il carattere probabilistico e ridondante<sup>25</sup> dei segnali vocali, che non costituiscono degli indici fedeli dell'emozione espressa: "as a consequence, decoders have to combine several different voice cues to achieve veridical perception. This is not a matter of pattern matching, however, since the voice cues contribute in an additive fashion to decoders' judgements- each cue is neither necessary nor sufficient, but the larger the number of cues used, the more reliable the communication" (ivi, p. 84). Alla complessità del processo, si sommano fattori ambientali (come la presenza di rumore) che possono distorcere e alterare la percezione di questi segnali da parte del *decoder*. Gli indicatori prossimali percepiti dal ricevente, anche nel caso in cui siano effettivamente genuini e non *distal cues* simulati dal parlante, possono essere interpretati erroneamente e attribuiti ad uno stato emotivo differente rispetto a quello espresso. Il processo di attribuzione di uno stato emotivo si fonda infatti sia su meccanismi biologici, in grado di cogliere indizi rilevanti per la sopravvivenza di un individuo, sia su meccanismi cognitivi di carattere inferenziale, basati sulle esperienze pregresse, sulle conoscenze condivise dai partecipanti in merito a convenzioni e norme sociali.

La natura incerta dei segnali vocali riflette dunque una serie di variabili quali: le differenze individuali tra i parlanti, le interazioni degli indizi acustici con il contenuto linguistico; la distorsione dei segnali acustici in ambienti naturali o rumorosi e infine le interazioni tra *push effect* e *pull effect*<sup>26</sup>.

Per concludere, il modello della lente, applicato alla comunicazione vocale delle emozioni, suggerisce che l'analisi dell'intero processo debba comprendere: 1) lo stato emotivo espresso, 2) la verifica acustica dei segnali vocali; 3) la verifica percettiva degli indizi vocali e 4) il processo cognitivo che integra tutti i segnali vocali determinando un'inferenza rispetto all'emozione codificata (*ibid.*). Inoltre, poiché ogni indizio vocale è probabilistico, Scherer suggerisce di prendere in considerazione nelle analisi un numero

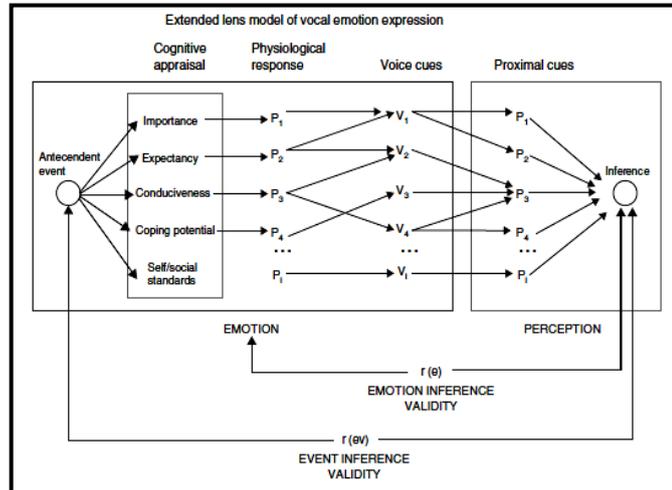
---

<sup>25</sup> Come spiegano Juslin e Scherer (2008, p. 84): "multiple cues that are partly redundant yield a robust communicative system that is 'forgiving' toward deviation from optimal code usage. However, robustness comes with a price: the redundancy of the cues means that the same information is conveyed by many cues. A system of this kind involves 'compromise and a falling short of precision, but also the relative infrequency of drastic error' (Brunswik 1956, p. 145)".

<sup>26</sup> "Pull effects, on the other hand, reflect external conditions that may lead to strategic posing of emotional expressions for manipulative purposes [...]. This may involve social processes such as *self-presentation* (wanting to create certain impressions of oneself in the perceiver in terms of identity, ability, moral, and prestige) and *conventionalization* (stereotyping to enhance the clarity and shared meaning of signals" (Juslin, Scherer, 2008, p.82, cfr, anche CAP. 1 ).

consistente di indici acustici, al fine di catturare il codice "completo" ed elaborare una previsione affidabile sull'emozione espressa nonché sull'emozione inferita (ivi, p. 87).

**FIG.1- 1. MODELLO DELLA LENTE DI BRUNSWIK (1956) APPLICATO ALLA COMUNICAZIONE VOCALE DELLE EMOZIONI.**  
**FONTE: JUSLIN, SCHERER, 2008, p.84.**



## 2.4 Lo scenario metodologico

La grande eterogeneità che caratterizza le metodologie applicate nello studio dell'espressione vocale delle emozioni genera una conseguente mancanza di unità e coesione all'interno dell'ambito di ricerca. Ciò rende molto più complessa l'individuazione di criteri di analisi unanimemente condivisi e il confronto dei risultati ottenuti sulla base di approcci metodologici differenti.

### 2.4.1 Raccolta e selezione del materiale emotivo: alcune problematiche

La creazione di un database di produzioni emotive ecologicamente valide rappresenta al tempo stesso una necessità e una sfida per i ricercatori. Un requisito essenziale è che la qualità della registrazione sia elevata, in quanto alcuni indici acustici sono molto sensibili alla qualità della registrazione e potrebbero essere distorti se la qualità del suono è insufficiente (Juslin, Scherer, 2008, p. 95). In secondo luogo, risulta essenziale che i ricercatori possano stabilire- al di là di ogni ragionevole dubbio- se una data emozione sia stata effettivamente espressa o intenzionalmente simulata, in un determinato contesto comunicativo (*ibid.*), occorre quindi un criterio oggettivo, molto

spesso rappresentato dalla validazione del corpus, attraverso la somministrazione di test percettivi a valutatori esterni. In terzo luogo, è importante che i campioni vocali siano "naturalsounding", ovvero, "that they are truly representative of real-world vocal affect expressions (*ibid.*); inoltre, se lo scopo dello studio sono le emozioni primarie, è necessario che le produzioni emotive siano sufficientemente intense in modo da rilevare differenze indicative tra uno stato emotivo e l'altro.

Infine, un campione eterogeneo e numeroso di parlanti è in grado di ridurre il peso di variabili individuali, sia nelle ricerche di *encoding* sia in quelle di *decoding*: "the reason for this is that there are considerable individual differences among speakers, and if one relies on a single speaker (as, in fact, several studies have done in the past), there is a serious risk that obtained data (whether in terms of decoding accuracy or patterns of voice cues) are at least partly an artifact of the idiosyncratic features of the speaker" (*ibid.*).

Uno degli aspetti più controversi riguarda la definizione e la scelta di produzioni emotive valide per l'analisi, sia nelle indagini di *decoding* sia in quelle di *encoding*. L'impostazione delle ricerche e le tecniche impiegate seguono delle direttive specifiche, in base al tipo di materiale che si intende analizzare. In linea generale, i campioni vocali utilizzati sono riconducibili a tre tipologie : a) produzioni naturali; b) produzioni elicitate mediante induzione sperimentale; c) produzioni simulate. Ciascun approccio presenta, in misura diversa, dei punti di forza e dei limiti.

#### *a) Produzioni emotive naturali*

Il parlato emotivo ricavato da questi contesti è essenzialmente genuino e autentico, in quanto prodotto in situazioni spontanee, prive di qualsiasi stimolazione. Tuttavia, come ricordano Douglas-Cowie *et al.* (2003, p. 37), "the price of naturalness is lack of control". Infatti, i dati ottenuti mediante tale procedura sono soggetti a variabili di carattere individuale (sesso ed età dei soggetti coinvolti) che incidono anche sulla durata e sul contenuto del parlato emotivo (cfr. D'Urso, Cavicchio, Magno Caldognetto, 2008). Risulta pertanto difficile isolare l'aspetto emozionale dagli altri fattori che potenzialmente influiscono sulla sua espressione vocale.

Inoltre, le maggiori criticità si riscontrano nel momento in cui è necessario operare una definizione oggettiva del materiale emotivo. Secondo D'Urso, Cavicchio e Magno Caldognetto (2008, p. 190), "Il grosso limite di questa procedura è che non si esce dalla soggettività", in quanto è lo stesso ricercatore ad attribuire, in seguito ad una

valutazione personale, una specifica etichetta emotiva a un dato elemento del corpus, senza che vi sia un riscontro esterno. Le produzioni spontanee, in virtù del loro carattere naturale e ibrido, si sottraggono all'interpretazione obiettiva dei ricercatori. Non sempre, infatti, risulta possibile discernere uno stato emotivo dall'altro; gli errori sistematici che solitamente emergono nella fase di *decoding* (cfr. PAR. 2.6) rivelano una certa tendenza da parte dei valutatori a confondere stati emotivi caratterizzati da un livello di attivazione simile (Anolli, Ciceri, 1992). Dunque, non è sempre netta la separazione tra emozioni affini come la rabbia e il disprezzo, o la sorpresa e la gioia.

Nonostante i limiti sopra citati, diversi studi<sup>27</sup> hanno prediletto l'impiego di materiale emotivo spontaneo (cfr. Toshinori Ishi, Campbell, 2004); in particolare Campbell (2000) ha suggerito il ricorso a materiale emotivo già esistente, ovvero non necessariamente destinato allo studio sperimentale (*talk show*, conversazioni telefoniche ecc.). L'autenticità che caratterizza questo tipo di produzioni potrebbe consentire un'analisi approfondita delle alterazioni fisiologiche che incidono sull'espressione vocale delle emozioni.

#### *b) Induzione sperimentale*

L'induzione sperimentale si avvale di procedure specifiche, definite MIPs (*Mood Induction Procedures*), per stimolare una reazione emotiva semi-naturale in un contesto di laboratorio. Ogni MIP ricorre ad una serie di tecniche per innescare l'*arousal* fisiologico e l'autovalutazione del soggetto. Queste procedure si dividono in cinque tipologie (cfr. Gerrards-Hesse, Spies, Hesse, 1994):

1. *MIPs based on the free mental generation of emotional states.* L'immedesimazione dei soggetti avviene essenzialmente in modo autonomo, attraverso la rappresentazione mentale degli stati emotivi. In questo caso, gli stimoli attivanti non sono introdotti da chi conduce l'esperimento, ma sono innescati dal soggetto stesso. In genere, i partecipanti sono invitati a immaginare e rivivere situazioni ed esperienze personali legate ad un determinato stato emotivo. Si tratta di una tecnica molto semplice, impiegata principalmente in ambito psicologico (cfr. DeSteno *et al.*, 2000; Gasper 2003; Zeelenberg, Pieters 2004), che ha riportato tuttavia ottimi risultati, in particolare nell'elicitazione di emozioni di base.

---

<sup>27</sup> Un esempio è costituito dall'analisi della voce di un presentatore radiofonico che annuncia in diretta l'avvicinamento di un dirigibile, che di colpo prende fuoco. Questo studio fu condotto nel 1972 da Williams e Stevens, i quali analizzarono anche la conversazione di un pilota che perde il controllo dell'aereo e muore nell'impatto con il suolo. In questi casi, ovviamente, la qualità delle registrazioni viene meno e risulta più complesso decodificare il tipo di emozione espressa.

2. *MIPs based on the guided mental generation of emotional states.* Tale tecnica si differenzia dalla prima, in quanto gli stimoli attivanti sono somministrati dai ricercatori. Il meccanismo rimane essenzialmente lo stesso, viene meno però l'autonomia dei partecipanti: ogni stato emotivo è elicitato mediante un percorso mentale guidato (e non più libero). Una delle tecniche maggiormente diffuse è quella di Velten (1968): i partecipanti ascoltano un enunciato dal contenuto emotivo (positivo, negativo o neutro), successivamente sono invitati a leggerlo (in silenzio, la prima volta; ad alta voce la seconda), cercando di immedesimarsi nella situazione emotiva suggerita dalla frase. Tale procedura è stata ampiamente utilizzata (sempre in ambito psicologico) da varie ricerche, sebbene alcuni autori abbiano avanzato dubbi sulla sua efficacia.

3. *MIPs based on the presentation of emotion-inducing material.* Si tratta di una serie di tecniche che prevede l'impiego di differenti tipologie di stimoli (verbali, visivi, sonori ecc.), in base ai quali il soggetto reagirà senza l'esplicita sollecitazione da parte dei ricercatori. Ad esempio, nel caso di stimoli verbali, l'immedesimazione dei partecipanti avviene mediante la somministrazione di storie, articoli di giornale ecc. dal contenuto emotivo (cfr. De Steno *et al.*, 2000; Brenner 2000). Gli studi che ricorrono a strumenti visivi si avvalgono, in genere, di immagini o di videoclip (cfr. Lang, Bradley, Cuthbert 2005), preposti all'attivazione di stimoli piacevoli (legati alla famiglia, alla natura) o spiacevoli (malattie, incidenti ecc.). Infine, per quanto concerne gli stimoli sonori, l'impiego della musica classica risulta particolarmente utile nell'elicitazione della gioia e della tristezza (cfr. Halberstadt, Niedenthal, 1997).

4. *MIPs based on the presentation of need-related emotional situation.* Queste tecniche mirano all'elicitazione di emozioni mediante la soddisfazione o l'insoddisfazione dei bisogni individuali. Il soggetto è inserito in un contesto che stimola l'attivazione di particolari bisogni, ad esempio la necessità di raggiungere un obiettivo specifico. Il *feedback* che gli sarà dato, positivo in caso di successo, negativo in seguito ad un insuccesso, scatenerà delle reazioni diverse, che saranno oggetto di indagine.

5. *MIPs aiming at the generation of emotional relevant physiological states.* Tali procedure ricorrono all'impiego di sostanze che alterano l'*arousal* fisiologico (ad esempio, alcuni tipi di droghe) suscitando l'insorgere di condizioni psicofisiche determinanti per alcuni stati emotivi.

Indipendentemente dalle tecniche impiegate, l'induzione sperimentale rappresenta un metodo efficace per ottenere materiale emotivo semi-spontaneo e di

buona qualità, tuttavia, alcune procedure di cui si avvale sollevano questioni etiche e morali. L'induzione di stati emotivi negativi può ledere psicologicamente il soggetto coinvolto: è necessario, pertanto, che ogni stimolazione tenga conto di una serie di fattori di natura etica, nel rispetto della dignità della persona.

*c) Produzioni simulate, scenari emotigeni, etichette lessicali*

Gli studi sul parlato emotivo condotti negli ultimi anni si sono concentrati prevalentemente sull'analisi di produzioni vocali simulate (*emotion portrayals*). Buona parte delle ricerche ha affidato l'elicitazione delle emozioni alla lettura o all'interpretazione di scenari da parte di attori o parlanti *naïf*, in grado di rappresentare una serie di situazioni emotive riconducibili al set di emozioni discrete selezionate per l'analisi. Generalmente, all'interno di ogni testo è presente un elemento invariabile (numeri, lettere dell'alfabeto, sillabe *non-sense* o anche parole, frasi, interi paragrafi, cfr. *standard content paradigm*<sup>28</sup>, Davitz 1964). Questo tipo di tecnica non lascia nessuno spazio all'improvvisazione, in quanto si richiede ai soggetti coinvolti (generalmente attori o parlanti *naïf*) di pronunciare delle frasi, associando a ciascuna di esse il tipo di emozione richiesta.

Il contesto simulato presenta molteplici aspetti positivi, in quanto consente l'elicitazione specifica di uno stato emotivo, evitando qualsiasi implicazione di tipo etico. Come spiegano Juslin e Scherer (2008, p. 96): "Advantages of this method include experimental control and the production of strong effects on voice cues. [...] One further advantage is that it is rather easy to achieve a good sound quality on the recordings".

Avvalendosi di attori professionisti e talvolta di soggetti *naïf*, il rischio che si corre, in entrambi i casi, è che le produzioni emotive risultino particolarmente enfatizzate, perciò meno naturali: "actors in a laboratory may involve vocal affect expressions that differ in important ways from those that occur in real life" (*ibid.*). Uno stato emotivo artificialmente riprodotto può rappresentare diversi svantaggi: in mancanza di uno stimolo genuino e (o indotto) esiste la possibilità che si registri una risposta fisiologica stereotipata e una produzione vocale altrettanto enfatizzata. Conseguentemente, anche l'analisi dei correlati acustici potrebbe essere falsata dalla performance dei soggetti. Ad esempio: "actors may overemphasize some voice cues,

---

<sup>28</sup> Come spiegano Juslin e Laukka (2003, p. 775): "the assumption is that, because the verbal/musical material remains the same in different portrayals, whatever effects that appear in listeners' judgments or acoustic measures should primarily be the result of the encoder's expressive intention".

such as speech rate, at the expense of more subtle cues that are harder to control, such as jitter" (*ibid.*).

Come argomenta Campbell (2000), l'impiego di attori solleva dubbi sulla naturalezza delle loro produzioni<sup>29</sup>, in quanto "[...] it is questionable whether they faithfully represent the characteristics of speech used by ordinary people when they naturally experience similar emotions" (p.114). Inoltre, nella realtà quotidiana, la manifestazione di alcuni stati emotivi può prescindere dalla volontà del soggetto ed essere quindi involontaria. A volte, invece, alcune emozioni vengono sopresse o addirittura simulate, poiché le circostanze esterne e i codici sociali lo richiedono (cfr. *display rules*, CAP.1).

In molti studi, ciò che manca, secondo Douglas-Cowie *et al.* (2003, p. 36), è l'interazione, la dinamicità della produzione: "the typical form is a non-interactive monologue, and so interpersonal effects are not represented. The context is typically minimal, so the material does not indicate how vocal signs of emotion build and fade over time, or relate to other kinds of signal" (ivi, p. 37).

Ci sono diversi modi per rendere i contesti simulati più vicini alle produzioni naturali, tra cui la tecnica di Stanislavski (Stanislavski 1988) o la ricreazione di scenari realistici che guidano l'immedesimazione dell'*encoder* nella situazione comunicativa proposta. Ad esempio, in uno studio di Anolli e Ciceri (1992), condotto sul parlato emotivo in italiano, l'elicitazione delle emozioni è stata affidata alla lettura di una serie di testi emotigeni induttori. Ogni brano conteneva delle informazioni stimolo-pertinenti-ovvero, "informazioni su stimoli esterni attivanti coerenti con ogni emozione" (ivi, p. 255), informazioni risposta-pertinenti, cioè "informazioni concernenti le risposte emozionali appropriate al contesto" (*ibid.*), e informazioni di significato, ossia "informazioni che definiscono il significato degli stimoli e delle risposte" (ivi, p. 256). I testi, di tipo narrativo, raccontavano le vicende di un protagonista maschile e i relativi stati d'animo in risposta ad un evento iniziale (atteso/inatteso; piacevole/spiacevole, rilevante/irrilevante; controllabile/incontrollabile). La seconda parte del brano consisteva in un monologo del protagonista "come reazione ed espressione emozionale appropriata e credibile ai fattori [...] presenti nell'episodio iniziale" (ivi, p. 257).

---

<sup>29</sup> Come riconoscono Douglas-Cowie *et al.* (2003, p. 36) "it is certainly true that good actors can generate speech that listeners classify reliably. Material studied by Banse and Scherer (1996), for example, produced recognition rates of 78% for hot anger, 76% for boredom and 75% for interest, though scores for other emotions were lower with an average recognition rate of 48% across 14 emotions", tuttavia l'alta percentuale di riconoscimenti non chiarisce in che misura il parlato simulato si avvicini al parlato spontaneo (*ibid.*).

All'interno di ogni testo è stata inserita una frase standard, *non è possibile, non ora*, in grado di assumere una specifica sfumatura emozionale in funzione del contesto emotigeno in cui veniva pronunciata, evidenziando la variazione dei tratti paralinguistici.

Questa soluzione, spiegano i due autori, pur presentando degli ovvi elementi di artificialità, "costituisce una risorsa, perché consente di raccogliere dati comparabili, empiricamente controllabili, dotati di un valore semantico definito e accertato, meno ambiguo di quanto succeda con l'espressione spontanea di reazioni emotive" (ivi, p. 249).

Secondo alcuni autori (cfr. D'Urso, Cavicchio, Magno Caldognetto, 2008) l'utilizzo di etichette lessicali sarebbe preferibile agli scenari, in quanto questi ultimi "hanno il limite di evocare nella mente di chi deve simulare l'emozione stati emotivi che possono variare molto a causa dell'interazione tra scenario e persona" (p. 191). Infatti, in contesti controllati, le parole "hanno una rappresentazione mentale che si avvicina al loro significato lessicale, rintracciabile nei vocabolari", al contrario "gli scenari presentano situazioni emotigene la cui valenza dipende da molte circostanze di vita, quali età, sesso, condizione economica eccetera" (*ibid.*); inoltre, le situazioni proposte sono, nella maggior parte dei casi, una trasposizione stereotipata della realtà quotidiana, per cui risulta difficile conoscere la "reale risonanza emotiva" (*ibid.*) di questo tipo di produzioni.

Il quadro metodologico delineato fin qui appare quanto mai controverso. L'induzione sperimentale e la simulazione (affidata all'interpretazione di un testo) sembrano rappresentare due soluzioni ottimali e soprattutto complementari, in quanto possono trovare entrambe una loro applicabilità nello stesso contesto empirico. Da un lato, l'induzione sperimentale garantisce un grado maggiore di immedesimazione dei soggetti, dall'altro, la lettura di uno scenario o di un testo emotigeno offre un controllo sul contenuto linguistico e sulle emozioni da elicitare; infine, riprendendo il suggerimento di Douglas-Cowie *et al.* (2003), l'adozione degli scenari è auspicabile nella misura in cui la situazione comunicativa proposta sia legata ad un contesto interattivo, in cui la dinamicità del parlato e l'empatia tra gli interlocutori possano contribuire ad una maggiore naturalezza delle produzioni emotive.

#### 2.4.2 Il processo di decodifica

Le ricerche sul processo di decodifica costituiscono una parte fondamentale nello studio dell'espressione vocale delle emozioni, per una serie di ragioni. In primo luogo, queste ricerche restituiscono un dato importante circa la capacità degli esseri umani nel riconoscere gli stati emotivi a partire dagli indizi vocali. Un'altra ragione che rende gli studi sul processo di decodifica indispensabili è che essi consentono di verificare, attraverso l'implementazione di test percettivi, l'effettiva corrispondenza tra l'emozione espressa dal parlante e l'interpretazione data dal valutatore, fornendo una prova indiretta sulla naturalezza e sull'efficacia delle produzioni emotive. A questo scopo, una batteria di stimoli vocali viene in genere sottoposta ad un campione di giudici valutatori, chiamati a *discriminare* (cioè a selezionare una tra le opzioni proposte) o a *riconoscere* lo stato emotivo espresso attraverso un'interpretazione personale, non vincolata ad un'etichetta emotiva (*free labeling* o *free description*).

Il modo in cui le emozioni espresse dal parlante vengono decodificate dall'ascoltatore è ancora oggetto di discussione. Secondo alcuni autori, i giudizi espressi si basano su categorie emotive discrete, secondo altri su dimensioni variabili. Come spiegano Juslin e Scherer: "A discrete-emotions approach to vocal expression [...] would predict that emotions are perceived as discrete categories" (2008, p. 111) e in effetti, alcuni risultati preliminari (de Gelder, Vroomen 1996; Laukka 2004) sembrano supportare questa ipotesi (in linea con i risultati ottenuti negli studi sulle espressioni facciali). Tuttavia, "it cannot be ruled out that categorical responses are explicitly or implicitly influenced by the existence of discrete, verbal emotion concepts" (Juslin, Scherer *ibid.*).

##### 2.4.2.1 I test percettivi

La selezione del tipo di test da sottoporre ai giudici e la scelta del formato delle risposte sollevano una serie di problematiche. I test a risposta chiusa (*forced-choice format*) consentono di confrontare i dati con i risultati delle ricerche che hanno adottato questo formato. Tuttavia, l'accuratezza di questi test "could be artificially inflated or deflated if there is a small number of response options simply because the participants are unable to choose other, potentially more applicable response options" (Juslin, Scherer, 2008, p. 111). Il problema potrebbe essere parzialmente risolto, suggeriscono Frank e Stennett (2001), inserendo un'ulteriore opzione ("other emotion"), per dare ai

valutatori la possibilità di esprimere il proprio disaccordo rispetto alle etichette emotive proposte.

Un'alternativa ai test a risposta chiusa è il ricorso ad un elenco di aggettivi in luogo delle etichette emotive. La scelta dei valutatori, in questo caso, non è limitata ad una sola opzione, ma essi sono liberi di scegliere tra gli aggettivi in elenco quelli che qualificano meglio lo stimolo percepito (a prescindere se questo venga riconosciuto o meno). Tuttavia, "adjective checklists do not offer a standard measure of decoding accuracy, and the results may thus be more complicated to analyze" (*ibid.*). Alcune ricerche associano alle liste di aggettivi un giudizio quantitativo (ad esempio, su una scala da 1 a 7), ma anche in questo caso, i risultati sono difficilmente confrontabili con gli studi precedenti.

I test a risposta aperta, che consentono una descrizione personale e non vincolata ad un set di etichette lessicali, non rappresentano una soluzione auspicabile (se non come indagine preliminare), soprattutto a ragione della grande variabilità che caratterizza le risposte e i giudizi forniti dai valutatori. Un ulteriore problema, a questo proposito, è rappresentato dalle etichette emotive: "there are several different words for the same emotion, and it is hard to decide whether or not different judges are perceiving the same emotion (i.e. regardless of the particular verbal label used)" (ivi, p. 112). In ogni caso, sostengono Juslin e Scherer (*ibid.*), "it is desirable to use a wider variety of response formats in future research on vocal expression [...] perhaps combining forced-choice with free description (as proposed by Rosenthal 1982) or using computer tools for continuous measurement of emotion perception from voice cues (Cowie *et al.* 2000)".

## **2.5 I profili acustici delle emozioni primarie**

Alla luce delle ricerche condotte da Scherer e colleghi (cfr. Scherer 1989; Pittam, Scherer, 1993; Banse, Scherer, 1996), è possibile ricostruire un quadro orientativo dei profili acustici che caratterizzano cinque emozioni di base (collera, paura, gioia, tristezza e disgusto). I risultati empirici riportati derivano dall'analisi delle produzioni emotive di attori e parlanti *naïf*, per cui in alcuni casi (in particolare per quel che concerne il disgusto) si osservano delle differenze sostanziali nella variazione dei tratti soprasedimentali. I parametri indagati sono: la frequenza fondamentale ( $f_0$ ) e i suoi

valori relativi all'altezza, al *range* e al contorno intonativo, l'intensità, la distribuzione dell'energia nello spettro delle frequenze; il tempo e le pause.

- Collera (*anger*): si tratta di un fenomeno emotivo caratterizzato da un alto livello di attivazione, pertanto si registra un aumento della media di  $f_0$ , della media dell'intensità e dell'energia ad alta frequenza. I profili intonativi sono discendenti e si rileva inoltre un incremento del ritmo di articolazione (*articulation rate*). In genere, gli studi condotti tendono a distinguere collera fredda da collera calda, di conseguenza alcuni parametri subiscono una certa variabilità in relazione all'intensità dello stato emotivo indagato. Come spiegano Anolli e Ciceri (1992, p. 244): "la voce della collera è caratterizzata da una voce stretta, notevolmente tesa, ma estremamente ampia".

- Paura (*fear*): anche questa emozione si distingue per un'elevata attivazione. Da un punto di vista acustico ciò trova corrispondenza nell'aumento della media di  $f_0$ , nel *range* e nell'energia ad alta frequenza. Inoltre, il ritmo di articolazione si presenta particolarmente sostenuto. Ne risulta un tipo di vocalizzazione "stretta, estremamente tesa ed esile" (*ibid.*).

- Tristezza (*sadness*): si tratta di un'emozione caratterizzata da una bassa attivazione, la voce risulta pertanto rilassata. I valori relativi alla media  $f_0$ , alla sua gamma di variazione, all'intensità e all'energia ad alta frequenza risultano particolarmente bassi. A ciò si associa inoltre un profilo intonativo solitamente discendente (e un ritmo di articolazione ridotto).

- Gioia (*joy*): essendo un'emozione attivante, i valori acustici rispecchiano un incremento della media di  $f_0$ , del *range*, della media dell'intensità e dell'energia ad alta frequenza. Inoltre, il ritmo dell'articolazione tende a salire. Per gioia, in questo caso, si intende uno stato emotivo euforico ed esaltato (e non un sentimento più equilibrato come la felicità). La voce della gioia risulta "ampia, piena e mediamente tesa" (*ibid.*).

- Disgusto (*disgust*): le tecniche impiegate per l'elicitazione di questa emozione hanno prodotto profili acustici discordanti. Alcune ricerche si sono avvalse di procedure di induzione mediante film (cfr. Scherer *et al.* 1985), altre invece si sono servite di attori per la simulazione. Nel primo caso, si è rilevato un incremento della media di  $f_0$ , mentre nel secondo si è registrata una diminuzione della media di  $f_0$  (Scherer, 1986; 1989). Secondo Scherer (1986) questa differenza potrebbe essere attribuita al processo di autoregolazione dettato dalle norme sociali, le quali tendono a modulare l'espressione vocale.

Le conclusioni tratte da queste analisi mostrano che emozioni caratterizzate da un grado di attivazione simile condividono parametri acustici analoghi. In particolare, tale associazione risulta evidente in presenza di stati emotivi corredati da un'elevata attivazione come paura, collera e gioia. Queste emozioni, infatti, presentano degli andamenti intonativi che si estendono nelle fasce frequenziali più alte (indicando un coinvolgimento emotivo importante). Anche i valori relativi all'intensità riflettono l'incremento generale degli indici soprasegmentali. La durata delle produzioni emotive, invece, si riduce visibilmente,

Per quanto riguarda gli stati emotivi caratterizzati da una bassa attivazione, come la tristezza e il disgusto, il loro contorno intonativo appare abbastanza statico e si situa soprattutto nelle fasce frequenziali più basse. L'escursione tonale è molto ridotta, mentre la durata viene enfatizzata, soprattutto nelle espressioni di disgusto.

Questi correlati sono stati rinvenuti per lo più in produzioni emotive in lingua inglese e costituiscono un punto di riferimento indicativo per quel che concerne la variabilità degli indici acustici coinvolti nell'espressione vocale delle emozioni di base<sup>30</sup>. Naturalmente, le ricerche condotte negli ultimi anni sono orientate verso la creazione di database multilingue in grado di fornire un confronto a livello interculturale di queste variazioni, e di gettare luce sulle capacità degli esseri umani nell'inferire uno stato emotivo espresso in una lingua diversa da quella nativa.

## **2.6 Tendenze comuni nel riconoscimento delle emozioni a livello intralinguistico**

La grande variabilità che caratterizza le metodologie impiegate nelle ricerche di *decoding* non consente un confronto sistematico e oggettivo dei dati ottenuti. È possibile comunque trarre una serie di conclusioni dalle ricerche che hanno seguito un percorso comune o simile.

In linea generale, la decodifica delle emozioni (a livello intra-linguistico) ha ottenuto mediamente un indice di accuratezza superiore alla casualità (oltre il 50%, in alcuni casi il 60%). Tuttavia, non tutte le emozioni hanno registrato le medesime percentuali di riconoscimento. Ad esempio, sembra che le emozioni negative (come

---

<sup>30</sup> Per quanto riguarda l'italiano, le ricerche condotte si sono focalizzate sia sul parlato simulato/indotto (Anolli, Cicero, 1992) sia sul parlato teatrale (Gili Fivela, Grimaldi, Stefano, 2005) e cinematografico (Pettorino, 2008), confermando sostanzialmente la presenza di variazioni simili a quelle riportate in altri studi.

rabbia e tristezza) siano decodificate con un grado di accuratezza più alto rispetto a quelle positive; l'emozione comunemente meno riconosciuta è invece il disgusto (Pittam, Scherer 1993).

Un'altra questione interessante riguarda gli errori di decodifica: alcune emozioni sono sistematicamente confuse con altre. È il caso del disprezzo- confuso con la rabbia o il disgusto- e la gioia, interpretata spesso come sorpresa. Alcuni ricercatori (cfr. Van Bezooijen 1984) sostengono che tale confusione sia generata dall'associazione di emozioni con un grado di attivazione simile (ad esempio rabbia e paura). Secondo altri invece, alcuni stati emotivi rientrano nella stessa sfera semantica e tale "vicinanza" provoca una sorta di osmosi tra emozioni diverse (Anolli, Ciceri, 1992).

Infine, per quanto concerne l'abilità dei valutatori nel decodificare gli stimoli emotivi, non sono emerse differenze sostanziali tra uomini e donne. L'età sembra invece essere determinante in tal senso: alcune ricerche (cfr. Van Bezooijen, 1984) hanno dimostrato che la capacità di *decoding* aumenta nel corso della vita, per cui i parlanti affinano i meccanismi di riconoscimento non solo a livello vocale, ma anche sul piano visivo. Tuttavia, studi recenti hanno sostenuto che in età matura tale capacità sia meno efficiente (Mitchell, Kingston, 2011), in quanto la riduzione delle facoltà uditive comporterebbe una percezione più limitata delle frequenze del parlato emotivo. Infatti, secondo Mitchell e Kingston, "older adults are also less able to identify nonverbal intonation cues, and have particular difficulty with the pragmatic functions of these intonational cues in conveying communicative intents such as speaker emotion" (2011, p. 156).

### *2.6.1 Il ruolo dei tratti paralinguistici nel riconoscimento delle emozioni*

L'efficacia dei segnali paralinguistici nella comunicazione vocale è stata documentata da quelle ricerche che hanno manipolato il segnale vocale, isolando o distorcendo i singoli parametri acustici ( $f_0$ , intensità, durata) al fine di stabilire il ruolo giocato da ognuno di essi nel favorire o ostacolare il riconoscimento delle emozioni (Anolli, Ciceri, 1992, p. 221). Se il contorno intonativo contribuisce a veicolare informazioni di natura emotiva soltanto in combinazione con altri parametri (tra cui l'escursione tonale, l'intensità, la durata), sembra invece che le qualità vocali contribuiscano in maniera considerevole al riconoscimento degli stati emotivi (Lieberman, Micheals 1962).

A questo proposito, occorre specificare che molti studi si sono focalizzati su quei fenomeni vocali che Scherer definisce *affect burst* (*pure non-linguistic affective bursts*<sup>31</sup>, Scherer 1986; 1994), ovvero le interiezioni primarie (Poggi, 1981), in virtù del fatto che in queste produzioni "le caratteristiche prosodiche e intonative paralinguistiche si presentano in un contesto fonetico semplice che esclude o riduce l'interazione con le caratteristiche fonologiche segmentali e soprasegmentali del parlato emotivo" (Magno Caldognetto, Cavicchio, Così, 2008, p. 187).

Schröder (2003) ha comprovato l'efficacia degli *affect bursts* nel veicolare specifici significati emotivi, in modo indipendente rispetto al contesto. Nella sua definizione, gli *affect burst* sono: "short, emotional nonspeech expressions, comprising both clear nonspeech sounds (e.g., laughter) and interjections with a phonemic structure (e.g., 'Wow!'), but excluding 'verbal' interjections that can occur as a different part of speech (like 'Heaven!,' 'No!,' etc.)", (Schröder, 2003, p. 103). Gli indici di riconoscimento riportati per le 10 emozioni indagate (ammirazione, paura, disgusto, gioia, noia, calma, sorpresa, preoccupazione, disprezzo, rabbia) sono piuttosto elevati (in media oltre l'80%). In effetti, questi dati lasciano intuire che alcune emozioni (come ad esempio il disgusto) siano veicolate più facilmente attraverso un'interiezione ("Ih!) e meno attraverso le risorse prosodiche e le qualità vocali. Gli indici di riconoscimento del disgusto a partire dai tratti prosodici e dalla qualità vocale risultano in effetti di gran lunga inferiori (intorno al 15%, cfr. Banse, Scherer 1996).

## **2.7 Il parlato emotivo in una prospettiva transculturale**

### *2.7.1 Universalità e specificità culturale: verso un approccio interazionista*

La dibattuta questione sulla universalità vs. specificità culturale dei fenomeni emotivi investe anche la loro espressione vocale. In linea generale, le ricerche compiute in questo ambito hanno tentato, più o meno esplicitamente, di rintracciare prove tangibili a sostegno dell'una o dell'altra ipotesi, molto spesso giudicando le due prospettive inconciliabili tra loro. La continuità filogenetica nella comunicazione vocale, postulata già da Darwin nel 1872, costituisce una prova inconfutabile per i sostenitori dell'universalità (cfr. CAP. 1). Sul versante opposto, tale impostazione,

---

<sup>31</sup> Nella definizione data da Scherer, gli *affect bursts* sono "very brief discrete non verbal expressions of affect in both face and voice as triggered by clearly identifiable events" (Scherer, 1994, p. 170).

tacciata spesso di etnocentrismo, viene rigettata in favore di una visione maggiormente attenta alle diversità culturali in grado di incidere profondamente sulle modalità di espressione e percezione delle emozioni.

Le ricerche attuali sono orientate verso un approccio interazionista, che tenta di conciliare entrambe le posizioni (Rosenthal *et al.* 1979; Matsumoto, 1989, Russell, 1994; Mesquita, Frijda, Scherer, 1997; Elfenbein, Ambady, 2002b). La prospettiva interazionista si fa carico di spiegare l'ampia variabilità che caratterizza il riconoscimento delle emozioni a livello interlinguistico. Infatti, i risultati empirici finora raccolti sembrano confermare che "members of different cultures appear to vary substantially in their emotion recognition accuracy" (Elfenbein, Ambady, 2003, p. 92).

In questo tipo di approcci, sostengono Zinken, Knoll, Panksepp (2008), la relazione tra universalità e diversità culturale è stata rappresentata lungo un *continuum*: un polo del *continuum* rappresenterebbe i meccanismi universali sottesi all'espressione vocale delle emozioni, il polo opposto costituirebbe il successivo adattamento dell'espressione in funzione del particolare, ovvero della specificità culturale. Questo modello è rintracciabile nelle due metafore proposte da Scherer (1986) per conciliare entrambe le prospettive: "l'effetto spingere/tirare" (*push* e *pull effects*) nella comunicazione vocale. Come precedentemente illustrato, questi due fattori entrano in gioco durante l'interazione umana, e uno dei può avere il sopravvento sull'altro, determinando produzioni emotive molto diverse.

Il modello del *continuum* suggerisce che si possano rintracciare meccanismi universali nell'espressione vocale delle emozioni individuando tendenze comuni nel comportamento dei membri di culture diverse (in particolare, nell'interpretazione degli stati emotivi). Tuttavia, secondo Zinken e colleghi, il modello non tiene in considerazione il fatto che la comunicazione emotiva intenzionale, soggetta dunque a regole culturalmente determinate, e le vocalizzazioni affettive che Scherer definisce *affect bursts* appartengono a due sistemi semiotici differenti (Zinken, Knoll, Panksepp, 2008). Le vocalizzazioni condivise da specie diverse appartengono ad un sistema di tipo sintomatico; al contrario, le emozioni in quanto tali appartengono ad un sistema iconico, che richiede abilità avanzate nella gestione e nell'interpretazione dei segni (a maggior ragione in un contesto simulato, in cui i partecipanti devono comunicare intenzionalmente il proprio stato emotivo). Entrambi i sistemi, sostengono Zinken, Knoll e Panksepp, presentano degli aspetti universali, tuttavia la relazione tra le

tendenze comuni all'interno di un sistema, e la struttura e le funzioni all'interno dell'altro, non è così semplice come suggerisce il modello del *continuum*.

Alla luce di ciò, gli autori propongono un modello in grado di chiarire, almeno parzialmente, tale relazione. Alla base, vi è una separazione tra il sistema emozionale (*affect bursts*) e il sistema comunicativo intenzionale. Panksepp (1982; 2005) definisce le vocalizzazioni emotive presenti in tutti i mammiferi *evolved action systems*. Negli animali, molti di questi suoni hanno un valore comunicativo durante l'interazione, e sono modulati in funzione del contesto e del tipo di relazione (attacco/difesa; dominanza/sottomissione). Per cui, non è difficile immaginare che anche nella specie umana, a dispetto dell'elevata acculturazione e delle conseguenti norme sociali, tali suoni abbiano un impatto emotivo distintivo durante l'interazione con gli altri membri (*ibid.*). Può darsi tuttavia che le vocalizzazioni emotive non siano strettamente connesse agli aspetti cognitivi di emozioni quali collera o gioia, ma che siano piuttosto parte integrante degli *arousal* motivazionali condivisi da molte specie, (Panksepp, 2005). Ad esempio, la tensione dei muscoli respiratori e laringali può influire sul parlato, incrementandone l'intensità e la frequenza fondamentale. In questo caso, risulta complesso chiarire se i suoni in questione siano essenzialmente comunicativi o piuttosto effetti secondari delle espressioni emotive (Zinken, Knoll, Panksepp, *ibid.*).

Lo studio delle vocalizzazioni sintomatiche nell'uomo risulta molto più complesso, in quanto è difficile analizzare le produzioni vocali non verbali senza ricorrere alle categorie lessicali. Inoltre, i vincoli sociali e culturali impongono, in misura diversa, un controllo cosciente dell'emotività (in particolare negli adulti) in funzione di ciò che è ritenuto accettabile o meno nella cultura di riferimento. In quest'ottica, il contributo di questi indici acustici all'interno della comunicazione dipenderebbe essenzialmente dal modo in cui ogni cultura intende l'espressione emotiva in termini di utilità e accettabilità. Ciò non esclude che l'espressione vocale- di tipo sintomatico- degli stati emotivi di base, presenti delle affinità sul piano acustico a livello interculturale (Juslin, Laukka, 2003; Panksep, Bernatzky, 2002), essendo associata a determinati mutamenti fisiologici (incremento della tensione muscolare e attivazione del sistema simpatico).

Al di là dell'aspetto puramente sintomatico, l'espressione vocale delle emozioni assume un valore funzionale all'interno della comunicazione umana (Scherer, 1992). I membri di un sistema culturale, in quanto parlanti di una lingua, ricorrono intenzionalmente alle modulazioni della voce per raggiungere determinati obiettivi . La

velocità del parlato, il volume, l'intonazione, ecc. subiscono delle alterazioni in base al contesto (formale/informale), al tipo di interlocutore e in funzione degli scopi che si intende perseguire (fare una richiesta, scusarsi, persuadere, ecc). In virtù del suo carattere intenzionale, la comunicazione vocale delle emozioni altera la struttura acustica in misura tale da non poter essere spiegata soltanto in termini di controllo culturale, ma implica anche aspetti cognitivi e valutazioni individuali (cfr. CAP. 1)

I limiti imposti dal carattere funzionale della comunicazione- affermano Zinken, Knoll e Panksepp- sono di per sé universali, tuttavia la loro relazione con le vocalizzazioni affettive varia in base al contesto linguistico e culturale. Alla luce di ciò, universalità e diversità non si escludono a vicenda. Gli autori postulano l'esistenza di un numero di emozioni condivise, sul piano pre-concettuale, da culture diverse che va al di là delle sei emozioni di base generalmente considerate universali. Tale universalità non esclude tuttavia che l'espressione delle emozioni, nella misura in cui implichi aspetti intenzionali e creativi, possa rivelarsi estremamente diversificata a livello interculturale. La categorizzazione della realtà operata dal linguaggio a sua volta dà origine ai processi cognitivi e comunicativi propri di ogni cultura<sup>32</sup> (cfr. CAP. 1).

### *2.7.2 Stato dell'arte degli studi interculturali sul parlato emotivo*

Gli studi sul parlato emotivo di carattere interculturale sono relativamente pochi rispetto alle ricerche effettuate nell'ambito delle espressioni facciali (cfr. Ekman 1972; Ekman *et al.* 1987). Elfenbein e Ambady (2002) hanno condotto una ricognizione di 97 esperimenti legati al riconoscimento interculturale delle emozioni attraverso stimoli visivi (espressioni facciali) e uditivi (porzioni di parlato). Questi stimoli sono stati sottoposti a membri di culture diverse (per lo più occidentali). In generale, il

---

<sup>32</sup> Cionondimeno la diversità non preclude l'esistenza di aspetti universali. Si pensi a due culture che abbiano codificato – convenzionalmente- l'espressione vocale della gioia attraverso modalità estremamente diverse. Nella cultura A, l'emozione in questione è espressa- sul piano vocale- mediante l'incremento della variazione frequenziale. Nella cultura B, la manifestazione della gioia è codificata attraverso un aumento della velocità d'eloquio. Sul piano acustico, le due espressioni vocali sono dunque estremamente diverse. Tuttavia, ciò non toglie che un membro della cultura B possa identificare correttamente l'emozione espressa da un membro della cultura A, individuando in essa una possibile simulazione delle vocalizzazioni affettive della gioia (Zinken, Knoll, Panksepp, *ibid.*). In altre parole, l'emozione in questione potrebbe essere decodificata in base ad altre caratteristiche acustiche (variazioni d'intensità e di timbro, incidenza del sorriso sul parlato, ecc.). Le probabilità di riconoscimento aumentano se la risposta è limitata ad un certo numero di opzioni (gioia, tristezza, collera, ecc.). In un certo senso, la diversità - pur essendo estrema - non intacca l'aspetto potenzialmente universale rappresentato dalla struttura acustica delle vocalizzazioni affettive (come il sorriso, il pianto, il sospiro, lo sbadiglio, ecc.).

riconoscimento delle emozioni affidato alle espressioni del volto e ai movimenti del corpo ha un indice di accuratezza che supera di gran lunga le percentuali di decodifica relative alla prosodia emotiva (Thompson, Balkwill, 2006). Tuttavia sono molto pochi gli studi che hanno coinvolto culture non occidentali; un confronto di questo tipo sarebbe auspicabile, in quanto "there is a long history of close contact between speakers of Western languages that might explain similar uses of prosody [...]. Similar uses of prosody in Western and non-Western languages, on the other hand, would provide compelling evidence for universal principles of emotional prosody (ivi, p. 408).

Buona parte delle ricerche sul parlato emotivo ha tentato di individuare tendenze comuni nell'interpretazione delle espressioni vocali delle emozioni a livello interculturale. Tale approccio dovrebbe fornire una prova indiretta dell'esistenza di elementi universali (o specificamente culturali) nella comunicazione emotiva. Alcuni studi, meno numerosi, si sono focalizzati sull'analisi dei parametri acustici e della loro variazione a livello interlinguistico. Entrambi gli approcci si fondano essenzialmente sulla raccolta, sulla validazione uditiva, sull'analisi e la comparazione di diversi corpora, contenenti le emozioni prodotte da parlanti di lingue differenti.

#### *2.7.2.1 Le ricerche sull'encoding*

Al fine di offrire un quadro esauriente circa lo stato dell'arte delle ricerche sull'*encoding* a livello interculturale, in questo paragrafo saranno presentati e discussi i risultati di alcuni esperimenti condotti negli ultimi anni.

In ambito italiano, Pettorino (2004) ha proposto un'indagine acustico-percettiva sui correlati acustici delle emozioni nel parlato cinematografico, attraverso un'analisi contrastiva di alcune frasi emotive pronunciate da un attore americano (Al Pacino, nel film *Scent of Woman*) e delle corrispondenti frasi doppiate in italiano e in giapponese. Come spiega l'autore, "la scelta del parlato doppiato come oggetto di analisi è stata motivata dalla considerazione che una delle maggiori difficoltà che si incontrano nelle indagini rivolte a individuare i correlati acustici delle emozioni, è quella di riprodurre, con l'ausilio di un attore, un determinato stato d'animo del parlante" (ivi, p. 45). Al contrario, "una frase cinematografica, nella sua versione originale e doppiata, è costretta a veicolare uno stesso stato emozionale in quanto restano immutati tutti i livelli che intervengono nella trasmissione del messaggio" (*ibid.*). Le frasi selezionate sono state ritenute rappresentative di sei stati emotivi di base (gioia, paura, sorpresa, tristezza, disgusto, collera). Inoltre, al fine di garantire un confronto con l'eloquio normale, sono

state selezionate alcune battute prive di sfumature emotive, e quindi neutre. Un test percettivo a risposta chiusa è stato somministrato ad un campione di 20 studenti universitari, per individuare le frasi con le percentuali di riconoscimento più elevate. L'analisi acustica si è focalizzata quindi soltanto sulle battute identificate con maggiore accuratezza e rappresentative di quattro stati emotivi: collera, gioia, tristezza e neutro. I parametri indagati sono stati i seguenti: durata delle sillabe toniche e atone, rapporto tra sillabe toniche e sillabe atone, valori massimi minimi e medi di  $f_0$ , *pitch range*, velocità d'articolazione.

Dalle analisi è emerso che tutti e tre i parametri indagati contribuiscono a differenziare l'eloquio sul piano emozionale, sia tra uno stato emotivo e l'altro, sia nel confronto tra le tre lingue (ivi, p. 55). In relazione alla frequenza fondamentale, la collera ha evidenziato i valori più elevati e variati, mentre la gioia valori intermedi e più stabili. I valori più bassi sono stati rilevati negli enunciati di tristezza. In conclusione, i risultati hanno evidenziato alcune similarità nelle produzioni di collera, gioia e tristezza in tutte e tre le lingue, sebbene la durata sillabica<sup>33</sup> sia stato il parametro più sensibile alle differenze prosodiche esistenti tra di esse.

Uno studio condotto da Anolli *et al.* (2008) ha messo a confronto il parlato emotivo di parlanti di madrelingua italiana e cinese. Le emozioni (otto in totale, insieme all'equivalente neutro della frase standard: gioia, collera, paura, tristezza, vergogna, orgoglio, disprezzo e senso di colpa) sono state elicitate mediante la lettura di scenari emotigeni tradotti nelle due lingue, contenenti una frase standard *non è possibile, non ora* (il cui significato è stato mantenuto nella traduzione cinese).

Dai risultati sono emerse differenze sostanziali tra i due corpora (e tra le emozioni considerate). In relazione al parametro temporale, la velocità d'eloquio dei cinesi è risultata inferiore rispetto a quella degli italiani (anche le pause avevano una durata maggiore). Sul piano frequenziale, è stata riscontrata una variazione minore del *pitch* nelle produzioni emotive dei parlanti cinesi rispetto alle espressioni neutre; per contro, nel parlato degli italiani tale variazione risultava maggiore. Anche in relazione

---

<sup>33</sup> Per quanto riguarda la durata delle sillabe atone, sia in inglese che in italiano, si è registrata una scarsa variabilità. La durata delle sillabe toniche sembrerebbe invece essere discriminante, consentendo una distinzione in due classi di durata: lunghe (collera e tristezza) e brevi (neutro e gioia) per quanto riguarda l'inglese; mentre in italiano, lunghe per tristezza e gioia, brevi per il neutro e la collera (*ibid.*). Il giapponese si è invece caratterizzato per una grande variabilità, tranne per la tristezza (in cui si allungano sia le toniche che le atone, *ibid.*). Per quel concerne la velocità di articolazione, secondo l'autore "la presenza di uno stato emotivo primario determina un rallentamento dell'eloquio. La caratterizzazione emotiva del parlato richiede sempre una aggiunta di qualcosa che a sua volta provoca una decelerazione" (*ibid.*).

all'intensità, i risultati hanno evidenziato un andamento simile: il parlato degli italiani presentava valori più alti e variazioni più ampie rispetto a quello dei cinesi.

Gli autori hanno interpretato i dati ottenuti proponendo una duplice conclusione: da un lato, i risultati hanno supportato l'ipotesi evuzionista, e dunque la continuità filogenetica della comunicazione emotiva non verbale; dall'altro, le differenze riscontrate tra i due corpora hanno evidenziato modalità di espressione vocale strettamente connesse al contesto culturale dei parlanti. In base all'interpretazione data dagli autori, la cultura cinese- al pari delle altre culture asiatiche- "stresses relational harmony and invites individuals to take their proper place, discouraging them from occupying too much space in relationships" (ivi, p. 583). Tale impostazione culturale si rifletterebbe nella tendenza a ridurre l'espressività nella comunicazione emotiva, poiché un atteggiamento di questo genere sarebbe in contrasto con le norme sociali che tentano, al contrario, di preservare lo spazio di ciascun membro.

Piot (1999) ha condotto uno studio sugli indici prosodici impiegati nella manifestazione di emozioni e atteggiamenti in quattro lingue diverse: inglese americano, coreano, finnico e vietnamita. I partecipanti (4 parlanti *naïf* per ogni gruppo) sono stati invitati a pronunciare due frasi, una delle quali era un nome proprio di persona (*Natacha*), tradotto nelle rispettive lingue; l'altra invece era costituita da un numero maggiore di sillabe (al fine di evidenziare differenze sul piano ritmico) e faceva riferimento ad uno spostamento del parlante (*OK, now, I'm going to the station to get that packet*). Ogni frase era associata ad un contesto che specificava un particolare atteggiamento o un'emozione da comunicare, ad esempio, ad un amico ("to a friend, after a deception, "sad") o al proprio figlio ("to your child who has just disobeyed you, warning before acting severely"). I parametri misurati sono stati: ampiezza, media e profilo di f0, intensità, durata media di ogni sillaba. I risultati delle analisi hanno portato l'autore ad affermare che, a dispetto delle differenze sul piano prosodico, le lingue considerate mostravano notevoli somiglianze nell'espressione delle emozioni e degli atteggiamenti presi in esame, in linea con le precedenti ricerche. Ad esempio, la tristezza era caratterizzata da una velocità di eloquio molto lenta in tutte e quattro le lingue e da una riduzione dei valori di f0 e intensità.

Tra i vari studi in ambito interlinguistico, ricordiamo anche il corpus elaborato da Goudbeek e Broersma (2010), il DEMO (DutchEmotion)/ KEMO (KoreanEmotion) Corpus. Quest'ultimo comprende otto emozioni (gioia, collera, irritazione, tenerezza, paura, orgoglio, sollievo e tristezza) espresse da parlanti di origine coreana e olandese.

Le emozioni sono state veicolate da una frase standard priva di significato, elaborata selezionando fonemi presenti in entrambe le lingue. Il corpus è stato validato attraverso la somministrazione di un test uditivo (contenente le produzioni emotive) ad un campione di ascoltatori coreani e olandesi, chiamati a giudicare il parlato emotivo in termini di naturalezza. Una volta validato il corpus, gli autori hanno confrontato la durata dei fonemi nelle rispettive lingue, ipotizzando una diversa gestione temporale dei segmenti fonetici dovuta alla specificità linguistica e culturale. I risultati hanno confermato la presenza di differenze sostanziali tra le due lingue. Il parlato dei coreani ha evidenziato, in media, una durata inferiore rispetto a quello degli olandesi; i segmenti fonetici hanno presentato dunque un pattern temporale diverso, sia nel confronto tra le due lingue sia all'interno delle emozioni espresse in una sola lingua. I risultati hanno portato gli autori ad affermare che "language as well as emotion has a significant influence on the duration of a segment. [...] these two factors often interact with one another, thus creating language specific effect of emotion on phonetic segment duration" (ivi, p.4).

L'interesse per il parlato emotivo ha inoltre indirizzato l'attenzione dei ricercatori sull'analisi e il confronto di un altro tipo interazione emotiva, ovvero il parlato prodotto nell'interazione madre-bambino (*child-directed speech, CDS*). Le variazioni prosodiche associate a questo tipo di parlato (*pitch* più alto, contorni intonativi accentuati, iper-articolazione delle vocali, frasi brevi, pause più lunghe) sono comunemente ritenute universali, sebbene alcuni studi abbiano dimostrato tendenze discordanti a livello cross-linguistico. Ad esempio, Kitamura *et al.* (2002) hanno confrontato il parlato emotivo prodotto da madri australiane e thailandesi, registrando differenze sostanziali nell'eloquio delle madri thailandesi, le quali tenderebbero a ridurre l'escursione tonale per non compromettere l'informazione tonale<sup>34</sup>, ricorrendo ad altre modificazioni vocali (ad esempio, particelle finali per esprimere stati d'animo).

Diversi studi hanno confermato una tendenza simile in altre lingue tonali. Ad esempio, in cinese questo tipo di parlato presenta dei valori frequenziali ridotti nel confronto con l'inglese americano (Toda, Fogel, Kawai, 1990; Papoušek, Hwang, 1991). Anche in giapponese la riduzione degli indici prosodici risulta compensata sul piano linguistico da un maggiore ricorso al lessico affettivo. Le differenze riscontrate a livello cross-linguistico dimostrano che le proprietà acustiche del CDS non sono

---

<sup>34</sup> Il thai infatti è una lingua tonale, per cui l'altezza dei suoni (il *pitch*) assume un valore distintivo sul piano semantico.

universalmente condivise (sebbene presentino degli aspetti in comune); inoltre, soprattutto nelle lingue tonali, i valori frequenziali sembrano avere funzioni linguistiche diverse rispetto a lingue non tonali (Zinken, Knoll, Panksepp, 2008). Ne consegue che "universality of the vocal expression of emotions per se can therefore not be completely maintained" (ivi, p. 189).

#### 2.7.2.2 *Le ricerche sul decoding*

Come accennato più sopra, la maggior parte delle ricerche condotte a livello interlinguistico ha inteso verificare l'universalità o la diversità culturale dell'espressione vocale delle emozioni attraverso un metodo indiretto, ovvero attraverso la decodifica uditiva delle emozioni (cfr. Albas, McCluskey, Albas, 1976; Scherer, Banse, Wallbott, 2001; Chen, Gussenhoven, Rietveld, 2004; Thompson, Balkwill 2006; Barrett, Bryant, 2008). In linea generale, sembrerebbe che gli esseri umani siano in grado di riconoscere gli stati emotivi prodotti in una lingua diversa dalla propria con una percentuale che supera la pura casualità (cfr. Elfenbein, Ambady 2002). Tuttavia, questo dato non restituisce una prova tangibile dell'universalità delle emozioni sul piano dell'espressione vocale, in quanto diversi fattori sembrano minimizzare o ridurre il peso della componente biologica all'interno della comunicazione emotiva. Soprattutto negli studi che hanno coinvolto culture molto diverse tra loro, l'indice di accuratezza si è dimostrato sensibile alla distanza culturale (cfr. Elfenbein, Ambady, 2002 Scherer, Banse, Wallbott 2001). Questo secondo dato, altrettanto significativo, ridimensiona le interpretazioni universaliste e solleva questioni tuttora ampiamente dibattute.

Ad esempio, la variabilità culturale nell'accuratezza del riconoscimento delle emozioni è stata attribuita a differenze linguistiche (*language-distance hypothesis*). I termini usati per tradurre le etichette emotive potrebbero veicolare significati diversi rispetto alla cultura in cui ha avuto origine l'esperimento (Elfenbein, Ambady, 2002, p. 204, cfr. anche Mesquita *et al.*, 1997).

Scherer, Banse e Wallbott (2001) hanno suggerito che alcuni elementi prosodici, in particolare i contorni intonativi e il ritmo, potrebbero ostacolare il riconoscimento delle emozioni a livello interlinguistico. Mesquita e Frijda (1992) hanno suggerito invece che le qualità vocali potrebbero "oscurare" la percezione degli altri indici acustici attivati per veicolare l'emozione, determinando un indice di accuratezza più basso nei giudizi espressi dai valutatori che non appartengono al sistema culturale del parlante.

Recentemente, Chen, Gussenhoven e Rietveld (2004) hanno esplorato la percezione degli andamenti intonativi paralinguistici legati al codice della frequenza e al codice dello sforzo (cfr. PAR. 2.2.2) nell'inglese britannico e in olandese. A questo proposito, sono state condotte due indagini percettive al fine di rilevare eventuali differenze tra le culture considerate nella percezione di 4 atteggiamenti (sicuro, amichevole, enfatico e sorpreso) legati all'attivazione dei due codici. I partecipanti, parlanti nativi inglesi e olandesi, sono stati invitati ad ascoltare alcuni stimoli<sup>35</sup> (manipolati mediante Praat) nella propria lingua e ad esprimere un giudizio in base a 4 scale semantiche derivanti dai due codici: *self-confident* vs. *not self-confident*; *friendly* vs. *not friendly* (codice della frequenza); *surprised* vs. *not surprised*; *emphatic* vs. *not emphatic* (codice dello sforzo). Nel primo esperimento gli stimoli differivano per: contorno intonativo, media di f0 (*pitch register*), e *pitch span* ("the difference between the highest F0 and the lowest F0", *ivi*, p.1); nel secondo, si distinguevano per: media di f0, altezza dei picchi (*peak height*), allineamento (*peak alignment*) e contorno finale. In base alle ipotesi formulate dagli autori, le due lingue, che utilizzano di base un *pitch range* differente (cfr. de Bot, 1982; de Pijper, 1983; Willems, 1982), avrebbero potuto in teoria ricorrere ad un'estensione tonale altrettanto diversa per veicolare significati paralinguistici; in questo caso, un incremento o un decremento dell'estensione tonale avrebbe influenzato la percezione degli andamenti intonativi e dunque l'interpretazione di questi codici. I risultati ottenuti hanno evidenziato alcune differenze interessanti tra le due lingue. Infatti, "languages may make different choices whenever conflicting meanings are derived from the biological codes" (*ivi*, p. 342). Questi significati contrastanti, secondo gli autori, si verificano quando i codici biologici si affidano ad un unico parametro, cioè f0. Ad esempio, "pitch register would appear to be used to signal friendliness in British English on the basis of the Frequency Code, but emphasis in Dutch on the basis of the Effort Code. When pitch register was raised, British English listeners perceived a lower degree of emphasis, as a result of perceiving a higher degree of friendliness, while Dutch listeners perceived a higher degree of emphasis" (*ibid.*). I risultati ottenuti hanno portato gli autori ad affermare che "contrary to the traditional view that the paralinguistic usage of intonation is similar across languages, [...] the present findings support a theory of paralinguistic meaning based on the universality of

---

<sup>35</sup> Gli stimoli consistevano in un enunciato standard tradotto in entrambe le lingue, realizzato da parlanti nativi in modalità neutra e successivamente manipolato con Praat.

biological codes, which however acknowledges a language-specific component in the implementation of these codes"<sup>36</sup> (ivi, p. 311).

Oltre a fattori linguistici, anche la condivisione dei medesimi codici sociali e di stili comunicativi simili gioca un ruolo importante nel riconoscimento delle emozioni. Come spiegano Elfenbien e Ambady (2002, p. 204): "It is possible that recognition accuracy is higher when emotions are both expressed and perceived by members of the same cultural group" (cfr. *ingroup-advantage hypothesis*<sup>37</sup>). A questo proposito, Elfenbien e Ambady (2003) hanno suggerito che il contatto e l'interazione tra le culture, nonché la condivisione di alcuni tratti culturali possano favorire la decodifica delle emozioni (*cultural proximity hypothesis*). In base a questa ipotesi, "members of cultures who share cultural elements such as degree of individualism or collectivism, power structure, and gender roles, should be more successful at decoding each other's emotional expressions than members of cultures that are less similar" (Thompson, Balkwill, 2006, p. 409).

Infatti, secondo gli autori, la variabilità degli indici di riconoscimento non è dovuta esclusivamente a fattori etnici, geografici o linguistici, ma anche all'assenza di un coinvolgimento reciproco tra le culture considerate. A supporto di ciò, alcune ricerche hanno dimostrato che l'indice di accuratezza nel riconoscimento delle emozioni a livello cross-linguistico risulta più alto nel momento in cui il contatto (anche telefonico) tra le culture prese in esame è maggiore (Elfenbein, Ambady, 2002). Allo stesso modo, i membri appartenenti a minoranze linguistiche all'interno di un dato paese sarebbero in grado di decodificare con maggiore accuratezza le espressioni vocali prodotte dai membri appartenenti alla cultura maggioritaria rispetto a questi ultimi nella situazione inversa (*ibid.*).

Alle ricerche condotte da Scherer e colleghi si deve indubbiamente il merito di aver condotto un'analisi sistematica a livello interlinguistico, prendendo in esame anche culture distanti tra loro. A tal proposito, uno studio condotto nel 2001 da Scherer, Banse, Wallbott ha coinvolto soggetti provenienti da nove paesi diversi (europei, statunitensi e asiatici), al fine di testare la loro abilità nel riconoscimento delle emozioni

---

<sup>36</sup> A sostegno della specificità linguistica, gli autori aggiungono che "the fact that British English listeners made a larger meaning distinction than Dutch listeners between the lowest and the highest pitch registers in the perception of friendliness is consistent with this explanation. The association of pitch register with emphasis precludes the use of pitch register for the expression of friendliness, or at least makes such use less effective" (*ibid.*).

<sup>37</sup> Tale espressione si riferisce all'abilità dei soggetti nel riconoscere con maggiore accuratezza le emozioni espresse da membri appartenenti al medesimo sistema culturale (cfr. Elfenbein, Ambady, 2002).

espresse in un'altra lingua (il tedesco). Il corpus di stimoli emotivi- veicolati attraverso una frase *non sense* e pronunciati da attori professionisti- comprendeva cinque emozioni: gioia, tristezza, paura, collera e disgusto (insieme alle produzioni neutre). I risultati hanno evidenziato un indice di accuratezza nell'identificazione delle emozioni pari al 50%. Tuttavia, la percentuale di decodifiche non si è rivelata omogenea in tutti i campioni considerati, in quanto il gruppo indonesiano ha totalizzato la percentuale più bassa, mentre il campione di uditori tedeschi ha registrato la percentuale più alta in assoluto (*in group-advantage*), seguito dal gruppo di inglesi. Gli autori hanno ipotizzato che la vicinanza linguistica abbia giocato un ruolo importante nel processo di decodifica (*language distance hypothesis*). Ciò ha reso conto della percentuale molto più alta ottenuta dagli inglesi e dagli olandesi rispetto ai parlanti di lingue romanze (italiani e francesi) e soprattutto rispetto agli indonesiani (non appartenenti al ceppo indoeuropeo).

Pur tenendo conto del fattore linguistico, gli autori hanno interpretato la percentuale globale di decodifiche corrette (50%) come una prova dell'esistenza di meccanismi universali nella comunicazione emotiva, in grado di veicolare lo stato emotivo del parlante attraverso gli indici acustici. Tuttavia, come rilevano Zinken, Knoll e Panksepp (2008, p. 187), lo studio in questione presenta una serie di limiti. In primo luogo, il campione di parlanti appare limitato e soprattutto poco rappresentativo in termini di diversità linguistica (su sette lingue considerate, sei erano di origine indoeuropea). Inoltre, il numero di emozioni ritenute primarie appare anch'esso riduttivo, e soprattutto, sostengono gli autori "given this restriction to a few basic emotions, the ambiguity of the results from recognition studies seems quite surprising. We would not expect such a result if the vocal expression of these emotions was itself grounded in universal mechanisms"(*ibid.*).

La riduzione dell'indice di accuratezza in funzione della distanza culturale è stata registrata anche in altri studi. Thompson e Balkwill (2006) hanno sottoposto il parlato emotivo prodotto da parlanti di diversa origine (inglese, tedesca, cinese, giapponese e tagalog) ad un campione di inglesi. Anche in questo caso, i soggetti sono stati in grado di decodificare con maggiore accuratezza le emozioni espresse nella propria lingua (l'inglese), mentre le percentuali di riconoscimento sono state molto inferiori per le altre lingue (soprattutto per il cinese e il giapponese). Sebbene il *decoding* in tutte e cinque le lingue abbia comunque ottenuto un indice di accuratezza superiore al caso (50%), gli autori hanno attribuito le differenze rilevate al peso

specifico esercitato dalla cultura all'interno della comunicazione emotiva, affermando che "our results support the view that emotional prosody is decoded by a combination of universal and culture-specific cues" (ivi, p. 407).

In ambito italiano<sup>38</sup>, uno studio condotto da Origlia, Galatà e Ludusan (2010), ha elaborato un database di produzioni emotive (€motion) in quattro lingue (italiano, francese, inglese e tedesco), allo scopo di verificare l'abilità dei parlanti nativi nel decodificare le emozioni in una lingua straniera. Inoltre, l'obiettivo dei ricercatori è stato quello di appurare se l'indice di accuratezza nel riconoscimento delle emozioni fosse sensibile alle differenze tra i sistemi linguistici considerati. Gli stati emotivi presi in esame (*happiness, anger, fear, sadness, disgust, surprise*, più l'eloquio neutro) sono stati veicolati da attori e da parlanti *naïf*, attraverso la lettura di scenari emotigeni che guidavano l'immedesimazione dei soggetti e contenevano una frase standard (tradotta nelle quattro lingue). Il corpus di produzioni emotive è stato successivamente validato attraverso la somministrazione di un test a risposta chiusa che includeva sette opzioni (le sei emozioni considerate e il neutro) ad un campione di ascoltatori. I parametri considerati per l'estrazione dei dati sono stati i seguenti: media, deviazione standard, valori massimi e minimi di  $f_0$  e intensità, nonché i valori relativi al grado di periodicità del segnale (Harmonics-to-Noise-Ratio) e quelli relativi a *shimmer* e *jitter*. Dai risultati dei test percettivi è emerso che le produzioni emotive dei soggetti *naïf* sono state riconosciute con maggiori difficoltà rispetto a quelle degli attori. Inoltre "from the test on the actors subset, we understand that emotions are cross-lingual because the accuracy level we achieved is similar to the one obtained on the monolingual test" (ivi, p. 4).

Risultati simili sono stati riportati anche da uno studio precedente di Galatà e Romito (2010) condotto sul medesimo corpus (€motion). La ricerca si è focalizzata sulla capacità degli italiani di decodificare le emozioni espresse nelle quattro lingue considerate, tenendo in considerazione anche il livello di competenza. A questo proposito, una sezione del test percettivo chiedeva ai partecipanti di inserire

---

<sup>38</sup> Tra i database elaborati per la lingua italiana, ricordiamo anche il corpus EMOVO (Iadarola 2009; Costantini *et al.* 2014), contenente produzioni emotive realizzate da attori professionisti (nello specifico, le sei emozioni primarie e l'eloquio neutro). Particolare rilievo va dato inoltre al contributo offerto dal Laboratorio per le ricerche sperimentali sul parlato all'interno del Centro di Studio per le Ricerche di Fonetica del CNR di Padova, che negli ultimi anni ha elaborato modelli di sintesi del parlato (cfr. Drioli *et al.*, 2003), focalizzandosi sull'implementazione di facce espressive parlanti, generate da programmi di sintesi bimodale uditivo-visiva del parlato, il cui obiettivo è quello di simulare "le reazioni causali tra movimenti articolatori e prodotto acustico, come metodo di indagine dei rapporti tra percezione uditiva dei segnali acustici del parlato ed estrazione di informazioni fonologica dai movimenti articolatori visibili e come banco di prova delle interazioni tra movimenti articolatori e configurazioni facciali espressive" (Magno Caldognetto, Cavicchio, Cosi, 2003, p. 199).

informazioni relative al proprio livello di competenza dell'inglese, del francese e del tedesco. I risultati hanno confermato una maggiore abilità nel riconoscere le produzioni emotive nella lingua nativa (*in-group advantage*) piuttosto che in lingue diverse dalla propria (ivi, p. 211). Le emozioni maggiormente riconosciute sono state la tristezza e la rabbia, seguite dalla paura, confermando i risultati delle precedenti ricerche (cfr. Scherer *et al.* 2003). Anche in questo caso, le percentuali di riconoscimento ottenute dai soggetti *naïf* sono state di gran lunga inferiori rispetto a quelle degli attori professionisti<sup>39</sup>. Al di là di queste differenze, l'indice di accuratezza registrato supera quello della pura casualità. I risultati hanno inoltre evidenziato "una certa sistematicità nella confusione di alcune emozioni" (ivi, p. 216), ad esempio, per il francese, l'emozione di rabbia è stata confusa in modo sistematico con gioia e sorpresa, sia nelle produzioni dell'attore, sia in quelle del soggetto *naïf*. In relazione al grado di conoscenza delle lingue straniere, dai risultati è emerso che "a livelli di competenza maggiore nelle tre lingue non corrisponde un numero maggiore di risposte corrette. Si può perciò affermare che tra numero di risposte corrette complessivamente fornite e livello di competenza nelle lingue straniere presentate vi sia assoluta indipendenza" (ivi, p. 224).

Laukka *et al.* (2013) hanno invece indagato il ruolo degli *affect bursts* nel favorire o ostacolare il riconoscimento delle emozioni a livello interlinguistico. I partecipanti, attori professionisti di diversa provenienza (India, Kenya, Singapore e Stati Uniti), sono stati invitati a veicolare nove emozioni positive e nove emozioni negative ricorrendo soltanto alle vocalizzazioni. Le produzioni emotive sono state sottoposte ad un campione di ascoltatori di origine svedese, attraverso un test a risposta chiusa. I risultati hanno confermato la capacità dei giudici nel decodificare le emozioni negative e positive con un indice di accuratezza che ha superato la soglia della casualità. Nel dettaglio, tra le emozioni positive, il sollievo (*relief*) è stato decodificato con maggiore accuratezza (seguito da: *lust*, *interest*, *serenity*, *positive surprise*), mentre affetto (*affect*) e orgoglio (*pride*) hanno ottenuto l'indice di riconoscimento più basso. Per quanto concerne le emozioni negative, le percentuali più alte si sono registrate per la collera, il disgusto, la paura e la tristezza, mentre colpa (*guilt*) e vergogna (*shame*) hanno ottenuto percentuali di gran lunga inferiori. Le conclusioni a cui sono giunti gli autori suggeriscono che "the voice can reveal both basic emotions and several positive

---

<sup>39</sup> "Si assiste, di fatto, ad una oscillazione del 7,5% ca. tra stimoli prodotti da soggetti attori e stimoli prodotti da soggetti *naïf* per la lingua italiana, oscillazioni ancor più evidenti ed apprezzabili nel caso di lingue diverse da quella degli ascoltatori" (Galatà, Romito, 2010, p. 214).

emotions other than happiness across cultures, but self-conscious emotions such as guilt, pride, and shame seem not to be well recognized from non-linguistic vocalizations" (ivi, p. 353).

## 2.8 Conclusioni

Le questioni discusse in questo capitolo rappresentano tuttora l'aspetto più stimolante della ricerca sul parlato emotivo. Il quadro teorico e lo scenario metodologico si caratterizzano per un'estrema eterogeneità, che in un certo senso, rende conto della complessità del fenomeno emotivo, la cui osservazione richiede l'adozione di molteplici punti di vista, da quello biologico a quello cognitivo. Pur ammettendo che universalità e specificità culturale possano coesistere all'interno dei modelli proposti, rimane indubbia la generale complessità di un sistema comunicativo- come quello umano- in cui si intersecano processi biologici e aspetti linguistici e culturali in grado di influenzarsi mutuamente. Alla luce dei risultati empirici riportati dagli studi interculturali, è ragionevole supporre che il linguaggio e la cultura possano agire come un potente filtro, modulando l'espressione vocale delle emozioni in funzione del contesto comunicativo in cui si realizza. Si suppone inoltre che la comunicazione emotiva- nel momento in cui è intenzionale e mediata dal linguaggio- sia regolata anche da processi cognitivi legati alla valutazione dello stimolo in accordo con le norme sociali e l'immagine di sé.

La variabilità nel riconoscimento interculturale delle emozioni è stata attribuita, di volta in volta, a fattori linguistici e culturali (o a condizionamenti dovuti al *setting* stesso degli esperimenti condotti). Inoltre, le ultime ricerche sugli andamenti intonativi paralinguistici (cfr. Chen, Gussenhove, Rietveld, 2004) hanno suggerito che l'interpretazione di questi codici biologici sia influenzata da fattori linguo-specifici, che rendono conto forse della variabilità degli indici di accuratezza a livello interculturale.

Il panorama finora delineato costituisce un punto di partenza per la comprensione delle dinamiche che regolano la comunicazione emotiva in un contesto non nativo, o meglio in un contesto di acquisizione di una lingua seconda. Oltre ad aspetti prettamente linguistici, le difficoltà sembrano rappresentate innanzitutto dall'acquisizione delle strutture intonative della lingua target e delle funzioni che esse svolgono sul piano pragmatico; mentre sul piano comunicativo, la distanza tra i sistemi culturali può innescare una serie di incomprensioni legate ad una diversa gestione della

dimensione emotiva nelle culture considerate. Questi aspetti verranno trattati più esaurientemente nel capitolo 3.

# Capitolo 3

---

## La comunicazione delle emozioni in una lingua non nativa

### 3.1 introduzione

Alla luce di quanto illustrato nei primi due capitoli, è ragionevole supporre che la percezione e la manifestazione di uno stato emotivo dipendano in buona misura non solo da meccanismi universali, ma anche da processi cognitivi individuali che variano in base al contesto linguistico, culturale e sociale. In quest'ottica, la comunicazione emotiva in una lingua non nativa, e in generale tra membri di culture distanti, che ricorrono cioè a modalità espressive e stili comunicativi diversi, può risultare quantomeno ambigua e irta di ostacoli.

In questo capitolo, saranno introdotti e discussi gli aspetti più problematici legati alla dimensione emotiva in una lingua seconda<sup>40</sup> (L2) partendo dal concetto di competenza emotiva proposto da Saarni (cfr. PAR. 3.2) e concludendo con le osservazioni e i risultati empirici delle indagini che hanno esplorato la relazione tra emozioni e lingua seconda (PAR. 3.2.3). Successivamente, ampio spazio sarà dato alla ricognizione degli studi sul parlato emotivo in L2 e agli aspetti più controversi evidenziati dalle indagini in questione.

### 3.2 La dimensione emotiva nel processo di acquisizione di una L2

#### 3.2.1 *La competenza comunicativa*

Comunicare in una lingua non nativa implica in primo luogo una competenza linguistica. Essa si fonda sulla capacità di dominare una lingua sia sul piano ricettivo sia

---

<sup>40</sup> Per lingua seconda o L2 si intende una lingua appresa successivamente a quella materna (L1) all'interno del contesto culturale in cui viene utilizzata. Per lingua straniera (LS) si intende invece la lingua imparata in un contesto scolastico (come ad esempio l'inglese nelle scuole italiane), avulsa quindi dal contesto culturale in cui si parla.

su quello produttivo. E' un tipo di conoscenza che si estende ai vari livelli della lingua (fonologico, lessicale, morfologico, ecc.), ma che da sola non è in grado di accedere a quei meccanismi interni, propri di ogni scambio comunicativo, che si fondano sull'uso concreto della lingua in funzione del suo contesto. In altre parole, quando comunichiamo ricorriamo non solo ad un'ampia gamma di conoscenze legate alla lingua stessa, ma anche ad una serie di conoscenze più o meno implicite che ci consentono di sapere come rivolgerci al nostro interlocutore, quale stile (formale/informale) adottare, il tipo di richiesta che è possibile avanzare o meno, ecc. In un certo senso, la competenza linguistica, da sola, non è sufficiente a spiegare le molteplici sfumature di significato che uno scambio comunicativo assume in base al contesto in cui si svolge.

La competenza linguistica, infatti, "circoscrive la descrizione a un parlante-ascoltatore 'ideale'; che conosce 'perfettamente' la propria lingua; che non è sottoposto ai 'condizionamenti' di ordine psicologico e sociologico nell'applicazione delle proprie conoscenze ai fini dell'esecuzione; che appartiene, infine, a una comunità linguistica 'omogenea'"(Zuanelli Sonino, 1981, p. 40). La comprensione e la gestione di fattori extra-linguistici (il ruolo dei parlanti, il contesto sociale e culturale, le varietà diatopiche e diastratiche della lingua, ecc.) rientrano infatti nella definizione stessa di competenza comunicativa, che implica, in un certo senso, l'interazione di altre competenze. Come spiega Zorzi (1996, p. 47):

Per formare una competenza comunicativa, concorrono, quindi, altre componenti: la conoscenza psicolinguistica (sapere se o in che misura qualcosa è fattibile in virtù dei mezzi di implementazione di cui si dispone, cioè in virtù della capacità dei parlanti di trasformare una realtà mentale (il significato) in una realtà sociale ai fini della comprensione (Zuanelli Sonino, 1981: 68)); la conoscenza socioculturale (se e in che misura qualcosa è appropriato in relazione al contesto in cui è usato e valutato); la conoscenza de facto, (sapere se e in che misura qualcosa è di fatto realizzato dalla comunità parlante quella lingua, e non soltanto possibile). La competenza comunicativa non solo richiede che il parlante abbia queste conoscenze, ma anche che sviluppi l'abilità d'usarle (Hymes, 1971).

Il modello di competenza comunicativa proposto nel Quadro Comune di Riferimento Europeo per le Lingue (pubblicato dal Consiglio d'Europa nel 2001) si fonda su approccio orientato verso l'uso linguistico e l'azione, per cui chi apprende e dunque usa una lingua, viene considerato un agente sociale, e in quanto tale deve portare a termine alcuni compiti, in determinate circostanze, in un contesto specifico, attraverso una serie di strategie.

Nel QCER ritroviamo due tipi di competenza: la competenza linguistico-comunicativa e le competenze generali. La competenza comunicativa consta di tre macro-componenti: linguistica, sociolinguistica e pragmatica<sup>41</sup>. Le competenze linguistiche "includono conoscenze ed abilità riferite al lessico, alla fonologia e alla sintassi, e ad altre dimensioni del linguaggio visto come sistema, indipendentemente dalla valenza sociolinguistica delle sue variabili e dalle funzioni pragmatiche delle sue realizzazioni" (Consiglio d'Europa, 2002, p.16). Le competenze sociolinguistiche "si riferiscono a fattori socioculturali dell'uso linguistico. Poiché è sensibile alle convenzioni sociali (regole di cortesia, norme che definiscono i rapporti tra le generazioni, i sessi, le classi e i gruppi sociali, codificazione linguistica di determinati rituali basilari per il funzionamento di una comunità) la componente sociolinguistica condiziona fortemente ogni comunicazione linguistica tra rappresentanti di culture diverse, benché spesso i protagonisti possano non esserne consapevoli"(ivi p. 17). Infine, la competenza pragmatica riguarda "l'uso formale delle risorse linguistiche [...] la padronanza del discorso, coesione e coerenza, l'identificazione di tipi e forme testuali, ironia e parodia"(*ibid.*).

L'insieme di queste competenze è fortemente connesso con altri tipi di competenze (che il Consiglio d'Europa inserisce all'interno delle "competenze generali") che un individuo possiede in virtù delle proprie esperienze precedenti. Si tratta di competenze strettamente legate alla natura stessa dell'essere umano: la conoscenza del mondo, la conoscenza socioculturale, la consapevolezza interculturale, ecc. "Gli individui utilizzano le proprie competenze in contesti e in condizioni differenti e con vincoli diversi per realizzare delle attività linguistiche [...] con l'attivazione delle

---

<sup>41</sup> Nel modello comunicativo proposto invece da Balboni (1999, p. 31) la competenza comunicativa assume la forma di una piramide a tre lati "ciascuno dei quali indica un sapere o un saper fare". La distinzione che opera è tra saper fare lingua (inteso come abilità linguistiche diversificate: leggere, scrivere, fare un monologo, partecipare a un dialogo); sapere fare con la lingua, ovvero «la capacità funzionale, che permette di utilizzare la lingua come strumento di azione in un determinato contesto», per raggiungere dunque determinati scopi. Questa faccia della piramide include diverse dimensioni: da quella sociale (ovvero la capacità, ad esempio, di identificare e rispettare i rapporti di ruolo, gestire situazioni in cui si richiede la conoscenza implicita dello status sociale dei partecipanti all'evento comunicativo, ecc.) a quella pragmatica (comunicare in modo efficace equivale a raggiungere i propri scopi attraverso degli atti comunicativi che sono intenzionali, ovvero con un obiettivo preciso). Il modello proposto da Balboni ingloba inoltre una serie di altre abilità e conoscenze che si estendono anche a canali non verbali e che rientrano in quella che definisce competenza extra-linguistica. Ad esempio, la competenza cinesica, "cioè l'uso comunicativo del corpo, delle sue posture e dei suoi movimenti", quella prossemica ovvero "l'uso comunicativo delle distanze interpersonali", vestemica e oggettistica" che consentono di utilizzare per la comunicazione l'abbigliamento e altri oggetti di vario tipo e natura" (ivi, p. 34). In quest'ottica, il concetto di competenza comunicativa si estende alla lingua inserita nel suo contesto d'uso, e tiene soprattutto conto dei soggetti che la utilizzano, attraverso una forte interazione tra sistemi linguistici e non.

strategie che sembrano più adatte a portare a buon fine i compiti previsti"(Consiglio d'Europa, 2002: 12).

In conclusione, ciò che emerge da queste definizioni è una visione della lingua utilizzata nel suo contesto, e che dunque ingloba sia aspetti puramente linguistici sia fattori extra-linguistici (socioculturali, pragmatici, non verbali, ecc.) che descrivono la comunicazione umana nel suo insieme.

### 3.2.2 *Intelligenza emotiva e competenza emotiva in L2*

Negli ultimi anni, l'ampia riflessione sulle competenze e sulle caratteristiche dell'apprendente ha riguardato anche altri tipi di conoscenze, che in un certo senso interessano la dimensione emotiva della comunicazione in L2. La componente motivazionale dei fenomeni emotivi lascia presupporre che le emozioni giochino un ruolo importante anche all'interno di un contesto di apprendimento, in cui ansia, paura e imbarazzo possono prendere il sopravvento e rallentare il percorso individuale di un apprendente. Ricordiamo l'ipotesi del filtro affettivo formulata da Krashen nel 1982, che tuttavia, prende in considerazione soltanto il peso dell'ansia sul processo di acquisizione<sup>42</sup>, escludendo altri tipi di emozione, in grado invece di motivare l'apprendente. La prospettiva adottata negli ultimi anni si propone al contrario di indagare quali abilità riconducibili all'intelligenza emotiva possano costituire un forte stimolo per l'apprendimento, fornendo una spinta motivazionale verso il superamento di ostacoli progressivamente più impegnativi e il raggiungimento di determinati obiettivi. Salovey e Mayer (1990, p. 185) hanno definito l'intelligenza emotiva come "the subset of social intelligence that involves the ability to monitor one's own and other's feelings, to discriminate among them and to use this information to guide one's thinking and actions". Goleman (1995) ha descritto queste abilità in base a cinque dimensioni che concernono l'autoconsapevolezza e l'autoregolazione, la motivazione, l'empatia e la socialità.

Da qui la necessità di misurare, in qualche modo, il grado di intelligenza emotiva individuale, e di studiarne gli effetti sull'apprendimento. Il questionario autovalutativo

---

<sup>42</sup> Secondo Krashen (1982) apprendimento e acquisizione sarebbero due processi distinti. Il termine acquisizione viene utilizzato in genere in riferimento alla lingua materna, tuttavia anche l'appropriazione delle strutture della L2 (o lingua target) implica un processo di acquisizione naturale simile a quello della L1 (ma che si caratterizza per alcune differenze, cfr. più avanti), per cui si parla di apprendimento spontaneo (*Naturalistic Second Language Acquisition*). Il termine "acquisizione" viene qui utilizzato in questa accezione, per indicare cioè il processo di apprendimento di una lingua seconda in un contesto spontaneo.

elaborato da Schutte, insieme ad altri ricercatori (1998) si propone di fornire una scala dell'intelligenza emotiva, attraverso la valutazione di una serie di abilità, tra le quali la capacità di riconoscere ed esprimere le emozioni, di monitorare e gestire le proprie emozioni e quelle degli altri. Molti studi hanno adottato questa scala per misurare le abilità degli apprendenti in relazione alla sfera emotiva, suggerendo una possibile correlazione tra una maggiore competenza emotiva e il grado di successo nell'apprendimento di una lingua seconda.

Come ricorda Oz (2015) le definizioni di intelligenza emotiva elaborate dai diversi autori tendono ad essere complementari e compatibili. Per questa ragione, mi soffermerò sulla definizione di competenza emotiva elaborata da Saarni (1999), in quanto essa si presta meglio alla descrizione delle diverse abilità che compongono un tipo di conoscenza che potenzialmente può essere sviluppata e migliorata, tanto in un contesto nativo quanto in un contesto di L2.

Il concetto di competenza emotiva proposta da Saarni (1999) è il risultato dell'interazione di una serie di abilità in grado di promuovere l'auto-efficacia (ovvero la convinzione e la fiducia nelle proprie capacità rispetto al raggiungimento di un risultato) all'interno delle interazioni sociali di natura emotiva. Il risultato desiderato riflette i valori e le credenze del sistema culturale di riferimento, sebbene ciascun individuo faccia propri questi valori, trasformandoli in significati e credenze personali (*ibid.*).

La prima di queste abilità, ovvero la consapevolezza degli stati emotivi interni, si traduce nella capacità di comprendere le emozioni esperite e le cause scatenanti, tenendo conto anche delle dinamiche inconsce che spesso portano un individuo a non essere consapevole dei propri sentimenti. Le altre abilità strettamente connesse alla sfera intra-personale riguardano la capacità di affrontare emozioni negative o angoscienti, ricorrendo a strategie di autoregolazione per ridurre l'intensità e la durata (rientra in questi comportamenti di regolazione anche l'inibizione dell'emozione stessa, soprattutto in contesti in cui la sua manifestazione non sarebbe socialmente accettata); la consapevolezza che uno stato emotivo interno non deve necessariamente avere una sua corrispondenza nell'espressione esteriore (sia in se stessi sia negli altri); la capacità di ricorrere ad un vocabolario di termini emotivi o di espressioni comunemente in uso presso il sistema culturale di riferimento, tenendo in considerazione le gerarchie e i ruoli sociali. Possiamo estendere questa consapevolezza all'impiego di risorse non verbali e paraverbali codificati culturalmente, nel rispetto del contesto e dell'interlocutore.

Le altre abilità che investono la sfera interpersonale si traducono nella capacità di interpretare le emozioni altrui, sulla base di indizi situazionali ed espressivi (verbali e non verbali); di partecipare empaticamente alle esperienze emotive degli altri, tenendo in considerazione l'influenza che può avere il proprio comportamento emotivo ed espressivo su quello altrui e proiettando questa consapevolezza nell'elaborazione di strategie comunicative ed espressive per il raggiungimento di uno scopo<sup>43</sup>.

Le ricerche sull'intelligenza emotiva in L2 hanno suggerito che quest'ultima giochi un ruolo importante nel processo di apprendimento, favorendone lo sviluppo o rallentandolo. Ad esempio, Alavinia, Bonyadi, e Razavi (2012) hanno evidenziato la profonda correlazione tra motivazione e intelligenza emotiva. Anche Pishghadam (2009) ha dimostrato che i fattori strettamente legati alla sfera emozionale (competenze interpersonali e abilità nel gestire situazioni di stress) incidono sul processo di apprendimento, così come la propensione a comunicare nella lingua seconda (*Willingness to Communicate* o WTC<sup>44</sup>) è stata associata ad un alto grado di intelligenza emotiva (Oz, 2015).

Domakani *et al.* (2014) in uno studio condotto con apprendenti iraniani di inglese hanno evidenziato una correlazione tra il grado di intelligenza emotiva e le competenze pragmatiche, in particolare nelle donne. Come spiegano gli autori (ivi, p. 1): " Results [...] indicated that the females with more EI and intrapersonal skills performed better on the pragmatics test than the males. Findings suggest that pragmatic development and performance are intricately linked to learner affect, subjectivity, and emotions, and that this interplay, in turn, is not immune to gender influence".

In conclusione, è ragionevole ipotizzare che il potenziamento di queste abilità interpersonali possa contribuire alla crescita dell'apprendente nel suo ruolo di agente sociale. Un alto grado di intelligenza/competenza emotiva si traduce in una maggiore

---

<sup>43</sup> Essere emotivamente competenti (tanto in una L1 quanto in una L2) significa inoltre comprendere che la struttura o la natura delle relazioni è definita sia dal grado di immediatezza emotiva (o di genuinità della manifestazione espressiva) sia dal grado di reciprocità o simmetria all'interno della relazione (ivi, p. 7). Ad esempio, come spiega Saarni, una relazione matura si fonda sulla condivisione reciproca di emozioni autentiche, mentre è possibile che una relazione genitore-figlio si basi su una condivisione asimmetrica di emozioni genuine.

<sup>44</sup> Come spiega Oz (2015, p. 425) "Willingness to communicate is defined as an individual's "readiness to enter into discourse at a particular time with a specific person or persons, using a L2" (MacIntyre et al., 1998, p. 547). Dörnyei (2005) maintains that "WTC is a composite ID variable that draws together a host of learner variables that have been well established as influences on second language acquisition and use, resulting in a construct in which psychological and linguistic factors are integrated in an organic manner" (p. 210). Thus, WTC is a multi-faceted construct that integrates affective, social-psychological, linguistic, and communicative variables and is potentially capable of describing, explaining and predicting language learners' communicative behavior in an L2".

flessibilità nella gestione delle emozioni e infonde una maggiore fiducia nelle proprie capacità (Oz, 2015).

### *3.2.3 La comunicazione emotiva in una lingua non nativa: alcune problematiche*

Sebbene l'interesse verso la dimensione emotiva in L2 sia in graduale aumento, lo studio di queste problematiche è appena agli esordi. Dai risultati empirici riportati da alcune indagini qualitative e quantitative, è emerso che la comunicazione emotiva in L2 rappresenta un caso speciale di interazione, in cui molte variabili possono determinare insuccessi e fallimenti comunicativi, in grado di demotivare l'apprendente.

In primo luogo, l'interpretazione di contesti, situazioni e messaggi potenzialmente ambigui o semplicemente lontani dalla lingua e dalla cultura d'origine, può risultare difficoltoso per chi si affaccia in una realtà culturale nuova. Da un lato, c'è la tendenza generale a interpretare le situazioni comunicative in una lingua altra in base ai propri parametri culturali, molto spesso andando incontro a spiacevoli fraintendimenti. Dall'altro, la tendenza a interpretare le reazioni (o le mancate reazioni) altrui sulla base dei propri standard culturali è spesso adottata anche dai parlanti nativi nei confronti dei non nativi; il risultato, in entrambi i casi, è un doppio fallimento. Alla luce di queste osservazioni, è possibile rintracciare alcuni aspetti più "sensibili" ai processi di acquisizione di una lingua seconda e alle dinamiche interazionali che si instaurano con i parlanti nativi. I risultati empirici di alcune ricerche hanno posto l'accento su una serie di fattori che nel complesso restituisce un quadro indicativo sulla relazione tra dimensione emotiva ed L2 (cfr. Pavlenko, Dewaele 2004; Dewaele 2005; 2010; 2011; Pavlenko 2006).

In linea generale, sembrerebbe che l'acquisizione di una lingua seconda sia in grado di incidere profondamente sulla personalità di un individuo (cfr. Gonzalez-Regiosa, 1976; Bond; Lai 1986; Amati-Mehler, Argentieri, Canestri; 1993; Pavlenko 2002; Dewaele 2004). Come scrive Dewaele (2011, p. 29): "The acquisition of a new language can potentially be a mind-changing experience". Diverse ricerche (cfr. Pavlenko 2006; Wilson 2013) hanno tentato di approfondire, attraverso interviste e questionari, il rapporto con le altre lingue vissuto da parlanti plurilingue. Pavlenko (2006) ha investigato il rapporto tra identità e multilinguismo attraverso un questionario a risposta aperta sottoposto a 1039 soggetti plurilingue. La maggior parte dei partecipanti (65%) ha dichiarato di percepire un cambiamento nella propria personalità in relazione alla lingua parlata.

In uno studio condotto su apprendenti di inglese L2, Wilson (2013) ha riportato una serie di osservazioni interessanti. Alcuni studenti hanno affermato di sentirsi più sicuri ed estroversi parlando la L2, altri ancora hanno riscoperto dei mutamenti sul piano non verbale e un diverso uso del linguaggio del corpo. Secondo questo studio, variabili quali l'età e il genere non incidono sulla sensazione di "sentirsi diversi" quando si parla un'altra lingua. Secondo altri, quando si apprende una L2 dopo il periodo della pubertà, è possibile che si crei un legame differente tra le due lingue; la L1 rimane intimamente connessa con il coinvolgimento personale e la L2 diventa la lingua della distanza e del distacco (Dewaele, 2004; Harris, Gleason, Aycicegi, 2006). Come scrive Dewaele (2005, p. 374) "The L1 is preferred to express emotional involvement whereas the L2 is experienced as colder, more distant, and more detached from the L2 user and less appropriate for the expression of emotions".

In generale, risulta evidente che apprendere un'altra lingua inneschi un meccanismo molto delicato nella mente di un individuo, in grado di dare gradualmente forma a molteplici identità in relazione alle lingue parlate (cfr. Wilson, 2013).

Dewaele e Pavlenko (2002) hanno riportato la tendenza dei parlanti di una lingua straniera a ricorrere ad un numero limitato di termini emotivi rispetto ai parlanti nativi. La scarsa conoscenza del lessico emotivo è stata attribuita a diversi fattori, tra cui il livello di competenza, le differenze di genere e il grado di estroversione.

Altri autori hanno posto l'accento sull'influenza del contesto socioculturale della lingua target, suggerendo che gli apprendenti di una lingua straniera o seconda abbiano un accesso limitato alle competenze socioculturali, in altre parole: "the learner's lack of cultural competence, the ability to recognize, internalize, and practice the speech community's verbal and nonverbal ways of behaviors (Dewaele, Pavlenko, 2002), may constrain ways in which one navigates his or her new meaning-making systems" (Kim, Dorner, 2013, pp. 46-47).

Uno studio recente condotto da Kim e Dorner (2013) ha investigato il peso dei fattori linguistici e culturali nella comunicazione emotiva in L2. I dati sono stati raccolti mediante interviste e questionari indirizzati ad apprendenti di inglese di origine coreana. Le domande riguardavano sia questioni specificamente linguistiche (ad esempio, l'esistenza di termini della L1 difficilmente traducibili in L2) sia problematiche legate alle interazioni con i nativi. I risultati ottenuti hanno evidenziato il disagio provato da alcuni soggetti per una serie di fattori, tra i quali l'assenza di un lessico emotivo (in inglese) idoneo alle loro esigenze comunicative. Infatti, gli apprendenti hanno

dichiarato, in molti casi, che alcuni concetti legati alla cultura coreana non trovavano nessuna corrispondenza nella lingua inglese. Inoltre, il ricorso alla L2 per esprimere stati d'animo o per parlare di questioni intime e personali risultava più problematico e limitava la loro spontaneità; in alcuni casi, preferivano evitare di trattare argomenti che riguardavano la sfera emotiva nelle interazioni con i nativi, in quanto ciò avrebbe presupposto una serie di spiegazioni non sempre del tutto esaustive. Anche in termini di espressività, non soltanto sul piano linguistico ma anche su quello vocale non verbale, alcuni soggetti hanno dichiarato di essere meno comunicativi e di risultare più distaccati e freddi nelle interazioni con i nativi. I risultati di questa indagine hanno portato le autrici ad affermare che "a full communication of emotions across cultures and languages was challenging because of the lack of shared cultural contexts among speakers"(2013: 43).

In conclusione, la dimensione emotiva in L2 rappresenta una realtà caleidoscopica e mutevole che assume tonalità e gradazioni differenti in base al legame che si instaura con la lingua seconda e il suo tessuto culturale. Diversi autori hanno sottolineato la necessità di integrare questi aspetti nelle pratiche didattiche e in generale negli studi sull'acquisizione di una lingua seconda (MacIntyre, 2002; Dörnyei, Skehan 2003, Dewaele 2005).

Comunicare efficacemente in una L2 implica inoltre una competenza prosodica, e dunque il ricorso alle risorse prosodiche del parlato per veicolare significati rilevanti sul piano pragmatico. In quest'ottica, risulta necessario introdurre brevemente le problematiche legate all'acquisizione della prosodia in L2 e le osservazioni riportate da alcune ricerche.

### **3.3. L'acquisizione delle strutture prosodiche in una L2**

#### *3.3.1 La prosodia dell'interlingua<sup>45</sup>*

Come accennato nel CAP. 2, l'acquisizione delle strutture ritmiche e melodiche della lingua materna si verifica già nelle prime fasi di sviluppo, e secondo molti autori la sensibilità degli infanti verso queste strutture è rintracciabile sin dai primi giorni di

---

<sup>45</sup> Per interlingua (termine coniato da Selinker nel 1972) si intende il sistema linguistico provvisorio posseduto dall'apprendente, da intendere non come una lingua a metà tra L1 ed L2 quanto piuttosto come un processo che consente all'apprendente di formulare delle ipotesi sulla lingua d'arrivo e di ricostruirne gradualmente il sistema linguistico, attraverso una sorta di sillabo mentale.

vita (Speer, Ito, 2008). Alcune ricerche hanno dimostrato che i neonati (fino a tre anni di età) sono in grado di discriminare due lingue sulla base della diversa componente prosodica (cfr. Mehler *et al.* 1988; Jusczyk *et al.* 1993), così come è stata riportata la capacità degli infanti di sei mesi di ricorrere ad alcuni aspetti prosodici della lingua materna per determinare la posizione delle parole nel flusso del parlato (Jusczyk *et al.* 1993; Morgan, Saffran 1995; Morgan 1996; Johnson, Jusczyk 2001). Come ricordano Speer e Ito (2008, p. 92):

Consistent with these early abilities, young children are demonstrably sensitive to correspondences among the acoustic aspects of prosodic phrasing when they listen to sentences, and they tend to pronounce their own utterances with appropriate affective and phrasal prosody. Young infants have been repeatedly shown to perceive prosodic phrasing, and to prefer speech that conforms to the phrasing pattern of their native language.

Inoltre, studi recenti suggeriscono che intorno ai 14-18 mesi gli infanti siano già in grado di inferire le intenzioni degli adulti basandosi sulle variazioni prosodiche (Sakkalou, Gattis, 2012). Lo sviluppo delle capacità percettive legate alla prosodia della lingua materna sembra inoltre proseguire fino all'età adulta, affinandosi fino ai venti anni di età.

Diverso è il caso della lingua seconda, in cui l'acquisizione delle strutture prosodiche della L2 appare in buona misura ostacolata dal profondo legame con la prosodia della L1 e dal ruolo marginale che viene solitamente assegnato ai fenomeni soprasedimentali nella comunicazione (e in generale nelle pratiche didattiche). Gli studi in ambito fonologico suggeriscono l'ipotesi di un periodo critico oltre il quale l'acquisizione della fonologia della lingua target sarebbe ostacolata dalla conclusione del processo di lateralizzazione celebrale (Lenneberg, 1967). Non c'è accordo sulla conclusione di questo periodo, tuttavia, per quel che concerne il livello fonologico, alcuni studi hanno evidenziato una forte correlazione tra l'età di esposizione alla L2 e il livello di competenza fonologica degli apprendenti; in questo senso, una maggiore permanenza nel contesto linguistico e culturale della L2 non sembra migliorare le prestazioni degli apprendenti, mentre il fattore dell'età in cui avviene l'esposizione alla L2 appare discriminante (Oyama, 1976; Johnson, Newport, 1989). Ciò non esclude che variabili individuali (come la motivazione, le attitudini fonologiche) e sociali (come ad esempio, il prestigio linguistico della lingua target o la vicinanza tipologica tra L1 e L2) possano favorire lo sviluppo di una competenza fonologica pari a quella di un nativo, come hanno dimostrato studi recenti (cfr. Bongaerts *et al.* 1997; Palmen, *et al.* 1997).

In relazione all'acquisizione delle strutture prosodiche, il quadro appare ancora oscuro, sebbene molti studi abbiano evidenziato la complessità che caratterizza questo processo in un contesto di L2 (Anderson-Hsieh, Johnson, Koehler 1992; Wennerstrom, 1998). Come spiega Marotta (2008, pp. 7-8): "è assodato che un parlante di L2, specialmente a livelli di competenza non elevati, presta più attenzione ai livelli lessicale, *in primis*, e morfosintattico *in secundis*, il che sembra implicare che, tranne alcune eccezioni, il percorso di apprendimento collochi gli aspetti prosodici in posizione relativamente tardiva".

Per quanto concerne gli studi sulla produzione prosodica da parte di apprendenti di varie L2 e con differenti L1 è stato osservato che gli apprendenti hanno difficoltà con la prosodia della lingua seconda e dunque con l'intonazione, l'accento, il ritmo, ecc. (Chiang, 1979; Grosser, 1993; Heiderscheidt, Hiligsmann, 2000, Guion, 2005; Caspers, Van Santen 2006).

Per quel che riguarda l'inglese, la lingua più studiata, è stato osservato che soprattutto nelle prime fasi del processo di acquisizione (cfr. Grosser 1993; Wieden 1993), gli apprendenti hanno difficoltà nel distinguere le informazioni nuove rispetto a quelle già note, per tale ragione tenderebbero a enfatizzare ogni parola all'interno della frase (Rasier, Hiligsmann, 2007 p. 48). Nelle fasi successive, si assiste a una sorta di ristrutturazione dell'interlingua, durante la quale gli apprendenti si avvicinano gradualmente al modello nativo (in questa fase, sono frequenti i casi di iper-generalizzazione). Gli studi prosodici in ambito italiano, con riferimento all'ambito dell'insegnamento/apprendimento dell'italiano come lingua seconda, sono in graduale aumento negli ultimi anni, come dimostrano le varie ricerche condotte sulla prosodia dell'interlingua (Gamal, 2006; De Meo, Pettorino, 2012; Mocchiari, 2014; Pellegrino *et al.* 2014), con particolare riferimento agli aspetti pragmatici (cfr. De Marco, Sorianello, Mascherpa, 2014; Chini, 2015).

L'ipotesi dell'interferenza della lingua materna nell'acquisizione delle strutture soprasegmentali della L2 è ancora oggetto di discussione. In linea generale, mentre alcuni studi confermano un effettivo influsso della L1 a livello segmentale e ritmico (Archibald, 2000), l'accentazione prosodica e la melodia sembrano seguire *pattern* diversi da quelli delle L1, mettendo in discussione l'ipotesi del *transfer* prosodico-fonologico (Gamal, 2006). Studi recenti sulle produzioni di bambini bilingui turco-tedesco (cfr. Queen, 2001) hanno evidenziato che, a parte alcuni casi di interferenza, i profili intonativi presentavano dei tratti ibridi, non riconducibili né alla L1 né alla L2,

ma derivanti dalla fusione di tratti prosodici delle due lingue (Marotta, 2008). Chen e Fon (2011) hanno esplorato invece le differenze tra il parlato spontaneo e letto prodotto da apprendenti di origine taiwanese in inglese, confrontandolo con le produzioni di un gruppo di nativi inglesi. Le analisi acustiche hanno evidenziato che "only native speakers merge word boundaries; non-native speakers were less fluent, and chunked their speech into more IPs" (ivi, p. 459).

L'analisi contrastiva di L1 ed L2 rimane dunque un punto di partenza fondamentale per poter prevedere eventuali fenomeni di transfer. Tuttavia, secondo Rasier e Hiligsmann (2007), la mera analisi della L1 degli apprendenti o della loro interlingua, appare comunque limitativa. Per questa ragione propongono un approccio più complesso, basato un modello di analisi integrata (*Integrated Contrastive Model*) che prevede una serie di confronti incrociati tra L1-L1, L1-L2 e L2-L2. Inoltre - secondo gli autori - il confronto tra le interlingue degli apprendenti (IL-IL) permette di individuare fattori specificamente linguistici in grado di incidere sul processo di acquisizione, e di osservare fenomeni di *transfer* comuni in apprendenti con un diverso background linguistico (cfr. ivi, p. 45).

In conclusione, l'acquisizione delle strutture prosodiche della lingua target può rappresentare un processo molto lento. La scarsa consapevolezza delle funzioni linguistiche e pragmatiche dell'intonazione può rallentare l'apprendimento e in alcuni casi generare incomprensioni e fallimenti sul piano comunicativo nelle interazioni con i nativi. In quest'ottica, è ragionevole supporre che l'acquisizione delle funzioni paralinguistiche dell'intonazione possa destare ulteriori difficoltà, soprattutto alla luce delle evidenze empiriche che hanno ridimensionato l'ipotesi dell'universalità degli andamenti intonativi paralinguistici (cfr. Chen, Gussenhoven, Rietveld, 2004; Chen, 2005).

### **3.4 Il parlato emotivo in L2: stato dell'arte**

Le ricerche sull'espressione vocale e la percezione delle emozioni in una lingua seconda non sono numerose. Come già accennato, l'interesse verso questo aspetto della comunicazione in L2 è relativamente recente, tuttavia le indagini acustico-percettive hanno evidenziato alcune problematiche legate alla gestione del parlato emotivo in una lingua seconda, ponendo l'accento su fattori linguistici, culturali e sociali.

Come suggerisce Dewaele (2005, p.375), la distanza tipologica e culturale dalla lingua target può influire sulla capacità degli apprendenti di gestire il parlato emotivo in L2: "SLA<sup>46</sup> research shows that learners from "distant" cultures experience significantly greater difficulties in identifying emotion in the L2 and in judging the intensity of that emotion than do fellow learners from "closer" cultures with similar levels of proficiency" (*ibid.*), per questa ragione richiama l'attenzione degli insegnanti e dei ricercatori sul peso di questi fattori all'interno di un contesto di apprendimento. A questo proposito, i primi studi in L2 (Rintell;1984, 1990) hanno dimostrato che la competenza legata alla comunicazione vocale delle emozioni dipende in buona misura dal grado di familiarità e di conoscenza del contesto culturale della lingua target (Kim, Dorner, 2013).

Le difficoltà sussistono anche sul piano prettamente prosodico. Uno studio condotto da Holden e Hogan (1993) ha esplorato la percezione degli andamenti intonativi paralinguistici in apprendenti di madrelingua russa e inglese, con l'obiettivo di valutare "the emotional and attitudinal 'confusion' that may arise in the use of foreign intonation in L2 (p. 70). Gli autori sono partiti dalla considerazione che le differenze esistenti tra i due sistemi linguistici in relazione all'altezza e all'estensione tonale<sup>47</sup> sono spesso causa di equivoci nelle interazioni tra apprendenti russi di inglese e parlanti nativi inglesi, evidenziando la tendenza di questi ultimi a interpretare come segnali di collera o irritazione variazioni intonative giudicate come neutrali da parlanti nativi russi. Per questa ragione, è stato chiesto ai gruppi coinvolti di valutare diverse tipologie di enunciato (interrogative e esclamative) in base alla loro valenza emotiva, al fine di verificare "whether there would be a significant change in the judgment of a selection of 10 emotions and attitudes with a change of intonation, while keeping other phonetic factors constant" (ivi, p. 67). Gli enunciati sono stati manipolati in modo da riprodurre l'intonazione inglese nelle frasi in russo e viceversa. Ai partecipanti è stato chiesto inoltre di valutare le differenze tra le produzioni originali e quelle manipolate. I risultati hanno evidenziato che in relazione alle emozioni positive, i parlanti nativi inglesi hanno giudicato l'intonazione della propria lingua più "alta" rispetto a quella dei parlanti nativi russi nella realizzazione delle interrogative polari. Inoltre, i parlanti nativi inglesi si sono mostrati più sensibili all'intonazione dei russi rispetto a questi ultimi, che al

---

<sup>46</sup> *Second Language Acquisition (SLA)*

<sup>47</sup> In base a quanto riportato dagli autori, il russo sarebbe caratterizzato da un'estensione tonale maggiore rispetto alla lingua inglese.

contrario hanno evidenziato una sensibilità minore alle variazioni di tono della lingua inglese. Tuttavia, in relazione al *pitch range*, entrambi i gruppi "reacted more negatively to the greater pitch range of Russian intonation in exclamations and yes-no questions" (p.84), attribuendo cioè agli enunciati originali in lingua russa una valenza negativa maggiore rispetto agli originali inglesi, definiti invece come più "passivi" nelle domande WH- (Ladd, 2014, p. 94). Come fa notare Ladd (*ibid.*): "English and Russian listeners interpreted *differences* in sentences in broadly the same way, regardless of their specific cultural norms about what count about 'neutral'".

Sulla percezione degli andamenti intonativi paralinguistici in una lingua seconda si è focalizzato lo studio di Chen (2005), condotto su apprendenti di olandese di origine inglese (con un basso livello di competenza) e su apprendenti di inglese di madrelingua olandese (con un livello avanzato), allo scopo di verificare se le variazioni prosodiche legate ai codici biologici veicolassero sul piano emotivo significati universalmente condivisi. La metodologia adottata è stata ripresa da Chen, Gussenhoven e Rietveld (2004), cfr. CAP. 2, PAR. 2.7. Anche in questo caso, i gruppi coinvolti hanno dimostrato una sensibilità linguo-specifica agli andamenti intonativi paralinguistici prodotti in una lingua straniera (in questo caso lingua seconda). Nel confronto tra i due studi, è emerso che "L2 English listeners differ from L1 English listeners in the perception of 'emphatic' and 'surprised' largely in the same way that L1 Dutch listeners differ from L1 English listeners" (Chen, 2005, p. 163). Secondo l'autrice, l'influenza della L1 potrebbe aver giocato un ruolo importante nell'interpretazione delle variazioni prosodiche:

First, there is strong evidence from both L2 English and L2 Dutch listeners that L1 transfer plays an important role in interpreting paralinguistic intonational meaning in L2, as in interpreting linguistic intonational meaning in L2. Also there is an indication that L2 listeners may activate their knowledge about intonational universals embodied in the biological codes (in particular, Gussenhoven's Effort Code), which accounts for L2 Dutch listeners' native-like behaviour in the perception of 'emphatic' as signalled by pitch register.

In relazione al livello di competenza, Chen ha suggerito che sin dalle prime fasi di acquisizione gli apprendenti riescano a cogliere alcune differenze sul piano paralinguistico tra L1 e L2, per poi affinare questa conoscenza nelle fasi più avanzate dell'interlingua<sup>48</sup>.

---

<sup>48</sup> Come riporta Chen "L2 learners appear to begin to gain knowledge of the difference in paralinguistic intonational meaning at the beginning and intermediate stages of language learning, and

Sulla percezione del parlato emotivo spontaneo si è focalizzato invece lo studio di de Abreu e Mathon (2005), i quali hanno investigato il ruolo della prosodia nel favorire o ostacolare il riconoscimento delle emozioni da parte di un gruppo di apprendenti portoghesi di francese L2. Gli stimoli consistevano in alcune porzioni di parlato spontaneo tratte da un corpus di scherzi telefonici realizzati da un conduttore radiofonico ai danni di istituzioni pubbliche e attività commerciali, allo scopo di innescare nelle vittime una reazione di collera. Il contenuto linguistico degli enunciati è stato oscurato introducendo rumore bianco, "in order to keep only the prosodic information" (ivi, p. 2). I soggetti, parlanti nativi francesi e apprendenti portoghesi di francese L2 con un livello di competenza intermedio, sono stati coinvolti in due task<sup>49</sup>: il primo mirava al riconoscimento degli enunciati di collera (*decision task*), il secondo chiedeva di valutare le produzioni in termini di intensità (*evaluation task*). Le percentuali di riconoscimento registrate per entrambi i gruppi (50% per i portoghesi e 62% per i francesi) hanno portato gli autori ad affermare che "prosodic information represents enough information to allow subjects recognizing anger" (ivi, p. 4). Tuttavia, sono emerse alcune differenze nelle interpretazioni degli enunciati di collera. In alcuni casi gli apprendenti portoghesi hanno valutato gli enunciati che contenevano pause, ripetizioni, errori come produzioni emotive diverse dalla collera ("not anger") a differenza dei francesi. In altri termini: "it seems that Portuguese do not consider a sentence said with anger when there are disfluencies, unlike French listeners" (*ibid.*).

Sul piano della produzione, Komar (2005) ha proposto un'analisi contrastiva del parlato emotivo prodotto da parlanti sloveni in inglese e da parlanti nativi inglesi. L'elicitazione delle emozioni è stata affidata alla lettura di un dialogo in lingua inglese. I risultati hanno evidenziato la tendenza degli sloveni ad impiegare un'intonazione "piatta", meno dinamica rispetto ai nativi, dovuta principalmente alle diversità dei due sistemi intonativi. Come spiega Komar (ivi, p. 4):

There are two main reasons for Slovenes sounding flat in English. First, the Slovenes produce the falling tones in a much narrower pitch range than the English, and second, the step up in pitch from the end of the falling pre-tonic segment and the beginning of the falling tone is significantly smaller compared to the step-up in pitch made by the English speakers.

---

obtain a grasp of what the differences may be at the advanced stage. They adjust their interpretation of 'emphatic' and 'surprised' accordingly"(2005, p. 179).

<sup>49</sup> Le istruzioni del test percettivo sono state tradotte in portoghese; inoltre, per facilitare gli apprendenti, le opzioni proposte "anger", "not anger" erano accompagnate da *emoticon* (espressioni facciali stilizzate).

Questi due fattori, insieme allo stato d'ansia e disagio che si avverte quando non si conosce bene una lingua, potrebbero essere responsabili di alcuni fallimenti comunicativi nelle interazioni con i parlanti nativi inglesi, i quali, secondo l'autrice, sarebbero inclini a giudicare l'intonazione meno dinamica degli sloveni come un segnale di disinteresse e di scarsa partecipazione (o di scortesia).

Sul piano percettivo, alcuni studi hanno evidenziato che l'accesso al contenuto verbale e vocale del messaggio emotivo non è così automatico in L2 così come lo è nella lingua nativa (Anooshian, Hertel, 1994; Harris, 2004; Harris, Gleason, Aycicegi 2006). A questo proposito, Chua Shi Min e Schirmer (2011) hanno esplorato il processo di integrazione del contenuto linguistico con gli indici prosodici in parlanti nativi e non nativi di inglese di diversa origine. Gli stimoli consistevano in una serie di termini con valenza positiva, negativa e neutra, pronunciati con tono felice, neutro o triste. In due esperimenti separati, i partecipanti hanno espresso un giudizio circa la valenza emotiva dello stimolo basandosi esclusivamente sul contenuto lessicale; in seguito, si sono focalizzati sul tono della voce, escludendo dunque il piano linguistico. I risultati hanno evidenziato alcune similarità tra i due gruppi, in particolare per quel che riguarda i tempi di risposta: "both native and non-native listeners responded faster and more accurately when verbal and vocal emotional expressions were congruent as compared to when they were incongruent" (ivi, p.). Nel caso degli stimoli incongruenti sul piano verbale e vocale, la reazione dei partecipanti è stata simile, a prescindere dal livello di competenza posseduto. Alla luce di ciò, gli autori hanno suggerito che "the integration of verbal and vocal emotional expressions occurs as readily in one's second language as it does in one's native language" (*ibid.*).

Graham, Hamblin e Feldstein (2001) hanno verificato l'abilità di un gruppo di parlanti nativi e non nativi di inglese nel riconoscere alcune produzioni emotive realizzate da parlanti nativi inglesi. I risultati hanno evidenziato che gli apprendenti con un livello di competenza più alto non sono stati in grado di riconoscere con maggiore accuratezza gli stimoli proposti. Secondo gli autori, la capacità di decodifica non può essere acquisita in assenza di un'esposizione intensiva al contesto culturale della lingua target o senza un intervento didattico mirato allo sviluppo di queste abilità.

Altrov (2013) ha indagato il peso delle influenze culturali sulle abilità di decodifica in una lingua seconda. I soggetti coinvolti appartenevano a tre gruppi distinti: parlanti nativi di origine estone, apprendenti di origine russa residenti in Estonia con un livello di competenza elevato (C1), parlanti nativi russi senza alcuna conoscenza della

lingua estone. A ciascun gruppo è stato somministrato un test uditivo contenente produzioni vocali in lingua estone, corrispondenti a tre emozioni (gioia, collera e tristezza) e all'eloquio neutro. Dai risultati è emersa una disparità nelle percentuali di riconoscimento in base alla tipologia dei soggetti coinvolti. I parlanti di origine estone hanno totalizzato la percentuale di decodifiche corrette più alta; nel caso degli apprendenti russi residenti in Estonia, si è registrata una riduzione dell'indice di accuratezza, tuttavia la percentuale di riconoscimento ha superato la soglia della casualità (oltre il 50%); infine, i parlanti nativi russi (senza alcuna conoscenza della lingua estone), hanno identificato con una percentuale superiore alla casualità soltanto l'espressione della tristezza; le altre emozioni sono state riconosciute con un indice di accuratezza molto inferiore. In base alle conclusioni tratte da Altrov, gli apprendenti russi, vivendo a contatto con la cultura estone, avrebbero "imparato" a decodificare con maggiore accuratezza le emozioni vocali espresse dai nativi rispetto ai parlanti russi residenti in Russia. L'autore conclude affermando: "Our results [...] suggest that understanding emotions is dependent on cultural factors and social interactions. That is, the social norms of a culture are learnt during practical interaction" (ivi, p. 172).

Zhu (2013) ha esplorato la percezione del parlato emotivo in lingua cinese da parte di 3 gruppi di ascoltatori di origine diversa: un campione di 20 ascoltatori di madrelingua cinese, 20 di lingua olandese e 20 apprendenti di cinese L2 di origine olandese. Gli stimoli veicolavano nello specifico: gioia, collera, sorpresa, sarcasmo (più il neutro). I risultati del test percettivo non hanno evidenziato nel gruppo di nativi cinesi una maggiore accuratezza nel riconoscimento delle emozioni rispetto agli altri gruppi (contrariamente all'ipotesi dell'*ingroup-advantage*, cfr. PAR. 2.7), che al contrario hanno totalizzato percentuali più elevate. Infatti, "the results showed that advanced Dutch L2 learners of Chinese recognized Chinese emotional prosody significantly better than Chinese native listeners and Dutch naïve listeners" (ivi, p. 23). L'ipotesi iniziale di Zhu prevedeva infatti un'attenzione minore agli usi paralinguistici del *pitch* da parte degli ascoltatori di madrelingua cinese (una lingua tonale) rispetto a parlanti di lingue non tonali (come l'olandese). L'eloquio neutro è stato riconosciuto con maggiore accuratezza da tutti e tre i gruppi, i quali hanno identificato altrettanto facilmente anche l'espressione della collera.

Bhatara *et al.* (2016) hanno esplorato la relazione tra la competenza in L2 e l'abilità nel riconoscimento di emozioni positive e negative nella lingua target, nello specifico l'inglese americano. I soggetti coinvolti erano di origine francese, con un

livello di competenza variabile (stabilito in base all'autovalutazione dei partecipanti). L'esperimento percettivo prevedeva l'ascolto di alcuni enunciati in lingua inglese realizzati da attori professionisti e la decodifica delle emozioni espresse mediante un test a risposta chiusa. In seguito, ai partecipanti è stato chiesto di valutare (su una scala) la piacevolezza (*pleasantness*), la potenza (*power*), l'attivazione (*alertness*) e l'intensità (*intensity*) degli stimoli emotivi. Gli stimoli comprendevano, oltre a interi enunciati, anche semplici vocalizzazioni (o *affect burst*, cfr. CAP. 2). Dai risultati è emerso che gli apprendenti con un livello di competenza elevato non sono stati facilitati nel riconoscimento delle emozioni, in particolar modo di quelle positive (gioia, orgoglio, interesse e sollievo). Secondo gli autori, l'aumento della competenza della lingua inglese potrebbe aver compromesso la percezione e il riconoscimento degli stati emotivi positivi anziché facilitarla, concludendo che: "increasing understanding of the L2 may be accompanied by a slight decrease in ability to understand subtle differences between positive emotions among other speakers" (Bathara *et al.*, 2016, p. 11).

La correlazione tra il livello di competenza della L2 (percepito dall'apprendente) e l'abilità nel riconoscimento delle emozioni appare dunque messa in dubbio. Tuttavia, studi recenti (cfr. Lorette, Dewaele 2015) hanno evidenziato che a un livello di competenza elevato nella lingua target corrisponde una maggiore accuratezza nella decodifica di stimoli audiovisivi di natura positiva; inoltre, la distanza culturale sembra influenzare le percentuali di riconoscimento (come già riportato negli studi cross-linguistici, cfr. cap. 2).

### 3.4.1 *Gli studi in italiano L2*

In ambito italiano, sono ancora piuttosto esigui gli studi che hanno indagato questo particolare aspetto della comunicazione nel processo di acquisizione dell'italiano L2. Uno studio recente condotto da Maffia, Pellegrino e Pettorino (2014) ha analizzato il parlato emotivo prodotto da apprendenti cinesi in italiano con un livello di competenza elevato (pari a C1 del CEFR). Nell'esperimento è stato coinvolto anche un campione di parlanti nativi di italiano. Le emozioni sono state elicitate mediante un *Card Task*, ovvero un'attività finalizzata all'interazione verbale tra due partecipanti impegnati in un gioco di carte<sup>50</sup>. Come spiegano gli autori, questa procedura è stata

---

<sup>50</sup> La struttura del task è la seguente: i soggetti, disposti in modo da evitare il contatto visivo, svolgono compiti diversi: il primo tenta di descrivere in due minuti le carte che ha ricevuto (cinque in

adottata allo scopo di suscitare "emotional linguistic reactions in the players and to arouse five different emotions (anger, anxiety, disgust, fear and surprise)" (ivi, p.). Le produzioni emotive raccolte sono state analizzate mediante Praat, tenendo in considerazione i seguenti parametri: durata e numero di sillabe, pause piene e pause silenti, valori massimi e minimi di f0. In base a queste misurazioni sono stati calcolati tempo, velocità di articolazione ed estensione tonale (in semitoni). Le analisi hanno tenuto conto anche della presenza di *affect bursts*.

In relazione all'italiano nativo i risultati hanno confermato la nota distinzione tra emozioni ad alta e bassa attivazione. Le espressioni di collera, paura e sorpresa si sono caratterizzate per valori di f0 elevati e un'estensione tonale ampia; al contrario l'espressione del disgusto ha presentato valori inferiori, essendo un'emozione a bassa attivazione. Il parlato emotivo dei cinesi non ha mostrato, invece, questa variabilità: "as a matter of fact, F0 height and tonal range are quite steady in the whole corpus. The only exception is represented by anger and fear that are expressed with slightly higher values. These data seems to suggest that Chinese learners do not vary their pitch contour to distinguish different emotional states as in the case of native Italian speakers." (ivi, p.). Le produzioni nell'italiano non nativo erano caratterizzate in linea generale da un rallentamento della velocità d'eloquio (dovuto al tentativo di articolare con maggiore precisione le singole parole).

Il ricorso ad un'intonazione poco modulata e variabile nell'espressione delle emozioni sembra dunque essere una tendenza comune negli apprendenti di una lingua seconda. Tale dato è stato riportato anche in alcuni studi pilota (De Marco, Paone, 2015; De Marco, Paone, in stampa) condotti su studentesse indonesiane e polacche, le cui produzioni hanno evidenziato contorni intonativi simili in tutte le emozioni indagate (gioia, collera, tristezza, paura, sorpresa, disgusto ed eloquio neutro). Inoltre, le indagini percettive condotte in questi studi hanno suggerito l'ipotesi che il parlato emotivo prodotto dagli apprendenti in italiano L2 possa risultare poco efficace sul piano comunicativo, andando incontro a fraintendimenti con i nativi (cfr. De Marco, Paone, in stampa).

Il ruolo del transfer della L1 nel parlato emotivo in italiano L2 è stato esplorato in uno studio recente (De Marco, Sorianello, Paone, in stampa), che ha coinvolto nello parlanti nativi italiani e apprendenti di origine russa e persiana. Si trattava nello

---

totale), il secondo prova a individuare le immagini descritte in base alle carte che gli sono state fornite (25 in totale) e che differiscono in piccoli dettagli.

specifico di studenti Erasmus aventi un livello di competenza pari a B2. L'elicitazione delle emozioni è stata affidata alla lettura di testi emotigeni induttori, all'interno dei quali è stata inserita una frase standard (*non è possibile*, ripresa e adattata da Anolli, Ciceri 1992; Anolli *et al.* 2008 e tradotta nelle tre lingue native). Per favorire l'immedesimazione dei soggetti, è stato adottato un metodo impiegato in alcuni studi di psicologia (cfr. Laird, Strout, 2007): i soggetti sono stati invitati ad esprimere in modo simultaneo, l'emozione vocale richiesta e l'espressione facciale equivalente. I partecipanti hanno avuto modo di ripetere più volte le loro performance, al fine di ridurre eventuali inibizioni o insicurezze dovute alla presenza del microfono. Gli apprendenti sono stati invitati ad esprimere le emozioni nella loro lingua e successivamente in italiano. I parametri indagati sono stati ripartiti in due macro-categorie: temporali (computo in ms della durata totale dell'enunciato bersaglio e della velocità di eloquio; intonativi (il rilievo dell'andamento e dei valori della frequenza fondamentale (f0), frequenza iniziale (*onset*), frequenza finale (*offset*), frequenza media, escursione melodica dell'enunciato (calcolata in semitoni). Dai risultati è emerso che in relazione ai parametri temporali (durata totale e velocità d'eloquio), a parità di numero di sillabe, tutti gli enunciati emotivi prodotti dagli apprendenti nella loro lingua nativa erano più brevi rispetto a quelli articolati in italiano L1, con la sola eccezione dell'enunciato neutro. Questa tendenza ha suggerito l'ipotesi che russi e persiani, quando si esprimono in italiano non riescano a controllare adeguatamente questo parametro elocutivo, specie quando il comportamento dell'emozione si discosta in modo significativo da quello dell'italiano nativo, determinando così produzioni emotive "diverse" da quelle attese. In effetti, nel confronto tra le produzioni in L1 e quelle in L2, si è osservato come in italiano L2 gli enunciati emotivi prodotti dagli apprendenti non si discostavano da quelli prodotti nella loro L1. L'effetto del *transfer* sembra, dunque, in questo caso, essere confermato. In relazione ai parametri intonativi, l'escursione frequenziale degli enunciati emotivi prodotti nell'italiano nativo, è risultata tendenzialmente più estesa rispetto a quella realizzata in L2. L'analisi degli enunciati prodotti nella L1 degli apprendenti ha mostrato, tanto in russo quanto in persiano, un'escursione tonale più ridotta rispetto all'italiano nativo, determinando realizzazioni nettamente più monotone. Un'intonazione meno dinamica ha caratterizzato infatti le produzioni in italiano di entrambi i gruppi (in misura maggiore, gli enunciati prodotti dai partecipanti russi hanno evidenziato un *range* frequenziale più ristretto rispetto ai persiani, i quali invece sono riusciti a modulare meglio l'intonazione, avvicinandosi in

alcuni casi alla lingua *target*). In questo caso, è stato difficile attribuire la poca dinamicità del parlato degli apprendenti ad un vero e proprio *transfer* prosodico, in quanto, come già accennato, tale comportamento può trovare una giustificazione nel disagio che si prova quando si parla una lingua straniera, a maggior ragione in contesti che investono la sfera emozionale (*ibid.*).

Da uno studio recente condotto da Sorianello e De Marco (in stampa), è emerso un quadro decisamente eterogeneo in relazione alla L1 degli apprendenti coinvolti (nello specifico, russi, tunisini e spagnoli). La verifica acustica delle produzioni emotive realizzate dagli studenti nella loro L1 e in italiano ha mostrato prove evidenti di *transfer*, soprattutto nel parlato dei tunisini, i quali tendevano a riprodurre in italiano contorni intonativi molto simili a quelli della loro lingua nativa. Diverso il caso delle produzioni dei russi e degli spagnoli, in cui i contorni, parzialmente sovrapposti e piuttosto monotoni, perdevano parte della loro distintività uditiva. Secondo le autrici, le produzioni in L2, per certi versi confuse e molto spesso non congruenti sul piano pragmatico rappresentano un indizio della difficoltà cui vanno incontro gli apprendenti quando realizzano in una lingua diversa un fenomeno paralinguistico complesso come quello delle emozioni (*ibid.*).

### **3.5 Conclusioni**

Alla luce di quanto esposto finora, possiamo concludere che lo sviluppo della competenza emotiva in L2 costituisca un processo delicato nel percorso linguistico di un apprendente. Variabili individuali (come l'età, la motivazione, il livello di socializzazione), insieme a fattori culturali e sociali (la distanza culturale tra L1 ed L2, ad esempio), possono accelerare o rallentare questo processo (cfr. Lorette, Dewaele 2015); come abbiamo osservato, le difficoltà maggiori investono il livello prosodico della lingua, dando luogo a fenomeni di transfer dalla L1 (Chen, 2005, De Marco, Sorianello, Paone, in stampa) e generando produzioni poco efficaci sul piano comunicativo (il ricorso, ad esempio, ad un'intonazione piatta e meno modulata, cfr. Komar, 2005, Maffia, Pellegrino, Pettorino, 2014, De Marco, Paone, 2015; in stampa) e incomprensioni sul piano percettivo. L'abilità nel riconoscimento delle emozioni a livello vocale sembrerebbe dipendere non tanto dal livello di competenza linguistica di L2, quanto piuttosto da una conoscenza più profonda del contesto culturale della lingua target (cfr. Graham, Hamblin e Feldstein, 2001, Altrov, 2005). Inoltre, alcuni studi

hanno posto l'accento sulla variabilità linguo-specifica che caratterizza l'interpretazione degli andamenti intonativi paralinguistici (Chen, Gussenhove, Rietveld, 2004, Chen, 2005). Tutti questi fattori assumono un peso specifico rilevante all'interno di un contesto di apprendimento guidato, in cui l'intervento didattico può contribuire a sensibilizzare gli apprendenti ad una presa di coscienza dei meccanismi legati alla comunicazione emotiva in L2, con particolare riferimento alla prosodia emotiva (cfr. De Marco, Paone, 2015; in stampa; De Marco, Soriano, Paone, in stampa).

L'abilità nel riconoscere ed esprimere- in modo efficace- le proprie emozioni e quelle altrui costituisce una delle competenze necessarie ad un agente sociale- che si muove in una realtà altra- per cogliere le infinite sfumature di senso che uno scambio comunicativo comporta e raggiungere obiettivi individuali e sociali.

# Capitolo 4

---

## Obiettivi e metodi della ricerca

### 4.1 Oggetto della ricerca

Le questioni trattate nei primi tre capitoli costituiscono lo sfondo teorico e metodologico necessario all'introduzione dell'oggetto della presente ricerca. La linea di congiunzione tra il fenomeno emotivo e il contesto linguistico e culturale rappresenta il punto di partenza di questo lavoro, ovvero la dimensione emotiva nel processo di acquisizione di una lingua seconda, nelle sue modalità di espressione e percezione.

Il presupposto di base è che nell'acquisizione di una lingua seconda la componente emozionale, concepita come un bisogno dell'apprendente per poter comunicare in modo efficace, sia un elemento da tenere in considerazione nelle pratiche didattiche, soprattutto quando la distanza culturale è notevole.

Alla luce degli studi precedenti (cfr. De Marco, Paone, 2016) si ipotizza, in primo luogo, che gli apprendenti di italiano sperimentino difficoltà nell'espressione e nella percezione delle emozioni espresse dai parlanti nativi, soprattutto quando la distanza linguistica e culturale tra L1 ed L2 è considerevole; in secondo luogo, che gli strumenti e le pratiche messi in atto dagli apprendenti nella loro produzione possano non essere efficaci sul piano comunicativo, dando luogo ad equivoci e fraintendimenti nell'interazione con i nativi.

Le domande di ricerca sono pertanto le seguenti:

1. A quali difficoltà vanno incontro gli apprendenti di una lingua seconda in contesti comunicativi che investono la dimensione emotiva intra e interpersonale?
2. Gli apprendenti di una lingua seconda sono in grado di esprimere efficacemente la proprie emozioni attraverso il ricorso alle strutture prosodiche della lingua *target*?
3. In che misura gli apprendenti sono in grado di decodificare la prosodia emotiva dei parlanti nativi?

4. In che misura la L1 degli apprendenti influenza i processi di codifica e decodifica delle emozioni sul piano vocale?

Per dare una risposta a questi quesiti, è stato necessario operare una selezione in merito alla L1 degli apprendenti. La necessità di coinvolgere membri di culture distanti da quella italiana<sup>51</sup>, sia sul piano linguistico sia sul piano culturale, ha indirizzato la scelta verso la cultura russa. Un'ulteriore motivazione si deve al fatto che, alla luce degli studi pilota condotti in precedenza (cfr. De Marco, Paone, 2014; 2015; 2016; in stampa, De Marco, Soriano, Paone, in stampa) in cui hanno preso parte apprendenti di diversa origine (russa, spagnola, inglese, iraniana, polacca, ecc.), si è valutata la possibilità di condurre un'indagine ad ampio raggio che coinvolgesse un numero maggiore di apprendenti russi, in modo da approfondire le difficoltà evidenziate da questi ultimi.

In base ai dati forniti dall'ISTAT<sup>52</sup>; la presenza di russi in Italia è in graduale aumento (si registrano oltre 35mila russi in Italia). In Calabria, regione in cui è stata condotta la presente ricerca, si registra la presenza di circa un migliaio di russi. Le motivazioni che spingono a scegliere come meta l'Italia sono di carattere professionale, accademico e turistico. In particolare, gli scambi internazionali favoriscono la mobilità degli studenti e il loro inserimento all'interno dei corsi di studio.

Per questa ragione, l'indagine si è focalizzata nello specifico su studenti internazionali di origine russa iscritti all'Università della Calabria. Nei paragrafi successivi, saranno illustrati gli obiettivi e i metodi della ricerca.

## 4.2 Obiettivi della ricerca

Gli obiettivi del presente lavoro sono riconducibili alle domande di ricerca formulate nel PAR. 4.1. I paragrafi successivi illustreranno nel dettaglio le finalità dell'indagine acustica e gli obiettivi della verifica uditiva.

---

<sup>51</sup> Il russo è una lingua indoeuropea e appartiene al ceppo slavo orientale insieme all'ucraino e al bielorusso. In base ai dati forniti dal Ministero dell'Educazione e della Scienza della Federazione Russa, si stima che il russo sia parlato da 180 milioni di persone come lingua madre e da 120 milioni di persone come L2, infatti si tratta della L1 della maggioranza della popolazione della Russia.

<sup>52</sup> <http://dati.istat.it>

#### 4.2.1 Il processo di encoding

In relazione al processo di codifica delle emozioni, l'obiettivo generale è quello di analizzare e confrontare le produzioni emotive realizzate da due gruppi linguistici aventi una differente L1 (l'italiano e il russo). Nel dettaglio, le produzioni prese in esame sono riconducibili a tre tipologie:

- a) parlato emotivo in italiano L1, prodotto cioè da parlanti nativi (da ora Ita L1)
- b) parlato emotivo italiano L2, prodotto da apprendenti di origine russa (R-Ita L2);
- c) parlato emotivo in russo L1, prodotto dai medesimi apprendenti (Ru L1)

L'analisi e il confronto dei tre contesti linguistici forniranno un quadro globale circa le tendenze prosodiche che caratterizzano il parlato emotivo dei soggetti coinvolti.

In secondo luogo, si intende verificare se e in che misura la L1 degli apprendenti influisca sul parlato emotivo in italiano L2. Come già accennato in precedenza, sebbene in molti casi non sia possibile parlare di un transfer vero e proprio, è indubbio il fatto che le strutture della lingua materna fungono da battistrada per la costruzione, e dunque l'acquisizione, di nuove strutture. Alcuni studi (cfr. Queen, 2001) condotti sulla prosodia dell'interlingua, hanno registrato una mescolanza di pattern non riconducibili né alla L1 degli apprendenti né alla L2, che tuttavia presentavano tratti di entrambe le lingue<sup>53</sup>. L'ipotesi del transfer prosodico (cfr. CAP. 3) verrà quindi verificata all'interno di un contesto specificamente emotivo.

Le tendenze rilevate nel parlato degli italiani nativi dovrebbero rappresentare il quadro di riferimento per comprendere i punti in comune e le differenze rispetto all'italiano L2 e al russo L1. In questo senso, l'indagine uditiva appare necessaria per verificare in modo puntuale l'efficacia comunicativa del parlato emotivo in italiano L2.

---

<sup>53</sup> Studi recenti sulle produzioni di bambini bilingui turco-tedesco (cfr. Queen, 2001) hanno evidenziato che, a parte alcuni casi di interferenza, i profili intonativi presentavano dei tratti ibridi, non riconducibili né alla L1 né alla L2, ma derivanti dalla fusione di tratti prosodici delle due lingue (Marotta, 2008), cfr. anche CAP. 3.

#### 4.2.2 Il processo di *decoding*

In relazione al processo di *decoding*, l'indagine uditiva rivolta agli apprendenti russi intende verificare in primo luogo le abilità di decodifica di questi ultimi rispetto alle produzioni emotive realizzate dagli italiani nativi e dai parlanti russi. In particolare, l'indagine tenterà di mettere in luce eventuali difficoltà associate al riconoscimento di una o più emozioni in relazione all'italiano nativo.

Per quanto concerne la verifica uditiva rivolta agli uditori italiani, il primo obiettivo è quello di stabilire se e in che misura le produzioni emotive realizzate dagli apprendenti in italiano L2 trovino corrispondenza nei giudizi affettivi espressi da un campione di nativi. In secondo luogo, l'indagine uditiva si propone di verificare se e in che misura il livello di competenza della lingua e di conoscenza della cultura target incidano sulla capacità di gestione e modulazione delle risorse prosodiche in contesti emotivi differenti. In altri termini, se ad un livello di competenza (linguistica e sociopragmatica) più elevato corrisponda una maggiore efficacia comunicativa nella lingua target all'interno di un contesto emotivo.

In terzo luogo, la ricerca si propone di testare la capacità degli italiani nel decodificare le produzioni emotive realizzate in italiano L1 e in russo L1. Indirettamente, questo confronto permetterà di verificare l'ipotesi dell'*in-group advantage* (cfr. CAP. 2, PAR. 2.7.2.2), in base alla quale le probabilità di riconoscimento aumentano se gli stimoli sono pronunciati nella lingua parlata dagli uditori.

### 4.3 L'impianto metodologico

Il disegno sperimentale si articola in due fasi: la prima relativa alla raccolta e all'analisi del materiale emotivo (*encoding*), la seconda relativa alla verifica uditiva di quest'ultimo (*decoding*).

#### 4.3.1 La codifica delle emozioni: strumenti e procedure

##### 4.3.1.1 Partecipanti

Hanno preso parte alla fase di *encoding* 10 apprendenti russi (5 uomini e 5 donne) tra i 22 e 27 anni (età media: 25 anni), studenti presso l'Università della Calabria, grazie al programma di scambio internazionale<sup>54</sup> e 10 parlanti nativi di

---

<sup>54</sup> Si trattava, nello specifico, di studenti di Economia.

italiano, di origine calabrese (5 uomini e 5 donne) dai 28 ai 37 anni (età media: 32 anni). Il livello di competenza degli apprendenti è intermedio (B1/B2)<sup>55</sup>, tuttavia esistono alcune differenze in merito al periodo di permanenza in Italia maturato da ciascuno di loro. In linea generale, al momento delle registrazioni la maggior parte degli apprendenti (7 su 10) si trovava in Italia da circa 15 mesi e aveva seguito regolarmente un corso di italiano presso l'università ospitante; 3 apprendenti su 10 invece si trovavano in Italia da un periodo di tempo inferiore (circa 8 mesi), sebbene avessero già seguito un corso di italiano nel loro paese d'origine. La TAB. 4-1 sintetizza i dati relativi ai partecipanti coinvolti in questa fase della ricerca.

**TAB. 4-1. PROSPETTO PARTECIPANTI**

<i>Prospetto partecipanti</i>						
	<b>N.</b>	<b>L1</b>	<b>Età (M)</b>	<b>Provenienza geografica</b>	<b>Livello competenza</b>	<b>Periodo di permanenza in Italia</b>
Gruppo (R)	10	Russo	25	Mosca, San Pietroburgo	B1/B2	8-15 mesi
Gruppo (I)	10	Italiano	32	Cosenza, Catanzaro	-	-

#### 4.3.1.2 Le emozioni oggetto di indagine

In relazione alle emozioni considerate, il presente studio si è focalizzato su tre stati emotivi: collera, tristezza e gioia. In una prospettiva dimensionale, si tratta di emozioni caratterizzate da livelli di attivazione, valutazione e potere differenti. La collera, così come la gioia, è uno stato emotivo ad alta attivazione, che riflette una condizione di dominanza/controllo (potenza) rispetto all'evento emotigeno. A differenza della gioia, la valutazione associata alla collera è di natura negativa, in quanto deriva dal mancato raggiungimento o dalla compromissione di uno scopo; viceversa, la valenza

<sup>55</sup> In base alle direttive del Framework, il livello B1: " Comprende i punti chiave di argomenti familiari che riguardano la scuola, il tempo libero ecc. Sa muoversi con disinvoltura in situazioni che possono verificarsi mentre viaggia nel paese in cui si parla la lingua. E' in grado di produrre un testo semplice relativo ad argomenti che siano familiari o di interesse personale. E' in grado di descrivere esperienze ed avvenimenti, sogni, speranze e ambizioni e spiegare brevemente le ragioni delle sue opinioni e dei suoi progetti"; il livello B2, invece: " Comprende le idee principali di testi complessi su argomenti sia concreti che astratti, comprese le discussioni tecniche nel suo campo di specializzazione. E' in grado di interagire con una certa scioltezza e spontaneità che rendono possibile un'interazione naturale con i parlanti nativi senza sforzo per l'interlocutore. Sa produrre un testo chiaro e dettagliato su un'ampia gamma di argomenti e spiegare un punto di vista su un argomento fornendo i pro e i contro delle varie opzioni" (Consiglio d'Europa, 2001).

positiva della gioia è data dalla realizzazione delle proprie aspettative. La tristezza, infine, condivide con la collera la valenza negativa, ma non il livello di attivazione e il grado di controllo, in quanto si tratta di uno stato emotivo derivante da un senso di impotenza e di sottomissione.

Per cogliere queste variazioni all'interno dell'eloquio, si è ritenuto opportuno ottenere una *base line* neutra rispetto alla quale osservare l'andamento dei parametri all'interno delle tre tipologie emotive. Per questa ragione, nel corso della trattazione, si farà riferimento all'eloquio neutro (o eloquio normale) per indicare un tipo di parlato privo di sfumature emozionali importanti<sup>56</sup>. La TAB. 4-2 mostra i termini emotivi utilizzati rispettivamente per il russo e l'italiano.

**TAB. 4-2. TERMINI EMOTIVI IN ITALIANO E IN RUSSO**

<i>Termini emotivi</i>	
Italiano	Russo
Tristezza	Грусть (Grust)
Collera	Гнев (Gnief)
Gioia	Радость (Radast)

#### 4.3.1.3 I dialoghi induttori

Il materiale emotivo e quello neutro sono stati elicitati mediante dei dialoghi induttori, tradotti in entrambe le lingue (cfr. Appendice). I dialoghi di carattere esplicitamente emotivo sono stati strutturati in modo da simulare il contesto comunicativo di una telefonata informale tra due amici legati da un rapporto di familiarità e confidenza. Per cogliere le variazioni rispetto all'eloquio normale dei partecipanti, è stato elaborato un dialogo privo di reazioni emotive, in cui prevale invece il carattere informativo della telefonata, che si svolge tra un tecnico e un cliente.

A differenza degli scenari emotigeni, la realizzazione di questo tipo di dialogo consente di proiettare lo scenario sperimentale in una situazione più naturalistica. In primo luogo, l'interazione tra i partecipanti rende più fluido ed empatico il contesto comunicativo (Douglas-Cowie *et al.*, 2003), il quale risulta meno artificiale e meccanico. In secondo luogo, il dialogo telefonico implica che gli interlocutori non si

<sup>56</sup> Di fatto è impossibile ottenere un parlato neutro, privo di qualsiasi sfumatura emozionale, per cui gli enunciati in questione sono da intendersi come frasi realizzate in assenza di un'attivazione fisiologica importante all'interno di un contesto non marcato emotivamente.

vedano (come in effetti è stato), e che ricorrano alla voce per segnalare e comunicare significati importanti. Ed è proprio l'intenzione comunicativa che è stata elicitata: ogni dialogo è stato costruito in modo da dare l'opportunità al parlante di comunicare esplicitamente il proprio stato d'animo.

L'intenzione comunicativa è stata associata ad una frase bersaglio (cfr. PAR. 4.2.3), la stessa per ogni dialogo. In una prospettiva dimensionale (cfr. CAPP. 1 e 2), le emozioni elicitate nei dialoghi si caratterizzano per diversi livelli di attivazione, valutazione e potere. I primi due dialoghi, relativi alla tristezza e alla collera, hanno una valenza negativa, tuttavia, mentre il primo riproduce un'emozione con un basso grado di attivazione e di dominanza/controllo, il secondo introduce un'emozione ad alta attivazione (la collera), che si distingue per un alto grado di potenza/dominanza. Il terzo dialogo, l'unico con valenza positiva, si caratterizza per un'emozione ad alta attivazione (la gioia) e con un grado elevato di controllo.

Per cogliere le diverse componenti dell'emozione, tre tipologie di informazioni sono state date ai partecipanti in sede di registrazione (cfr. Lang, 1979; Anolli, Ciceri, 1992; Anolli *et al.* 2008):

a) informazioni stimolo-pertinenti (ovvero informazioni circa l'evento emotigeno e il contesto in cui si è verificato);

b) informazioni risposta- pertinenti, cioè "informazioni concernenti le risposte emozionali appropriate al contesto" (Anolli, Ciceri, 1992, p. 255).

c) informazioni di significato, ovvero "informazioni che definiscono il significato degli stimoli e delle risposte" (ivi, p. 256).

Per favorire ulteriormente l'immedesimazione, è stato chiesto ai partecipanti di rievocare un evento passato<sup>57</sup> che riflettesse le dimensioni di attivazione, valutazione e potere associate agli stati emotivi in questione.

### *La struttura dei dialoghi*

Nei primi tre dialoghi, la conversazione si sviluppa a partire da una serie di conoscenze implicite condivise dai parlanti (A e B). A è in attesa di una notizia importante per il raggiungimento dei suoi obiettivi personali; B ne è a conoscenza e contatta A per accertarsi che tutto sia andato secondo le aspettative di A; a questo punto,

---

<sup>57</sup> Cfr. cap.2, par. 2.4.1, *MIPs based on the free mental generation of emotional states* (Gerrards-Hesse, Spies, Hesse, 1994).

la reazione di A rispetto all'evento cambia in base al grado di attivazione, valutazione e potere che il soggetto sperimenta nel corso della situazione emotiva proposta:

- nel primo dialogo (*tristezza*), la notizia ricevuta da A, di valenza negativa, lo proietta in una condizione di impotenza e di rassegnazione nei confronti di un avvenimento irreversibile. La valenza comunicativa dell'enunciato assume qui (ma anche negli altri due dialoghi) un valore espressivo- emotivo, e non solo informativo: A comunica intenzionalmente lo stato emotivo innescato dall'avvenimento in questione, suggerendo dunque a B un'inferenza sulla propria reazione (la tristezza) rispetto a ciò che è accaduto.

- nel secondo dialogo, la notizia (anch'essa di valenza negativa) scatena una reazione di collera da parte di A, i cui piani sono stati compromessi da terzi, verso i quali crede di esercitare una certa autorità (suggerendo dunque un grado di controllo maggiore su quanto è accaduto). A comunica la notizia a B e la sua collera si manifesta sin dalle prime battute ("Ah, non sai cos'è successo?").

- nel terzo dialogo, la notizia è in linea con le aspettative di A, che vengono dunque realizzate. A comunica a B la lieta notizia, con l'intenzione di veicolare contestualmente anche la propria reazione di gioia rispetto a quanto avvenuto.

In tutti e tre i casi l'evento in questione non viene esplicitato del tutto durante la telefonata, poiché fa parte delle conoscenze implicite condivise dai due interlocutori, mentre viene resa nota la reazione emotiva che caratterizza il comportamento vocale dell'*encoder* e il significato dell'evento verificatosi.

L'ultimo dialogo riproduce un contesto comunicativo leggermente più formale e non contrassegnato da particolari sfumature emotive, che vede come protagonisti il tecnico di una ditta (A) e un cliente (B). A chiama B per accertarsi che il nuovo apparecchio sia stato consegnato presso il domicilio del cliente. In questo caso, la risposta di B alla domanda di A ("le hanno portato il nuovo apparecchio") è essenzialmente di carattere informativo ("lo hanno portato"), privato cioè di sfumature emotive importanti.

La traduzione dei testi dall'italiano al russo è stata affidata ad un linguista di origine italo-russa, parlante bilingue, con una competenza quindi, non solo delle strutture linguistiche dei due sistemi considerati ma anche dei contesti situazionali in cui vengono utilizzate, nonché dei valori e delle norme culturali che regolano i comportamenti di entrambe le culture. Le indicazioni fornite dal traduttore bilingue sono state fondamentali per la costruzione di un dialogo che mantenesse una valenza

socio-pragmatica simile nei due contesti culturali sotto osservazione. Naturalmente, il passaggio da una lingua all'altra implica spesso la perdita di sfumature sottili, difficilmente traducibili. Tuttavia, la traduzione dei dialoghi risulta abbastanza fedele al testo originale.

#### 4.3.1.4 *La frase bersaglio*

All'interno di ogni dialogo è stata inserita una frase bersaglio, tradotta in entrambe le lingue: *Lo hanno portato*, per l'italiano, *Eto privezli*, per il russo. La scelta di questo enunciato si deve principalmente a due ragioni: la prima è che sul piano semantico esso non è connotato da termini esplicitamente emotivi e si presta a diverse interpretazioni in base al contesto e al modo in cui viene pronunciato (sia in italiano che in russo); il pronome personale *lo* in italiano (*eto* in russo) si riferisce ad un oggetto o ad una persona la cui identità non viene rivelata nel corso della telefonata, ma che fa parte delle conoscenze implicite condivise dagli interlocutori. Soltanto nell'ultimo dialogo la natura dell'oggetto viene esplicitata (cfr. PAR. 4.2.2). La seconda ragione, di carattere pratico, è dovuta alla necessità di distinguere con maggiore precisione l'inizio della produzione e in generale i confini frasali, per cui il primo e l'ultimo fono dell'enunciato sono entrambi sonori<sup>58</sup> (cfr. Anolli, Ciceri, 1992, Anolli *et al.* 2008; Galatà, Romito, 2010).

Nei primi tre dialoghi (D1\_TRIS; D2\_COLL; D3\_GIO) la frase bersaglio viene ripetuta tre volte nel corso della telefonata, in fasi diverse della conversazione. In posizione iniziale, viene pronunciata dal parlante in seguito ad una domanda esplicita dell'interlocutore che chiede delucidazioni sulle notizie ricevute dall'altro (es. "Ma cos'è successo?"). In posizione mediana, cioè all'interno del turno, la frase è preceduta e seguita da un enunciato. In questo caso, la ripetizione dell'enunciato può essere interpretata come la volontà del parlante di ribadire l'evento verificatosi e la propria reazione rispetto ad esso. In posizione finale, preceduta da un enunciato, segnala la volontà di confermare la fondatezza dell'informazione data, nonché la conclusione del turno di parola e della telefonata.

---

<sup>58</sup> Inoltre, il numero di vocali e di consonanti di cui è composto l'enunciato risulta abbastanza bilanciato in entrambe le lingue. Come ricordano Anolli e Ciceri (1992, p. 268), le consonanti contribuiscono in misura maggiore a veicolare informazioni linguistiche, mentre le vocali quelle non linguistiche (cfr. Stankiewicz, 1964; Studdert-Kennedy, 1975). Nella versione italiana, le 6 sillabe che compongono l'enunciato si distinguono per la presenza ripetuta di due vocali (o/a) per un totale di 6 vocali e di 6 consonanti. Nella traduzione russa, le 5 sillabe sono caratterizzate dalla presenza di 5 vocali, e di 6 consonanti.

Nell'ultimo dialogo (D4\_NEU) la frase bersaglio viene ripetuta due volte: la prima in posizione iniziale all'interno del turno di parola e l'altra in posizione finale. In entrambi i casi, viene enunciata in seguito ad una domanda esplicita dell'interlocutore. La scelta di ripetere più volte l'enunciato bersaglio si deve principalmente alla necessità di valutare, in sede di analisi, quale tra le tre posizioni risultasse maggiormente incisiva e soprattutto più genuina e spontanea.

#### *4.3.1.5 Strumenti e procedure di registrazione*

Le sedute di registrazione si sono svolte nell'arco di tre mesi, all'interno di un'abitazione privata dotata di un locale insonorizzato. Questa scelta si deve principalmente a questioni di carattere metodologico. Si è ipotizzato che la selezione di un ambiente familiare e privato avrebbe contribuito a ridurre l'impatto iniziale con il *setting* sperimentale e l'eventuale imbarazzo dei partecipanti, chiamati a interagire in un contesto specificamente emotivo. Per questa ragione, ogni seduta di registrazione è stata preceduta da uno scambio di battute informali, per consentire ai partecipanti di familiarizzare con l'ambiente circostante e con le altre persone presenti all'esperimento (la sottoscritta, un tecnico audio e un parlante bilingue italo-russo che ha svolto il ruolo di interprete). A questo proposito, è bene sottolineare che sia i parlanti russi che i parlanti italiani erano legati tra loro da rapporti di amicizia. Questa conoscenza preliminare non solo ha contribuito a ridurre il carico d'ansia e imbarazzo dei partecipanti, ma anche a rendere più fluide e naturali le interazioni dialogiche di natura emotiva, caratterizzate da un alto grado di immediatezza e informalità.

La registrazione dei dialoghi è stata effettuata mediante un microfono cardioide (AKG C 3000 B<sup>59</sup>) e una scheda audio (Digidesign Mbox2) collegata ad un computer portatile. Il microfono, agganciato ad un'asta mediante una sospensione elastica, è stato posizionato accanto ad un leggio che ha facilitato la lettura dei dialoghi e la dinamicità dell'interazione, riducendo eventuali rumori ambientali. È stato inoltre chiesto ai partecipanti di mantenere una distanza non inferiore ai 30 cm rispetto al microfono (cfr. Frank, Juslin, Harrigan, 2005; Anolli *et al.*, 2008). Un secondo leggio, destinato all'altro interlocutore (non dotato di microfono), è stato situato a circa 3 metri di distanza. Al fine di simulare una conversazione telefonica, si è ritenuto opportuno

---

<sup>59</sup> La strumentazione è stata offerta a titolo gratuito da EchoMobile Studio, Bisignano (Cs).

limitare il contatto visivo tra i due partecipanti, ai quali è stato chiesto di posizionarsi di spalle, affinché il contributo maggiore alla conversazione provenisse dal canale vocale<sup>60</sup>.

Ogni registrazione è stata preceduta da una fase di prove, durante la quale i partecipanti hanno avuto modo di familiarizzare con il testo e di proporre eventuali modifiche qualora la situazione comunicativa proposta non fosse in linea con il modo di comportarsi e le norme culturali di riferimento. Sia i parlanti italiani che quelli russi hanno dichiarato che la manifestazione delle emozioni negative (collera e tristezza) risultava più immediata e semplice, mentre la manifestazione della gioia è stata più difficoltosa per entrambi i gruppi, in particolare per gli uomini. Le donne, invece, non hanno mostrato alcuna titubanza o imbarazzo.

Tra una registrazione e l'altra è stata osservata una pausa di 10 minuti per consentire ai partecipanti di calarsi nella situazione comunicativa proposta. L'ordine di registrazione è stato il seguente:

- a) dialogo 1: tristezza
- b) dialogo 2: collera
- c) dialogo 3: gioia
- d) dialogo 4: neutro/normale

Il dialogo neutro è stato registrato per ultimo per dare modo ai partecipanti di sviluppare una maggiore consapevolezza riguardo all'assenza di attivazione emotiva che caratterizzava lo scambio di battute con il tecnico rispetto alle conversazioni precedenti. I partecipanti hanno avuto modo di modulare il proprio comportamento paraverbale in funzione della situazione comunicativa proposta. Per quanto riguarda gli apprendenti russi, i dialoghi sono stati interpretati prima nella loro lingua nativa e successivamente in italiano L2, allo scopo di introdurre gradualmente gli studenti in una situazione sperimentale che richiedeva avanzate capacità di gestione del parlato. In questo senso, si è ipotizzato che sarebbe stato molto più difficoltoso per gli apprendenti cimentarsi nell'interpretazione dei dialoghi emotivi ricorrendo alla L2, senza aver compreso le dinamiche della situazione sperimentale *in primis* attraverso il ricorso alla L1. Indubbiamente, l'ordine di registrazione adottato può aver influenzato le performance degli studenti, tuttavia, tra la prima sessione (in L1) e la seconda (in

---

<sup>60</sup> Gli apprendenti russi hanno ricevuto istruzioni relative alla preparazione e alla realizzazione dell'esperimento sia in italiano sia nella loro lingua nativa.

L2) è stata osservata una pausa maggiore (di circa 45 minuti), per dare agli apprendenti l'opportunità di prepararsi ad affrontare l'interazione dialogica in L2.

Complessivamente sono stati registrati 90 dialoghi di carattere emotivo, ai quali si sommano 30 produzioni neutre, per un totale di 120 dialoghi<sup>61</sup>.

#### 4.3.1.6 La selezione degli stimoli

I dati raccolti durante le sedute di registrazione sono stati sottoposti ad un processo di selezione e validazione. Per ogni dialogo, sono stati isolati<sup>62</sup> gli enunciati bersaglio ripetuti nelle diverse posizioni. Per verificare quale posizione (iniziale, mediana o finale) fosse maggiormente efficace sul piano comunicativo/emotivo e, soprattutto, quali fossero le produzioni più rappresentative in italiano L1 e in russo L1, è stato somministrato un test uditivo ad un gruppo di 10 uditori (5 di origine italiana e 5 di origine russa). Il test conteneva complessivamente 60 stimoli<sup>63</sup> (30 in italiano L1 e 30 in russo), selezionati in base alla posizione. Gli uditori sono stati invitati ad ascoltare di volta in volta il file audio<sup>64</sup> e a scegliere una tra le 7 opzioni proposte (collera, tristezza, gioia, sorpresa, paura, neutro, altro<sup>65</sup>). Gli enunciati in posizione iniziale hanno ottenuto globalmente una percentuale di riconoscimenti maggiore (65%) rispetto a quelli in posizione mediana e finale (cfr. FIG. 4-1), che al contrario sono stati oggetto di diverse interpretazioni. Per questa ragione si è ritenuto opportuno sottoporre all'analisi acustica e uditiva soltanto gli enunciati realizzati all'inizio del turno di parola. In relazione alle produzioni emotive realizzate dagli apprendenti in italiano (escluse dal test), sono state selezionate, anche in questo caso, le frasi in posizione iniziale.

---

<sup>61</sup> Tutti i partecipanti coinvolti nelle registrazioni hanno dato il proprio consenso al trattamento e alla diffusione dei dati in forma anonima e per fini scientifici.

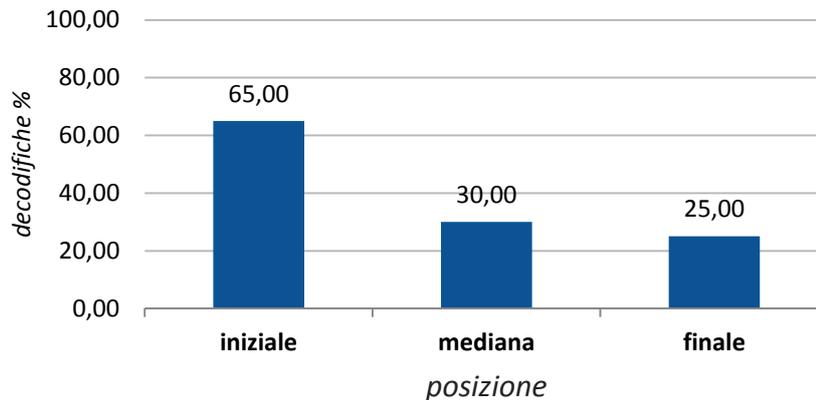
<sup>62</sup> La segmentazione del segnale è stata effettuata mediante Praat, un programma *open source*, che permette di individuare con precisione i punti di attacco e di fine del segnale selezionato (*zero crossing*) e di estrarre la porzione corrispondente all'enunciato. Cfr. <http://www.fon.hum.uva.nl/praat/>.

<sup>63</sup> I partecipanti hanno ascoltato gli stimoli attraverso una cuffia collegata ad un computer portatile. Poiché la batteria di stimoli era abbastanza consistente, si è ritenuto opportuno invitare i partecipanti ad osservare una pausa di 5 minuti durante la compilazione del test.

<sup>64</sup> Si è ritenuto opportuno inserire nel test una batteria di stimoli non eccessivamente ampia per evitare di influenzare negativamente le capacità di giudizio degli uditori.

<sup>65</sup> L'ultima opzione del test, "altro", dava ai partecipanti la possibilità di suggerire un'etichetta emozionale diversa tra quelle proposte.

**Fig-4-1- Media % decodifiche corrette degli enunciati emotivi i (prodotti sia in italiano L1 che in russo L1) in base alla posizione della frase nel turno di parola (iniziale, mediana, finale)**



#### 4.3.1.7 L'analisi acustica: parametri indagati

Complessivamente, sono state sottoposte ad analisi acustica 60 produzioni emotive e 20 neutre realizzate dagli apprendenti in italiano e in russo, e 30 produzioni emotive e 10 neutre realizzate dagli italiani nativi, per un totale di 120 enunciati. I parametri indagati sono stati suddivisi in tre macro-categorie (cfr. De Marco, Sorianello, Paone, in stampa; Sorianello, De Marco, in stampa):

- a) Ritmico-temporali<sup>66</sup>: ovvero durata totale dell'enunciato bersaglio (in millisecondi, da ora ms) e la velocità di articolazione, intesa come numero di sillabe al secondo (sill/sec), escluse le pause.
- b) Intonativi<sup>67</sup>: frequenza iniziale (*onset*), frequenza finale (*offset*), frequenza media (f0x), Escursione Melodica<sup>68</sup> dell'enunciato (da ora EM calcolata in semitoni, ST).
- c) Relativi all'ampiezza: ovvero rilievo dei valori medi di intensità<sup>69</sup> (espressa in decibel, dB).

L'analisi acustica è stata condotta mediante Praat. Le tipologie di eloquio analizzate sono principalmente due: italiano nativo e italiano non nativo, tuttavia, con la

<sup>66</sup> In via sperimentale, oltre alla durata totale dell'enunciato, è stata computata anche la durata della vocale tonica, sede di accento nucleare, nelle produzioni in Ita L1 ed Ita L2.

<sup>67</sup> I valori frequenziali sono stati convertiti in logaritmi naturali (nlog) per ridurre le differenze tra i locutori.

<sup>68</sup> La conversione in semitoni è stata effettuata adottando l'approccio di 't Hart *et al.* (1990).

<sup>69</sup> L'intensità costituisce il parametro acustico maggiormente sensibile ai condizionamenti esterni, *in primis* la distanza dalla fonte di registrazione. Tuttavia, è possibile anticipare che i dati raccolti mostrano una certa omogeneità in relazione a questo parametro in tutti e tre i corpora.

dovuta cautela, sono state oggetto di analisi acustica anche le produzioni emotive realizzate dagli apprendenti nella loro lingua nativa, come riportato nella TAB. 4-3.

TAB. 4-3. TIPOLOGIE DI ELOQUIO E RELATIVI ENCODER

		<i>Encoder</i>	<i>Lingua</i>
Corpus (A)	Ita L1	Italiani nativi	italiano
Corpus (B)	R-Ita L2	Apprendenti russi	italiano
Corpus (C):	R-Ru L1	Apprendenti russi	Russo

#### 4.3.1.8 Le analisi statistiche

Al fine di verificare in che misura le emozioni prese in esame incidessero sulla variazione dei parametri osservati all'interno di ogni corpus, è stato adottato il test statistico ANOVA per misure ripetute, considerando le 3 condizioni di collera, tristezza e gioia come variabili indipendenti e il parametro misurato come variabile dipendente.

In secondo luogo, poiché l'obiettivo principale del presente lavoro era quello di stabilire se esistessero differenze significative tra i due gruppi in relazione ai parametri sopradescritti, si è proceduto a confrontare in prima battuta la variazione dei parametri all'interno delle produzioni emotive di ogni gruppo rispetto a quelle neutre. Per ogni parametro, è stata computata la media aritmetica (M) e la deviazione standard (DS). Per stabilire se la variazione di ogni parametro fosse significativa rispetto all'enunciato neutro, è stata adottata una statistica t, nello specifico un test t a due code per dati appaiati (con un livello di significatività dello 0.05). Il valore del parametro osservato viene considerato come la variabile dipendente rispetto alle due condizioni esaminate, ovvero condizione 1 (neutro)- condizione 2 (emozione).

In seconda battuta, si è proceduto a un confronto incrociato tra i due gruppi e le tre tipologie di eloquio analizzate, ovvero:

1) Ita L1-R-Ita L2, al fine di stabilire quali parametri variassero in modo significativo nelle produzioni realizzate in italiano da nativi e non nativi.

2) Ita L1- R-Ru L1, allo scopo di stabilire le differenze tra il parlato emotivo in italiano e quello in russo.

3) R-Ita L2- R-Ru L1, al fine di cogliere affinità e discrepanze tra le produzioni emotive realizzate dagli apprendenti in italiano e quelle prodotte nella loro lingua nativa.

Per il confronto tra gruppi linguistici diversi, è stata adottata una procedura di normalizzazione dei dati (cfr. Hinde, Herrmann, 1977; Nasoz *et al.* 2003), precedentemente utilizzata in altri studi (cfr. Anolli *et al.* 2008, Sorianello, De Marco, in stampa) al fine di ridurre le differenze legate alla lingua dei locutori. Questa procedura permette di confrontare i valori di ciascuna emozione con quelli del neutro, che costituisce una sorta di *baseline* di riferimento ( $B_{(N)}$ ). Calcolando la differenza tra il valore assoluto rilevato per l'enunciato emotivo ( $X_{(E)}$ ) e il valore assoluto rilevato per la *baseline* neutra ( $B_{(N)}$ ), la normalizzazione consiste nel rapporto tra queste differenze e la *baseline* neutra (Anolli *et al.*, 2008, p. 577), come mostrato nell'esempio (i):

$$(i) \text{ valore normalizzato} = \frac{X(E) - B(N)}{N}$$

Questa procedura ha permesso di osservare all'interno delle produzioni emotive di ogni gruppo la variazione dei parametri rispetto al neutro e di operare un confronto tra i valori oggetto di normalizzazione.

Successivamente, al fine di verificare se e in che misura il parlato emotivo degli italiani differisse da quello prodotto dagli apprendenti in italiano L2, è stato utilizzato un test t di Student a due code per dati indipendenti (con un livello di significatività dello 0,05) per confrontare i valori dei parametri sopradescritti all'interno delle produzioni emotive dei due gruppi.

Per il confronto tra l'italiano L1 e il russo L1 è stata adottata la medesima statistica test, in quanto si trattava di campioni indipendenti.

Per confrontare le produzioni emotive realizzate dagli apprendenti in italiano e in russo, è stato utilizzato un test t per dati appaiati (in quanto si trattava dello stesso campione).

#### 4.3.2 *La decodifica delle emozioni: strumenti e procedure*

La seconda parte della ricerca si è focalizzata sulla verifica uditiva delle produzioni emotive (e neutre) oggetto di analisi. Questo passaggio ha implicato l'elaborazione e l'implementazione di due test percettivi destinati rispettivamente ad un gruppo di apprendenti russi e ad un gruppo di italiani nativi. La selezione degli stimoli è stata effettuata tenendo conto degli enunciati che avevano ottenuto una percentuale elevata di riconoscimenti nel test condotto in precedenza (cfr. PAR.4.4.5). Sono stati isolati 36 stimoli, 12 per ogni corpus (Ita L1, R-Ita L2, R-Ru L1). Per l'italiano nativo, sono stati selezionati 4 diversi encoder (2M e 2F), mentre per l'italiano non nativo e per

il russo sono state utilizzate le produzioni emotive di 3 apprendenti (2M e 1F). Entrambi i test sono stati elaborati e implementati attraverso GoogleModuli<sup>70</sup>, uno strumento *open source* che consente di creare e personalizzare questionari di diversa tipologia. La struttura è la medesima, ovvero un questionario a risposta multipla, tuttavia, il test rivolto agli apprendenti si componeva di un numero inferiore di stimoli rispetto al questionario dei nativi. Infatti, si è ritenuto opportuno escludere dal test rivolto agli apprendenti le produzioni realizzate in italiano L2, al fine di ridurre eventuali confusioni da parte degli uditori russi. Nei paragrafi successivi saranno illustrati gli strumenti e le procedure utilizzate per la realizzazione della verifica uditiva, nonché le caratteristiche dei due gruppi di uditori coinvolti.

#### 4.3.2.1 Il test uditivo degli apprendenti russi

Il test uditivo destinato agli apprendenti russi conteneva 24 stimoli, di cui 12 in italiano e 12 in russo. La pagina iniziale del test chiedeva ai partecipanti di inserire alcune informazioni anagrafiche, nonché alcune informazioni socio-linguistiche (cfr. TAB. 4-4), tra cui il livello di conoscenza dell'italiano e l'eventuale possesso di certificazioni di lingua. L'inserimento di questi ultimi due quesiti è stato necessario per stabilire una possibile correlazione tra il livello di competenza e la capacità di decodifica del parlato emotivo in L2.

TAB. 4-4. PAGINA INIZIALE DEL TEST RIVOLTO AGLI APPRENDENTI

INFORMAZIONI ANAGRAFICHE/SOCIOLINGUISTICHE
Genere
Età
Nazionalità
Lingua materna
Altre lingue parlate
Periodo di permanenza in Italia
Conoscenza della lingua italiana (sufficiente, buona, ottima)
Eventuali certificazioni di lingua italiana (A1-A2; B1-B2; C1-C2)

La sezione successiva del test conteneva alcune istruzioni che illustravano le modalità di svolgimento del questionario e il significato dei termini emotivi impiegati. Si è ritenuto opportuno mostrare attraverso un esempio concreto la procedura da

<sup>70</sup> <https://www.google.it/intl/it/forms/about/>

seguire, affinché i partecipanti familiarizzassero con il tipo di task richiesto. Per questa ragione, è stato inserito un file audio contenente una frase diversa da quella utilizzata nelle registrazioni. Il test chiedeva di riprodurre il file audio e individuare l'emozione veicolata dalla voce; le opzioni di risposta erano le seguenti: arrabbiato, felice, triste, spaventato, sorpreso, normale, altro (per un totale di 7). Per rendere più intuitiva la modalità di selezione, si è ritenuto opportuno sostituire le etichette emotive con degli aggettivi (tradotti sia in italiano che in russo) e associare ad ogni possibile risposta un'immagine raffigurante l'emozione corrispondente, come illustrato nella TAB. 4-5.

**TAB. 4-5 OPZIONI DI RISPOSTA DEL TEST Uditivo**

Opzioni di risposta (italiano/russo)	Emoticon
Felice (счастливы́й)	
Triste (грустны́й)	
Arrabbiato (сердиты́й)	
Spaventato (испуганны́й)	
Sorpreso (удивленны́й)	
Normale (нормальны́й)	
Altro...	

Invece dell'opzione "neutro", è stato inserito il termine "normale" per indicare un tipo di parlato privo di sfumature emozionali. È stata inoltre data la possibilità di esprimere un giudizio diverso da quelli proposti. L'opzione "altro" si è resa necessaria al fine di avere un riscontro concreto circa l'affidabilità del test, o meglio riguardo alla rappresentatività degli stimoli emotivi selezionati, come suggerito da Frank e Stennett (2001), cfr. CAP. 2, PAR. 2.4.2.1. Per la medesima ragione, le opzioni di risposta spaziavano tra diverse emozioni primarie, oltre a quelle effettivamente somministrate al giudizio degli uditori. In questo senso, le probabilità di un riconoscimento casuale erano basse, dato l'elevato numero di possibili risposte.

Infatti, il test rispecchiava le direttive suggerite da Juslin e Scherer (cfr. CAP. 2) in merito ai test di decodifica, presentando i seguenti punti di forza:

a) Tra le varie tipologie di test, quello a risposta multipla consente, in fase di analisi, un confronto diretto con i risultati riportati dalle ricerche che hanno adottato questo formato.

b) l'inserimento di un numero elevato di opzioni di risposta (7), garantisce una maggiore affidabilità del test.

c) La possibilità di esprimere un giudizio diverso da quelli proposti riduce inoltre la probabilità che i risultati del test vengano falsati semplicemente perché i partecipanti sono stati obbligati a scegliere una tra le etichette disponibili (cfr. Juslin, Scherer, 2008, p. 111).

Il test è stato suddiviso in 3 blocchi per evitare di appesantire il task con un carico eccessivo di file audio. Gli stimoli sono stati somministrati in modo casuale. Si è ritenuto opportuno non somministrare mai due stimoli realizzati dallo stesso locutore di seguito, per evitare che gli uditori potessero assuefarsi a quel tipo di eloquio (*learning effect*, cfr. Juslin, Scherer, 2008).

TAB. 4-6. STIMOLI SOMMINISTRATI PER TIPOLOGIA DI ENUNCIATO

<i>Stimoli</i>	Ita L1	Ru L1
Collera	3	3
Tristezza	3	3
Gioia	3	3
Neutro	3	3
Totale	12	12

Il test è stato distribuito a 100 studenti di origine russa ed è rimasto attivo per circa tre mesi. Tuttavia, il numero di responsi ottenuti è stato piuttosto basso (10) rispetto a quanto ipotizzato. Cionondimeno è stato possibile procedere ad un'analisi accurata dei dati raccolti; inoltre, gli apprendenti che hanno risposto al questionario possedevano un livello di competenza simile (B2) a quelli che hanno partecipato alle registrazioni. Per questa ragione, è stato possibile operare un confronto tra le abilità di produzione e di percezione avendo come riferimento questo specifico livello di competenza in L2.

I dati sono stati sottoposti ad analisi, tenendo conto dei seguenti punti:

- i. percentuali globali di decodifiche corrette

- ii. percentuali relative al riconoscimento degli stimoli in italiano
- iii. percentuali relative al riconoscimento degli stimoli in russo
- iv. percentuali di riconoscimento per tipologia emotiva in italiano e in russo
- v. presenza di errori ricorrenti nel riconoscimento delle emozioni

#### 4.3.2.2 Il test uditivo degli italiani

Il test rivolto agli italiani conteneva 36 stimoli (12 per ogni corpus). Rispetto al test dei russi, comprendeva infatti le produzioni realizzate dagli apprendenti in italiano L2. Per verificare se ad una maggiore conoscenza della cultura target corrispondesse una maggiore abilità nel veicolare uno stato emotivo attraverso le risorse prosodiche della lingua seconda, sono state inserite nel test le produzioni di 2 apprendenti di sesso maschile (da ora MR1 ed MR2) di livello B2, in Italia da circa 15 mesi, e di 1 apprendente di sesso femminile (da ora FR1) con un periodo di permanenza in Italia inferiore rispetto ai due uomini, ovvero 8 mesi e con un livello di competenza tra B1 e B2. Per ogni apprendente sono stati selezionati 4 enunciati (ovvero le tre emozioni e il neutro), come mostrato nella TAB. 4-6 che sintetizza anche le informazioni linguistiche degli *encoder*.

TAB. 4-7. INFORMAZIONI LINGUISTICHE RELATIVE AGLI ENCODER IN LINGUA RUSSA

<i>Encoder</i>	<i>Livello di Competenza</i>	<i>Periodo di Permanenza</i>	<i>N. Enunciati</i>
MRI	B2	15 mesi	4
MR2	B2	15 mesi	4
FR1	B1/B2	8 mesi	4

Come già accennato, la struttura del test uditivo degli italiani riflette in tutto e per tutto quella del test rivolto ai russi. La pagina iniziale chiedeva ai partecipanti di inserire informazioni anagrafiche e sociolinguistiche (età, genere, lingua materna, altre lingue parlate). La pagina successiva conteneva le istruzioni per lo svolgimento del test e l'esempio pilota per consentire ai partecipanti di familiarizzare con il task. Anche in questo caso, le opzioni di risposta erano 7 ed erano accompagnate da emoticon. Vista la mole di stimoli, si è proceduto a suddividere il test in 4 sezioni, in modo che fosse possibile osservare una pausa tra una sezione e l'altra.

Il test, distribuito via mail, è rimasto attivo per circa tre mesi. Il numero di risposte raccolte ammonta a 33. L'analisi dei dati ha inteso verificare i seguenti punti: 1

- i. percentuali globali di decodifiche corrette
- ii. percentuali relative al riconoscimento degli stimoli in italiano L1
- iii. percentuali relative al riconoscimento degli stimoli in italiano L2
- iv. percentuali di riconoscimento in italiano L2 in base al livello di competenza dei locutori
- v. percentuali relative al riconoscimento degli stimoli in russo
- vi. percentuali di riconoscimento per tipologia emotiva in italiano L1, italiano L2 e russo.
- vii. presenza di errori ricorrenti nel riconoscimento delle emozioni

#### **4.4. Conclusioni**

I metodi e gli strumenti illustrati in questo capitolo costituiscono il disegno sperimentale della presente ricerca. Nel CAP. 5, saranno illustrati i risultati relativi al processo di *encoding* e *decoding* delle emozioni oggetto di indagine.

# Capitolo 5

---

## Analisi e discussione dei dati

### 5.1 Introduzione

In questo capitolo, saranno illustrati i risultati relativi al processo di *encoding* e *decoding* delle produzioni emotive realizzate nell'italiano nativo (Ita L1), non nativo (R-Ita L2) e in russo (R-Ru L1). Il capitolo si compone di due sezioni, dedicate rispettivamente ai risultati dell'analisi acustica e a quelli dell'indagine uditiva.

### 5.2 *Encoding*: risultati

Al fine di comprendere pienamente le differenze osservate nelle produzioni oggetto di esame, si procederà, in primo luogo, alla descrizione dei parametri significativi all'interno di ogni corpus. Successivamente, saranno riportati i risultati relativi al confronto tra i corpora: a) Ita L1-Ita L2; b) Ita L1- Ru L1; c) Ita L2- Ru L1.

#### 5.2.1 *La variazione dei parametri all'interno di ogni corpus*

La TAB. 5-1 riporta, per ogni parametro osservato, la media (M) e la deviazione standard (DS) all'interno dei tre corpora in funzione della tipologia di enunciato.

**TAB. 5--1 PROSPETTO DEI PARAMETRI OSSERVATI  
ALL'INTERNO DI CIASCUN CORPUS IN FUNZIONE DELLA TIPOLOGIA DI ENUNCIATO**

Parametri	Corpus	Tipologia enunciato (E)							
		Collera		Tristezza		Gioia		Neutro	
		M	DS	M	DS	M	DS	M	DS
Dur. Tot.	Ita L1	982,100	108,870	816,000	128,434	1014,200	159,341	821,900	129,070
	Ita L2	1079,200	152,947	1011,500	136,668	1071,100	110,980	935,300	187,401
	Ru L1	952,400	117,490	889,300	143,422	915,300	99,034	753,400	145,506
VdA	Ita L1	6,182	0,736	7,516	1,172	6,054	0,938	7,488	1,336
	Ita L2	5,661	0,804	6,029	0,803	5,655	0,575	6,639	1,264
	Ru L1	5,327	0,702	5,763	0,975	5,520	0,587	26,834	1,160
Int. (M)	Ita L1	68,900	5,174	58,600	5,777	71,100	5,587	61,6	4,402
	Ita L2	71,900	3,665	62,400	2,989	72,500	5,126	63,200	6,286
	Ru L1	68,900	3,635	58,200	3,084	69,300	7,973	60,900	4,886
<i>Onset</i>	Ita L1	5,181	0,403	5,033	0,323	5,257	0,471	5,217	0,457
	Ita L2	4,953	0,309	4,810	0,255	5,032	0,361	4,923	0,340
	Ru L1	5,101	0,408	4,824	0,395	4,994	0,415	4,933	0,411
<i>Offset</i>	Ita L1	4,955	0,398	4,929	0,455	4,974	0,419	4,970	0,360
	Ita L2	4,766	0,348	4,668	0,220	4,796	0,392	4,586	0,217
	Ru L1	4,785	0,280	4,791	0,380	4,957	0,324	4,768	0,346
F0 (M)	Ita L1	5,248	0,363	5,057	0,354	5,408	0,302	5,178	0,311
	Ita L2	5,018	0,322	4,840	0,382	5,116	0,442	4,836	0,303
	Ru L1	5,032	0,302	4,847	0,337	5,209	0,422	4,897	0,390
EM	Ita L1	9,172	1,062	7,523	4,486	13,970	5,587	8,581	1,529
	Ita L2	6,850	2,753	6,723	4,938	8,979	3,544	8,049	4,255
	Ru L1	10,798	5,312	6,862	4,875	11,928	4,583	6,540	2,597
Dur. VT	Ita L1	181,000	27,909	146,700	43,889	208,900	57,241	151,300	51,041
	Ita L2	184,500	41,594	190,400	30,913	196,200	41,678	153,700	51,290

Note: M=media; DS= deviazione standard; Dur. Tot.= durata totale; VdA= *VdAate*; Int. (M)= intensità media; *Onset*= frequenza iniziale; *Offset* (M)= frequenza finale; F0 (M)= frequenza fondamentale media (espressa in semitoni); EM= escursione melodica (espressa in semitoni); Dur. VT= durata della vocale tonica.

TAB. 5-2. RISULTATI TEST ANOVA IN RELAZIONE AL FATTORE EMOZIONE ALL'INTERNO DEI TRE CORPORA

Parametri	Corpus					
	Ita L1		Ita L2		Ru L1	
	F (2,27)	P	F (2,27)	P	F (2,27)	p
<i>Dur. Tot.</i>	6,3	<b>0,005*</b>	0,7	0,48	0,6	0,51
<i>Dur. VT</i>	4,5	<b>0,01*</b>	0,2	0,79		
<i>VdA</i>	6,7	<b>0,004*</b>	0,8	0,43	0,7	0,46
<i>Int.</i>	14	<b>0,0004*</b>	19,7	<b>0,000005*</b>	13,7	<b>0,00007*</b>
<i>Onset</i>	0,7	0,46	1,3	0,28	1,1	0,32
<i>Offset</i>	0,02	0,97	0,4	0,66	0,8	0,43
<i>F0 (M)</i>	2,6	0,08	1,3	0,28	2,5	0,09
<i>EM</i>	6,4	<b>0,005*</b>	1,08	0,35	2,9	0,07

\*=risultati significativi,  $p < 0,05$ .

Per stabilire quali parametri variassero in modo significativo all'interno di ogni corpus, sono stati condotti tre serie di test ANOVA per misure ripetute (con un livello di significatività dello 0,05), considerando le 3 condizioni di collera, tristezza e gioia come variabili indipendenti e il parametro misurato come variabile dipendente. Se si osservano i risultati riportati nella TAB. 5.2, si noterà una certa discrepanza tra le produzioni realizzate dagli italiani e quelle realizzate dagli apprendenti russi (in entrambe le lingue). Nell'italiano nativo, il fattore "emozione" influenza in modo significativo la variazione dei parametri temporali (durata totale, durata della vocale tonica e velocità di articolazione) e quelli relativi all'intensità ( $p < 0,05$ ). In relazione ai parametri intonativi, soltanto l'escursione melodica varia in modo significativo ( $p < 0,05$ ), mentre *onset*, *offset* e *f0* media non sono statisticamente significativi ( $p > 0,05$ ).

La situazione cambia se si osservano i risultati riportati per Ita L2 e Ru L1: in entrambi i casi, l'unico parametro a variare in modo significativo è l'intensità ( $p < 0,05$ ).

### 5.2.2 Le produzioni emotive nell'italiano nativo (Ita\_L1)

Prendiamo in esame le produzioni emotive realizzate nell'italiano nativo, confrontando le variazioni dei parametri indagati rispetto all'eloquio neutro. La TAB. 5-3 riporta i risultati del test t e il livello di significatività per ciascun parametro.

TAB. 5-3 RISULTATI T TEST ITA L1: CONFRONTO NEUTRO-EMOZIONE

		Tipologia enunciato (E)						Test t (IC 95%)		
		Neutro			Gioia			T	gdl	P(T<=t)
		M	DS	N.	M	DS	N.			
<b>Dur. Tot.</b>	Ita L1	821,9	129,07	10	1014,20	159,34	10	5,1	9	<b>0,0006*</b>
<b>VdA</b>	Ita L1	7,488	1,336	10	6,054	0,938	10	5,1	9	<b>0,0006*</b>
<b>Int. (M)</b>	Ita L1	61,6	4,402	10	71,100	5,587	10	9,08	9	<b>0,000007*</b>
<i>Onset</i>	Ita L1	5,217	0,457	10	5,257	0,471	10	0,7	9	0,49
<i>Offset</i>	Ita L1	4,970	0,360	10	4,974	0,419	10	0,1	9	0,88
<b>F0 (M)</b>	Ita L1	5,178	0,311	10	5,408	0,302	10	3,4	9	<b>0,007*</b>
<b>EM</b>	Ita L1	8,581	1,529	10	13,970	5,587	10	2,6	9	<b>0,02*</b>
<b>Dur. VT</b>	Ita L1	151,300	51,041	10	208,9	57,241	10	2,7	9	<b>0,02*</b>

		Tipologia enunciato (E)						Test t (IC 95%)		
		Neutro			Tristezza			T	Gdl	P(T<=t)
		M	DS	N.	M	DS	N.			
Dur. Tot.	Ita L1	821,900	129,070	10	816,000	128,434	10	0,15	9	0,88
VdA	Ita L1	7,488	1,336	10	7,516	1,172	10	0,07	9	0,94
Int. (M)	Ita L1	61,6	4,402	10	58,600	5,777	10	2,2	9	0,05
<i>Onset</i>	Ita L1	5,217	0,457	10	5,033	0,323	10	1,9	9	0,07
<i>Offset</i>	Ita L1	4,970	0,360	10	4,929	0,455	10	0,7	9	0,45
<b>F0 (M)</b>	Ita L1	5,178	0,311	10	5,057	0,354	10	2,5	9	<b>0,03*</b>
EM	Ita L1	8,581	1,529	10	7,523	4,486	10	0,7	9	0,44
Dur. VT	Ita L1	151,300	51,041	10	146,700	43,889	10	0,3	9	0,72

		Tipologia enunciato (E)						Test t (IC 95%)		
		Neutro			Collera			T	gdl	P(T<=t)
		M	DS	N.	M	DS	N.			
<b>Dur. Tot.</b>	Ita L1	821,900	129,070	10	982,100	108,870	10	5,7	9	<b>0,0002*</b>
<b>VdA</b>	Ita L1	7,488	1,336	10	6,182	0,736	10	4,1	9	<b>0,002*</b>
<b>Int. (M)</b>	Ita L1	61,6	4,402	10	68,900	5,174	10	6,6	9	<b>0,0009*</b>
<i>Onset</i>	Ita L1	5,217	0,457	10	5,181	0,403	10	1,1	9	0,28
<i>Offset</i>	Ita L1	4,970	0,360	10	4,955	0,398	10	0,4	9	0,63
F0 (M)	Ita L1	5,178	0,311	10	5,248	0,363	10	2,1	9	0,06
EM	Ita L1	8,581	1,529	10	9,172	1,062	10	1,4	9	0,17
Dur. VT	Ita L1	151,300	51,041	10	181,000	27,909	10	2,06	9	0,06

\*= risultati statisticamente significativi: **p<0,05**

### *Collera*

#### a) Parametri temporali

Dal confronto delle medie, si evince che la *collera* (M=982,1 ms) ha una durata maggiore rispetto al *neutro* (M=821,9 ms) e viene realizzata con una velocità di articolazione inferiore (rispettivamente 6,1 vs. 7,4 sill/sec). In entrambi i casi, la differenza è statisticamente significativa (vd. TAB. 5-2). La durata della vocale tonica si allunga negli enunciati di *collera* (181 ms vs. 151,3 ms), tuttavia, questa differenza non è significativa ( $p > 0,05$ ).

#### b) Parametri frequenziali

In relazione ai parametri frequenziali, i risultati del t test non hanno riscontrato differenze significative ( $p > 0,05$ ). Dal confronto delle medie si evince che nelle produzioni di *collera* l'escursione melodica è più ampia (M=9,1 st; DS=1) rispetto al *neutro* (M=8,5 st, DS=1), tuttavia questa differenza non è supportata da significatività statistica. La  $f_0$  media si colloca intorno ai 5,2 st in entrambe le produzioni. *Onset* e *offset* non sembrano essere due parametri discriminanti. *L'onset* si colloca, in tutte e due i casi sullo stesso livello della  $f_0$  media, mentre *l'offset* è più basso in entrambe le produzioni (intorno ai 4,9 st).

#### c) Intensità

Per quanto riguarda l'intensità, i valori subiscono in media un brusco incremento nelle produzioni di *collera* rispetto al *neutro* (M=68,9 vs. M=61,6 dB). La differenza tra le due produzioni è statisticamente significativa ( $p < 0,05$ ).

### *Gioia*

#### a) Parametri temporali

Nelle produzioni analizzate, la *gioia* ha una durata maggiore rispetto al *neutro* (M=1014,1 vs. M=821,3) e viene realizzata con una velocità di articolazione più lenta (M=6 vs. M=7,4 sill/sec). Si verifica, inoltre, un consistente allungamento della vocale tonica (M=208,9 vs. M=151,3 ms). Dai risultati del test t, si evince che in tutti i casi osservati queste differenze sono significative ( $p < 0,05$ ).

#### b) Parametri frequenziali

In relazione ai parametri frequenziali, dal confronto delle medie si evince che nelle produzioni di *gioia* il *pitch* sia in media più alto (M=5,4 st) rispetto alle produzioni neutre (M=5,1), la differenza è significativa ( $p < 0,05$ ). L'escursione melodica della *gioia* è molto ampia (M=13,9 st), mentre il *neutro* viene realizzato con un *range* più ristretto (M=8,5 st), tuttavia i risultati del test t non hanno restituito evidenze significative

( $p > 0,05$ ). Nella gioia, l'*onset* è più basso rispetto alla  $f_0$  media (5,2 st), mentre nell'eloquio neutro l'*onset* si colloca più o meno sullo stesso livello della  $f_0$  (5,2 st). In entrambe le produzioni l'*offset* è più basso rispetto alla  $f_0$  media e si colloca intorno ai 4,9 st.

### c) Intensità

In relazione all'intensità, i valori subiscono in media un brusco incremento nelle produzioni di gioia rispetto al neutro ( $M=71,1$  vs.  $M=61,6$ ). La differenza tra le due produzioni è statisticamente significativa ( $p < 0,05$ ).

### *Tristezza*

#### a) Parametri temporali

Sul piano ritmico-temporale, le produzioni di tristezza non variano in modo significativo rispetto al neutro ( $p > 0,05$ ). L'estensione temporale dell'enunciato è abbastanza simile (neutro=821,9 ms; tristezza= 816 ms), così come la durata della vocale tonica (151 ms vs. 146 ms). La velocità di articolazione non subisce variazioni importanti, ma mantiene un ritmo di circa 7,5 sill/sec.

#### a) Parametri frequenziali

In relazione ai parametri frequenziali, la media di  $f_0$  è l'unico parametro a variare in modo significativo ( $p < 0,05$ ). Dal confronto con le medie si evince che il *pitch* è leggermente più basso nelle produzioni di tristezza ( $M=5$  st) rispetto ai valori riportati per eloquio neutro ( $M=5,1$  st). In modo analogo, l'escursione melodica è più ristretta (7,5 st) rispetto al neutro (8,5 st), tuttavia questa differenza non è supportata dai risultati delle analisi statistiche ( $p > 0,05$ ). *Offset* e *onset* non variano in modo significativo, in quanto si mantengono più o meno sullo stesso livello rispetto al neutro. Nelle produzioni di tristezza l'*onset* è in linea con la  $f_0$  media (intorno ai 5 st), mentre l'*offset* risulta inferiore (4,9 st).

### c) Intensità

Per quanto riguarda l'intensità, nella tristezza i valori tendono ad abbassarsi leggermente rispetto al neutro (58,6 vs. 61,6 db), anche se questa differenza non è statisticamente significativa ( $p > 0,5$ ).

Confrontando le produzioni emotive degli italiani, è possibile intuire quali parametri contribuiscano a discriminare con più precisione uno stato emotivo dall'altro. Ad esempio, sul piano temporale sono emerse ulteriori differenze tra le tipologie di enunciato analizzate che permettono di osservare il modo in cui si comportano i

parametri sotto esame in funzione dell'emozione veicolata. Ad esempio, sul piano temporale, la tristezza viene realizzata con un'estensione temporale minore e una VdA più elevata rispetto alle altre produzioni: la differenza con la collera e la gioia è significativa ( $p < 0,05$ ).

Sul piano frequenziale, se si confrontano i valori di  $f_0$  e dell'EM della gioia con quelli riportati per le altre produzioni emotive, si noterà che la gioia viene realizzata con una  $f_0$  più elevata e con una EM notevolmente più ampia, sia rispetto alla collera sia rispetto alla tristezza ( $p < 0,05$ ).

Infine, per quanto concerne l'intensità, questo parametro si comporta in modo speculare, contribuendo a differenziare le produzioni di gioia (che presentano i valori più alti in assoluto) da quelle di tristezza, che presentano invece valori significativamente più bassi rispetto alle altre emozioni.

### 5.2.3 Le produzioni emotive nell'italiano non nativo (Ita\_L2)

TAB. 5-4. RISULTATI TEST ITA L2: CONFRONTO NEUTRO-EMOZIONE

		Tipologia enunciato (E)								
		Neutro			Collera			Test t (IC 95%)		
		M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
<b>Dur. Tot.</b>	Ita L2	935,300	187,401	10	1079,200	152,947	10	4,2	9	<b>0,002*</b>
<b>VdA</b>	Ita L2	6,639	1,264	10	5,661	0,804	10	4,5	9	<b>0,001*</b>
<b>Int. (M)</b>	Ita L2	63,200	6,286	10	71,900	3,665	10	6,6	9	<b>0,00009*</b>
<i>Onset</i>	Ita L2	4,923	0,340	10	4,953	0,309	10	0,4	9	0,06
<i>Offset</i>	Ita L2	4,586	0,217	10	4,766	0,348	10	1,8	9	0,09
<b>F0 (M)</b>	Ita L2	4,836	0,303	10	5,018	0,322	10	4,1	9	<b>0,002*</b>
EM	Ita L2	8,049	4,255	10	6,850	2,753	10	0,7	9	0,49
Dur. VT	Ita L2	153,700	51,290	10	184,500	41,594	10	1,9	9	0,08

		Tipologia enunciato (E)								
		Neutro			Gioia			Test t (IC 95%)		
		M	DS	N.	M	DS	N.	t	gdl	P(T<=t)
<b>Dur. Tot.</b>	Ita L2	935,300	187,401	10	1071,100	110,980	10	3,6	9	<b>0,005*</b>
<b>VdA</b>	Ita L2	6,639	1,264	10	5,655	0,575	10	3,6	9	<b>0,005*</b>
<b>Int. (M)</b>	Ita L2	63,200	6,286	10	72,500	5,126	10	5,2	9	<b>0,005*</b>
<i>Onset</i>	Ita L2	4,923	0,340	10	5,032	0,361	10	1,3	9	0,20
<i>Offset</i>	Ita L2	4,586	0,217	10	4,796	0,392	10	2,09	9	0,06
<b>F0 (M)</b>	Ita L2	4,836	0,303	10	5,116	0,442	10	4,04	9	<b>0,002*</b>
EM	Ita L2	8,049	4,255	10	8,979	3,544	10	0,5	9	0,60
<b>Dur. VT</b>	Ita L2	153,700	51,290	10	196,200	41,678	10	2,3	9	<b>0,04*</b>

		Tipologia enunciato (E)								
		neutro			Tristezza			Test t (IC 95%)		
		M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
Dur. Tot.	Ita L2	935,300	187,401	10	1011,500	136,668	10	2,08	9	0,06
<b>VdA</b>	Ita L2	6,639	1,264	10	6,029	0,803	10	2,5	9	<b>0,03*</b>
Int. (M)	Ita L2	63,200	6,286	10	62,400	2,989	10	0,3	9	0,72
<i>Onset</i>	Ita L2	4,923	0,340	10	4,810	0,255	10	1,4	9	0,17
<i>Offset</i>	Ita L2	4,586	0,217	10	4,668	0,220	10	2,3	9	<b>0,04*</b>
F0 (M)	Ita L2	4,836	0,303	10	4,840	0,382	10	0,09	9	0,92
EM	Ita L2	8,049	4,255	10	6,723	4,938	10	1,3	9	0,21
Dur. VT	Ita L2	153,700	51,290	10	190,400	30,913	10	2,2	9	0,05

\* = risultati statisticamente significativi:  $p < 0,05$

### Collera

#### a) Parametri temporali

Nelle produzioni realizzate dagli apprendenti in italiano, durata e velocità di articolazione variano in modo significativo rispetto all'eloquio neutro ( $p < 0,05$ ). Se si confrontano le medie, si osserva che la collera ( $M=1079,2$  ms) ha una durata maggiore rispetto al neutro ( $M=935,3$  ms) e viene realizzata, in media, con una velocità di articolazione inferiore ( $M=5,6$  vs.  $M=6,6$ ). La durata della vocale tonica non varia in modo significativo ( $p > 0,05$ ), sebbene negli enunciati di collera è possibile notare un lieve allungamento della vocale tonica ( $153,7$  vs  $184,5$  ms).

#### a) Parametri frequenziali

In relazione ai parametri frequenziali, soltanto la  $f_0$  media differisce in modo ( $p < 0,05$ ). Dal confronto delle medie si evince che nelle produzioni di collera il *pitch* è leggermente più alto ( $M=5$  st) rispetto alle produzioni neutre ( $M=4,8$ ). Per quanto riguarda l'escursione melodica, le produzioni di collera sono realizzate con un *range* più ristretto ( $M=6,8$  st) rispetto al neutro ( $M=8$  st), tuttavia questa differenza non è statisticamente significativa ( $p > 0,05$ ). *Onset* e *offset* non variano in modo significativo: nell'eloquio neutro, i valori di *onset* ( $4,9$  st) risultano di poco più alti rispetto alla  $f_0$  media ( $4,8$  st), mentre l'*offset* è più basso ( $4,5$ st); nelle produzioni di collera, l'*onset* ( $4,9$  st) è più basso rispetto alla  $f_0$  media ( $5$  st); i valori si abbassano ulteriormente in prossimità dell'*offset* ( $4,7$  st).

### c) Intensità

Infine per quanto riguarda l'intensità, i valori subiscono in media un brusco incremento nelle produzioni di collera rispetto al neutro (M=71,9 vs. M=63,2). La differenza tra le due produzioni è statisticamente significativa ( $p < 0,05$ ).

### *Gioia*

#### a) Parametri temporali

In relazione ai parametri temporali, durata e velocità di articolazione variano in modo significativo ( $p < 0,05$ ). Dal confronto delle medie si evince che la gioia ha una durata maggiore rispetto al neutro (M=1071,1 vs. M=935,3) e viene realizzata, in media, con una velocità di articolazione inferiore (M=5 vs. M=6,6). Anche la durata della vocale tonica è più lunga negli enunciati di gioia (M=196,2 vs. M=153,7 ms) e la differenza è significativa ( $p < 0,05$ ).

#### b) Parametri frequenziali

In relazione ai parametri frequenziali, soltanto la  $f_0$  media differisce in modo significativo nelle produzioni di gioia rispetto a quelle neutre ( $p < 0,05$ ). Dal confronto delle medie si evince che nelle produzioni di gioia il *pitch* è in media più alto (M=5,1 st) rispetto all'eloquio neutro (M=4,8). Per quanto riguarda gli altri parametri, si può notare come l'escursione melodica della gioia (M=8,9 st) sia più ampia rispetto a quella del neutro (M=8 st), tuttavia questa differenza non è statisticamente significativa ( $p > 0,05$ ). Per quanto riguarda l'*onset*, nella gioia raggiunge in media i 5 st (leggermente più basso rispetto alla  $f_0$  media), mentre l'*offset* si colloca intorno ai 4,7 st. Nel confronto con l'eloquio neutro, non sono emerse differenze significative ( $p > 0,05$ ).

### c) Intensità

In relazione all'intensità, si può notare come i valori subiscano un brusco incremento nelle produzioni di gioia rispetto al neutro (M=72,5 vs. M=63,2). La differenza tra le due produzioni in termini di intensità è statisticamente significativa ( $p < 0,05$ ).

### *Tristezza*

#### a) Parametri temporali

Dai risultati del t test è emerso che in relazione ai parametri temporali, soltanto la velocità di articolazione differisce significativamente nelle due produzioni ( $p < 0,05$ ). La tristezza viene realizzata, in media, più lentamente rispetto al neutro (M=6 vs. M=6,6). La durata totale dell'enunciato è pari a 1011 ms, 76 ms in più rispetto

all'eloquio normale, tuttavia questa differenza non è statisticamente significativa ( $p < 0,05$ ). In modo analogo, la durata della vocale tonica risulta più lunga (190 ms), ma lo scarto non è significativo.

#### b) Parametri frequenziali

In relazione ai parametri frequenziali, la  $f_0x$  rimane stabile in entrambe le produzioni (intorno ai 4,8 st). L'escursione melodica è più ristretta (6,7 st) rispetto al neutro (8 st), tuttavia questa differenza non è significativa ( $p > 0,05$ ). Negli enunciati di tristezza i valori di *onset* sono simili a quelli della  $f_0$  media (4,8 st), mentre l'*offset* è in più basso (4,6 st). Nel confronto con l'eloquio neutro, l'unico parametro a variare in modo significativo è l'*offset*, che risulta leggermente più alto negli enunciati di tristezza (4,6 vs. 4,5 st).

#### c) Intensità

Per quanto riguarda l'intensità, i valori risultano leggermente più bassi rispetto all'eloquio neutro (62,4 vs. 63,2 db), tuttavia questa differenza non è significativa ( $p > 0,5$ ).

Confrontando le produzioni emotive degli apprendenti, sono emerse ulteriori differenze tra le tipologie di enunciato analizzate che permettono di osservare il modo in cui si comportano i parametri sotto esame in funzione dell'emozione veicolata. Ad esempio, sul piano temporale, la collera ha una durata maggiore rispetto alla tristezza e viene realizzata più lentamente (la differenza tra le medie è significativa,  $p < 0,05$ ). Sul piano frequenziale, la  $f_0$  media contribuisce a differenziare la tristezza dalla collera: infatti, i valori della prima risultano significativamente più bassi rispetto alla seconda (4,8 st vs. 5 st). L'escursione melodica invece appare discriminante soltanto nel confronto tra la gioia (8,9 st) e la collera (6,8 st): la differenza tra le medie è statisticamente significativa ( $p < 0,05$ ). Per quanto concerne l'intensità, questo parametro raggiunge valori significativamente più bassi negli enunciati di tristezza rispetto quelli di collera e gioia (in entrambi i casi,  $p < 0,05$ ).

### 5.2.4 Le produzioni emotive in russo L1 (Ru\_L1)

**TAB. 5-5. RISULTATI T TEST: CONFRONTO NEUTRO-EMOZIONE**

		Tipologia enunciato (E)								
		Neutro			Collera			Test t (IC 95%)		
		M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)
<b>Dur. Tot.</b>	Ru L1	753,400	145,506	10	952,400	117,490	10	6,07	9	<b>0,0001*</b>
<b>VdA</b>	Ru L1	6,834	1,160	10	5,327	0,702	10	6,2	9	<b>0,0001*</b>
<b>Int. (M)</b>	Ru L1	60,900	4,886	10	68,900	3,635	10	7,8	9	<b>0,00002*</b>
<b>Onset</b>	Ru L1	4,933	0,411	10	5,101	0,408	10	2,7	9	<b>0,02*</b>
<b>Offset</b>	Ru L1	4,768	0,346	10	4,785	0,280	10	0,3	9	0,76
<b>F0 (M)</b>	Ru L1	4,897	0,390	10	5,032	0,302	10	2,9	9	<b>0,01*</b>
<b>EM</b>	Ru L1	6,540	2,597	10	10,798	5,312	10	2,6	9	<b>0,02*</b>

		Tipologia enunciato (E)								
		Neutro			Gioia			Test t (IC 95%)		
		M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)
<b>Dur. Tot.</b>	Ru L1	753,400	145,506	10	915,300	99,034	10	4,3	9	<b>0,001*</b>
<b>VdA</b>	Ru L1	6,834	1,160	10	5,520	0,587	10	4,3	9	<b>0,001*</b>
<b>Int. (M)</b>	Ru L1	60,900	4,886	10	69,300	7,973	10	3,6	9	<b>0,004*</b>
<b>Onset</b>	Ru L1	4,933	0,411	10	4,994	0,415	10	1,2	9	0,24
<b>Offset</b>	Ru L1	4,768	0,346	10	4,957	0,324	10	2,2	9	0,05
<b>F0 (M)</b>	Ru L1	4,897	0,390	10	5,209	0,422	10	7,9	9	<b>0,00002*</b>
<b>EM</b>	Ru L1	6,540	2,597	10	11,928	4,583	10	3,6	9	<b>0,005*</b>

		Tipologia enunciato (E)								
		Neutro			Tristezza			Test t (IC 95%)		
		M	DS	N.	M	DS	N.	t	gdl	P(T<=t)
<b>Dur. Tot</b>	Ru L1	753,400	145,506	10	889,300	143,422	10	4,8	9	<b>0,0009*</b>
<b>VdA</b>	Ru L1	6,834	1,160	10	5,763	0,975	10	5,6	9	<b>0,0003*</b>
<b>Int. (M)</b>	Ru L1	60,900	4,886	10	58,200	3,084	10	1,9	9	0,08
<b>Onset</b>	Ru L1	4,933	0,411	10	4,824	0,395	10	2,3	9	<b>0,04*</b>
<b>Offset</b>	Ru L1	4,768	0,346	10	4,791	0,380	10	0,5	9	0,57
<b>F0 (M)</b>	Ru L1	4,897	0,390	10	4,847	0,337	10	1,2	9	0,24
<b>EM</b>	Ru L1	6,540	2,597	10	6,862	4,875	10	0,2	9	0,82

Note: \* = risultati statisticamente significativi: **p<0,05**

*Collera*

#### a) Parametri temporali<sup>71</sup>

Le produzioni di collera in lingua russa hanno in media una durata maggiore (M=952,1 ms) rispetto all'eloquio neutro (M=753,4 ms) e sono realizzate con una velocità di articolazione più lenta (M=5,3 vs. 6,8 sill/sec). I risultati del t test (cfr. TAB. 5-5) confermano che, in entrambi i casi, la differenza tra le medie è statisticamente significativa ( $p<0,05$ ).

#### b) Parametri frequenziali

In relazione ai parametri frequenziali, dal confronto delle medie si evince che nelle produzioni di collera l'escursione melodica è più ampia (M=5 st) rispetto al neutro (4,8st), la differenza tra le due produzioni è statisticamente significativa ( $p<0,05$ ). Anche in relazione all'escursione melodica, i risultati hanno evidenziato una differenza significativa, infatti nelle produzioni di collera, il *range* è in media più ampio (10,7 st) rispetto al neutro (6,5 st). In entrambe le produzioni l'*onset* è più alto rispetto alla f0 media, tuttavia nella collera i valori sono in media più elevati rispetto all'eloquio neutro (M=5,1 st vs. M=4,9 st): la differenza tra le medie è significativa ( $p<0,05$ ). L'*offset* si comporta in modo analogo, ovvero presenta valori più bassi rispetto alla f0 media, collocandosi intorno ai 4,7 st in entrambe le produzioni.

#### c) Intensità

Infine per quanto riguarda l'intensità, i valori subiscono in media un brusco incremento nelle produzioni di collera rispetto al neutro (M=68,9 vs. M=60,9 dB). La differenza tra le due produzioni è statisticamente significativa ( $p<0,05$ ).

### *Gioia*

#### a) Parametri temporali

Le produzioni di gioia hanno un'estensione temporale maggiore (M=915,3 ms) rispetto al neutro (M=753,4 ms) e sono realizzate più lentamente (M=5,5 vs. M=6,8 sill/sec). In entrambi i casi, la differenza è significativa ( $p<0,05$ ).

#### b) Parametri frequenziali

In relazione ai parametri frequenziali, la f0 media differisce in modo significativo nelle produzioni di gioia rispetto a quelle neutre ( $p<0,05$ ). Dal confronto delle medie si evince che nelle produzioni di gioia il *pitch* è in media più alto (M=5,2 st) rispetto alle produzioni neutre (M=4,8). Per quanto riguarda l'escursione melodica, le

---

<sup>71</sup> La sillabificazione degli enunciati russi è stata effettuata con una certa cautela, considerando l'esiguo numero di riferimenti circa i criteri di sillabificazione della lingua russa (cfr. Halmilton, 1980).

produzioni di gioia sono realizzate con un *range* notevolmente più ampio (M=11,9 st) rispetto al neutro (M=6,5 st), questa differenza è statisticamente significativa ( $p>0,05$ ). *Onset* e *offset* questi non sembrano essere discriminanti, tuttavia nella gioia i valori di *onset* (M=4,9) e *offset* (M=4,9) sono inferiori rispetto alla f0 media.

#### c) Intensità

Infine per quanto riguarda l'intensità, i valori subiscono in media un brusco incremento nelle produzioni di gioia rispetto al neutro (M=69,3 vs. M=60,9). La differenza tra le due produzioni è statisticamente significativa ( $p<0,05$ ).

#### *Tristezza*

##### a) Parametri temporali

Dai risultati del t test è emerso che in relazione ai parametri temporali, le differenze rispetto al neutro sono significative ( $p<0,05$ ). La durata media della tristezza è maggiore rispetto a quella del neutro (rispettivamente 889,3 ms vs. 753,4 ms), la velocità di articolazione diminuisce (5,7 sill/sec vs. 6,8 sill/sec).

##### b) Parametri frequenziali

In relazione ai parametri frequenziali, non sono emerse differenze sostanziali per quanto riguarda la media di f0, che si mantiene sullo stesso livello in entrambe le produzioni, intorno ai 4,8 st, né per quanto concerne l'escursione melodica che varia molto poco nelle produzioni di tristezza (6,8 st) rispetto a quelle neutre (6,5 st). L'*onset* è l'unico parametro a variare in modo significativo ( $p<0,05$ ): se si osservano le medie, si noterà che nelle produzioni di tristezza i valori sono più bassi (4,8 st vs. 4,9st) e si mantengono sulla stessa fascia frequenziale della f0 media. abbassandosi leggermente nelle produzioni di tristezza (4,8 st) rispetto a quelle neutre (4,9 st). L'*offset* invece scende fino a 4,7 st in entrambe le produzioni.

#### c) Intensità

Per quanto riguarda l'intensità, nella tristezza i valori tendono ad abbassarsi leggermente rispetto al neutro (58,2 vs. 60,9 db), tuttavia questa differenza non è statisticamente significativa ( $p>0,5$ ).

Confrontando le produzioni emotive in Ru L1, emergono alcune differenze nella variazione dei parametri osservati. Ad esempio, in relazione ai parametri temporali, la VdA della tristezza è più lenta rispetto a quella della collera (la differenza tra le medie è significativa). La f0 varia in modo significativo nelle produzioni di gioia, che

presentano i valori più elevati in assoluto, seguiti da quelli della collera. In modo speculare, la tristezza presenta valori inferiori a tutte le produzioni emotive. L'escursione melodica risulta più ampia nelle produzioni di collera rispetto a quelle di tristezza ( $p < 0,05$ ), in cui l'EM presenta valori significativamente più bassi anche rispetto alla gioia. Infine, per quel che concerne l'intensità, questo parametro si comporta in modo simile alle produzioni degli altri due corpora, contribuendo a discriminare emozioni come collera e gioia da emozioni a bassa attivazione come la tristezza. Infatti, anche in questo caso, i valori più bassi si sono registrati negli enunciati di tristezza (la differenza rispetto alla collera e alla gioia è in entrambi i casi significativa,  $p < 0,05$ ).

### 5.2.5 Il confronto tra i gruppi

**TAB. 5-6. PROSPETTO DEI VALORI NORMALIZZATI ALL'INTERNO DELLE PRODUZIONI EMOTIVE IN ITALIANO L1 (ITA L1), ITALIANO L2 (ITA L2) E RUSSO L1 (RU L1).**

Parametri	Corpus	Emozioni					
		Collera		Tristezza		Gioia	
		M	DS	M	DS	M	DS
<i>Dur. Tot.</i>	Ita L1	0,195	-0,156	-0,007	-0,005	0,234	0,235
	Ita L2	0,154	-0,184	0,081	-0,271	0,145	-0,408
	Ru L1	0,264	-0,193	0,180	-0,014	0,215	-0,319
<i>VdA</i>	Ita L1	-0,174	-0,449	0,004	-0,123	-0,191	-0,260
	Ita L2	-0,147	-0,364	-0,092	-0,365	-0,148	-0,545
	Ru L1	-0,221	-0,394	-0,157	-0,159	-0,192	-0,494
<i>Int.</i>	Ita L1	0,119	0,175	-0,049	0,312	0,154	0,269
	Ita L2	0,138	-0,417	-0,013	-0,525	0,147	-0,184
	Ru L1	0,131	-0,256	-0,044	-0,369	0,138	0,632
<i>Onset</i>	Ita L1	-0,007	-0,118	-0,035	-0,294	0,008	0,030
	Ita L2	0,006	-0,091	-0,023	-0,250	0,022	0,060
	Ru L1	0,034	-0,008	-0,022	-0,039	0,012	0,009
<i>Offset</i>	Ita L1	-0,003	0,104	-0,008	0,263	0,001	0,162
	Ita L2	0,039	0,605	0,018	0,013	0,046	0,809
	Ru L1	0,003	-0,191	0,005	0,098	0,040	-0,063
<i>FO (M)</i>	Ita L1	0,013	0,170	-0,023	0,140	0,044	-0,027
	Ita L2	0,038	0,063	0,001	0,259	0,058	0,456
	Ru L1	0,028	-0,225	-0,010	-0,135	0,064	0,082
<i>EM</i>	Ita L1	0,069	-0,306	-0,123	1,933	0,628	2,653
	Ita L2	-0,149	-0,353	-0,165	0,161	0,116	-0,167
	Ru L1	0,651	1,045	0,049	0,877	0,824	0,764
<i>Dur. VT</i>	Ita L1	0,196	-0,453	-0,030	-0,140	0,381	0,121
	Ita L2	0,200	-0,189	0,239	-0,397	0,277	-0,187

Per cogliere appieno le differenze osservabili tra gruppi linguistici diversi, è stata adottata la procedura di normalizzazione, precedentemente descritta nel CAP. 4 (PAR. 4.4.6). Nella TAB. 5-6 sono riportati i valori normalizzati (media e deviazione standard) di ciascun parametro indagato all'interno di ogni corpus. I valori ottenuti hanno un valore positivo o negativo in base alla variazione dei parametri osservati.

Nei paragrafi successivi saranno riportati i risultati relativi ad ogni confronto operato tra i corpora oggetto di indagine.

### 5.2.5.1 Ita L1- Ru L1

I grafici nelle figg. 5-1/5-3 permettono di osservare le differenze riscontrate tra l'italiano e il russo in funzione dell'emozione veicolata. La linea tratteggiata rappresenta la *baseline* neutra di riferimento. Dal confronto dei dati normalizzati, emergono alcune differenze in merito alla variazione dei parametri osservati. Queste variazioni si osservano sia all'interno di ciascun gruppo sia nel confronto tra i due gruppi. Ad esempio, sia nell'italiano nativo che in russo la collera e la gioia hanno una durata maggiore rispetto al neutro e, di conseguenza, sono realizzate con una velocità di articolazione inferiore (cfr. FIG. 5-1). Queste differenze, come già come già evidenziato in precedenza sono statisticamente significative per entrambi i gruppi ( $p < 0,05$ ). La tristezza segue invece un andamento differente: in italiano L1 la VdA subisce un lieve incremento rispetto al neutro, mentre in russo si verifica un netto rallentamento.

FIG. 5-1. CONFRONTO ITA L1- RU L1: DURATA E VdA, VALORI NORMALIZZATI

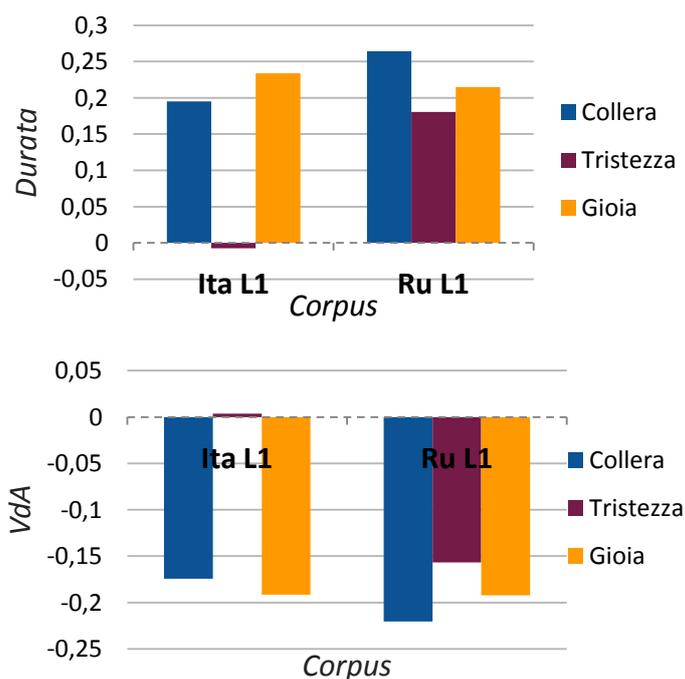


FIG. 5-2. PARAMETRI FREQUENZIALI (F0 MEDIA, EM)

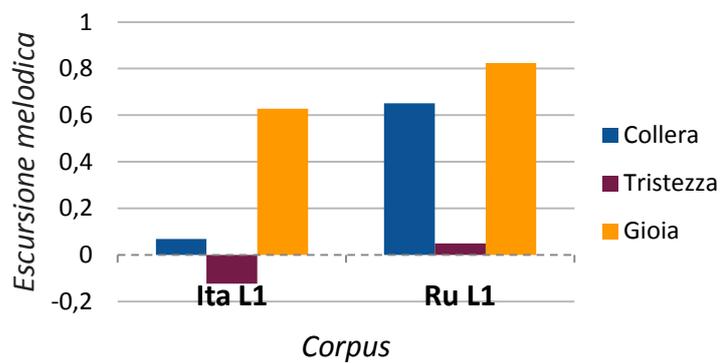
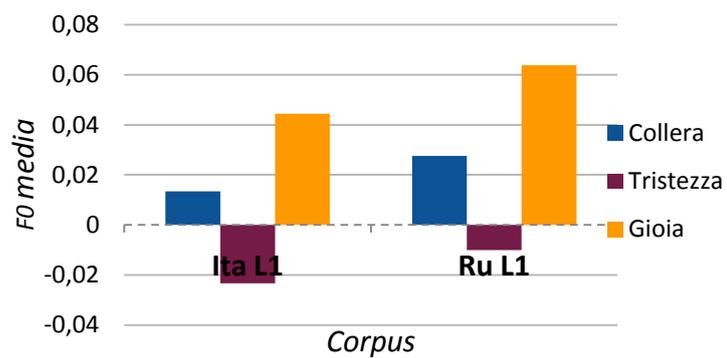
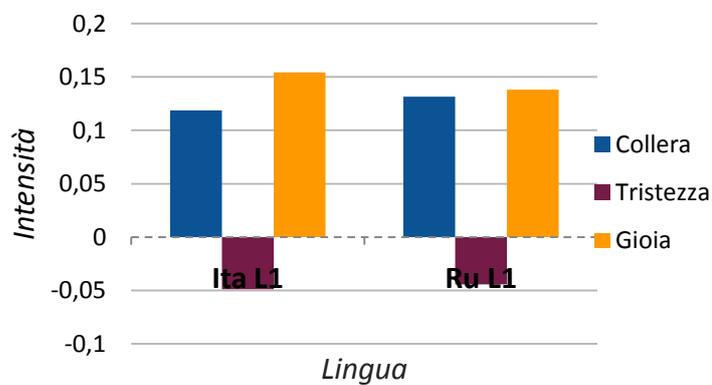


FIG. 5-3. INTENSITÀ



Per quanto concerne i parametri frequenziali, la  $f_0$  media tende ad innalzarsi notevolmente nelle produzioni di gioia di entrambi i gruppi, in misura maggiore nel russo. Come si può notare dalla FIG. 5-2 (istogramma in alto), anche le produzioni di collera mostrano un comportamento simile, ovvero un incremento della media di  $f_0$ , tuttavia sembrerebbe che nell'italiano nativo la variazione sia minore rispetto al russo. La tristezza mostra un andamento opposto, ovvero un abbassamento notevole rispetto al neutro, che si realizza in misura maggiore nell'italiano nativo. In relazione all'escursione melodica, lo scenario è più eterogeneo, in quanto nell'italiano nativo la collera viene realizzata con un'EM che si discosta poco dal neutro, mentre in russo l'escursione melodica è maggiore. La gioia viene realizzata in entrambi i gruppi con un EM notevolmente più ampia rispetto al neutro. La tristezza, al contrario, mostra un andamento opposto, come era prevedibile, si verifica, cioè, una consistente riduzione dello spazio melodico in entrambe le lingue.

Infine, per quel che concerne l'intensità (cfr. TAB. 5.3), la variazione dei valori rispetto al neutro è abbastanza omogenea in entrambi i gruppi: la gioia presenta i valori più elevati, seguita dalla collera, si registra quindi un netto incremento dell'intensità rispetto al neutro in misura maggiore nella gioia e in misura minore nella collera. Specularmente, negli enunciati di tristezza, si verifica un netto decremento dell'intensità rispetto al neutro.

#### *Risultati t test*

Nel confronto tra le produzioni in Ita L1 e Ru L1, i risultati del t test hanno evidenziato variazioni significative soprattutto sul piano ritmico-temporale, nello specifico nelle produzioni di collera e tristezza. Se si osservano i valori medi riportati per le due lingue (cfr. TAB. 5-7), si noterà che in italiano la velocità di articolazione della collera ( $M=6,18$ ,  $DS=0,73$ ) è maggiore rispetto a quella rilevata nelle produzioni in russo ( $M=5,32$ ;  $DS=0,7$ ). Parimenti, la tristezza viene realizzata con una VdA superiore in italiano ( $M=7,5$ ,  $DS=1,1$ ) rispetto al russo ( $M=5,7$ ,  $DS=0,9$ ). In entrambi i casi, la differenza tra le due lingue è significativa ( $p<0,05$ ). In relazione alla gioia non sono emerse differenze significative, in quanto le due lingue mostrano un andamento simile, ovvero un rallentamento della velocità di articolazione e un incremento della durata. Infatti, se si osservano le medie relative alla VdA si noterà che sia in russo ( $M=5,5$  sill/sec) sia in italiano ( $M=6,05$ ) l'eloquio rallenta di oltre 1 sill/sec rispetto al neutro.

TAB. 5-7. RISULTATI T TEST RELATIVI ALLA VdA: CONFRONTO ITA L1-RU L1

Collera								
VdA	Ita L1			Ru L1			Test t (IC 95%)	
	M	DS	N.	M	DS	N.	t	P(T<=t)
	6,182	0,736	10	5,327	0,702	10	2,6	<b>0,001*</b>

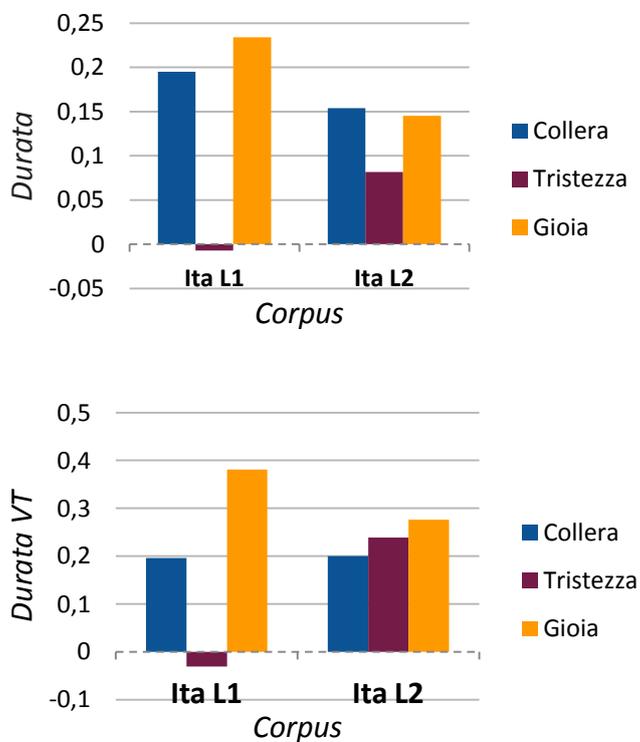
Tristezza								
VdA	Ita L1			Ru L1			Test t (IC 95%)	
	M	DS	N.	M	DS	N.	t	P(T<=t)
	6,182	0,736	10	5,763	0,975	10	3,6	<b>0,002*</b>

Gioia								
VdA	Ita L1			Ru L1			Test t (IC 95%)	
	M	DS	N.	M	DS	N.	t	P(T<=t)
	6,05	0,9	10	5,5	0,5	10	2,6	0,16

5.2.5.2 Ita L1- Ita L2

FIG. 5-4 CONFRONTO ITA L1- RU L1: DURATA TOT.; DURATA VT, E VdA



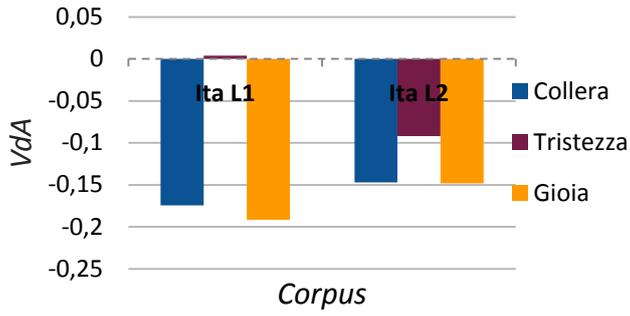


FIG. 5-5. VALORI FREQUENZIALI (ONSET E OFFSET): CONFRONTO ITA L1-ITA L2

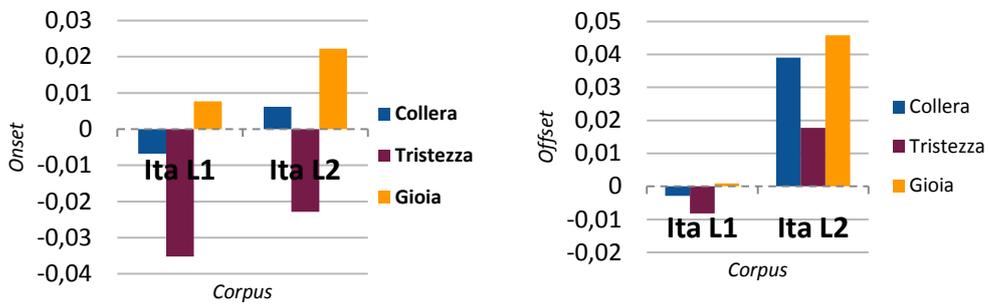


FIG. 5-6. VALORI FREQUENZIALI (F0 MEDIA; EM): CONFRONTO ITA L1-ITA L2

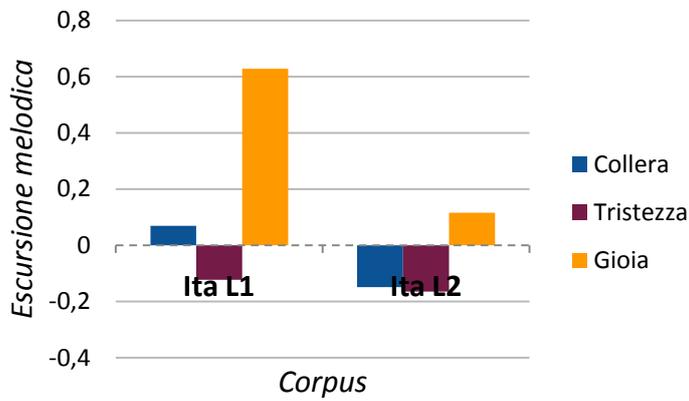
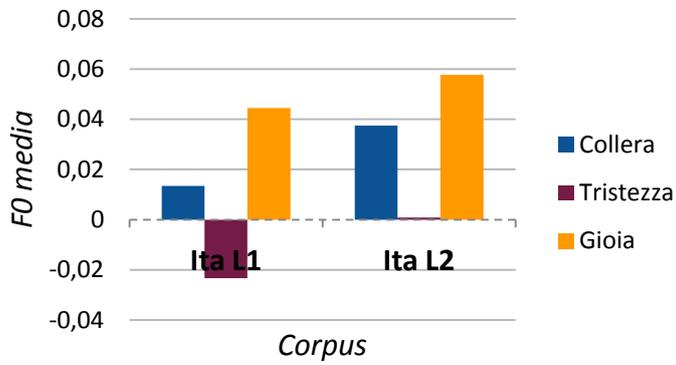
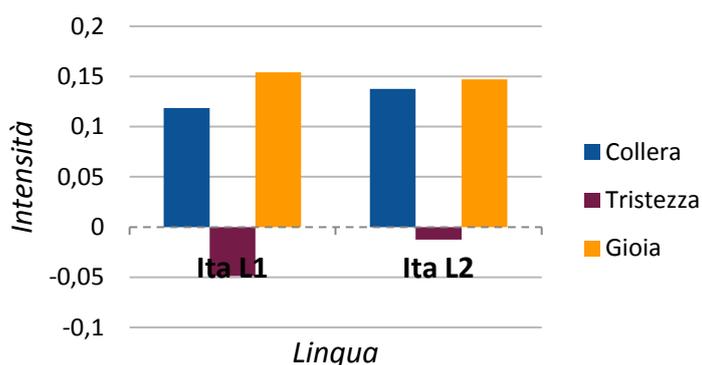


FIG. 5-6. VALORI RELATIVI ALL'INTENSITÀ: CONFRONTO ITA L1-ITA L2



Dal confronto dei dati normalizzati, si osservano alcune differenze in merito ai parametri temporali. In italiano L1, la durata della tristezza è più breve rispetto all'eloquio neutro (anche la durata della vocale tonica), mentre la velocità di articolazione aumenta. Al contrario, in Ita L2, la tristezza viene realizzata più lentamente rispetto al neutro. In entrambi i gruppi si registra un incremento della durata e una riduzione della velocità di articolazione negli enunciati di gioia e collera.

Per quanto concerne i parametri frequenziali, sia in italiano L1 che in italiano L2 i valori di  $f_0$  subiscono un lieve incremento negli enunciati di collera rispetto al neutro; anche l'escursione melodica è più ampia. Tuttavia, come è possibile osservare nella FIG.5-8, le produzioni degli apprendenti presentano valori di  $f_0$  molto più elevati rispetto al neutro, ma sono realizzati con un'escursione melodica più ridotta rispetto all'eloquio normale. La tristezza segue un andamento simile in entrambi i gruppi, ovvero: un abbassamento della media di  $f_0$  e una riduzione dell'escursione melodica. La gioia si caratterizza in entrambi i gruppi per una  $f_0$  molto più elevata rispetto al neutro, tuttavia nell'italiano L1 l'estensione tonale sembrerebbe di gran lunga maggiore rispetto alle produzioni dei non nativi, che invece realizzano entro un *range* piuttosto ristretto. I valori di *onset* si innalzano nelle produzioni di collera in entrambi i gruppi, e si abbassano in modo cospicuo in quelle di tristezza. In relazione alla gioia, questa presenta, in particolare nell'eloquio dei non nativi, un incremento importante rispetto ai valori riportati per il neutro. L'*offset*, invece, risulta più basso negli enunciati emotivi dei nativi (soprattutto, in quelli di tristezza); mentre, nelle produzioni degli apprendenti risulta più elevato rispetto all'eloquio neutro. Sembrerebbe dunque che gli apprendenti ricorrano ad una  $f_0$  tanto elevata da ridurre lo spazio melodico.

Infine, per quel che concerne l'intensità si registra in entrambi i gruppi un incremento dei valori negli enunciati di gioia, seguiti da quelli di collera e da un decremento dei valori negli enunciati di tristezza.

### *Collera*

**TAB. 5-8. RISULTATI T TEST (COLLERA): CONFRONTO ITA L1-ITA L2**

	Collera						Test t (IC 95%)	
	Ita L1			Ita L2			t	P(T<=t)
	M	DS	N.	M	DS	N.		
Dur. Tot.	982,100	108,870	10	1079,200	152,947	10	1,6	0,12
VdA	6,182	0,736	10	5,661	0,804	10	1,5	0,14
Int. (M)	68,900	5,174	10	71,900	3,665	10	1,4	0,15
<i>Onset</i>	5,181	0,403	10	4,953	0,309	10	1,4	0,17
<i>Offset</i>	4,955	0,398	10	4,766	0,348	10	1,1	0,27
F0 (M)	5,248	0,363	10	5,018	0,322	10	1,4	0,15
<b>EM</b>	9,172	1,062	10	6,850	2,753	10	2,4	<b>0,03*</b>
Dur. VT	181,000	27,909	10	184,500	41,594	10	0,2	0,82

Nel confronto tra Ita L1 e Ita L2 in relazione alla collera, emerge un unico parametro significativo, ovvero l'escursione melodica (cfr. TAB. 5-8). Nell'italiano nativo, le produzioni di collera sono realizzate in media con un'estensione tonale ampia (M=9,1 st), mentre i russi ricorrono tendenzialmente ad un *range* più ristretto (M=6,8 st). Questa differenza tra i due gruppi è statisticamente significativa ( $p < 0,05$ ). Sul piano ritmico temporale, i russi inoltre realizzano con una velocità di articolazione inferiore (5,6 sill/sec) rispetto agli italiani (6,1 sill/sec). Parimenti, osservando le medie relative alla durata totale si nota un'estensione temporale maggiore nell'eloquio dei non nativi (M=1079,2 ms) rispetto a quello dei non nativi (M=982,1 ms), tuttavia, le differenze osservate non sono significative ( $p > 0,05$ ). La durata della vocale tonica presenta valori simili in entrambe le produzioni (181 ms per Ita L1 e 184,5 ms per Ita L2). In relazione all'intensità, non emergono differenze importanti, in quanto l'eloquio oscilla intorno ai 68,9 db nelle produzioni degli italiani e ai 71,9 db nelle produzioni degli apprendenti.

## Tristezza

TAB. 5-9. RISULTATI T TEST (TRISTEZZA): CONFRONTO ITA L1-ITA L2

Tristezza	Ita L1			Ita L2			Test t (IC 95%)	
	M	DS	N.	M	DS	N.	T	P(T<=t)
<b>Dur. Tot.</b>	816,000	128,434	10	1011,500	136,668	10	3,2	<b>0,004*</b>
<b>VdA</b>	7,516	1,172	10	6,029	0,803	10	3,3	<b>0,004*</b>
Int. (M)	58,600	5,777	10	62,400	2,989	10	0,8	0,08
<i>Onset</i>	5,033	0,323	10	4,810	0,255	10	1,7	0,1
<i>Offset</i>	4,929	0,455	10	4,668	0,220	10	1,6	0,12
F0 (M)	5,057	0,354	10	4,840	0,382	10	1,3	0,2
<b>EM</b>	7,523	4,486	10	6,723	4,938	10	2,3	<b>0,03*</b>
<b>Dur. VT</b>	146,700	43,889	10	190,400	30,913	10	2,5	<b>0,02*</b>

La normalizzazione dei dati (cfr. PAR. 5.2.5.2) lascerebbe supporre nell'italiano nativo la tristezza abbia una durata inferiore (M= 816, DS=128,4) rispetto all'italiano non nativo e sia quindi realizzata con una velocità di articolazione più elevata (M=7,5 sill/sec vs. M=6 sill/sec). In effetti, i risultati del t test (cfr. TAB. 5-9) hanno evidenziato differenze significative in relazione ai parametri temporali ( $p < 0,05$ ); anche la durata della vocale tonica risulta più lunga nelle produzioni degli apprendenti (M=190,4 ms vs. M=146.7 ms).

Sul piano frequenziale, non sono emerse differenze significative in relazione alla f0 che raggiunge i 5 st nelle produzioni dei nativi e 4,8 st in quelle dei non nativi. L'estensione tonale dei russi è in media più ristretta (6,7 st) rispetto a quella dei nativi che ricorrono a un *range* più ampio (7,5 st). La differenza, in questo caso, è statisticamente significativa ( $p < 0,05$ ). I valori di *onset* e *offset* non variano in modo significativo, come pure l'intensità, i cui valori si mantengono bassi in entrambi i gruppi (58,6 dB per i nativi, 62,4 per i russi).

## Gioia

TAB. 5-10. RISULTATI T TEST (GIOIA): CONFRONTO ITA L1-ITA L2

	Gioia						Test t (IC 95%)	
	Ita L1			Ita L2			t	P(T<=t)
	M	DS	N.	M	DS	N.		
Dur. Tot.	1014,200	159,341	10	1071,100	110,980	10	0,9	0,36
VdA	6,054	0,938	10	5,655	0,575	10	1,1	0,2
Int. (M)	71,100	5,587	10	72,500	5,126	10	0,5	0,56
Onset	5,257	0,471	10	5,032	0,361	10	1,1	0,24
Offset	4,974	0,419	10	4,796	0,392	10	0,9	0,34
F0 (M)	5,408	0,302	10	5,116	0,442	10	1,7	0,1
EM	13,970	5,587	10	8,979	3,544	10	2.3	<b>0,03*</b>
Dur. VT	208,900	57,241	10	196,200	41,678	10	0,5	0,59

In merito alla gioia, non sono emerse differenze significative sul piano temporale (cfr. TAB. 5-10), come era prevedibile. In entrambi i casi si registra inoltre un allungamento della vocale tonica rispetto alle altre produzioni, soprattutto rispetto all'eloquio neutro (208 ms per i nativi e 196,2 ms per i non nativi).

In relazione ai parametri frequenziali, *onset* e *offset* presentano valori simili, così come la *f0* media, che si colloca intorno ai 5,1 st nelle produzioni degli apprendenti e raggiunge i 5,4 st in Ita L1. L'unico parametro a differire in modo significativo è l'escursione melodica: ancora una volta nell'italiano nativo, l'EM è più ampia (13,9 st), mentre l'estensione tonale dei russi si assesta intorno ai 6,7 st, uno spazio melodico nettamente dimezzato.

Infine, i valori relativi all'intensità non presentano differenze importanti, anzi sono piuttosto elevati in entrambe le produzioni (71,1 db in Ita L1 e 72,5 db in Ru L1), soprattutto se confrontate con l'eloquio neutro e con la tristezza.

### 5.2.5.3 Ita L2-Ru L1

FIG. 5.7 PARAMETRI TEMPORALI (DURATA E VdA): CONFRONTO ITA L2-RU L1

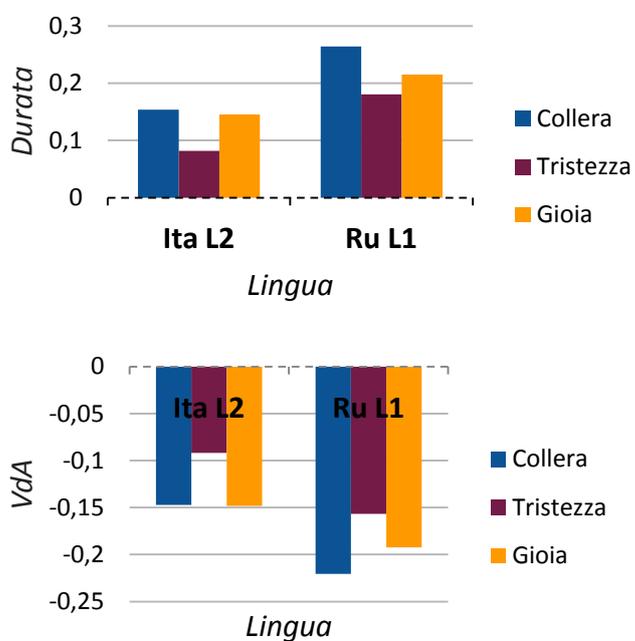


FIG. 5.8 PARAMETRI FREQUENZIALI (F0 MEDIA; EM): CONFRONTO ITA L2-RU L1

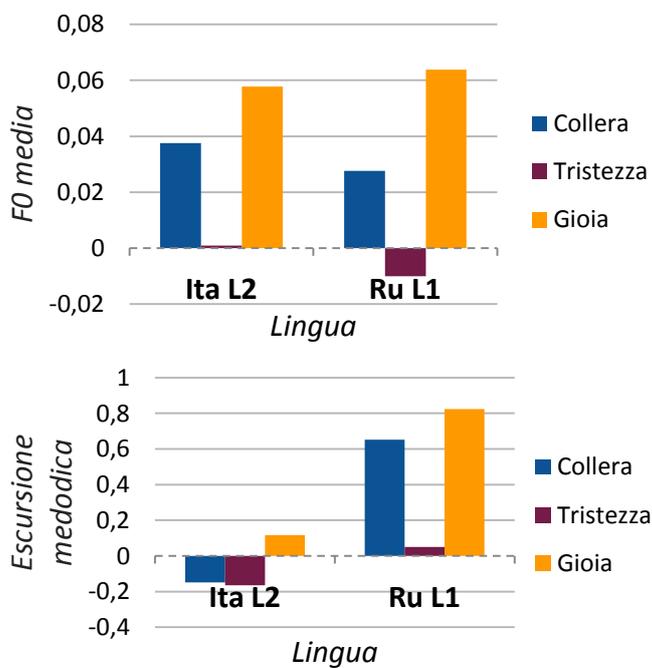
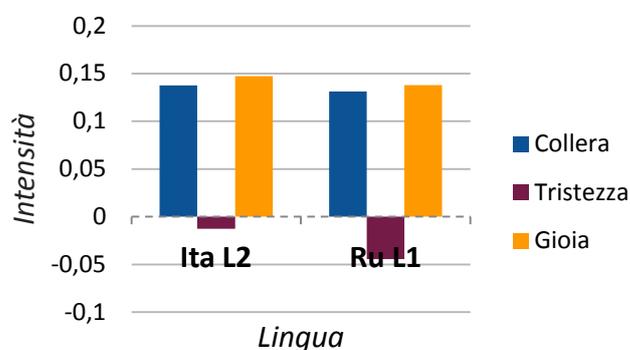


FIG. 5.9 PARAMETRI RELATIVI ALL'INTENSITÀ: CONFRONTO ITA L2-RU L1



Prendiamo in esame le produzioni realizzate dagli apprendenti in italiano e nella loro lingua nativa. Dal confronto dei dati normalizzati, si osserva una variazione simile in relazione ai parametri ritmico-temporali (cfr. 5-7). Gli apprendenti russi realizzano gli enunciati di collera con una velocità di articolazione inferiore rispetto al neutro sia nella loro lingua nativa sia in italiano L2. Lo stesso andamento seguono anche la gioia, in entrambe le produzioni e in misura minore la tristezza. In sostanza, sembrerebbe che gli apprendenti quando esprimono un'emozione tendono a rallentare la velocità di articolazione, in misura maggiore nella collera, seguita dalla gioia e dalla tristezza e questo si verifica soprattutto nelle produzioni in italiano, lingua che non padroneggiano ancora bene.

In relazione ai parametri frequenziali (cfr. TAB. 5-8), lo scenario appare più controverso. In italiano L2 e in russo L1 la collera viene realizzata con una  $f_0$  media più alta rispetto al neutro, tuttavia mentre in italiano L2 l'escursione melodica è più ristretta rispetto al neutro, in russo L1 è notevolmente più ampia. Anche la gioia viene realizzata con una  $f_0$  media più alta sia in ita L2 che in russo L1, tuttavia, l'escursione melodica si discosta di poco rispetto al neutro nelle produzioni in italiano L2, mentre è notevolmente ampia in russo L1. Questo potrebbe essere imputabile a qualche difficoltà nella gestione delle strutture intonative della lingua target. La tristezza infine, si discosta di poco dal neutro in entrambe le produzioni, non si osservano quindi differenze importanti.

Per quanto concerne l'intensità (cfr. FIG. 5-9), la variazione dei valori è simile in entrambe le produzioni: si registra un incremento dell'intensità rispetto al neutro negli enunciati di gioia, seguiti da quelli di collera e una riduzione dei valori nelle produzioni di tristezza.

TAB. 5-11. RISULTATI T TEST (TRISTEZZA, GIOIA, NEUTRO): CONFRONTO RU L1-ITA L2

Tristezza								
	Ru L1			Ita L2			Test t (IC 95%)	
	M	DS	N.	M	DS	N.	t	P(T<=t)
<b>Int.</b>	58,200	3,084	10	62,400	2,989	10	3,09	<b>0,006*</b>
Gioia								
	Ru L1			Ita L2			Test t (IC 95%)	
	M	DS	N.	M	DS	N.	t	P(T<=t)
<b>Dur. Tot.</b>	915,300	99,034	10	1071,100	110,980	10	3,3	<b>0,004*</b>
Neutro								
	Ru L1			Ita L2			Test t (IC 95%)	
	M	DS	N.	M	DS	N.	t	P(T<=t)
<b>Dur. Tot.</b>	753,400	145,506	10	935,300	187,401	10	2,4	<b>0,02*</b>

Nel confronto Ru L1-Ita L2, i risultati del t test (cfr. TAB. 5-11) hanno evidenziato differenze significative principalmente sul piano temporale, in relazione alla durata della gioia e del neutro: in russo la durata delle produzioni di gioia (M=915, 3 ms) e di quelle neutre (M=753,4) è inferiore rispetto a quelle realizzate in italiano dagli apprendenti(cfr. TAB. 5-11).

In merito ai parametri frequenziali, f0 media ed escursione melodica mostrano un andamento simile in russo L1 e in italiano L2, sebbene tendenzialmente gli apprendenti ricorrano ad una EM più ampia nella loro lingua nativa (soprattutto rispetto all'eloquio neutro), anche se questa differenza non è significativa (p>0,05).

Infine, l'intensità varia in modo significativo nelle produzioni di tristezza, che presentano valori più elevati in Ita L2 (M=62,4 db) rispetto a Ru L1 (M=58,2).

### 5.2.6 Discussione dei dati relativi al processo di encoding

Le analisi condotte sui tre corpora hanno restituito risultati per certi versi congruenti con l'ipotesi che l'espressione vocale delle emozioni si fondi su aspetti universali nonché su aspetti specificamente linguistici e culturali. Le produzioni analizzate hanno mostrato un comportamento analogo, o meglio, una tendenza similare

nella modulazione delle risorse prosodiche, tuttavia, la variazione dei parametri indagati assume un peso specifico diverso nel confronto tra i corpora oggetto di indagine.

In primo luogo, sembrerebbe che il fattore "emozione" non influenzi allo stesso modo la variazione dei parametri all'interno di ciascun corpus. Nell'italiano nativo, tutte le produzioni emotive sono interessate da una variazione significativa in termini di durata e di velocità di articolazione, mentre sul piano frequenziale, è l'escursione melodica il parametro maggiormente soggetto a variazioni significative. Anche l'intensità varia sensibilmente in tutte le produzioni emotive. Nelle produzioni realizzate dagli apprendenti in italiano e in russo, i parametri temporali e frequenziali non subiscono variazioni significative, soltanto l'intensità dimostra una certa variabilità ( $p < 0,05$ ). In linea generale, queste prime differenze permettono di intravedere nel parlato degli italiani una maggiore variabilità sul piano temporale e frequenziale rispetto a quello degli apprendenti (sia in L1 sia in L2).

#### *Le variazioni rispetto all'eloquio neutro nelle due lingue native*

Considerando, per il momento, soltanto le produzioni realizzate nelle due lingue native (italiano e russo), emergono alcuni punti in comune e alcune discrepanze. In base ai dati raccolti, sembrerebbe che i segnali acustici impiegati dagli *encoder* italiani e russi seguano *pattern* simili, congruenti con quelli riportati da altri studi condotti a livello intra e inter-linguistico sulle due lingue (cfr. Makarova, Petrushin 2003; Pettorino, 2004; Anolli et al., 2008). Nello specifico, sembrerebbe che emozioni come la gioia e la collera implicino un rallentamento dell'eloquio (come già evidenziato da Pettorino per l'italiano e da Makarova e Petrushin per il russo), infatti, come è emerso dalle analisi, tutte le produzioni emotive all'interno di ciascun corpus sono state realizzate con una velocità di articolazione inferiore rispetto alla *baseline* neutra. Un secondo parametro che mostra andamenti comuni a tutti e due i corpora è l'intensità, che raggiunge valori elevati negli enunciati di collera e gioia, e si riduce in quelli di tristezza, equiparando o scendendo al di sotto dei valori registrati per l'eloquio normale. In relazione ai parametri frequenziali, lo scenario è più eterogeneo. Ad esempio, sia nell'italiano nativo che in russo, la gioia si caratterizza per un incremento significativo dei valori relativi a  $f_0$  media ed escursione tonale. La collera, invece, viene realizzata in italiano attraverso una EM moderatamente più ampia rispetto all'eloquio normale, ma che non varia in modo significativo; inoltre, la  $f_0$  si colloca in uno spazio frequenziale simile a quello del neutro. In russo, sia la  $f_0$  che l'EM subiscono variazioni significative

rispetto all'eloquio normale, così come l'*onset*, che risulta sensibilmente più elevato negli enunciati di collera. La tristezza, infine, viene realizzata sfruttando in modo diverso il *pitch*: nell'italiano nativo questo stato emotivo si caratterizza per un abbassamento significativo dei valori di  $f_0$  e da una riduzione (anche se non significativa) dell'EM. In russo, si verifica un significativo abbassamento dei valori di *onset*, mentre  $f_0$  media ed EM variano poco rispetto all'eloquio normale.

#### *La variazione dei parametri nelle produzioni emotive*

Prendendo in esame la variazione dei parametri da uno stato emotivo all'altro, le similarità tra i due corpora appaiono maggiori, nel senso che alcuni parametri contribuiscono in modo analogo a discriminare emozioni ad alta attivazione e potenza (come gioia e collera) da emozioni a bassa attivazione e potenza (come la tristezza).

Ad esempio, sul piano temporale, la tristezza viene realizzata in italiano con un'estensione temporale minore e una VdA più elevata rispetto alle altre emozioni: la differenza con la collera e la gioia è significativa ( $p < 0,05$ ). Anche in russo, la VdA della tristezza è significativamente più veloce rispetto a quella della collera (ma non rispetto a quella della gioia).

Sul piano frequenziale, se si confrontano i valori di  $f_0$  e dell'EM della gioia con quelli riportati per le altre produzioni emotive, si noterà che questa viene realizzata in italiano con una  $f_0$  più elevata e con una EM notevolmente più ampia, sia rispetto alla collera sia rispetto alla tristezza ( $p < 0,05$ ). specularmente, la tristezza presenta i valori di  $f_0$  più bassi in assoluto. Anche in russo, la  $f_0$  varia in modo significativo nelle produzioni di gioia, che presentano i valori più elevati in assoluto, seguiti da quelli della collera; mentre la tristezza presenta valori di  $f_0$  inferiori a tutte le produzioni emotive. L'escursione melodica contribuisce a discriminare in russo le produzioni di collera rispetto a quelle di tristezza ( $p < 0,05$ ), in cui l'EM presenta valori significativamente più bassi anche rispetto alla gioia. Infine, per quel che concerne l'intensità, questo parametro si comporta in modo simile in italiano e in russo: infatti, i valori più bassi si sono registrati negli enunciati di tristezza (la differenza rispetto alla collera e alla gioia è in entrambi i casi significativa,  $p < 0,05$ ).

Nel confronto diretto operato tra le emozioni prodotte nelle due lingue native, le maggiori differenze sono emerse in relazione al piano ritmico-temporale, nello specifico le produzioni di collera e tristezza. In italiano la velocità di articolazione della collera è maggiore rispetto a quella rilevata per il russo. In modo analogo, la tristezza viene

realizzata con una VdA superiore in italiano ( $M=7,5$ ,  $DS=1,1$ ) rispetto al russo ( $M=5,7$ ,  $DS=0,9$ ). In entrambi i casi, la differenza tra le due lingue è significativa ( $p<0,05$ ). La gioia invece mostra una velocità di articolazione simile, per cui non sono emerse differenze significative.

### *Le produzioni emotive in italiano L2*

Se si prendono in esame le produzioni emotive degli apprendenti in italiano L2, il quadro delineato dalle analisi rivela, anche in questo caso, alcuni punti in comune con gli altri due corpora. Ad esempio, sul piano temporale, la collera e la gioia hanno una durata maggiore rispetto alla tristezza e sono realizzate più lentamente (la differenza tra le medie è significativa,  $p<0,05$ ), *trend* già riscontrato in Ita L1 Ru L1. Sul piano frequenziale, la  $f_0$  media contribuisce a differenziare la tristezza dalla collera: infatti, i valori della prima risultano significativamente più bassi rispetto alla seconda (4,8 st vs. 5 st). L'escursione melodica invece appare discriminante soltanto nel confronto tra la gioia (8,9 st) e la collera (6,8 st): la differenza tra le medie è statisticamente significativa ( $p<0,05$ ). Per quanto concerne l'intensità, questo parametro raggiunge valori significativamente più bassi negli enunciati di tristezza rispetto quelli di collera e gioia (in entrambi i casi,  $p<0,05$ ), come già riportato per Ita L1 e Ru L1.

Confrontando il parlato emotivo prodotto dagli apprendenti in italiano e nella loro lingua d'origine, la normalizzazione dei dati ha evidenziato variazioni simili rispetto all'eloquio neutro, sia sul piano temporale, sia su quello frequenziale e dell'intensità. Le uniche differenze significative sono state registrate negli enunciati di gioia, che risultano significativamente più brevi in russo rispetto all'italiano L2, e in quelli di tristezza che presentano valori di intensità più elevati in italiano L2 ( $p<0,05$ ).

Dal confronto Ita L1-Ita L2 sono emerse alcune discrepanze in relazione alle emozioni considerate. Ad esempio, nell'italiano nativo, le produzioni di collera sono realizzate in media con un'estensione tonale ampia ( $M=9,1$  st), mentre gli apprendenti russi ricorrono tendenzialmente ad un *range* più ristretto ( $M=6,8$  st). Questa differenza tra i due gruppi è statisticamente significativa ( $p<0,05$ ). Sul piano ritmico- temporale e dell'intensità non sono emerse invece differenze significative in merito alla collera. Le produzioni di tristezza presentano una velocità di articolazione maggiore nell'italiano nativo rispetto all'italiano non nativo ( $p<0,05$ ). Il rallentamento dell'eloquio dei non nativi è comprovato anche da una maggiore durata della vocale tonica ( $p<0,05$ ). Anche l'escursione melodica differisce in modo significativo: nell'eloquio degli italiani l'EM è

maggiore rispetto a quello dei russi, che invece ricorrono a un *range* più ristretto. Gli altri parametri frequenziali e l'intensità mostrano un andamento simile in entrambe le produzioni. Infine, in relazione alla gioia, le differenze sono meno importanti: sul piano temporale si verifica in entrambi i casi un cospicuo rallentamento della VdA e un allungamento della durata totale e della VT; anche l'intensità raggiunge livelli simili. L'unico parametro a differire in modo significativo è ancora una volta l'EM, che risulta nettamente più ampia nelle produzioni dei nativi (13,9 st) rispetto a quelle dei non nativi (8,9 st).

In conclusione, se da un lato le analisi hanno rivelato una variazione dei parametri simile all'interno dei tre corpora analizzati in relazione alle emozioni veicolate, dall'altro hanno evidenziato come l'entità di queste variazioni sia stata maggiore nel parlato degli italo-foni rispetto a quello dei russo-foni (sia in L1 sia L2). In questo senso, è ragionevole supporre che gli italiani abbiano modulato la propria prosodia emotiva adattandola al contesto comunicativo attraverso variazioni più significative rispetto ai russi. Come già evidenziato, l'escursione melodica rappresenta quasi sempre un parametro discriminante tra i due gruppi. Il *pitch range* più ampio che caratterizza il parlato emotivo degli italiani rispetto a quello dei russi potrebbe costituire in un certo senso un aspetto linguo-specifico, che andrebbe ulteriormente approfondito. Alla luce dei risultati discussi finora, la verifica uditiva delle produzioni emotive analizzate appare necessaria per chiarire in che misura i segnali acustici impiegati dagli *encoder* per veicolare una specifica emozione siano stati recepiti e decodificati dai valutatori esterni. Infatti, come già anticipato nel cap. 2, il Modello della Lente rielaborato da Scherer (1978; 82), sottolinea il carattere probabilistico e ridondante degli segnali vocali, che non costituiscono dunque degli indici fedeli dell'emozione espressa ma vengono filtrati attraverso l'esperienza percettiva e i processi cognitivi del ricevente.

### **5.3 Risultati relativi al processo di *decoding***

Come illustrato nel CAP. 4, la seconda parte della ricerca ha previsto l'elaborazione e la distribuzione di due test uditivi, rispettivamente destinati ad un gruppo di apprendenti di origine russa e ad un gruppo di parlanti nativi di italiano.

### 5.3.1 Risultati del test uditivo sottoposto agli apprendenti russi (U\_Ita\_L2)

Prendiamo in esame il test uditivo somministrato agli apprendenti russi. Il gruppo era composto di 10 studenti (6 F e 4 M), dai 22 ai 30 anni (età media=25) con un livello di competenza intermedio (B2) e in Italia da un periodo variabile (dai 12 ai 24 mesi). Oltre al russo e all'italiano, parlavano altre lingue tra cui l'inglese.

#### 5.3.1.1 Analisi dei riconoscimenti corretti

Il test uditivo somministrato al gruppo russofono conteneva 12 stimoli in lingua russa e 12 in italiano, ovvero 3 produzioni per ogni tipologia di enunciato (4). Nella descrizione dei risultati, saranno presi in esame in primo luogo le percentuali globali di riconoscimento in funzione della lingua, in secondo luogo le decodifiche corrette totalizzate in funzione dell'emozione espressa.

TAB. 5-12. STATISTICA RELATIVA ALLE DECODIFICHE CORRETTE TOTALIZZATE DAGLI APPRENDENTI NELLE DUE LINGUE

Statistica decodifiche corrette	<i>Ita L1</i>	<i>Ru L1</i>	<i>Totale</i>
Media	57,5	70	63,7
Errore standard	6,3	6,5	2,6
Mediana	58,3	66,6	62,5
Moda	66,6	91,6	62,5
Deviazione standard	20,2	20,8	8,3
Varianza	408,1	435,1	69,6
Range	75	58	25
Minimum	16,6	41,6	50
Maximum	91	100	75

La TAB. 5-12 sintetizza i risultati relativi alla media dei riconoscimenti in funzione della lingua e riporta inoltre la media dei riconoscimenti globali, che oscilla intorno al 63,7% (DS=8,3%). In un test a risposta multipla con 7 opzioni, la probabilità di un riconoscimento casuale è di circa il 14,3%. Dai risultati del test binomiale, risulta che la percentuale globale di decodifiche corrette supera soglia della pura casualità ( $p < 0,05$ ). Parimenti, le percentuali di riconoscimento per ciascuna emozione in funzione della lingua (Ita L1 e Ru L1) superano la soglia della pura casualità ( $p < 0,05$ ).

La FIG. 5-10 mette a confronto le decodifiche corrette totalizzate dagli apprendenti in italiano e in russo. La linea tratteggiata rappresenta la soglia del riconoscimento casuale (14,3%). Sembrerebbe che gli apprendenti siano riusciti a identificare con maggiore accuratezza produzioni realizzate nella loro lingua nativa (circa il 70% di decodifiche corrette); le percentuali relative all'italiano sono infatti più basse (57,5%). Per stabilire se la differenza tra le medie fosse statisticamente significativa è stato condotto un test t di Student per dati appaiati, che non ha restituito tuttavia risultati significativi ( $p > 0,05$ ).

La FIG. 5-11 illustra invece le percentuali di riconoscimento globale per ciascuna emozione, considerando cioè la totalità delle risposte corrette sia in italiano L1 che in russo L1. Come si può notare le percentuali relative alle tre emozioni considerate sono abbastanza eterogenee. La tristezza ha ottenuto in assoluto il numero di riconoscimenti più elevato (80%), seguita dalla gioia (65%). La percentuale della collera, invece, risulta nettamente inferiore (35%). In relazione all'eloquio neutro, le decodifiche corrette sono piuttosto elevate (75%). Naturalmente, queste percentuali non descrivono in modo chiaro le differenze osservabili nelle due lingue, che sono illustrate invece nella FIG. 5-12.

FIG. 5-10. PERCENTUALI DECODIFICHE TOTALIZZATE DAGLI APPRENDENTI NELLE DUE LINGUE

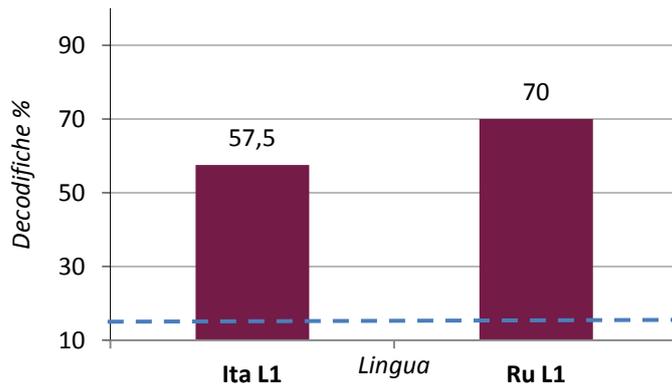


FIG. 5-11. PERCENTUALI DECODIFICHE GLOBALI TOTALIZZATE DAGLI APPRENDENTI IN BASE ALLA TIPOLOGIA DI ENUNCIATO

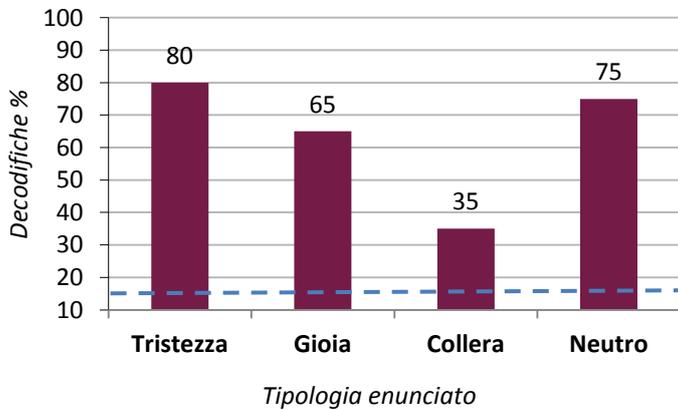
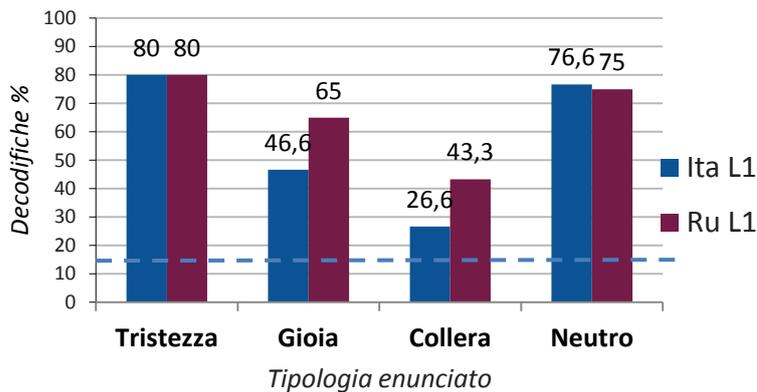


FIG. 5-12. PERCENTUALI DECODIFICHE TOTALIZZATE DAGLI APPRENDENTI NELLE DUE LINGUE IN BASE ALLA TIPOLOGIA DI ENUNCIATO



Mentre per la tristezza non emergono differenze, anzi la percentuale risulta in entrambe le lingua molto elevata (80%), in relazione alla gioia si assiste ad un leggero squilibrio: l'italiano ha ottenuto una percentuale inferiore (46,6%) rispetto al russo

(65%). Lo scarto aumenta nel caso della collera, che ha ottenuto soltanto il 26,6% di decodifiche corrette in italiano, mentre in russo la percentuale è nettamente superiore (43,3%). Per verificare se le differenze osservate fossero statisticamente significative, sono stati condotti due test t di Student (cfr. TAB. 5-13) per campioni appaiati sulle medie dei riconoscimenti totalizzati complessivamente in italiano L1 e in russo L1. I risultati hanno differenze significative in relazione al riconoscimento della gioia, che è stata decodificata con maggiore accuratezza in russo rispetto all'italiano. Per quanto riguarda la collera, invece, non sono emerse differenze significative ( $p > 0,05$ ), vd. TAB. 5-12. In relazione all'eloquio neutro, non si osservano differenze importanti. Da notare, tuttavia, che la percentuale di riconoscimento risulta in entrambe le lingue piuttosto elevata (76,6% per l'italiano e 75% per il russo).

**TAB. 5-13. RISULTATI T TEST: CONFRONTO DECODIFICHE CORRETTE IN RUSSO E IN ITALIANO IN RELAZIONE ALLA GIOIA E ALLA COLLERA.**

Tipologia enunciato (E)								
<i>Gioia</i>								
RU LI			ITA L1			Test t (IC 95%)		
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)
2,5	0,7	10	1,4	0,7	10	3,4	9	<b>0,006*</b>
<i>Collera</i>								
RU LI			ITA L1			Test t (IC 95%)		
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)
1,3	1,3	10	0,8	0,6	10	0,9	9	0,3

Alla luce dei risultati ottenuti, sembrerebbe che emozioni come la gioia e la tristezza siano state riconosciute con maggiore accuratezza in entrambe le lingue rispetto alla collera, che ha ottenuto invece percentuali inferiori. Per stabilire se le differenze osservate fossero significative, sono stati condotti due serie di test t di Student per dati appaiati confrontando le medie totalizzate da ciascuna emozione, sia per il russo (cfr. TAB. 5-13) che per l'italiano (TAB. 5-14). Dai risultati del t test, si evince che la media dei riconoscimenti relativa alla collera differisce in modo significativo da quella della tristezza, ciò si verifica sia in russo, in cui la tristezza ha ottenuto in media un maggior numero di riconoscimenti ( $M=2,4$ ;  $DS=0,5$ ) rispetto alla

collera (M= 1,3; DS=1,3), sia in italiano (rispettivamente: M=2,4; DS=0,7 per la tristezza; M=0,8; DS=0,6 per la collera). Inoltre, le decodifiche della collera differiscono in modo significativo anche da quelle totalizzate dalla gioia, in entrambe le lingue ( $p < 0,05$ ).

**TAB. 5-14. RISULTATI T TEST: CONFRONTO DECODIFICHE CORRETTE IN BASE ALLA TIPOLOGIA DI ENUNCIATO (IN RUSSO)**

Tipologia enunciato (E)									
COLLERA			TRISTEZZA			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
RU L1	1,3	1,3	10	2,4	0,5	10	3,9	9	<b>0,003*</b>
COLLERA			GIOIA			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
RU L1	1,3	1,3	10	2,5	0,7	10	3,08	9	<b>0,006*</b>
COLLERA			NEUTRO			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
RU L1	1,3	1,3	10	2,2	0,9	10	1,9	9	0,08
TRISTEZZA			GIOIA			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
RU L1	2,4	0,5	10	2,5	0,7	10	0,4	9	0,6

**TAB. 5-15. RISULTATI T TEST: CONFRONTO DECODIFICHE CORRETTE IN BASE ALLA TIPOLOGIA DI ENUNCIATO (IN ITALIANO)**

Tipologia enunciato (E)									
COLLERA			TRISTEZZA			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
ITA L1	0,8	0,6	10	2,4	0,7	10	7,2	9	<b>0,00004*</b>
COLLERA			GIOIA			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
ITA L1	0,8	0,6	10	1,4	0,7	10	3,6	9	<b>0,005*</b>
COLLERA			NEUTRO			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
ITA L1	0,8	0,6	10	2,3	1,2	10	3,7	9	<b>0,004*</b>
TRISTEZZA			GIOIA			Test t (IC 95%)			
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
ITA L1	2,4	0,7	10	1,4	0,7	10	3,8	9	<b>0,003*</b>

Queste differenze suggeriscono che gli apprendenti abbiano avuto maggiori difficoltà nel riconoscimento della collera rispetto alle altre emozioni ( $p < 0,05$ ) sia in

russo che in italiano. In relazione all'italiano L1, si osserva che anche la media relativa all'eloquio neutro (M=2,3; DS=1,2) differisce da quella della collera in modo significativo. Un'ulteriore differenza è emersa dal confronto tra la tristezza e la gioia: gli apprendenti hanno riconosciuto con maggiore accuratezza gli enunciati di tristezza (M=2,4; DS=0,7) rispetto a quelli di gioia (M=1,4;DS=0,7), la differenza è significativa (p<0,05). Mentre non hanno avuto difficoltà in tal senso nella loro L1, in cui hanno totalizzato una media di riconoscimenti molto alta per entrambe le emozioni.

### 5.3.1.2 Analisi dei riconoscimenti erronei

TAB. 5-16 MATRICE ERRORI IN ITA L1 E RU L1

Emozioni	Collera		Tristezza		Gioia		Neutro	
	Ita L1	Ru L1	Ita L1	Ru L1	Ita L1	Ru L1	Ita L1	Ru L1
Collera	<b>26,6</b>	<b>43,3</b>	0	0	6,6	3,3	0	0
Tristezza	3,3	13,3	<b>80</b>	<b>80</b>	0	0	3,3	26,6
Gioia	0	13,3	0	0	<b>46,6</b>	<b>65</b>	3,3	0
Neutro	33,3	6,6	0	6,6	16,6	0	<b>76,6</b>	<b>75</b>
Sorpresa	20	13,3	0	0	16,6	13,3	10	0
Paura	16,6	10	13,3	13,3	10	0	6,6	0
Altro	0	0	6,6	0	3,3	0	0	0

L'analisi dei riconoscimenti erronei, o per meglio dire, diversi da quelli ipotizzati, ha evidenziato alcune differenze in relazione alle emozioni considerate e alla lingua in cui sono state espresse. In linea generale, l'opzione "altro" è stata poco utilizzata, per cui le percentuali maggiori si sono concentrate sulle altre opzioni proposte oltre a collera, gioia, tristezza e neutro, ovvero sorpresa e paura. Analizziamo nel dettaglio i risultati riportati nella TAB. 5-16.

Per quanto concerne la collera, sembrerebbe che le confusioni più consistenti (in relazione all'italiano) si siano verificate con l'eloquio neutro (33,3%), la cui percentuale supera quella della collera (26,6%). Da una prima analisi, si potrebbe dedurre che gli apprendenti russi abbiano avuto qualche difficoltà nel discernere un enunciato di collera da uno privo di sfumature emozionali. La collera è stata inoltre confusa con la sorpresa (20%) e in misura minore con la gioia (16,6), tutte emozioni ad alta attivazione. Soltanto il 3,3% delle risposte si è orientato sulla tristezza. Per quanto riguarda invece le emozioni espresse nella lingua nativa degli uditori, ovvero il russo, le confusioni sono di

entità minore. La collera viene scambiata per, gioia, sorpresa e tristezza nel 13,3 % e per paura nel 10% dei casi. Le confusioni con l'eloquio neutro sono di gran lunga inferiori rispetto all'italiano (6,6%).

Per quanto riguarda la tristezza, questa emozione viene confusa con la paura nel 13% dei casi, sia in italiano sia in russo. Evidentemente c'è un'associazione con sensazione di impotenza derivante da questi due stati emotivi (cfr. Anolli, Ciceri, 1992). Tuttavia, la percentuale di riconoscimenti, come già detto in precedenza, è molto elevata (80%), per cui si osservano errori di natura più lieve.

In relazione alla gioia, le confusioni maggiori derivano principalmente dall'italiano, lingua in cui questa emozione è stata scambiata con la sorpresa e con l'eloquio neutro nel 16,6% dei casi. In misura minore, è stata confusa con la paura (10%). In russo, invece, la gioia è stata principalmente confusa con la sorpresa (13,3%).

Infine, per quanto riguarda l'eloquio neutro in relazione all'italiano, esso è stato confuso maggiormente con la sorpresa (10%) e in misura minore con la paura (6,6%), la gioia e la tristezza (3,3%). In russo, invece, l'eloquio neutro è stato confuso esclusivamente con la tristezza (26,6%). Gli apprendenti hanno dunque avuto meno dubbi e difficoltà in lingua russa, orientandosi su uno stato emotivo di tristezza, che riflette la mancanza di attivazione fisiologica dell'eloquio neutro, mentre in italiano lo spettro di risposte si amplia e include anche emozioni ad alta attivazione come la sorpresa, la gioia e la paura, segno che gli apprendenti hanno percepito nelle produzioni dei nativi uno stato di attivazione potenzialmente associabile alle suddette emozioni.

### *5.3.2 Risultati del test uditivo sottoposto agli italiani nativi*

Prendiamo in esame i risultati relativi al test uditivo degli italiani. Il gruppo si componeva di 33 uditori (10 M e 13 F) di origine italiana, dai 16 ai 62 anni ( $M=34$ ,  $DS=11$ ). Il test conteneva 36 stimoli, di cui 12 in italiano L1, 12 in italiano L2 e 12 in russo. La probabilità di riconoscimento casuale oscillava anche in questo caso intorno al 14,3%. La TAB.5-17 illustra la statistica relativa ai riconoscimenti corretti, da cui si evince che gli italiani hanno riconosciuto in media 16,7 stimoli su 36, pari al 46% circa, con una deviazione standard di 3,6 (pari al 10%). Dal test binomiale risulta che la percentuale di decodifiche corrette supera la soglia della pura casualità ( $p=0,00001$ ;  $p<0,05$ ).

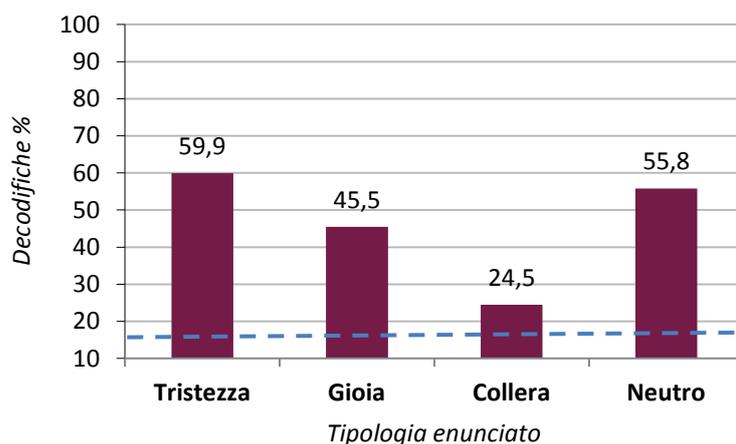
**TAB. 5-17. STATISTICA DECODIFICHE CORRETTE (GLOBALI) E RELATIVE PERCENTUALI**

<i>Statistica Decodifiche corrette</i>	M	%
Media	16,72	46,46
Errore standard	0,62	1,74
Mediana	17	47,2
Moda	18	50
Deviazione standard	3,61	10,04
Varianza	13,07	100,92
Range	15	41,66
Minimum	9	25
Maximum	24	66,66

### 5.3.2.1. Riconoscimenti corretti in base alla tipologia di enunciato

Se si osservano le percentuali globali relative alle tre produzioni emotive (tristezza, gioia e collera) e all'eloquio neutro (cfr. FIG. 5-13), si noter  che l'emozione riconosciuta con maggiore accuratezza   stata la tristezza (59,9%), seguita dalla gioia (45,5%). La collera, invece, ha ottenuto una percentuale nettamente inferiore (24,5%). L'eloquio neutro ha ottenuto in media il 55,8% di decodifiche corrette.

**FIG.5-13. MEDIA % DECODIFICHE GLOBALI TOTALIZZATE DAGLI UDITORI ITALIANI IN BASE ALLA TIPOLOGIA DI ENUNCIATO**



Al fine di verificare se le differenze osservate avessero significativit  statistica, sono stati condotti tre t test per campioni appaiati a due code. I confronti operati sono stati i seguenti:

- a. tristezza-gioia
- b. tristezza-collera

- c. gioia-collera
- d. neutro-gioia
- e. neutro-collera

I risultati ottenuti sono riportati nella TAB. 5-18. In sintesi, gli enunciati di tristezza (M=5,4; DS=1,6) sono stati riconosciuti con maggiore accuratezza rispetto alla collera (M=2,2; DS=1,6) e alla gioia (M=4; DS=1,6) e queste differenze sono significative ( $p < 0,05$ ). La media dei riconoscimenti della gioia, più elevata rispetto alla media della collera, differisce in modo significativo. Particolare attenzione va riservata all'eloquio neutro (M=5; DS=1,6) che è stato identificato mediamente con una percentuale più elevata (55,8%) rispetto alla collera e alla gioia: in entrambi i casi, i risultati dei due t test hanno evidenziato che la differenza è statisticamente significativa ( $p < 0,05$ ).

**TAB. 5-18. RISULTATI T TEST: CONFRONTO MEDIE RICONOSCIMENTI CORRETTI IN BASE ALLA TIPOLOGIA DI ENUNCIATO.**

Tipologia enunciato (E)									
TRISTEZZA			GIOIA			Test t (IC 95%)			
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)	
5,4	1,6	33	4	1,6	33	4,05	32	<b>0,0002*</b>	
TRISTEZZA			COLLERA			Test t (IC 95%)			
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)	
5,4	1,6	33	2,2	1,6	33	8,3	32	<b>0,00000001*</b>	
GIOIA			NEUTRO			Test t (IC 95%)			
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)	
4	1,6	33	5	1,6	33	2,1	32	<b>0,03*</b>	
GIOIA			COLLERA			Test t (IC 95%)			
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)	
4	1,6	33	2,2	1,6	33	5,2	32	<b>0,00001*</b>	
COLLERA			NEUTRO			Test t (IC 95%)			
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)	
2,2	1,6	33	5	1,6	33	6,4	32	<b>0,0000002*</b>	

### 5.3.2.2 Riconoscimenti globali in funzione della lingua dei locutori

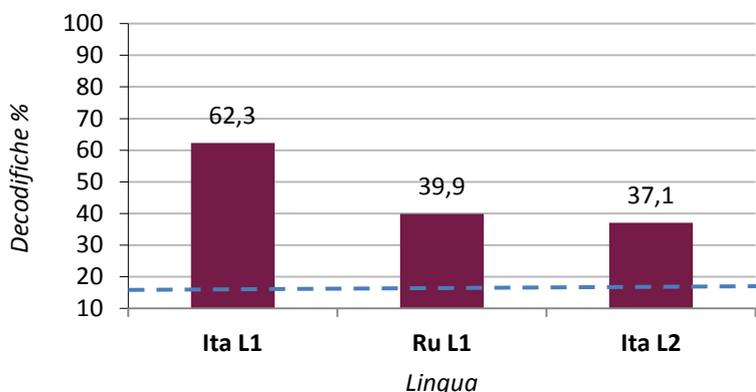
Da una prima analisi del computo delle medie totalizzate dagli uditori italiani in relazione alle tre modalità d'eloquio (italiano L1, italiano L2 e russo L1), emergono

alcune disparità (cfr. TAB.5-19). In media, gli stimoli pronunciati dagli italiani nativi sono stati riconosciuti con maggiore accuratezza (M=62,3; DS=17,1) rispetto alle produzioni degli apprendenti in italiano L2 (M=37,1; DS=11,2) e in russo (M=39,8; DS=14,8), come si evince dall'istogramma mostrato nella FIG. 5-14.

TAB. 5-19. STATISTICA DECODIFICHE CORRETTE IN FUNZIONE DELLA LINGUA DEI LOCUTORI

Statistica decodifiche corrette in base alla lingua dei locutori	<i>Ita L1</i>	<i>Ru L1</i>	<i>Ita L2</i>
Media %	62,3	39,8	37,1
Errore standard	2,9	2,55	1,9
Mediana	66,6	41,6	41,6
Moda	66,6	33,3	41,67
Deviazione standard	17,1	14,8	11,2
Varianza	295,6	220,31	126,2
Range	66,6	58,3	41,6
Minimum	25	8,3	8,3
Maximum	91,6	66,6	50

FIG.5-14. MEDIA % DECODIFICHE CORRETTE PER LINGUA (ITALIANO L1, ITALIANO L2 E RUSSO L1)



Al fine di verificare se le differenze osservate fossero statisticamente significative, sono stati condotti due t test per campioni appaiati a due code. Le medie considerate fanno riferimento ai riconoscimenti globali totalizzati per ciascuna lingua. I confronti operati sono stati i seguenti:

- Ita L1- Ita L2;
- Ita L1- Ru L1

Dai risultati del t-test (cfr. TAB. 5-20) è emerso che gli stimoli somministrati nella lingua nativa dei locutori sono stati riconosciuti con maggiore accuratezza sia

rispetto alle produzioni realizzate dagli apprendenti in italiano sia rispetto a quelle in russo. In entrambi i casi, la differenza è significativa ( $p < 0,05$ ). In effetti, se si osserva la media dei riconoscimenti in funzione della lingua, si noterà che per l'italiano nativo questa oscilla intorno a 7,4 decodifiche corrette su un totale di 12 stimoli. La media subisce un brusco decremento sia in italiano L2 ( $M=4,4$ ) che in russo ( $M=4,7$ ): in sostanza gli uditori italiani hanno identificato più facilmente gli stimoli presentati nella loro lingua nativa (*in group-advantage*) rispetto a quelli in russo, ma anche rispetto a quelli pronunciati dagli apprendenti in italiano L2.

TAB. 5-20. RISULTATI T TEST: CONFRONTO ITA L1-ITA L2; ITA L1-RU L1

Decodifiche corrette						Test t (IC 95%)		
ITA L1			ITA L2					
M	DS	N.	M	DS	N.	t	gdl	P(T<=t)
7,4	4,2	33	4,4	1,8	33	8,8	32	<b>0,000000004*</b>
ITA L1			RU L1			Test t (IC 95%)		
M	DS	N.	M	DS	N.	T	Gdl	P(T<=t)
7,4	4,2	33	4,7	3,1	33	6,8	32	<b>0,0000001*</b>

### 5.3.2.3 Riconoscimenti corretti per tipologia di enunciato in Ita L1, Ita L2 e Ru L1

Se si osservano le percentuali ottenute per ciascuna lingua (italiano nativo, russo e italiano non nativo), emergono ulteriori disparità. Come si può evincere dalla TAB. 5-21, in relazione alle produzioni realizzate dagli italiani, emozioni come la tristezza e la gioia sono state riconosciute con percentuali molto elevate (70,7%) rispetto alla collera, che ha ottenuto invece soltanto il 27,2% di riconoscimenti. L'eloquio neutro è stato identificato con maggiore accuratezza (80% di decodifiche corrette).

TAB. 5-21. MEDIA % DECODIFICHE CORRETTE PER TIPOLOGIA DI ENUNCIATO IN ITA L1, RU L1 E ITA L2

	% Lingua		
	Ita L1	Ru L1	Ita L2
Tristezza	70,7	58,5	50,5
Gioia	70,7	36,3	29,2
Collera	27,2	20,2	26,6
Neutro	80,8	44,4	42,2
<b>TOT.</b>	<b>62,3</b>	<b>39,9</b>	<b>37,1</b>

Per verificare se queste differenze fossero significative, sono stati condotti due test t per confrontare le medie dei riconoscimenti corretti totalizzate rispettivamente dagli enunciati di collera (M=0,8; DS=0,9) rispetto alle altre tipologie: tristezza (M=2,1; DS=1), gioia (M=2,1; DS=0,8) e neutro (M=2,4; DS=0,7). Dai risultati sono emerse differenze statisticamente significative ( $p < 0,05$ ) in tutti e tre i casi (cfr. TAB. 5-22), il che suggerisce che la decodifica di questo stato emotivo sia stata più problematica, o quanto meno che emozioni come la tristezza e in misura minore la gioia siano state più facilmente identificabili. Anche dal confronto con l'eloquio neutro, la differenza è significativa.

TAB. 5-22. RISULTATI T TEST ITA L1: CONFRONTO MEDIE RICONOSCIMENTI PER ENUNCIATO

		COLLERA			TRISTEZZA			Test t (IC 95%)		
Ita L1		M	DS	N.	M	DS	N.	t	gdl	P(T<=t)
		0,8	0,9	33	2,1	1,02	33	5,33	32	<b>0,000007*</b>
		COLLERA			GIOIA			Test t (IC 95%)		
Ita L1		M	DS	N.	M	DS	N.	t	gdl	P(T<=t)
		0,8	0,9	33	2,1	0,8	33	5,3	32	<b>0,000007*</b>
		COLLERA			NEUTRO			Test t (IC 95%)		
Ita L1		M	DS	N.	M	DS	N.	t	gdl	P(T<=t)
		0,8	0,9	33	2,4	0,7	33	6,9	32	<b>0,0000003*</b>

La situazione cambia per quanto riguarda le produzioni in italiano L2 (cfr. TAB. 5-23) che sono state riconosciute con maggiore difficoltà: le percentuali relative alla collera (26,6%) e alla gioia (29,2%) sono piuttosto basse, se confrontate con quelle totalizzate dalla tristezza (50,5%), e dall'eloquio neutro (42,2%). Dal confronto delle medie, sono emerse differenze statisticamente significative ( $p < 0,05$ ) in relazione alla tristezza (M=1,5; DS=0,7), che è stata identificata con maggiore accuratezza rispetto a collera (M=0,8; DS=0,4) e gioia (M=0,9; DS=0,7), ma non rispetto al neutro (M=1,3; DS=0,9), la differenza non è significativa ( $p > 0,05$ ).

TAB. 5-23. RISULTATI T TEST ITA L2: CONFRONTO MEDIE RICONOSCIMENTI PER ENUNCIATO

		Tipologia enunciato (E)							Test t (IC 95%)	
		TRISTEZZA			COLLERA					
Ita L2		M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
		1,5	0,7	33	0,8	04	33	4	32	<b>0,0003*</b>

Ita L2	TRISTEZZA			GIOIA			Test t (IC 95%)		
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
	1,5	0,7	33	0,9	0,7	33	3,5	32	<b>0,001*</b>

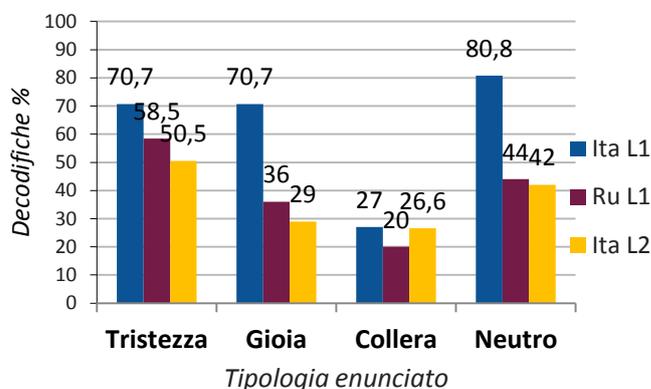
Se si osservano le percentuali relative al russo, si noterà che anche in questo caso, emozioni come tristezza (58,5%) e gioia (36,3%) sono state riconosciute con percentuali maggiori rispetto alla collera (20,2%). L'eloquio neutro ha ottenuto invece il 44,4% di decodifiche corrette. Queste percentuali riflettono in un certo senso lo stesso andamento osservato per l'italiano L2. Dal confronto delle medie (cfr. TAB. 5-24), è emerso che la tristezza (M=1,7; DS=0,9) differisce significativamente dalla collera (M=0,6; DS=0,8), dalla gioia (M=1; DS=0,9) e dal neutro (M=1,3; DS=0,8): in tutti e tre i casi  $p < 0,05$ . In sostanza l'emozione maggiormente riconosciuta è stata la tristezza, mentre la collera ha totalizzato una media di riconoscimenti nettamente più bassa rispetto alle altre tipologie di enunciato (compreso quello neutro), e questa differenza è statisticamente significativa ( $p < 0,05$ ).

TAB. 5-24. RISULTATI T TEST RU L1: CONFRONTO MEDIE RICONOSCIMENTI PER ENUNCIATO

Tipologia enunciato (E)									
Ru L1	TRISTEZZA			COLLERA			Test t (IC 95%)		
	M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)
	1,7	0,9	33	0,6	0,8	33	6,2	32	<b>0,000005*</b>
Tipologia enunciato (E)									
Ru L1	TRISTEZZA			GIOIA			Test t (IC 95%)		
	M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)
	1,7	0,9	33	1	0,9	33	3	32	<b>0,004*</b>
Tipologia enunciato (E)									
Ru L1	TRISTEZZA			NEUTRO			Test t (IC 95%)		
	M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)
	1,7	0,9	33	1,3	0,8	33	2,2	32	<b>0,02*</b>
Tipologia enunciato (E)									
Ru L1	COLLERA			GIOIA			Test t (IC 95%)		
	M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)
	0,6	0,8	33	1	0,9	33	2,3	32	<b>0,02*</b>
Tipologia enunciato (E)									
Ru L1	COLLERA			NEUTRO			Test t (IC 95%)		
	M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)
	0,6	0,8	33	1,3	0,8	33	3,3	32	<b>0,002*</b>

L'istogramma mostrato nella FIG. 5-15 permette di osservare le differenze nel riconoscimento delle emozioni rilevate all'interno delle tre produzioni in relazione alla tipologia di enunciato.

**FIG.5-15. MEDIA % DECODIFICHE CORRETTE PER TIPOLOGIA DI ENUNCIATO IN ITA L1, RU L1 E ITA L2**



Un caso speciale è costituito dalla collera, che ha ottenuto percentuali piuttosto basse in tutte e tre le produzioni: 27,2% per l'italiano L1, 20,2% per il russo e 26,6% per l'italiano L2. Sembrerebbe quindi che, alla base, gli uditori italiani abbiano avuto maggiori difficoltà nel discriminare questo stato emotivo dagli altri, come si vedrà più avanti nei paragrafi dedicati all'analisi dei riconoscimenti erronei. In questo caso, le percentuali più alte sono per l'italiano L1 ed L2, mentre in russo è stata riconosciuta con percentuali inferiori. Infine, per quanto concerne l'eloquio neutro, anche gli italiani, come i russi, hanno riconosciuto un enunciato privo di sfumature emozionali con percentuali molto elevate in italiano (80,2%). Mentre in russo (44,4%) e in italiano L2 (42,2%) le percentuali sono quasi dimezzate.

Per verificare se la media dei riconoscimenti per tipologia di enunciato differisse in modo significativo nell'italiano nativo, non nativo e in russo, sono stati condotti dei confronti incrociati: Ita L1- Ita L2; Ita L1-Ru L1, Ita L2-Ru L1.

a) *Ita L1-Ita L2*

TAB. 5-25. RISULTATI T TEST ITA L1-ITA L2: CONFRONTO MEDIE RICONOSCIMENTI (TRISTEZZA; GIOIA)

TRISTEZZA									
Ita L1			Ita L2			Test t (IC 95%)			
M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)	
2,1	0,7	33	1,5	0,5	33	3	32	<b>0,004*</b>	

GIOIA									
Ita L1			Ita L2			Test t (IC 95%)			
M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)	
2,1	0,7	33	0,9	0,3	33	6,2	32	<b>0,0000005*</b>	

Dal computo delle medie, le maggiori differenze si osservano per la tristezza che ha ottenuto percentuali più elevate in italiano L1 (70,7%) rispetto all'italiano L2 (50,5%). Nel caso della gioia, lo scarto è ancora maggiore: 70,7% per l'italiano nativo e soltanto il 29, 2% per l'italiano non nativo. Dai risultati del t test (cfr. TAB. 5-25) è emerso che, in entrambi i casi, la differenza tra le medie è statisticamente significativa ( $p < 0,05$ ). Per quanto concerne la collera e il neutro, non sono emersi discrepanze di particolare rilievo.

b) *Ita L1-Ru L1*

TAB. 5-26. RISULTATI T TEST ITA L1-RU L1: CONFRONTO MEDIE RICONOSCIMENTI (GIOIA; NEUTRO)

GIOIA									
Ita L1			Ru L1			Test t (IC 95%)			
M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)	
2,1	0,7	33	1	0,4	33	4,5	32	<b>0,00008*</b>	

NEUTRO									
Ita L1			Ru L1			Test t (IC 95%)			
M	DS	N.	M	DS	N.	t	Gdl	P(T<=t)	
2,4		33	1,3	0,4	33	6,1	32	<b>0,0000006*</b>	

Dal confronto delle medie totalizzate per ciascuna tipologia di enunciato nelle due lingue, sono emerse differenze significative ( $p < 0,05$ ) per quanto concerne la gioia (che in russo ha ottenuto soltanto il 36,3% di riconoscimenti, quasi la metà rispetto all'italiano) e l'eloquio neutro (le cui percentuali ruotano intorno al 44%), cfr. TAB. 5-26. In relazione alla collera e alla tristezza, i risultati del t test non hanno restituito evidenze significative ( $p > 0,05$ ).

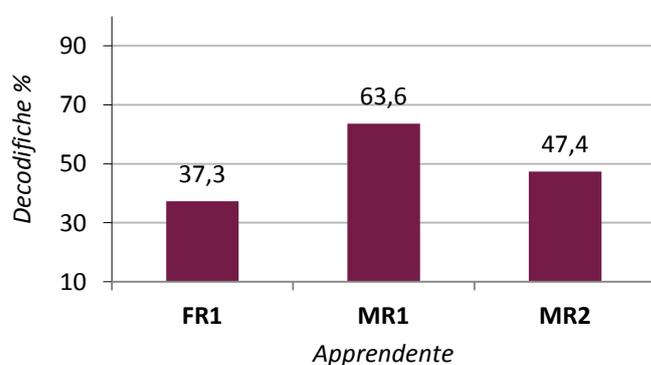
### *c) Ita L2-Ru L1*

In relazione al confronto Ita L2-Ru L1, sembrerebbe che gli italiani abbiano riconosciuto con maggiore accuratezza le emozioni espresse in russo che non quelle realizzate nella loro lingua madre da parlanti non nativi. Tuttavia, dai risultati del t test non sono emerse differenze significative in relazione alla tipologia di enunciato (in tutti i casi,  $p > 0,05$ ).

#### *5.3.2.4 Differenze di riconoscimento in funzione dei locutori russi*

Come illustrato nel CAP. 4, gli stimoli in italiano L2 presenti nel test sottoposto agli italiani contenevano le produzioni di tre apprendenti russi, 1 femmina (FR1), di livello B1 e 2 maschi (MR1 ed MR2), di livello B2. Mentre la prima era in Italia da 8 mesi al momento della registrazione, i secondi vivevano in Italia da 15 mesi e frequentavano regolarmente l'Università e gli amici di corso. Se si osserva la FIG. 5-16, si noterà che i due apprendenti russi hanno ottenuto globalmente le percentuali di riconoscimenti più elevate: 63,6% per MR1 e 47,4% per MR2. La donna, invece, ha ottenuto una percentuale inferiore: 37,3%. Queste differenze tra i locutori, tuttavia, non sono statisticamente significative ( $p > 0,05$ ).

**FIG.5-16. MEDIA % RICONOSCIMENTI TOTALIZZATI DALLE PRODUZIONI DEI TRE APPRENDENTI (FR1, MR1, MR2)**



TAB. 5-27 MEDIE % DECODIFICHE CORRETTI PER TIPOLOGIA DI ENUNCIATO

	% Apprendente		
	FR1	MR1	MR2
Tristezza	6	84,8	60,6
Gioia	9,09	15,1	63,3
Collera	39,3	36,3	3
Neutro	57,5	54,5	15,1
TOT.	37,3	63,6	47,4

Se si considerano invece le percentuali totalizzate per ciascuna emozione, emergono alcuni dati rilevanti. Come si evince dalla TAB.5-27, la tristezza è stata riconosciuta con percentuali molto alte sia nelle produzioni di MR1 (84,8%), sia nelle produzioni di MR2 (60,6%), mentre l'enunciato realizzato da FR1 ha ottenuto soltanto il 6% di riconoscimenti.

Sono stati condotti tre t test (cfr. TAB. 5-28) per confrontare rispettivamente le medie di MR1-MR2, MR1-FR1, FR1-MR2 (cfr. TAB. 5-27). Dal confronto è emerso che la media dei riconoscimenti totalizzata da FR1 differisce significativamente da quella di MR1 e di MR2 ( $p < 0,05$ ). Anche la differenza tra la media di MR1 ed MR2 è significativa. In sostanza, sembrerebbe che MR1 (che ha ottenuto mediamente 0,7 riconoscimenti) sia riuscito a comunicare con maggiore efficacia lo stato emotivo in questione.

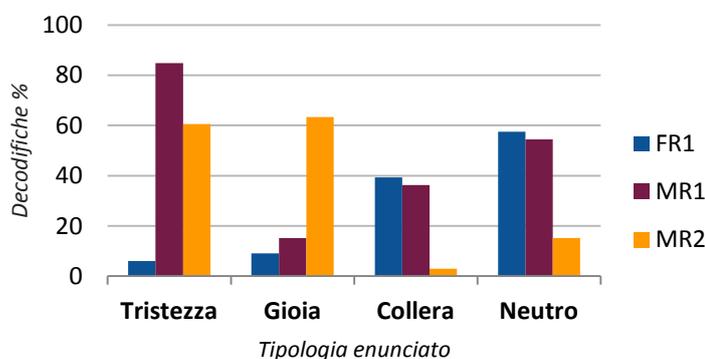
Per quanto riguarda la gioia, le produzioni di MR2 sono state riconosciute con percentuali sostanzialmente più elevate (63,3%) rispetto a MR1 (15,1%) e a FR1 (9%), la differenza è statisticamente significativa ( $p < 0,05$ ).

In relazione alla collera, la situazione si ribalta: le produzioni di FR1 sono state riconosciute nel 39,3% dei casi e in misura minore quelle di MR1 (36,3%). Il parlato di MR2 ha ottenuto invece soltanto il 3% di decodifiche, la differenza rispetto alle altre produzioni è statisticamente significativa ( $p < 0,05$ ).

Infine, in relazione all'eloquio neutro le produzioni di FR1 (57,5%) e di MR1 (54,5%) sono state identificate come neutrali con percentuali nettamente maggiori a quelle ottenute da MR2 (15,1%), la differenza in entrambi i casi è statisticamente significativa ( $p < 0,05$ ).

Queste differenze sono osservabili più intuitivamente nella FIG. 5-17, che mette a confronto le percentuali di riconoscimento ottenute dagli apprendenti in relazione alla tipologia di enunciato.

**FIG.5-17. MEDIA % RICONOSCIMENTI TOTALIZZATI DALLE PRODUZIONI DEI TRE APPRENDENTI (FR1, MR1, MR2) IN BASE ALLA TIPOLOGIA DI ENUNCIATO**



In base ai risultati ottenuti sembrerebbe che ciascuno dei tre apprendenti abbia avuto qualche difficoltà a veicolare almeno un'emozione. Ad esempio, l'enunciato di gioia realizzato da MR1 ha ottenuto percentuali di riconoscimento molto basse, il che lascerebbe supporre che sia stato poco efficace sul piano comunicativo. MR2 mostra difficoltà diverse da MR1, in quanto riesce ad esprimere la gioia con maggiore efficacia rispetto agli altri apprendenti, tuttavia, l'enunciato di collera prodotto da MR2 ha ottenuto le percentuali più basse in assoluto. Anche l'eloquio neutro è stato riconosciuto con minor accuratezza rispetto alle produzioni degli altri due. Infine, FR1 evidenzia chiare difficoltà a veicolare due emozioni su tre: le espressioni di tristezza e gioia hanno totalizzato percentuali nettamente inferiori a quelle di MR1 ed MR2. Tuttavia, riesce ad esprimere con maggiore efficacia la collera, ottenendo percentuali più alte rispetto agli altri due (in particolare rispetto ad MR2). Anche l'enunciato neutro ha totalizzato una media di riconoscimenti più elevata.

TAB. 5-28. T-TEST: CONFRONTO MEDIE RICONOSCIMENTI MR1-MR2;MR2-FR1

	MR1			MR2			Test t (IC 95%)		
	M	DS	N.	M	DS	N.	T	gdl	P(T<=t)
<b>Tristezza</b>	0,7	0,4	33	0,2	0,5	33	5,2	32	<b>0,00001*</b>
<b>Gioia</b>	0,2	0,4	33	0,6		33	4,9	32	<b>0,00002*</b>
<b>Collera</b>	0,4	0,5	33	0,03	0,2	33	3,5	32	<b>0,001*</b>
<b>Neutro</b>	0,5	0,5	33	0,2	0,4	33	4,5	32	<b>0,00007*</b>

	MR2			FR1			T	gdl	P(T<=t)
	M	DS	N.	M	DS	N.			
<b>Tristezza</b>	0,2	0,5	33	0,8	0,2	33	5,7	32	<b>0,000002*</b>
<b>Gioia</b>	0,6	0,5	33	0,09	0,3	33	5,5	32	<b>0,000003*</b>
<b>Collera</b>	0,03	0,2	33	0,4	0,5	33	3,8	32	<b>0,0006*</b>
<b>Neutro</b>	0,2	0,4	33	0,6	0,5	33	3,9	32	<b>0,0003*</b>

### 5.3.2.5 Analisi dei riconoscimenti erronei

TAB. 5-29. MATRICE ERRORI IN ITA L1, RU L1 E ITA L2

Lingua	Stimolo	Risposte						
		Collera	Tristezza	Gioia	Neutro	Sorpresa	Paura	Altro
Ita L1	Collera	<b>27,2</b>	6	7	21,2	21,2	17,1	0
	Tristezza	2	<b>70,7</b>	1	7	0	13,1	6
	Gioia	4	2	<b>70,7</b>	6	12,1	4	1
	Neutro	3	4	0	<b>80,8</b>	7	1	4
Ru L1	Collera	<b>20,2</b>	8	19,1	11,1	28,2	13,1	0
	Tristezza	2	<b>58,5</b>	1	16,1	12,1	5	5
	Gioia	10,1	2	<b>36,3</b>	2	28,2	21,2	0
	Neutro	9	34,3	1	<b>44,4</b>	4	7	0
Ita L2	Collera	<b>26,2</b>	5	7	34,3	19,1	8	0
	Tristezza	15,1	<b>50,5</b>	1	18,1	8	3	4
	Gioia	17,1	0	<b>29,2</b>	5	42,4	5	1
	Neutro	7	34,3	2	<b>42,4</b>	8	5	1

L'analisi dei riconoscimenti erronei (cfr. TAB. 5-29) ha evidenziato alcune confusioni ricorrenti per quanto riguarda l'italiano L1. Gli enunciati di collera sono stati confusi maggiormente con la sorpresa e con l'eloquio neutro (21,2%), in misura minore con la paura (17,1%). La tristezza è stata confusa con la paura nel 13,1% dei casi e con

l'eloquio neutro nel 7% dei casi, percentuali molto basse che fanno notare l'elevata percentuale di riconoscimenti corretti ottenuta da questo stato emotivo (70,7%). Anche la gioia è stata confusa in misura molto lieve con la sorpresa (12,1%). Per quanto riguarda l'eloquio neutro, le confusioni sono di entità ancora minore: nel 7% dei casi è stato confuso con la sorpresa o identificato come tristezza o altro (4%).

In sintesi, la collera è stata confusa con l'eloquio neutro nel 21,2% dei casi, ma non è avvenuto il contrario, cioè l'eloquio neutro non è stato interpretato come collera (soltanto nel 3% dei casi). La tristezza è stata confusa con la paura e la gioia con la sorpresa.

Per quanto riguarda le produzioni in russo L1, si può notare come gli enunciati di collera siano stati oggetto di diverse interpretazioni: sono stati confusi principalmente con la sorpresa (28,2%), con la gioia (19,1%), e in misura minore con la paura (13,1%) e con l'eloquio neutro (11,1%). Sembrerebbe che questa emozione, espressa in una lingua straniera, abbia destato maggiori problemi in fase di decodifica.

La tristezza è stata confusa con l'eloquio neutro nel 16,1% dei casi, e con la sorpresa nel 12,1%. Mentre la gioia, che ha ottenuto una percentuale di riconoscimenti inferiore rispetto alla tristezza (36,3%), è stata confusa maggiormente con la sorpresa (28,2%) e con la paura (21,2%). Da notare che il 10% delle risposte si è orientato sulla collera. Infine per quel che concerne l'eloquio neutro, questo è stato principalmente confuso con la tristezza (34,3%). Le altre percentuali sono di gran lunga inferiori (9% per la collera, 7% per la paura).

In relazione alle produzioni in italiano L2, anche in questo caso, le confusioni maggiori sono legate agli enunciati di collera, che viene scambiata con il neutro nel 34,3% dei casi e con la sorpresa nel 19,1%. La tristezza è stata scambiata con l'eloquio neutro nel 18,1% e con la collera nel 15,1%. La gioia è stata invece confusa in misura maggiore con la sorpresa (42,4%) e con la collera (17,1%). In relazione all'eloquio neutro, esso è stato confuso principalmente con la tristezza (34,3%).

### *5.3.3 Discussione dei dati relativi al processo di decoding*

I risultati dei test uditivi hanno contribuito a chiarire in che misura le produzioni emotive oggetto di indagine venissero decodificate dai due gruppi di uditori, aventi una diversa L1. In primo luogo, le percentuali di decodifiche corrette ottenute in entrambi i test superano la soglia della pura casualità. Questo significa che le tre emozioni considerate (e l'eloquio neutro) sono state discriminate a prescindere dalla lingua degli

*encoder*. Questo dato è in linea con gli studi precedenti che hanno confermato la capacità degli esseri umani nel decodificare il parlato emotivo anche in una lingua diversa dalla propria (cfr. Scherer, Banse, Wallbott 2001; Scherer *et al.* 2003; Thompson, Balkwill 2006; Galatà, Romito 2010). Ciò non toglie che, osservando le percentuali ottenute dalle due lingue, non sia possibile cogliere dati altrettanto significativi e coerenti con l'ipotesi che vi siano delle peculiarità linguo-specifiche in grado di spiegare le basse percentuali di riconoscimento e l'elevato numero di riconoscimenti erronei associati ad alcune emozioni.

#### *Il test uditivo degli apprendenti russi*

Facendo riferimento al test uditivo somministrato agli apprendenti russi, è possibile concludere che questi ultimi abbiano riconosciuto con maggiore accuratezza gli enunciati di tristezza, sia in italiano sia nella loro lingua nativa. In relazione alla gioia, le difficoltà sono state maggiori in italiano rispetto al russo, in cui le percentuali sono significativamente più alte. La collera è stata l'emozione meno riconosciuta in assoluto, in entrambe le lingue. Tuttavia, gli stimoli presentati nella L1 degli studenti sono stati identificati con percentuali significativamente più elevate. In sostanza, i partecipanti hanno evidenziato una maggiore capacità nel riconoscere gli stimoli nella propria lingua nativa (in linea con l'*ingroup-advantage hypothesis*, cfr. Elfenbien, Ambady, 2002). Da rilevare inoltre che l'eloquio neutro è stato identificato con percentuali molto elevate in entrambe le lingue, segno che gli stimoli somministrati riflettevano, in modo quasi unanime, enunciati privi di evidenti sfumature emozionali.

Considerando invece l'analisi dei riconoscimenti erronei, è possibile osservare ulteriori differenze. Ad esempio, gli enunciati di collera in italiano L1 sono stati confusi in misura maggiore con l'eloquio neutro: l'elevata percentuale associata a questa risposta (33%) sembra suggerire che le produzioni emotive degli italiani siano state poco salienti sul piano comunicativo secondo il giudizio degli apprendenti. È ragionevole supporre che gli indizi acustici sfruttati dagli *encoder* italiani non siano stati percepiti e associati a quei segnali che caratterizzano l'espressione della collera nella lingua nativa degli apprendenti. Per quanto riguarda la tristezza, questa emozione è stata confusa in entrambe le lingue con la paura (13%), un'associazione già evidenziata anche in altri studi (cfr. Anolli, Ciceri, 1992) che riflette la condizione di impotenza comune ad entrambi gli stati emotivi. In relazione alla gioia, le confusioni maggiori si sono

verificate con gli stimoli in italiano, che sono stati identificati principalmente con l'eloquio neutro e con la sorpresa nel 16,6% dei casi.

### *Il test uditivo degli italiani*

Considerando i risultati del test uditivo sottoposto al gruppo di italiani, emergono alcune tendenze simili in relazione ai riconoscimenti corretti ed erranei. Anche in questo caso, l'emozione identificata con maggiore accuratezza è stata la tristezza, seguita dalla gioia. La collera ha ottenuto percentuali nettamente inferiori rispetto alle altre emozioni, come già riportato nel test sottoposto agli apprendenti. Inoltre, anche gli italiani hanno riconosciuto l'eloquio neutro con percentuali molto elevate. Tuttavia come già evidenziato, l'accuratezza nel riconoscimento è significativamente inferiore se si considerano le percentuali totalizzate da ciascuna emozione in Ita L1, Ita L2 e Ru L1. Ne è un esempio la gioia, che ha totalizzato il 70% di decodifiche corrette in Ita L1, mentre in Ita L2 (29,2%) e in Ru L1 (36,3%) le percentuali sono praticamente dimezzate.

Prendendo in esame i riconoscimenti erranei, si è osservato che le maggiori confusioni si sono concentrate sulle produzioni emotive degli apprendenti in russo e in italiano L2. Ad esempio, gli enunciati di collera sono stati oggetto di diverse interpretazioni: sorpresa (28,2%), gioia (19,1%), e in misura minore, paura (13,1%) e neutro (11,1%). Mentre la gioia, che ha ottenuto una percentuale di riconoscimenti inferiore rispetto alla tristezza (36,3%), è stata confusa maggiormente con la sorpresa (28,2%) e con la paura (21,2%). In relazione alle produzioni in italiano L2, anche in questo caso, le confusioni maggiori sono legate agli enunciati di collera, confusa con il neutro e la sorpresa. Anche la gioia è stata confusa principalmente con la sorpresa (42,4%) e con la collera (17,1%).

L'alta variabilità di risposte, pur riflettendo alcune associazioni tra emozioni affini già rilevate da studi precedenti, è segno che vi siano state maggiori difficoltà nel definire ed etichettare il parlato emotivo degli apprendenti in italiano e nella loro L1 (un'elevata percentuale di risposte si è infatti concentrata sull'eloquio neutro). In relazione agli stimoli prodotti dagli italiani, soltanto la collera ha destato alcune difficoltà, dirottando le risposte principalmente sulla sorpresa e sul neutro (21%).

### *L'efficacia comunicativa delle produzioni degli apprendenti FR1, MR1 e MR2*

In relazione alle produzioni emotive dei tre apprendenti, sembrerebbe che ognuno di loro abbia avuto qualche difficoltà a veicolare almeno un'emozione. FR1, l'apprendente che al momento delle registrazioni si trovava in Italia da un periodo di tempo inferiore rispetto agli altri, ha evidenziato maggiori difficoltà nell'espressione della tristezza e della gioia. Può darsi che l'apprendente in questione non riesca ancora a modulare in modo efficace le risorse prosodiche della L2, in particolare l'EM. Infatti, le produzioni di FR1 si distinguevano, sul piano acustico e percettivo, per una scarsa variabilità sul piano tonale, mentre il parlato di MR1 ed MR2 mostrava maggiori variazioni e risultava più modulato. La conoscenza più profonda della cultura italiana da parte dei due apprendenti MR1 ed MR2 (in Italia da 15 mesi) può averli agevolati nella comprensione e nella gestione dei contesti emotivi proposti, permettendogli di sfruttare meglio le risorse della L2 rispetto a FR1, che ha mostrato difficoltà in più contesti. Ciò nonostante, l'enunciato di gioia realizzato da MR1 ha ottenuto percentuali di riconoscimento molto basse, il che lascerebbe supporre che sia stato poco efficace sul piano comunicativo, così come le produzioni di collera di MR2.

### *Il Modello della Lente e il Sistema delle Display Rules: una possibile interpretazione dei dati*

Come già illustrato nel CAP. 3, il modello della lente di Brunswik (1956), rielaborato da Scherer (1978; 1982) e adattato alla comunicazione vocale delle emozioni, fornisce una serie di spiegazioni sul perché alcune emozioni non vengano riconosciute o siano confuse con altre. Riprendendo la terminologia di Brunswik, Scherer afferma che i segnali vocali emessi dal parlante costituiscono degli indicatori distali dell'emozione (*distal cues*), non ancora percepiti dal ricevente. In seguito, diventano degli indicatori prossimali (*proximal cues*), e sono convogliati nel processo percettivo del ricevente che li elabora a partire da categorie soggettive legate alla discriminazione della voce e delle sue modificazioni (Juslin, Scherer, *ivi*, p. 83). La differenza tra i due indicatori è utile per comprendere la complessità dell'intero processo che, in alcuni casi, non si conclude con la decodifica corretta dell'emozione espressa dal parlante. I segnali vocali possono essere interpretati erroneamente e attribuiti ad uno stato emotivo differente rispetto a quello espresso. Il processo di attribuzione di uno stato emotivo si fonda infatti sia su meccanismi biologici sia su meccanismi cognitivi di

carattere inferenziale, basati sulle esperienze pregresse, sulle conoscenze condivise dai partecipanti in merito a convenzioni e norme sociali. Infatti, se da un lato le variazioni sul piano dell'espressione si devono a determinate alterazioni fisiologiche (tensione delle pliche vocali, modificazione del ritmo respiratorio, ecc.) dall'altro l'attivazione del comportamento emozionale viene preceduta dal sistema delle *display rules* (Ekman 1982), ovvero le regole di esibizione che monitorano la manifestazione delle emozioni, intensificando, riducendo, simulando o neutralizzando l'espressione emotiva (vocale e non verbale) in base a norme culturali apprese e socialmente condivise (cfr. CAP. 1).. Come ricordano Anolli e Ciceri (1992), la natura convenzionale di tali strategie facilita in un certo senso la comunità di parlanti, fornendo una sorta di codice di riconoscimento, che permette di individuare i "tratti pertinenti e salienti del piano del contenuto e del piano dell'espressione" (ivi, p. 391). A partire da essi, "viene modulata la corrispondente espressione vocale che apparirà ad un interlocutore che condivide le medesime regole di esibizione, come appropriata e immediatamente riconoscibile" (*ibid.*).<sup>72</sup> In questo senso, la maggiore accuratezza nel riconoscimento delle emozioni nella L1 degli uditori dovrebbe confermare questa capacità di cogliere le minime variazioni dell'eloquio nella lingua materna e di categorizzarle in base alle esperienze pregresse e alle norme culturali. D'altro canto, la riduzione dell'indice di accuratezza in una lingua non nativa sembra denotare quantomeno difficoltà maggiori nell'interpretazione di indizi vocali, forse perché alcuni elementi non rientrano nelle strategie e nei codici convenzionali adottati dal sistema culturale della L1. In effetti, la maggiore variabilità che caratterizzava il parlato degli italofoeni soprattutto sul piano melodico ha orientato le risposte degli apprendenti verso stati emotivi che condividevano livelli di attivazione simili, per questa ragione molto spesso la gioia è stata confusa con la paura o con la sorpresa<sup>73</sup>. Viceversa, la poca dinamicità che distingueva il parlato dei russi (nella loro L1 ma soprattutto in L2) ha portato gli uditori italofoeni a ricorrere più volte all'opzione "neutro", il che potrebbe suggerire che la ridotta escursione melodica abbia in qualche modo ostacolato la decodifica dell'emozione veicolata dal parlante proprio in virtù della scarsa salienza sul piano percettivo.

---

<sup>72</sup> Tuttavia, non va sottovalutato il fatto che le produzioni in italiano L2 degli apprendenti sono state realizzate con un *pitch range* più ridotto rispetto a quello rilevato per la L1, il che potrebbe essere dovuto anche alla difficoltà di modulare le strutture intonative della L2 in modo appropriato.

<sup>73</sup> Un caso speciale è rappresentato dalla collera, confusa sia con l'eloquio neutro sia con altre emozioni ad alta attivazione, come sorpresa e gioia.

Questi due aspetti, ovvero la maggiore accuratezza nella lingua nativa e l'alta variabilità di risposte in quella non nativa, sembrerebbero confermare non solo l'ipotesi dell'*ingroup-advantage*, ma anche quella di un possibile condizionamento linguo-specifico (come l'escursione melodica) strettamente legato dunque al contesto linguistico e culturale di provenienza. Come già esposto nei capp. 2 e 3, gli studi che hanno indagato la percezione degli andamenti intonativi paralinguistici legati al codice della frequenza e al codice dello sforzo in una lingua non nativa (Chen, Gussenhoven e Rietveld 2004; Chen, 2005) hanno suggerito l'ipotesi che l'implementazione di questi due codici biologici dipenda da fattori universali nonché linguo-specifici. Questo aspetto andrebbe dunque approfondito e supportato da ulteriori analisi.

In conclusione, l'indagine acustico-uditiva ha confermato che la comunicazione vocale delle emozioni non solo si fonda su meccanismi biologici, ma riflette anche una serie di variabili quali le differenze individuali tra i parlanti, le interazioni tra i bisogni individuali, il contesto sociale e culturale e l'immagine di sé (Juslin, Scherer, 2008). In questo senso, la comunicazione emotiva (cfr. cap. 2), dunque intenzionale, appare necessariamente legata al tessuto culturale e sociale in cui si realizza e risulta mediata da pratiche comunicative socialmente condivise che guidano il comportamento espressivo di un individuo affinando le sue capacità di esprimere e decodificare specifici significati emotivi attraverso i segnali vocali.



# Conclusioni

---

L'obiettivo della presente ricerca era quello di verificare le difficoltà sperimentate da un gruppo di apprendenti di origine russa con un livello di competenza intermedio in contesti comunicativi che implicassero la riflessione metalinguistica nonché il ricorso alla prosodia emotiva della lingua target per veicolare e interpretare significati emotivi.

In relazione al parlato emotivo degli apprendenti, i risultati ottenuti lasciano supporre che gli studenti siano riusciti a modulare le risorse prosodiche della L2, tuttavia, ad un attento esame delle differenze e delle similarità riportate, questa elasticità non sembra essersi manifestata in egual misura in tutte le emozioni. In tal senso, anche se si intravede il tentativo degli apprendenti di avvicinarsi alle produzioni dei nativi, non si rileva, tuttavia, una completa padronanza delle strutture prosodiche della L2. Questo fa sì che in alcuni contesti la prosodia emotiva della L1 venga utilizzata per sopperire alle difficoltà riscontrate dagli apprendenti (cfr. anche De Marco, Sorianello, Paone, in stampa). Ad esempio, come osservato in fase di analisi, le produzioni degli italiani si caratterizzavano per una maggiore estensione tonale, variabilità di  $f_0$  e intensità, nonché per una velocità di articolazione più elevata. Tutti questi *acoustic cues* costituiscono i segnali utilizzati principalmente dagli *encoder* italiani per veicolare le emozioni. D'altro canto, le produzioni in lingua russa si caratterizzavano per una minore variabilità, soprattutto in termini di estensione tonale, per cui i segnali acustici impiegati dagli apprendenti russi possono essere stati percepiti come poco salienti sul piano comunicativo o aver dirottato il giudizio degli italiani verso altre emozioni.

In relazione alle abilità di decodifica degli apprendenti, anche in questo caso il loro livello di competenza (B2) e di conoscenza della lingua e della cultura target ha favorito il riconoscimento di alcune emozioni, come la tristezza e in misura minore la gioia. Tuttavia, gli stimoli presentati nella L1 degli studenti sono stati identificati con percentuali significativamente più elevate. Inoltre, l'alta variabilità dei riconoscimenti erronei lascia supporre che il processo di decodifica abbia destato maggiori difficoltà in italiano rispetto alla loro L1.

Alla luce dei risultati ottenuti, è ragionevole supporre che la comunicazione vocale delle emozioni, pur confermando la natura filogenetica delle emozioni, sia profondamente legata al tessuto sociale e culturale in cui si realizza. Appare quindi riduttivo dare rilievo soltanto alle differenze riscontrate o esclusivamente alle similarità evidenziate in fase di *encoding* e *decoding*. I risultati ottenuti, pur presentando dei limiti, dati principalmente dal numero poco ampio di partecipanti e dall'impiego di materiale emotivo simulato/indotto, vanno forse interpretati in una prospettiva interazionista, che concilia la dicotomia universalità vs. specificità culturale, promuovendo una visione dinamica della comunicazione vocale, sensibile cioè sia agli aspetti creativi del linguaggio che ai vincoli sociali e culturali.

In quest'ottica, l'acquisizione di una lingua seconda implica l'accesso a queste peculiarità linguo-specifiche, sulle quali l'intervento didattico può orientare l'attenzione degli apprendenti, i quali, in qualità di agenti sociali, interagiscono in un contesto culturale che mette continuamente alla prova le loro capacità comunicative. Come già riportato in studi precedenti, le difficoltà in questo senso investono sia la percezione e il riconoscimento sia l'espressione e la manifestazione degli stati emotivi (cfr. Graham, Hamblin, Feldstein, 2001; Chen, 2005; Dewaele, 2005; Komar, 2005; Altrov, 2013; Bhatara *et al.* 2016; De Marco, Paone, 2016; De Marco, Sorianello, Paone, in stampa).

Come riconoscono Graham, Hamblin e Feldstein (2001), la capacità di decodifica non può essere acquisita in assenza di un'esposizione intensiva al contesto culturale della lingua target o senza un intervento didattico mirato allo sviluppo di queste abilità. Per questa ragione, sembrerebbe opportuno riservare nelle pratiche didattiche uno spazio alla comunicazione emotiva e in particolare alla prosodia emotiva della lingua target (cfr. De Marco, Paone, 2016), e quindi alle altre componenti non verbali, per stimolare e promuovere una riflessione metalinguistica e un uso consapevole di questi canali, che costituiscono il fulcro delle interazioni *vis-à-vis* con i nativi.

# Bibliografia

---

- ABERCROMBIE D. (1967), *Elements of General Phonetics*, Edinburgh University Press, Edinburgh.
- ALAVINIA A., BONYADI N., RAZAVI N. (2012), *On the correlation between teachers' emotional intelligence and learners' motivation: The case of Iranian EFL learners*, in "Journal of Education and Practice", 3, pp. 100-10.
- ALBAS D. C., MCCLUSKEY K. W., ALBAS C. A. (1976), *Perception of the Emotional Content of Speech: A Comparison of Two Canadian Groups*, in "Journal of Cross-Cultural Psychology", 7-4, pp. 481-90.
- ALTROV R. (2013), *Aspects of Cultural Communication in Recognizing Emotions*, in "Trames. Journal of the Humanities and Social Sciences", 17-2, pp. 159-74.
- AMATI-MEHLER J., ARGENTIERI S., CANESTRI J. (1993), *The Babel of the unconscious: mother tongue and foreign languages in the psychoanalytic dimension*, International Universities Press, Madison.
- ANDERSON-HSIEH J., JOHNSON R., KOEHLER K. (1992), *The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure*, in "Language Learning", 42, pp. 529-55.
- ANOLLI L. (2006), *La mente multiculturale*, Laterza, Roma-Bari.
- ID. (a cura di) (2011), *Fondamenti di Psicologia della Comunicazione*, Il Mulino, Bologna.
- ANOLLI L., CICERI R. (1992), *La voce delle emozioni: verso una semiosi della comunicazione vocale non-verbale delle emozioni*, Franco Angeli, Milano.
- ANOLLI L., WANG L., MANTOVANI F., DE TONI, A. (2008), *The Voice of Emotion in Chinese and Italian Young Adults*, in "Journal of Cross-Cultural Psychology", 39-5, pp. 565-98.
- ANOOSHIAN J.L., HERTEL P.T. (1994), *Emotionality in free recall: Language specificity in bilingual memory*, in "Cognition and Emotion", 8, pp. 503-54.
- ARCHIBALD J. (2000) (ed.), *Second Language Acquisition and Linguistic Theory*, Blackwell, Oxford.

- ARISTOTELE (1996), *Retorica*, Mondadori, Milano.
- ARNOLD M. B. (1960), *Emotion and Personality*, vol. 1, *Psychological Aspects*, Columbia University Press, New York.
- ASIMOV I. (2015), *Notturmo* in *Il meglio di Asimov* (1973), Mondadori, Milano.
- BALBONI P. (1999), *Parole comuni culture diverse. Guida alla comunicazione interculturale*, Marsilio, Venezia.
- BALCONI M. (2000), *Se il \kane\ non ringhia. Correlati percettivi, cognitivi e comunicativi nell'analisi dei fenomeni sinestesici e fisiognomici*, in "Psychofenia", 3-4/5, pp. 57–82.
- ID. (2002), *Fisiognomia e fonosimbolismo fisiognomico. Analisi dell'iconismo linguistico e dei correlati emotivi delle componenti fonemiche della lingua italiana*, in "Psychofenia", 5-8, pp. 1–22.
- BANSE R., SCHERER K. R. (1996), *Acoustic Profiles in Vocal Emotion Expression*, in "Journal of Personality and Social Psychology", 70-3, pp. 614–636.
- BÄNZIGER T., SCHERER K. R. (2005), *The role of intonation in emotional expressions*, in "Speech Communication", 46, pp. 252–267.
- BARD P. (1934), *Emotion: I. The Neuro-humoral Basis of Emotional Reactions*, in C. Murchison (ed.), *A handbook of general experimental psychology*, Clark University Press, Worcester, pp.264–311 (<http://doi.org/10.1037/11374-006>; consultato il 20 dicembre 2016).
- BARRETT H. C., BRYANT G. (2008), *Vocal Emotion Recognition Across Disparate Cultures*, in "Journal of Cognition and Culture", 8-1, pp. 135–48.
- BERTINETTO P. M. (1981), *Strutture prosodiche dell'italiano. Accento, quantità, sillaba, giuntura, fondamenti metrici*, Accademia della Crusca, Firenze.
- ID. (1987), *Lingue segrete e segreti delle lingue. Alcuni problemi di fonologia italiana studiati attraverso un gioco linguistico*, in "Annali della Scuola Normale Superiore", 3-17, pp. 889-920.
- BERTINETTO P. M., MAGNO CALDOGNETTO E. (1993), *Ritmo e intonazione*, in A.A. Sobrero (a cura di), *Introduzione all'italiano contemporaneo*, vol. 1, *Le strutture*, Laterza, Bari, pp. 141-92.
- BEZOOIJEN R. VAN. (1984), *Characteristics and recognizability of vocal expressions of emotion*, Foris Publications, Dordrecht.
- BHATARA A., LAUKKA P., BOLL-AVETISYAN N., GRANJON L., ELFENBEIN H. A., BÄNZIGER T. (2016), *Second Language Ability and Emotional Prosody Perception*, in "PloS

- One”, 11-6 (<http://dx.doi.org/10.1371/journal.pone.0156855>; consultato il 15 giugno 2016).
- BOLINGER D. (1984), *Intonation and its parts: Melody in spoken English*, Stanford University Press, Stanford.
- BOND M. H., LAI T. (1986). *Embarrassment and code-switching into a second language*, in “The Journal of Social Psychology”, 126-2, pp. 179–86.
- BONELLI L. (2013), *Comunicazione Emotiva :Problemi Di Pragmatica Linguistica*, Tesi di Laurea, Università degli Studi di Genova ([https://www.academia.edu/7550687/Comunicazione\\_emotiva\\_problemi\\_di\\_pragmatica\\_linguistica](https://www.academia.edu/7550687/Comunicazione_emotiva_problemi_di_pragmatica_linguistica), consultato il 15 ottobre 2015).
- ID. (2015), *Disaffiliation and Pragmatic strategies of Emotive Communication in a Multiparty Online Conflict Talk*, in F. D’Errico, I. Poggi, A. Vinciarelli, Laura Vinzce (eds.), *Conflict and multimodal Communication. Social research and Machine Intelligence*, Springer, Cham-Heidelberg- New York-Dordrecht-London, pp. 159-82.
- BONGAERTS T., VAN SUMMEREN C., PLANKEN B., SCHILS E. (1997), *Age and ultimate attainment in the pronunciation of a foreign language*, in "Studies in Second language Acquisition", 19, pp. 447-65.
- BOUCHER J. D., EKMAN P. (1975), *Facial Areas and Emotional Information*, in "Journal of Communication ", 25-2, pp. 21–29.
- BRENNER E. (2000), *Mood induction in children: Methodological issues and clinical implications*, in "Review of General Psychology", 4, pp. 264–83.
- BRUNSWIK E. (1956), *Perception and the representative design of psychological experiments*, University of California Press, Berkeley, CA.
- BUCK R. (1984), *The Communication of Emotion*, Guilford Press, New York.
- ID. (2007), *The Evolutionary Bases of Social and Moral Emotions: Dominance, Submission, and True Love*, in J. P. Forgas, M. G. Haselton, W. von Hippel (eds.), *Evolution and the Social Mind: Evolutionary Psychology and Social Cognition*, Psychology Press, New York, pp. 89–106.
- BUCK R., VANLEAR C. A. (2002), *Verbal and Nonverbal Communication: Distinguishing Symbolic, Spontaneous, and Pseudo-Spontaneous Nonverbal Behavior*, in "Journal of Communication", 52-3, pp. 522–41.
- CAFFI C. (2001), *La mitigazione. Un approccio pragmatico alla comunicazione in contesti terapeutici*. LIT Verlag, Münster- Hamburg- London.

- CAFFI C., JANNEY (1994), *Toward a pragmatics of emotive communication* in "Journal of Pragmatics", 22, pp. 325-73.
- CAMPBELL N. (2000), *Databases Of Emotional Speech*, in R. Cowie, E. Douglas-Cowie, M. Schroeder (eds.), *Proceedings of the ISCA Workshop on Speech and Emotion*, (Newcastle, 5-7 settembre 2000), Textflow, Belfast, pp. 114–21.
- CANNON W. B. (1927), *The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory*, in "The American Journal of Psychology", 39-1/4 (<http://doi.org/10.2307/1415404>; consultato il 12 ottobre 2013).
- CHEN A. (2005), *Universal and language-specific perception of paralinguistic intonational meaning*, LOT, Utrecht.
- CHEN A., GUSSENHOVEN C., RIETVELD T. (2004), *Language-specificity in the perception of paralinguistic intonational meaning*, in "Language and Speech", 47- 4, pp. 311–49 (<http://doi.org/10.1177/00238309040470040101>, consultato il 2 dicembre 2015).
- CHIANG T. (1979), *Some interferences of English intonation with Chinese tones*, in "IRAL - International Review of Applied Linguistics in Language Teaching", 17, pp. 245–50.
- CHINI M. (2015) (a cura di), *Il parlato in (italiano) L2. Aspetti pragmatici e prosodici*, Franco Angeli, Milano.
- CHUA SHI MIN, SCHIRMER A. (2011) *Perceiving verbal and vocal emotions in a second language*, in "Cognition and Emotion" 25, pp. 1376-92.
- CONSIGLIO D'EUROPA (2002), *Quadro comune europeo per le lingue: apprendimento, insegnamento, valutazione*, La Nuova Italia-Rizzoli, Firenze-Milano.
- count more than statistics*, in "Journal of Memory & Language", 44, pp. 548-67.
- CYTOWIC R. E. (1989), *Synesthesia: A union of the senses*, Springer-Verlag, New York.
- D'URSO V. D., CAVICCHIO F., MAGNO CALDOGNETTO E. (2008), *Le etichette lessicali nelle ricerche sperimentali sulle emozioni. Problemi teorici e metodologici*, in E. Magno Caldognetto, F. Cavicchio, P. Cosi (a cura di), *Atti del I Convegno Nazionale GSCP su "Comunicazione parlata e manifestazione delle emozioni"*, Liguori, Napoli, pp. 188–205
- D'URSO V., TRENTIN R. (1998), *Introduzione alla psicologia delle emozioni*, Laterza, Bari.
- DARWIN C. (1982), *L'espressione delle emozioni nell'uomo e negli animali. Taccuini M e N; Profilo di un bambino (1872)*, Bollati Boringhieri, Torino
- ID. (2011), *L'origine delle specie (1859)*, Bollati Boringhieri, Torino.
- DAVITZ J.R. (1964), *Auditory correlates of vocal expression of emotional feeling*, in Davitz

- J.R. (ed.) *The communication of emotional meaning*, McGraw–Hill, New York, pp. 101–12.
- DE ABREU S., MATHON C., PARIS U. (2005), *Can you hear I'm angry? Perception of anger in a spontaneous French corpus by Portuguese learners of French as a foreign language*, in *Proceedings of the Phonetics Teaching and Learning Conference (PTLC)*, (27-30 July 2005), University College of London, London.
- DE DOMINICIS A. (2010), *Intonazione. Una teoria della costituenza delle unità intonative*, Carocci, Roma.
- DE GELDER B., VROOMEN J. (1996), *Categorical perception of emotional speech* (abstract), in "Journal of the Acoustical Society of America", 100, p. 2818.
- DE MARCO A. (2010), *Sociopragmatica*, Editori Riuniti, Roma.
- DE MARCO A., PAONE E. (in stampa), *Uno studio sui correlati acustici delle emozioni vocali in apprendenti di italiano L2*, in A. Elia, C. Iacobini, M. Voghera (a cura di), *Livelli di analisi e fenomeni di interfaccia. Atti del XLVII Congresso Internazionale SLI 2013*, Bulzoni, Roma.
- EAD. (2015), *The Acquisition Of Emotional Competence In L2 Learners Of Italian Through Specific Instructional Training*, in S. Gesuato, F. Bianchi, W. Cheng (eds.) *Teaching, Learning and Investigating Pragmatics*, Cambridge Scholar Publishing, Cambridge, pp. 443–70.
- EAD. (2016), *Dalle emozioni alla voce. Nuovi orizzonti della comunicazione in italiano L2*, Carocci, Roma.
- DE MARCO A., SORIANELLO P., MASCHERPA E. (2014), *L'acquisizione dei profili intonativi in apprendenti di italiano L2 attraverso un'unità di apprendimento in modalità blended learning*, in De Meo A., D'Agostino M., Iannaccaro G., Spreafico L. (a cura di), *Varietà dei contesti di apprendimento linguistico*, AItLA, Milano, pp. 189-210.
- DE MARCO A., SORIANELLO P., PAONE E. (in stampa), *L'acquisizione delle emozioni nell'italiano non-nativo. Un percorso didattico longitudinale*, in C. Bardel (ed.) *Romance Languages: Multilingualism and Language Acquisition*, Peter Lang, Bern.
- DE MEO A., PETTORINO M. (2012), *Prosodic and Rhythmic Aspects of L2 Acquisition: The Case of Italian*, Cambridge Scholar Publishing, Cambridge.
- DESCARTES R. (2003), *Le passioni dell'anima* (1649), Bompiani, Milano.
- DESTENO D., PETTY R. E., RUCKER D. D., WEGENER D. T. (2000), *Beyond Valence in the*

- Perception of Likelihood: The Role of Emotion Specificity*, in "Journal of Personality and Social Psychology", 78- 3, pp. 397–416.
- DEVÌS HERRAIZ A. A. (2010), *Aspetti prosodici dell'acquisizione linguistica: il caso delle interrogative polari tra italiano e spagnolo*, in "Studi e Saggi Linguistici XXLVIII", pp. 239-70.
- DEWAELE J.-M. (2004), *Blistering barnacles! What language do multilinguals swear in?!*, in "Sociolinguistic Studies", 5-1, pp. 83–105.
- ID. (2005), *Investigating the psychological and emotional dimensions in instructed language learning: obstacles and possibilities*, in "The Modern Language Journal", 89-3, pp. 367–380.
- ID. (2010), *Emotions in Multiple Languages*, Palgrave Macmillan, Basingstoke.
- ID. (2011), *Reflections on the Emotional and Psychological Aspects of Foreign Language Learning and Use*, in "Anglistik: International Journal of English Studies", 22-1, pp. 23–42.
- DEWAELE J.-M., PAVLENKO A. (2001), *Web questionnaire bilingualism and emotions*, University of London, London, ([https://www.academia.edu/2134800/Web\\_questionnaire\\_bilingualism\\_and\\_emotions](https://www.academia.edu/2134800/Web_questionnaire_bilingualism_and_emotions), consultato il 15 ottobre 2015).
- EAD. (2002), *Emotion vocabulary in interlanguage*. *Language Learning*, 52, pp. 263-322.
- DI RUSSO D. (2011). *Intonazione e modalità nella significazione*, Tesi di Dottorato, Università degli Studi di Roma, La Sapienza.
- DOMAKANI M. R., MIRZAEI A., ZERAATPISHEH S. (2014). "L2 Learners ' Affect and Pragmatic Performance: A Focus on Emotional Intelligence and Gender Dimensions", in "RALs", 5, pp. 149-73.
- DÖRNYEI Z. (2005), *The psychology of the language learner: Individual differences in second language acquisition*, Erlbaum, Mahwah, NJ.
- DÖRNYEI Z., SKEHAN P. (2003), *Individual differences in second language learning*, in Doughty C. J., Long M.H. (eds.), *The handbook of second language acquisition*, Blackwell, Oxford, pp. 589-630.
- DRIOLI C., TISATO G., COSI P., TESSER F. (2003), *Emotions and voice quality: Experiments with sinusoidal modeling*, in C. d'Alessandro, B. Doval, K. Scherer (eds), *Proceedings of the ISCA VOQUAL'03, Workshop on Voice Quality: Functions, Analysis and Synthesis* (Geneva, 27-29 agosto 2003), pp. 127–132.
- EAD. (1976), *Measuring Facial Movement*, in "Environmental Psychology and Nonverbal

- Behavior", 1-1, pp. 56–75.
- EAD. (1978), *Facial Action Coding Systems (FACS): A technique for the measurement of facial action*, Consulting Psychologists Press, Palo Alto.
- EKMAN P. (1972), *Universals and cultural differences in facial expression of emotion*, In J. Cole (ed.), *Nebraska Symposium on Motivation*, University of Nebraska Press, Lincoln, pp. 207–83.
- ID. (1982) (ed.), *Emotion in the Human Face*, Oxford University Press, Oxford.
- ID. (1992), *An Argument for Basic Emotions*, in "Cognition and Emotion", 6, 3-4, pp. 169–200.
- ID. (2004), *Emotional and Conversational Nonverbal Signals*, in J. M. Larrazabal, L. Pérez Miranda, *Language, Knowledge and Representation*, Kluwer Academic Publisher, Dordrecht, pp. 39–50.
- EKMAN P., FRIESEN W. V. (1969), *The Repertoire of Nonverbal Behavior: Categories, Origins, and Coding*, in "Semiotica" 1-1, pp. 49–98.
- EKMAN P., FRIESEN W. V., O'SULLIVAN M., CHAN A., DIACOYANNI-TARLATZIS I., HEIDER K. RICCI-BITTI P. E. (1987), *Universals and cultural differences in the judgments of facial expressions of emotion*, in "Journal of Personality and Social Psychology", 53-4, pp. 712–17.
- ELFENBEIN H. A., AMBADY N. (2002), *Is there an in-group advantage in emotion recognition?* in "Psychological Bulletin", 128-2, pp. 243–49.
- FÓNAGY I. (1972), *Motivation and remotivation. Comment se dépasser*, in "Poétique", 3, pp. 414–31.
- ID. (1987), *Semantic Diversity in Intonation*, in T. Gamqrelize (ed.) *Proceedings of the XIth ICPHS*, (Tallin 1-7 agosto 1987), vol. 2, Academy of Sciences of the Estonian, Tallin, pp. 468–71.
- ID. (1989), *La vive voix. Essais de psychophonétique*, Payot, Paris.
- FÓNAGY I., MAGDICS K. (1963), *Emotional Patterns in Intonation and Music*, in "STUF - Language Typology and Universals", 16, pp. 1-4.
- FRANK M.G., STENNETT J. (2001), *The forced-choice paradigm and the perception of facial expressions of emotion*, in "Journal of Personality and Social Psychology", 80, pp. 75–85.
- FRICK R. W. (1985), *Communicating emotion: The role of prosodic features*, in "Psychological Bulletin", 97, pp. 412–29.
- FRIJDA N. H. (2007), *The Laws of Emotion*, Erlbaum, Mahwah.

- ID. (1986), *The emotions*, Cambridge University Press, Cambridge.
- GALATÀ V., ROMITO L. (2010), *Un corpus sperimentale per lo studio crosslinguistico europeo delle emozioni vocali*, in S. Schmid, M. Schwarzenbach, D. Studer (a cura di), *La dimensione temporale del parlato*, Atti del 5° Convegno Internazionale AISV (Zurich, 4-6 febbraio 2009) , EDK,Torriana, pp. 603– 42.
- GALATI D. (a cura di) (1993), *Le emozioni primarie*, Bollati Boringhieri, Torino.
- GAMAL D. (2006), *La Prosodia Direttiva in Italiano L2. Studio pilota*, in R. Savy, C. Crocco (a cura di), *Analisi prosodica. Teorie, modelli e sistemi di annotazione. Atti del II Convegno Nazionale dell’AISV* (Salerno, 30 novembre-2 dicembre 2005), EDK Editore, Padova, pp. 189–202.
- GASPER K. (2003), *When necessity is the mother of invention: Mood and problem solving*, in "Journal of Experimental Social Psychology", 39-3, pp. 248–62.
- GERRARDS-HESSE A., SPIES K., HESSE F. W. (1994), *Experimental inductions of emotional states and their effectiveness: A review*, in "British Journal of Psychology", 85-1, pp. 55–78.
- GILES H., SCHERER K. R., TAYLOR D. M. (1979), *Speech Markers in Social Interaction*, in K. R. Scherer, H. Giles (eds.) *Social Markers in Speech*, Cambridge University Press, Cambridge, pp. 343–81.
- GILI FIVELA B., GRIMALDI M., STEFÀNO M. (2004), *Emozioni a teatro: voce e gesto nello spazio*, in Magno Caldognetto E., Cavicchio F., Cosi P. (a cura di), *Atti del Convegno Comunicazione Parlata e Manifestazione delle Emozioni*, Liguori Multimedia, Napoli, pp. 383-423.
- GOFFMAN E. (1981), *Relazioni in Pubblico* (1971), Bompiani, Milano.
- GOLDSMITH J.A. (1976), *Autosegmental phonology*, Ph.D Dissertation (published by Garland, New York 1979).
- GOLEMAN D. (1996), *Intelligenza emotiva*, Rizzoli, Milano (1995).
- GONZALEZ-REGIOSA F. (1976), *The anxiety arousing effect of taboo words in bilinguals*, in C.D. Spielberger, R. Diaz-Guerrero (eds.), *Cross-cultural anxiety*, Hemisphere, Washington, pp. 89–105.
- GORDON M.K. (1999), *The intonational structure of Chickasaw*, in Ohala J., Hasegawa Y., Ohala M., Granville D., Bailey A.C. (eds.) *Proceedings of 14th ICPHS* (San Francisco, 1-7 Agosto 1999), pp. 1993-6.
- GOUDBEEK M., BROERSMA M. (2010), *Language specific effects of emotion on phoneme duration*, in *Proceedings of the 11th Annual Conference of the International*

- Speech Communication Association (Interspeech 2010)*, pp. 2026–29.
- GRAHAM CR., HAMBLIN AW., FELDSTEIN S. (2001), *Recognition of emotion in English voices by speakers of Japanese, Spanish and English*, in "IRAL—International Review of Applied Linguistics in Language Teaching", 9-39, pp. 19–37.
- GROSSER W. (1993), *Aspects of intonation L2 acquisition*, in B. Kettemann, W. Wieden (eds.), *Current Issues in European Second Language Acquisition research*, Gunter Narr Verlag, Tübingen, pp. 81–94.
- GUION S. G. (2005), *Knowledge Of English Word Stress Patterns In Early And Late Korean-English Bilinguals*, in "Studies in Second Language Acquisition", 27-4, pp. 503–533.
- GUMPERZ J. J. (1992), *Contextualization and Understanding*, in A. Duranti and C. Goodwin (eds.) *Language as an Interactive Phenomenon*, Cambridge University Press, Cambridge, pp. 229–52
- GUSSENHOVEN C. (2002), *Intonation and interpretation: Phonetics and Phonology*, in "Natural Language and Linguistics Theory", 9, pp. 46-96.
- GUSSENHOVEN C., CHEN A. J. (2000), *Universal and language-specific effects in the perception of question intonation*, in B. Yuan, T. Huang, X. Tang (eds.), *Proceedings of The Sixth International Conference on Spoken Language Processing*, vol. 1, China Military Friendship Publish, Beijing, pp. 91-4.
- HALBERSTADT J. B., NIEDENTHAL P. M. (1997), *Emotional state and the use of stimulus dimensions in judgment*, in "Journal of Personality and Social Psychology", 72- 5, pp. 1017–33.
- HALL E. T. (1968). *La dimensione nascosta* (1966), Bompiani, Milano.
- ID.. (1969). *Il linguaggio silenzioso* (1959), Garzanti, Milano.
- ID. (1976), *Beyond culture*, Anchor Press, Garden City.
- HALLIDAY M. A. (1970), *A Course in Spoken English Intonation*, Oxford University Press, London.
- HARRIS C.L., GLEASON J.B., AYCICEGI A., (2006), *When is a first language more emotional? Psychophysiological evidence from bilingual speakers*, in Pavlenko A. (ed.), *Bilingual minds: Emotional experience, expression, and representation*, Multilingual Matters, Clevedon, UK.
- HARRIS C.L. (2004), *Bilingual speakers in the lab: psychophysiological measures of emotional reactivity*, in "Journal of Multilingual and Multicultural Development" 25, pp. 223-47.

- HEIDERSCHIEDT S., HILIGSMANN P. (2000), *De accentuering in de tussentaal van Franstalige leerders van het Nederlands*, in "Leuvense Bijdragen", 89-1/2, pp. 117-31.
- HINDE R. A., HERRMANN J. (1977), *Frequencies, durations, derived measures and their correlations in studying dyadic and triadic relationships*, in Schaffer H. R. (ed.) *Studies in mother-infant interaction*, Academic Press, London, pp. 19-46.
- HOLDEN K., HOGAN J. (1993), *The emotive impact of foreign intonation: An experiment in switching English and Russian intonation*, in "Language and Speech", 36, 1, pp. 67-88.
- HUTTAR G.L. (1968), *Relations between prosodic variables and emotions in normal American English utterances*, in "Journal of Speech and Hearing Research" 11, pp. 481-7.
- HYMES D. H. (1966), *On Communicative Competence*, in J. B. Pride, J. Holmes (eds.), *Sociolinguistics*, Penguin, Baltimore, pp. 269-293.
- IZARD C. E. (1977), *Human emotions*, Plenum Press, New York.
- JAMES W. (1884), *What is an emotion?*, in "Mind Association", 9-34, pp. 188-205 (<http://www.jstor.org/stable/2246769>; consultato il 5 dicembre 2016).
- JARMAN E., CRUTTENDEN A. (1976), *Belfast intonation and the myth of the fall*, in "JIPA", 6, pp. 4-12.
- JOHNSON E., JUSCZYK P. W. (2001), *Word segmentation by 8 month olds: when speech cues*
- JOHNSON J. S., NEWPORT E. L. (1989), *Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a Second Language*, in "Cognitive Psychology", 21, pp. 60-99.
- JUSCZYK P. W., FREDERICI A., WESSELS J. M., SVENKERUD V. Y., JUSCZYK A. M. (1993), *Infants'*
- JUSLIN P.N. (2000), *Cue utilization in communication of emotion in music performance: relating performance to perception*, in "Journal of Experimental Psychology: Human Perception and Performance", 26, pp. 1797-813.
- JUSLIN P.N., LAUKKA P. (2003), *Communication of emotions in vocal expression and music performance: different channels, same code?* in "Psychological Bulletin", 129, pp. 770-814.
- JUSLIN P.N., SCHERER K. (2008), *Vocal expression of affect*, in Harrigan J.A., Rosenthal R., Scherer K. (eds.) *The New Handbook of Methods in Nonverbal Behavior Research*, Oxford University Press, Oxford, pp. 65-135.

- KIM S., DORNER L. M. (2013), *"I won't talk about this here in America" : Sociocultural context of Korean English learners' emotion speech in English*, in "L2 Journal", 5-2, pp. 43–67.
- KITAMURA C., THANAVISHUTH C., BURNHAM D., LUKSANEYANAWIN S. (2002), *Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language*, in "Infant Behavior and Development", 24-4, pp. 372–92.
- KOMAR S. (2005), *The Impact of Tones and Pitch Range on the Expression of Attitudes in Slovene Speakers of English*, in *Proceedings of PTLC2005 (London 27- 30 luglio 2005)*, University College London, London, CD- rom.
- KRASHEN S. (1982), *Principles and Practice in Second Language Acquisition*, Pergamon Press, Oxford.
- LADD D. R. (1981), *On Intonational Universals*, in Myers T. et al. (eds.) *The Cognitive Representation of Speech*, North Holland Publishing, Amsterdam, pp. 389-97.
- ID. (1996), *Intonational Phonology*, Cambridge University Press, Cambridge.
- ID. (2014), *Simultaneous structure in phonology*, Oxford University Press, Oxford.
- LADD D. R., SILVERMAN K. E. A., TOLKMITT F., BERGMANN G., SCHERER K. R. (1985), *Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect*, in "The Journal of the Acoustical Society of America", 78-2, pp. 435 -44.
- LAIRD D. J., STROUT S. (2007), *Emotional Behaviours as Emotional Stimuli*, in J. A. Coan, J. B. Allen (eds.), *Handbook of emotion elicitation and assessment*, Oxford University Press, Oxford, pp. 54–64.
- LANG P. J., BRADLEY M. M., CUTHBERT B. N. (2005), *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*, Technical Report A-6, Center for Research in Psychophysiology, University of Florida, Gainesville.
- Language" 35, pp. 666-88.
- LAUKKA P. (2004), *Vocal expression of emotion: discrete-emotions and dimensional accounts*, Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences, Uppsala University Library, Uppsala, Sweden.
- LAUKKA P., ELFENBEIN H. A., SÖDER N., NORDSTRÖM H., ALTHOFF J., CHUI W. (2013), *Cross-cultural decoding of positive and negative non-linguistic emotion vocalizations*, in "Frontiers in Psychology" 4-353 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3728469/>, consultato il 15

dicembre 2015).

- LAVIER J. (1980), *The Phonetic Description of Voice Quality*, Cambridge University Press, Cambridge.
- LAVIER J., TRUDGILL P. (1979), *Phonetic and linguistic markers in speech*, in Scherer K.R., Giles H. (eds.), *Social Markers in speech*, Cambridge University Press, Cambridge, pp. 1-32.
- LEDOUX J. E. (2000), *Emotion circuits in the brain*, in "Annual Review of Neuroscience", 23, pp.155–84.
- ID. (2003), *The Emotional Brain, Fear, and the Amygdala*, in "Cellular and Molecular Neurobiology", 23-45, pp. 727–38.
- LENNEBERG. E.H. (1967), *Biological foundation of language*, Wiley, New York.
- LIEBERMAN P., MICHAELS S. B. (1962), *Some Aspects of Fundamental Frequency and Envelope Amplitude as Related to the Emotional Content of Speech*, in "Journal of the Acoustical Society of America", 34, pp. 922-927.
- LORETTE P., DEWAELE J.-M. (2015), *Emotion recognition ability in English among L1 and LX users of English*, in "International Journal of Language and Culture", 2-1, pp. 62–86.
- MACINTYRE P. D., CLÉMENT R., DÖRNYEI Z., NOELS K. A. (1998), *Conceptualizing willingness to communicate in a L2: A situated model of confidence and affiliation*, in "The Modern Language Journal", 82, pp. 545-62.
- MACINTYRE, P.D. (2002), *Motivation, Anxiety and Emotion in Second Language Acquisition*, in Robinson P. (ed.) *Individual Differences and Instructed Language Learning*, John Benjamins, Amsterdam, pp. 45-68.
- MACLEAN P. D. (1990), *The triune brain in evolution : role in paleocerebral functions*, Plenum Press, New York.
- MAFFIA M., PELLEGRINO E., PETTORINO M. (2014), *Labeling expressive speech in L2 Italian: the role of prosody in auto-and external annotation*, in A. W. Campbell, D. Gibbon, D. Hirst (eds.), *Proceedings of the 7th international conference on Speech Prosody* (Dublin, 20-23 maggio 2014), Speech Prosody Special Interest Group (SProSIG), Urbana, Illinois, pp. 81–84.
- MAGNO CALDOGNETTO E., CAVICCHIO F., COSI P. (2008), *La faccia e la voce delle emozioni*, in Poggi I. (a cura di) *La mente del cuore : le emozioni nel lavoro, nella scuola, nella vita*, Armando, Roma.
- MAKAROVA V. (2000), *Cross-linguistic aspects of intonation perception*, in *Proceedings of*

- the 6th International Conference on Spoken Language Processing*, vol. 3, China Military Friendship Publish, Beijing, pp. 452-55.
- MAKAROVA V., PETRUSHIN V.A (2003), *Phonetics of emotion in Russian speech*, in Solé D. R., Recasens D., Romero J. (eds.), *Proceedings Of The Xvth International Conference Of Phonetic Sciences*, Barcellona 3-5 agosto 2003 ([https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/p15\\_intro.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/p15_intro.pdf), consultato il 5 gennaio 2017).
- MAROTTA G. (2008), *Sulla percezione dell'accento straniero*, in Lazzeroni R., Banfi E., Bernini G., Chini M., Marotta G. (a cura di), *Diachronica et Synchronica. Studi in onore di Anna Giacalone Ramat*, ETS, Pisa, 327-347.
- MARTY A. (1908), *Untersuchungen Zur Grundlegung Der Allgemeinen Grammatik Und Sprachphilosophie*, Niemeyer, Halle an der Salle.
- MATSUMOTO D. (1989), *Cultural influences on the perception of emotion*, in "Journal of Cross-Cultural Psychology", 20, pp. 92-105.
- MAYER J. D., SALOVEY P. (1997), *What Is Emotional Intelligence?*, in Salovey P., Sluyter D. J. (eds.), *Emotional development and emotional intelligence: Educational implications*, Harper Collins, New York, pp. 3-34.
- MEHLER J., JUSZYK P. W., LAMBERTZ G., HALSTED N., BERTONCINI J., AMIEL-TISON C. (1988), *A precursor of language acquisition in young infants*, in "Cognition" 29, pp. 143-78.
- MESQUITA B., FRIJDA N. H. (1992), *Cultural variations in emotions: A review*, in "Psychological Bulletin", 112, pp. 197-204.
- MESQUITA B., FRIJDA N. H., SCHERER K. R. (1997), *Culture and emotion*, in Berry J. W., Dasen P. R., Saraswathi T. S. (eds.), *Handbook of cross-cultural psychology*, vol. 2, Allyn & Bacon, Boston, pp. 255-97.
- MITCHELL R. L. C., KINGSTON R. (2011), *Is Age-related Decline in Vocal Emotion Identification an Artefact of Labelling Cognitions?*, in "International Journal of Psychological Studies", 3-2, pp. 156-63.
- MOCCIARO E. (2014), *Aspetti della prosodia nell'interlingua di apprendenti vietnamiti di italiano L2: un'ipotesi di trattamento task-based in contesto guidato*, in De Meo A., D'Agostino M., Iannaccaro G., Spreafico L. (a cura di), *Varietà dei contesti di apprendimento linguistico*, AItLA, Milano, pp. 163-78.
- MORGAN J. L. (1996), *A rhythmic bias in preverbal speech segmentation*, in "Journal of Memory &

- MORGAN J. L., SAFFRAN J. R. (1995), *Emerging integration of sequential and suprasegmental information in preverbal speech segmentation*, in "Child Development" 66, pp. 911-36.
- NASOZ F., ALVAREZ K., LISETTI C. L., FINKELSTEIN N. (2004), *Emotion recognition from physiological signals using wireless sensors for presence technologies*, in "Cognition, Technology & Work", 6, pp. 4-14.
- NESPOR M. (1981), *Fonologia*, Il Mulino, Bologna.
- O'CONNOR J. D., ARNOLD G. F. (1961), *Intonation of colloquial English: a practical handbook*, Longmans, London (2<sup>a</sup> ed.).
- OATLEY K. (2007), *Breve storia delle emozioni* (2004), Il Mulino, Bologna.
- Occasion of his Eightieth Birthday*, Longman, London, pp. 271-279.
- OHALA J.J. (1984), *An ethological perspective on common cross-language utilization of f0 voice*, in "Phonetica", 41, pp. 1-16.
- ORIGLIA A., GALATA V., LUDUSAN B. (2010), *Automatic classification of emotions via global and local prosodic features on a multilingual emotional database*, in *Proceedings of Speech Prosody 2010*, Chicago, 11-14 maggio 2010 (<http://speechprosody2010.illinois.edu/>, consultato il 20 dicembre 2015).
- OSGOOD C. E. (1952), *The Nature And Measurement Of Meaning*, in "Psychological Bulletin" 49-3, pp.
- ID. (1980), *The cognitive dynamics of synesthesia and metaphor*, in R. Honeck, R. R. Hoffman (eds.), *Cognition and figurative language*, Erlbaum, New Jersey, pp. 203-38.
- OSGOOD C. E., SUCI G. J., TANNENBAUM P. H. (1957), *The measurement of meaning*, University Of Illinois Press, Urbana-Chicago.
- OYAMA S. (1976), *A sensitive period for the acquisition of a nonnative phonological system*, in "Journal of Psycholinguistic Research", 5, pp. 261-85.
- OZ H. (2015), *Emotional Intelligence as A Predictor of L2 Communication*, in "Procedia. Social and Behavioral Sciences" 186, pp. 424-30.
- PAKOSZ M. (1983), *Attitudinal judgments in intonation: Some evidence for a theory*, in "Journal of Psycholinguistic Research", 12-3, pp. 311-326.
- PALMEN M.-J., BONGAERTS T., SCHILS E. (1997), *L'authenticité de la prononciation dans l'acquisition d'une langue étrangère au delà de la période critique : des apprenants néerlandais parvenus à un niveau très avancé en français*, in "Acquisition et Interaction en Langue Etrangère", 9, pp. 173-91.

- PANKSEPP J. (1982), *Towards a general psychobiological theory of the emotions*, in "Behavioral and Brain Sciences", 5, pp. 407-67.
- ID. (2005), *On the embodied neural nature of core emotional affects*, in "Journal of Consciousness Studies", 12, pp. 158-84.
- PAPOUŠEK M., HWANG, S.-F. C. (1991), *Tone and intonation in Mandarin babytalk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction*, in "Applied Psycholinguistics", 12-04, pp. 481-504.
- PAVLENKO A. (2002), *Bilingualism and emotions*, in "Multilingua - Journal of Cross-Cultural and Interlanguage Communication", 21-1, pp. 45-78.
- ID. (2006) (ed.) *Bilingual minds: Emotional experience, expression, and representation*, Multilingual Matters, Clevedon.
- PELLEGRINO E., VITALE M., SALVATI L., VITALE G., LIPARI S. (2014), *La competenza prosodica nella classe di lingue. L'italiano in contesto L2, LS ed e-learning*, in De Meo A., D'Agostino M., Iannaccaro G., Spreafico L. (a cura di), *Varietà dei contesti di apprendimento linguistico*, AItLA, Milano, pp. 153-68.
- PEREIRA C. (2000), *Dimensions of emotional meaning in speech*, in Cowie R., Douglas-Cowie E., Schroeder M. (eds.) *Proceedings of the ISCA Workshop on Speech and Emotion* (CD-ROM), International Speech Communication Association.
- PETTORINO M. (2004), *Inglese italiano e giapponese: analisi dei correlati acustici delle emozioni nel parlato cinematografico*, in Atti del I Convegno Nazionale GSCP, Padova, 30 ottobre 2004, Liguori, Napoli, pp. 45-57.
- PIERREHUMBERT J.B. (1980), *The phonology and phonetics of English Intonation*, Ph.D. Dissertation, MIT, Cambridge.
- PIKE K. L. (1945), *The Intonation of American English*, University of Michigan Press, Ann Arbor.
- PIOT O. (1999), *Experimental study of the expression of emotions and attitudes in four Languages*, in Ohala J., Hasegawa Y., Ohala M., Granville D., Bailey A.C. (eds.) *Proceedings of 14th ICPHS*, San Francisco, 1-7 Agosto 1999, pp. 369-370.
- PISHGHADAM R. (2009), *A quantitative analysis of the relationship between emotional intelligence and foreign language learning*, in "Electronic Journal of Foreign Language Teaching", 6, pp. 31-41.
- PITTAM J., GALLOIS C., CALLAN V. (1990), *The long-term spectrum and perceived emotion*, in "Speech Communication", 9, pp. 177-87.
- PITTAM J., SCHERER K. R. (1993), *Vocal expression and communication of emotion*, in M.

- Lewis, J. Haviland (eds.), *Handbook of Emotions*, Guilford Press, New York, pp. 185-97.
- POGGI I. (1981), *Le interiezioni: studio del linguaggio e analisi della mente*, Bollati Boringhieri, Torino.
- ID. (2008) (a cura di), *La mente del cuore : le emozioni nel lavoro, nella scuola, nella vita*, Armando, Roma.
- POGGI I., MAGNO CALDOGNETTO E. (2004), *Il Parlato Emotivo. Aspetti Cognitivi, Linguistici E Fonetici*, in F. Albano Leoni, F. Cutugno, M. Pettorino, R. Savy (a cura di), *Atti Del Convegno "Italiano Parlato"*, (Napoli 14-15 febbraio 2003), D'Auria Editore, Napoli, CD-rom.
- QUEEN R. M. (2001), *Bilingual intonation patterns: Evidence of language change from Turkish-German bilingual children*, in "Language in Society " 30, pp. 55-80.
- RASIER L., HILIGSMANN P. (2007), *Prosodic transfer from L1 to L2. Theoretical and methodological issues*, in "Nouveaux Cahiers de Linguistique Française", 28, pp. 41–66.
- RINTELL E. (1984), *But how did you feel about that? The learner's perception of emotion in speech*, in "Applied Linguistics", 5, pp. 255-64.
- ID. (1990), *That's incredible: Stories of emotion told by second language and native speakers*, in Scarcells R., Andersen E., Krashen S. (eds.), *Developing communicative competence in a Second Language*, Heinle and Heinle Publishers, Boston, MA, pp. 75-94.
- ROMITO L. (2000), *Manuale di fonetica articolatoria, acustica e forense*, Centro Editoriale e Librario, Università della Calabria.
- ROSENTHAL R., HALL J. A., DIMATTEO M. R., ROGERS P. L., ARCHER D. (1979), *Sensitivity to Nonverbal Communications: The PONS Test*, The Johns Hopkins University Press, Baltimore, MD.
- RUSSELL J. A. (1980), *A Circumplex Model of Affect*, in "Journal of Personality and Social Psychology", 39-6, pp.1161–178.
- ID. (1991), *Culture and the Categorization of Emotions*, in "Psychological Bulletin", 110-3, pp. 426–50.
- ID. (1994), *Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies*, in "Psychological Bulletin", 115, pp. 102-41.
- SAARNI C. (1999), *The Development of Emotional Competence*, Guildford Press, New York.

- SAKKALOU E., GATTIS M. (2012), *Infants Infer Intentions from Prosody*, in "Cognitive Development" 27, pp. 1-16.
- SAMOVAR L. A., PORTER R. E., MCDANIEL E. R. (2007), *Communication between cultures*, Wadsworth Thomson Learning, Belmont.
- SCHACHTER S., SINGER J. (1962), *Cognitive, Social, and Physiological Determinants of Emotional State*, in "Psychological Review", 69, pp. 379–99.
- SCHERER K.R. (1974), *Acoustic concomitants of emotional dimensions: judging affect from synthesized tone sequences*, in Weitz S. (ed.) *Nonverbal communication*, Oxford University Press, New York, pp. 105–11.
- ID. (1978), *Personality inference from voice quality: the loud voice of extroversion*, in "European Journal of Social Psychology", 8, pp. 467-87.
- ID. (1982), *Methods of research on vocal communication: paradigms and parameters*, in Scherer K.R., Ekman P. (eds.) *Handbook of methods in nonverbal behavior research* (eds.) Cambridge University Press, Cambridge, pp. 136–98.
- ID. (1984), *Emotion as a Multicomponent Process. A model and some cross-cultural data*, in P. Shaver (eds.), *Review of personality and social psychology: Emotions, relationships and health*, vol. 5, Sage, Beverly Hills, pp. 37–63.
- ID. (1986), *Vocal affect expression: a review and a model for future research*, in "Psychological Bulletin", 99-2, pp. 143–65.
- ID.. (1989), *Vocal correlates of emotional arousal and affective disturbance*, in H. Wagner, A. Manstead (eds.), *Handbook of Social Psychophysiology*, John Wiley & Sons, New York, pp. 165–97.
- ID. (1992), *What does facial expression express*, in K. Strongman (ed.), *International review of studies on emotion*, vol. 2, Wiley, Chichester, pp. 139–65.
- ID. (1993), *L'espressione vocale delle emozioni*, in Galati D. (a cura di), *Le emozioni primarie*, Bollati Boringhieri, Torino, pp. 47-70.
- ID. (1994), *Affect bursts*, in van Goozen S., van de Poll N.E, Sergeant J.A. (eds.) *Emotions: essays on emotion theory*, Erlbaum, Hillsdale, pp. 161-96.
- ID. (2009), *The dynamic architecture of emotion: Evidence for the component process model*, in "Cognition and Emotion", 23-7, pp. 1307–351. <http://doi.org/10.1080/02699930902928969>.
- SCHERER K. R., BANSE R., WALLBOTT H. G. (2001), *Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures*, in "Journal of Cross-Cultural Psychology", 32-1, pp. 76–92.

- SCHERER K.R., FELDSTEIN S., BOND R.N., ROSENTHAL R. (1985), *Vocal cues to deception: a comparative channel approach*, in "Journal of Psycholinguistic Research", 14, pp. 409-25.
- SCHERER K., LADD D., SILVERMAN K. E. A. (1984), *Vocal cues to speaker affect: Testing two models* in "The Journal of the Acoustical Society of America", 76-5, pp. 1346-56.
- SCHERER K.R., OSHINSKY JS. (1977), *Cue utilisation in emotion attribution from auditory stimuli*, in "Motivation and Emotion", 1, pp. 331-46.
- SCHLOSBERG H. (1952), *The description of facial expressions in terms of two dimensions*, in "Journal of Experimental Psychology", 44-4, pp. 229-37.
- SCHRÖDER M. (2003), *Experimental study of affect bursts*, in "Speech Communication. Special Issue Speech and Emotion", 40,1-2, pp. 99-116.
- SCHROEDER M., COWIE R., DOUGLAS-COWIE E., WESTERDIJK M., GIELEN S. (2001), *Acoustic correlates of emotion dimensions in view of speech synthesis*, in *Proceedings of Eurospeech 2001*, Vol. 1, International Speech Communication Association, Aalborg, pp. 87-90.
- SCHUTTE N.S., MALOUFF J. M., HALL L.E., HAGGERTY D.J., COOPER J.T., GOLDEN C.J. DORNHEIM L. (1998), *Development and validation of a measure of emotional intelligence*, in "Personality and Individual Differences" 25, pp. 167-77.
- SELINKER L. (1972), *Interlanguage*, in "IRAL-International Review of Applied Linguistics in Language Teaching", 10-1/4, pp. 209-32.
- sensitivity to the sound patterns of native language words*, in "Journal of Memory & Language", 32, pp. 402-20.
- SORIANELLO P. (2006), *Prosodia. Modelli e ricerca empirica*, Carocci, Roma.
- SORIANELLO P., DE MARCO A. (in stampa), *Sulla realizzazione prosodica in italiano nativo e non nativo*, in *Atti del XII Convegno Nazionale AISV, La fonetica sperimentale nell'insegnamento e nell'apprendimento delle lingue straniere* (Salerno, 27-29 gennaio 2016).
- SPEER S. R., ITO K. (2009), *Prosody in First Language Acquisition. Acquiring Intonation as a Tool to Organize Information in Conversation*, in "Linguistics and Language Compass" 3, pp. 90-110.
- STANISLAVSKI C. (1988), *An actor prepares*, Methuen, London (1937.)
- THOMPSON W. F., BALKWILL L. L. (2006), *Decoding speech prosody in five languages*, in "Semiotica", 158, pp. 407-24.
- TODA S., FOGEL A., KAWAI M. (1990), *Maternal speech to three-month-old infants in the*

- United States and Japan*, in "Journal of Child Language", 17-2, pp. 279–94.
- TOMKINS S. (1962), *Affect imagery consciousness: The positive affects*, vol. 1, Springer, New York.
- TOSHINORI ISHI C., CAMPBELL N. (2004), *Analysis of acoustic-prosodic features of spontaneous expressive speech*, in "Revista de Estudos Da Linguagem" 12-2, pp.37-49 (<http://doi.org/10.17851/2237-2083.12.2.37-49>, consultato il 5 novembre 2015).
- TRAGER L. (1958), *Paralanguage: a first approximation*, in "Studies in Linguistics" 13, pp. 1-12.
- TROJAN F. (1948), *Der Ausdruck der Sprechstimme: Eine phonetische*, Maudrich, Wien.
- ID. (1962). *Von Primär - und Sekundärfunktionen. Ein Beitrag zur Philosophie der Natur*, in "Studium Generale", 15, pp.1–8.
- TRUBECKOJ N. S. (1939), *Grundzüge der Phonologie*, Praga (trad. it. *Principi di fonologia*, Einaudi, Torino, 1971).
- ULDALL E. (1964), *Dimension of meaning in intonation*, in Abercrombie D., Fry D.B., MacCarthy
- VELTEN E. (1968), *A laboratory task for induction of mood states*, in "Behaviour Research and Therapy", 6-4, pp. 473–82.
- WENNERSTROM A. (1998), *Intonation as cohesion in academic discourse: A study of Chinese speakers of English*, in "Studies in Second Language Acquisition", 20, pp. 1-15.
- WHORF B. L. (1956), *Language, thought and reality*, MIT Press, Cambridge, MA.
- WIEDEN W. (1993), Aspects of acquisitional stages, in Kettemann B., Wieden W. (eds), *Current Issues in European Second Language Acquisition research*, Gunter Narr Verlag, Tübingen, pp. 125-35.
- WILLIAMS C. E., STEVENS K. N. (1972), *Emotion on speech: some acoustical correlates*, in "Journal of Acoustical Society of America", 52, pp. 1238–50.
- WILSON R. (2013), *Another language is another soul*, in "Language and Intercultural Communication", 13-3, pp. 298–309.
- WUNDT W. M. (1897), *Outlines of Psychology* (1907), Engelmann, Leipzig.
- ZEELLENBERG M., PIETERS R. (2004), *Beyond valence in customer dissatisfaction: A review and new findings on behavioral responses to regret and disappointment in failed services*, in "Journal of Business Research", 57,4, pp. 445–55.
- ZHU Y. (2013) *Expression and recognition of emotion in native and foreign speech : the*

*case of Mandarin and Dutch*, LOT, Utrecht.

ZINKEN J., KNOLL M., PANKSEPP J. (2008), *Universality and diversity in the vocalization of emotions*, in K. Izdebski (ed.), *Emotions of the human voice*, vol. 1, Plural, San Diego, pp. 185–202.

ZORZI D. (1996), *Dalla competenza comunicativa alla competenza interculturale*, in "Babilonia" 2-96, pp. 48-52.

ZUANELLI SONINO E. (1981), *La competenza comunicativa: precondizioni, conoscenze e regole per la comunicazione*, Bollati Boringhieri, Torino.

# Appendice

---

## Testi dei dialoghi induttori e del dialogo neutro nella versione italiana e in quella russa

### Dialogo 1- *Tristezza* (italiano)

A: Pronto?

B: Ciao, sei a casa?

A: Sì, sono a casa.

B: Tutto bene?

A: Non sai cos'è successo?

B: Cosa?

A: Luca è appena passato. Mi ha detto tutto. Non so cosa fare. Mi sento perso.

B: Ma cos'è successo?

A: **Lo hanno portato.**

B: Lo hanno portato?

A: Sì, poco fa. Non sapevo nulla. **Lo hanno portato.** Non posso fare niente.

B: Non posso crederci.

A: E' così invece, credimi. **Lo hanno portato.**

### Dialogo 1- *Tristezza* (russo)

A: Алло!

B: Привет, ты дома?

A: Да, я дома.

B: Всё хорошо?

A: ты не знаешь, что случилось?

B: что?

A: Лука только что был. Он мне всё сказал. Не знаю, что делать. Чувствую себя растерянной.

В: Но что случилось?

А: **Это привезли.**

В: Это привезли?

А: Да, только что. Я ничего не знал(а). **Это привезли.** Я ничего не могу сделать.

В: Не могу поверить.

А: Это так, поверь. **Это привезли.**

### **Dialogo 2- Collera** (italiano)

А: Pronto?

В: Ciao, tutto bene?

А: Ah, sei tu! Stavo per telefonarti!

В: Cos'è successo?

А: **Lo hanno portato!** Ti rendi conto? L'ho appena scoperto.

В: Non posso crederci.

А: E' così invece! **Lo hanno portato!** E io non sapevo niente.

В: Ma sei sicuro?

А: Sì che sono sicuro! **Lo hanno portato!**

### **Dialogo 2- Collera** (russo)

А: Алло!

В: Привет, всё хорошо?

А: это ты! Я в бешенстве!

В: Что случилось?

А: **Это привезли!** Ты понимаешь? Я только что заметила.

В: Не могу поверить.

А: Я тебе говорю, это так. **Это Привезли.** И я ничего не знал(а).

В: ты уверен(на)?

А: Да, уверен(на). **Это Привезли!**

**Dialogo 3- Gioia (italiano)**

A: Pronto?

B: Ciao, tutto bene?

A: Ciao!! Stavo per telefonarti! Non sai cos'è successo!!!

B: Cosa?

A: **Lo hanno portato!**

B: Quando?

A: Proprio ora!! Ancora non ci credo!! **Lo hanno portato!** Finalmente!!

B: Davvero?? Ne sei sicuro??

A: Sicurissimo! Me l'hanno appena detto! **Lo hanno portato!**

**Dialogo 3- Gioia (russo)**

A: Алло!

B: Привет, всё хорошо?

A: Привет!! Как раз собиралась(собирался) позвонить! Ты не знаешь, что случилось!!!

B: что?

A: **Это Привезли!**

B: Когда?

A: Прямо сейчас!! Ещё не могу поверить!! **Это Привезли!** Наконец!

B: Правда?? Ты увен(на)??

A: Точно уверен(на)! Мне только что сказали! **Это Привезли!**

**Dialogo 4- Eloquio normale/neutro (italiano)**

A: Pronto?

B: Salve, sono il tecnico. Chiamo per fare un controllo.

A: Sì, mi dica.

B: Hanno portato il nuovo apparecchio?

A: Un attimo che controllo. Sono appena tornato a casa.

B: Prego.

A: **Lo hanno portato.** C'è tutto mi pare.

B: Le hanno consegnato anche il nuovo cavo?

A: Un attimo, non lo vedo. Sì, c'è. **Lo hanno portato.**

**Dialogo 4-** *Eloquio normale/neutro* (russo)

A: Алло!

B: Здравствуйте, это техник. Звоню для проверки.

A: Да, говорите.

B: Вам привезли новое устройство?

A: Один момент, сейчас проверю. Я только вернулась домой.

B: жду.

A: **Это Привезли.** Мне кажется всё есть.

B: Вам выдали тоже новый кабель?

A: Момент, не вижу. Да, есть. **Это Привезли.**

## Valori relativi ai parametri osservati all'interno dei tre corpora (Ita L1, Ita L2 e Ru L1)

TAB. A-1. VALORI RELATIVI AI PARAMETRI TEMPORALI, FREQUENZIALI (IN HZ) E DI INTENSITÀ ALL'INTERNO DELLE  
PRODUZIONI EMOTIVE E DI QUELLE NEUTRE DEL CORPUS ITA L1 (ITALIANI NATIVI)

nome	emozione	Durata t	N sill	VdA	Intensità	Onset	Offset	F0x	F0 mx	F0 mn
I1	collera	883	6	6,795017	64	108	86	112	133	84
I2	collera	900	6	6,666667	59	85	75	115	131	75
I3	collera	1058	6	5,671078	69	130	101	136	165	102
I4	collera	1055	6	5,687204	70	148	108	148	182	109
I5	collera	1003	6	5,982054	72	243	195	277	336	189
I6	collera	1107	6	5,420054	68	235	190	239	273	170
I7	collera	1061	6	5,655042	71	220	189	246	310	161
I8	collera	783	6	7,662835	76	259	202	271	338	199
I9	collera	893	6	6,718925	75	260	216	253	302	187
I10	collera	1078	6	5,565863	65	213	154	213	254	143
I1	gioia	1157	6	5,185825	66	107	87	160	218	85
I2	gioia	879	6	6,825939	62	95	79	120	228	79
I3	gioia	1261	6	4,758128	78	158	102	223	294	102
I4	gioia	976	6	6,147541	75	147	98	215	246	98
I5	gioia	1112	6	5,395683	69	244	186	236	284	181
I6	gioia	1121	6	5,352364	72	185	179	242	295	178
I7	gioia	1106	6	5,424955	73	247	219	283	318	219
I8	gioia	872	6	6,880734	75	430	222	352	511	188
I9	gioia	761	6	7,884363	77	324	228	278	350	97
I10	gioia	897	6	6,688963	64	185	158	211	245	152
I1	tristezza	665	6	9,022556	50	93	75	92	116	75
I2	tristezza	622	6	9,646302	50	118	76	120	130	77
I3	tristezza	791	6	7,585335	64	117	101	110	120	96
I4	tristezza	856	6	7,009346	65	123	92	112	123	92
I5	tristezza	792	6	7,575758	59	175	155	174	185	147
I6	tristezza	963	6	6,23053	64	168	163	183	207	161
I7	tristezza	808	6	7,425743	60	182	161	170	189	146
I8	tristezza	777	6	7,722008	58	145	244	243	296	124
I9	tristezza	821	6	7,308161	63	265	257	256	297	209
I10	tristezza	1065	6	5,633803	53	223	188	201	224	90
I1	neutro	885	6	6,779661	60	104	87	110	131	86
I2	neutro	612	6	9,803922	52	86	85	136	295	85
I3	neutro	866	6	6,928406	67	120	118	135	162	103
I4	neutro	830	6	7,228916	66	153	106	125	153	106
I5	neutro	871	6	6,888634	63	254	177	233	284	153
I6	neutro	946	6	6,342495	60	228	182	211	253	160
I7	neutro	924	6	6,493506	62	277	191	238	288	175
I8	neutro	706	6	8,498584	66	296	231	255	309	194
I9	neutro	622	6	9,646302	61	291	198	221	302	192
I10	neutro	957	6	6,269592	59	200	147	185	242	130

**TAB. A-2. VALORI RELATIVI AI PARAMETRI TEMPORALI, FREQUENZIALI (IN HZ) E DI INTENSITÀ  
 ALL'INTERNO DELLE PRODUZIONI EMOTIVE E DI QUELLE NEUTRE DEL CORPUS ITA L2 (PARLATO EMOTIVO DEGLI  
 APPRENDENTI IN ITALIANO)**

nome	emozione	durata t	N sill	VdA	intensità	onset	offset	fox	mx	mn
R1	collera	988	6	6,072874	74	99	82	122	145	82
R2	collera	1132	6	5,300353	73	150	119	139	153	119
R3	collera	1120	6	5,357143	73	99	89	124	152	89
R4	collera	843	6	7,117438	72	108	87	112	123	87
R5	collera	942	6	6,369427	76	139	91	145	171	91
R6	collera	1035	6	5,797101	67	238	154	239	265	154
R7	collera	966	6	6,21118	64	234	262	307	334	231
R8	collera	1332	6	4,504505	72	134	130	141	160	127
R9	collera	1274	6	4,709576	73	146	121	132	146	118
R10	collera	1160	6	5,172414	75	135	114	133	149	114
R1	gioia	991	6	6,05449	78	146	87	151	196	87
R2	gioia	1166	6	5,145798	71	129	113	138	188	108
R3	gioia	983	6	6,103764	77	124	93	140	172	93
R4	gioia	921	6	6,514658	78	126	96	133	154	96
R5	gioia	964	6	6,224066	78	141	95	150	190	95
R6	gioia	1044	6	5,747126	64	360	219	370	419	210
R7	gioia	1082	6	5,545287	66	240	281	392	433	240
R8	gioia	1213	6	4,946414	70	120	110	122	137	108
R9	gioia	1097	6	5,469462	71	129	108	126	137	108
R10	gioia	1250	6	4,8	72	129	111	127	146	110
R1	tristezza	927	6	6,472492	60	112	94	92	114	82
R2	tristezza	1111	6	5,40054	62	130	116	118	130	93
R3	tristezza	898	6	6,681514	58	103	87	86	103	76
R4	tristezza	817	6	7,343941	60	95	87	93	104	83
R5	tristezza	961	6	6,243496	60	92	90	91	100	81
R6	tristezza	920	6	6,521739	65	225	163	250	272	161
R7	tristezza	944	6	6,355932	62	113	87	233	277	87
R8	tristezza	1173	6	5,11509	66	127	129	133	147	119
R9	tristezza	1200	6	5	67	133	131	142	161	119
R10	tristezza	1164	6	5,154639	64	138	105	121	138	104
R1	neutro	860	6	6,976744	70	113	84	113	133	80
R2	neutro	1222	6	4,909984	65	126	98	118	136	98
R3	neutro	874	6	6,864989	67	111	85	108	131	81
R4	neutro	684	6	8,77193	62	101	79	93	102	78
R5	neutro	803	6	7,47198	60	94	85	92	104	92
R6	neutro	914	6	6,564551	56	234	169	228	285	164
R7	neutro	1252	6	4,792332	68	121	110	125	158	101
R8	neutro	956	6	6,276151	68	123	98	116	127	98
R9	neutro	1034	6	5,802708	66	202	102	121	208	102
R10	neutro	754	6	7,95756	50	227	95	205	248	94

**TAB. A-3. VALORI RELATIVI AI PARAMETRI TEMPORALI, FREQUENZIALI (IN HZ) E DI INTENSITÀ  
 ALL'INTERNO DELLE PRODUZIONI EMOTIVE E DI QUELLE NEUTRE DEL CORPUS RU L1 (PARLATO EMOTIVO DEGLI  
 APPRENDENTI IN RUSSO)**

nome	emozione	durata t	N sill	VdA	Intensità	Onset	Offset	F0x	F0mx	F0mn
R1	collera	940	5	5,319149	72	115	93	124	243	93
R2	collera	966	5	5,175983	70	114	99	115	150	90
R3	collera	916	5	5,458515	72	137	84	129	175	84
R4	collera	885	5	5,649718	70	165	85	115	168	85
R5	collera	1072	5	4,664179	64	143	125	145	176	124
R6	collera	747	5	6,69344	73	431	178	277	431	118
R7	collera	1027	5	4,868549	66	161	141	152	177	109
R8	collera	1064	5	4,699248	67	150	125	146	171	124
R9	collera	1101	5	4,541326	63	128	129	154	185	126
R10	collera	806	5	6,203474	72	252	181	246	303	181
R1	tristezza	828	5	6,038647	60	80	82	88	99	79
R2	tristezza	909	5	5,50055	60	90	83	97	125	80
R3	tristezza	864	5	5,787037	60	81	92	102	129	81
R4	tristezza	751	5	6,65779	62	93	86	96	112	85
R5	tristezza	1019	5	4,906771	59	133	136	128	144	105
R6	tristezza	674	5	7,418398	58	229	232	218	249	78
R7	tristezza	1034	5	4,83559	54	128	114	124	141	101
R8	tristezza	1047	5	4,775549	54	127	114	124	139	106
R9	tristezza	1045	5	4,784689	61	137	132	132	148	118
R10	tristezza	722	5	6,925208	54	243	221	238	264	207
R1	gioia	976	5	5,122951	58	128	192	165	228	117
R2	gioia	906	5	5,518764	74	125	97	135	201	97
R3	gioia	877	5	5,701254	73	94	95	136	285	92
R4	gioia	839	5	5,959476	70	118	104	136	153	95
R5	gioia	793	5	6,30517	74	128	139	156	190	124
R6	gioia	807	5	6,195787	80	379	220	398	497	192
R7	gioia	1035	5	4,830918	63	126	192	177	237	121
R8	gioia	1056	5	4,734848	61	125	116	148	180	112
R9	gioia	1021	5	4,89716	61	143	133	156	178	124
R10	gioia	843	5	5,931198	79	251	202	402	498	183
R1	neutro	650	5	7,692308	63	96	91	96	114	79
R2	neutro	766	5	6,527415	69	118	93	123	147	93
R3	neutro	679	5	7,36377	63	104	93	96	110	87
R4	neutro	646	5	7,739938	64	104	93	96	110	87
R5	neutro	810	5	6,17284	55	126	108	118	132	106
R6	neutro	1083	5	4,616805	55	130	115	143	166	116
R7	neutro	805	5	6,21118	55	126	108	118	132	106
R8	neutro	856	5	5,841121	58	120	105	118	140	80
R9	neutro	649	5	7,70416	65	289	214	259	358	197
R10	neutro	590	5	8,474576	62	303	232	279	384	226

