

**UNIVERSITA' CATTOLICA DEL SACRO CUORE
MILANO**

**Dottorato di ricerca in Psicologia della comunicazione
e dei processi linguistici
ciclo XIX
S.S.D: M-PSI/01**

**L'INTEGRAZIONE CROSS-MODALE
DELLE EMOZIONI:
COMPONENTE MIMICA E VOCALE.
CORRELATI PSICOFISIOLOGICI (ERPs)**

**Tesi di Dottorato di: Alba Carrera
Matricola: 3280123**

Anno Accademico 2005/2006



**UNIVERSITA' CATTOLICA DEL SACRO CUORE
MILANO**

**Dottorato di ricerca in Psicologia della comunicazione
e dei processi linguistici
ciclo XIX
S.S.D: M-PSI/01**

**L'INTEGRAZIONE CROSS-MODALE
DELLE EMOZIONI:
COMPONENTE MIMICA E VOCALE.
CORRELATI PSICOFISIOLOGICI (ERPs)**

Coordinatore: Ch.mo Prof. Paola Di Blasio

**Tesi di Dottorato di: Alba Carrera
Matricola: 3280123**

Anno Accademico 2005/2006

A Michela: grazie di ciò che mi hai insegnato

Ai miei genitori: grazie perché da voi ho appreso l'impegno e la perseveranza

A Cate: grazie perché con nessun altro è come con te

A Claudio, a Davide, a Serafino, a Simona: grazie per il supporto e la compagnia

A Diego: grazie perché ora ci sei, per rimanere

L'INTEGRAZIONE CROSS-MODALE DELLE EMOZIONI:
COMPONENTE MIMICA E VOCALE.
CORRELATI PSICOFISIOLOGICI (ERPs)

RIASSUNTO

Il decoding delle emozioni costituisce un caso specifico di elaborazione multimodale di componenti provenienti da differenti canali sensoriali. Precedenti ricerche comportamentali e neuropsicologiche hanno evidenziato che l'elaborazione di stimoli emotivi sulla base di informazioni sensoriali multiple implica un processo di integrazione cross-modale. Il presente studio si prefigge di indagare il decoding simultaneo degli elementi vocali e della mimica facciale delle emozioni mediante i potenziali evocati corticali (ERPs), utilizzando un'ampia gamma di emozioni (gioia, tristezza, paura, rabbia, sorpresa e disgusto). Stimoli emotivi vocali (una parola a contenuto neutro pronunciata con tono emotivo) e patterns mimici (espressioni facciali delle emozioni) sono stati accoppiati in condizioni di congruenza (la medesima emozione per entrambi i canali sensoriali) o di incongruenza (emozioni differenti). Le variazioni ERPs e i tempi di risposta (TR) rilevati sono stati sottoposti ad analisi univariata della varianza per misure ripetute (ANOVA). Specificamente, sono stati analizzati quattro intervalli temporali (0-150; 150-250; 250-350; 350-500 ms post-stimulus), al fine di sondare i processi cognitivi (variazioni ERP a lunga latenza) oltre che quelli strettamente percettivi (a breve latenza), già ampiamente indagati dalle precedenti ricerche. L'ANOVA ha consentito di rilevare la presenza di numerosi effetti ERP con funzioni cognitive differenti. Alcuni di tali fenomeni sono altamente sensibili alla condizione di congruenza/incongruenza del pattern (con una maggiore ampiezza per stimoli congruenti rispetto a stimoli incongruenti) e costituiscono marker specifici dell'integrazione intersensoriale. In particolare, i dati permettono di riconoscere l'indice di media latenza P200 come un marker dell'integrazione intersensoriale di stimoli emotivi. Altri fenomeni, invece, maggiormente sensibili al contenuto emotivo, segnalano la presenza di processi cognitivi legati più in generale al decoding emotivo. I risultati indicano inoltre che l'integrazione, che nelle prime fasi di processamento è un fenomeno automatico, coinvolge successivamente processi decisionali intenzionali. Infine, è stato riscontrato che la condizione di congruenza provoca un effetto di riduzione dei TR per alcune delle emozioni analizzate (tristezza) ed un effetto inverso per un secondo gruppo di emozioni (paura, rabbia e sorpresa). Tale risultato viene discusso in riferimento al significato adattivo dei diversi correlati emotivi e dei rispettivi processi di decodifica cross-modale.

**CROSS-MODAL INTEGRATION OF EMOTIONS:
MIMIC AND VOCAL COMPONENTS.
PSYCHOPHYSIOLOGICAL CORRELATES (ERPs)**

ABSTRACT

Emotional decoding constitutes a case of multimodal processing of cues from multiple channels. Previous behavioral and neuropsychological studies indicated that the recognition of emotions on the basis of multiple perceptive information implies a cross-modal integration has place. The present study investigates the simultaneous processing of emotional tone of voice and emotional facial expression by event-related potentials (ERPs), through an ample range of different emotions (happiness, sadness, fear, anger, surprise, and disgust). Auditory emotional stimuli (a neutral word pronounced in an affective tone) and visual patterns (emotional facial expressions) were matched in congruous (the same emotion in face and voice) and incongruous (different emotions) pairs. ERPs variations and behavioral data (response time, TR) were submitted to repeated measures analysis of variance (ANOVA). We considered four time intervals (0-150; 150-250; 250-350; 350-500 ms post-stimulus), in order to explore the cognitive processes (long latency ERP variations) in addition to the perceptive (early ERP variations), fully investigated by previous researches. ANOVA showed numerous ERP effects, with different cognitive functions. Some of them, in particular the medium-latency P200, are highly sensible to pattern congruent/incongruent condition (with more intense amplitude for congruent rather than incongruent stimuli) and constitute intersensory integration specific markers. The other ERP effects, instead, are more sensible to the emotional content and signal the presence of cognitive processes that are more generally tied to the emotional decoding. Furthermore results show that, in the first processing phase, integration is an automatic and obliged phenomenon, while later it implies intentional decisional processes. Finally, a TR reduction was found for some congruous patterns (i.e. sadness) and an inverted effect for a second group of emotions (i.e. fear, anger, and surprise). Finally, behavioural results indicate that congruence causes a RT reduction for some emotions (sadness) and, on the contrary, an inverse effect for other emotions (fear, anger, surprise). This result is discussed with reference to different emotional correlates adaptive function and their respective cross-modal decoding processes.

INDICE

INTRODUZIONE	p. 1
1. Premessa	p. 1
2. Il viaggio è appena iniziato	p. 2
PARTE I: LO STATO ATTUALE DELLA RICERCA	
CAPITOLO 1: IL DECODING DELL'ESPRESSIONE FACCIALE DELLE EMOZIONI	p. 5
1. La comunicazione delle emozioni come processo	p. 5
2. I correlati emotivi: la prospettiva dimensionale	p. 6
3. Le funzioni delle emozioni	p. 7
4. Il contributo delle neuroscienze	p. 8
5. Il volto delle emozioni	p. 10
6. Breve storia dello studio delle espressioni facciali delle emozioni	p. 10
6.1. <i>Le teorie discrete delle emozioni</i>	p. 10
6.2. <i>La prospettiva dell'ecologia comportamentale</i>	p. 11
6.3. <i>I modelli dimensionali</i>	p. 12
7. L'apporto degli studi sul riconoscimento delle emozioni in età evolutiva	p. 13
8. La specificità del volto	p. 15
8.1. <i>Il modello gerarchico di Ellis e Young</i>	p. 16
9. I processi neurali sottesi al riconoscimento delle emozioni	p. 17
9.1. <i>L'indice N400: l'elaborazione delle anomalie semantiche</i>	p. 18
9.2. <i>Gli indici N230 e N270: il marker del decoding della mimica emotiva</i>	p. 18
9.3. <i>L'incidenza del contenuto emotivo</i>	p. 19
9.3.1. <i>Il contributo dell'elettromiografia facciale</i>	p. 20
10. Il ruolo dell'amigdala	p. 21
10.1. <i>La natura automatica del processamento dell'espressione emotiva</i>	p. 21

CAPITOLO 2: LA COMUNICAZIONE DELLE EMOZIONI ATTRAVERSO IL CANALE VOCALE NON VERBALE	p. 25
1. Introduzione	p. 25
2. Gli indicatori vocali implicati	p. 25
3. Gli studi empirici sulla comunicazione vocale delle emozioni	p. 27
3.1. <i>Problemi metodologici</i>	p. 27
3.2. <i>Le ricerche sull'encoding</i>	p. 29
3.2.1. <i>Il modello di Scherer</i>	p. 32
3.3. <i>Le ricerche sul decoding</i>	p. 34
4. Il contributo della neuropsicologia	p. 38
CAPITOLO 3: LA CONVERGENZA CROSS-MODALE AUDIO-VISIVA E LA SPECIFICITA' DEI PATTERN EMOTIVI	p. 41
1. Introduzione	p. 41
2. Il contributo degli studi comportamentali	p. 42
2.1. <i>L'integrazione multimodale come risposta alla complessità ambientale</i>	p. 42
2.2. <i>I principi che regolano la percezione multimodale</i>	p. 42
2.3. <i>La natura dell'integrazione: una questione aperta</i>	p. 44
3. L'apporto della neuropsicologia alla comprensione del processo di integrazione	p. 46
3.1. <i>Circuiti neurali implicati nel decoding intersensoriale</i>	p. 49
3.2. <i>I neuroni multisensoriali</i>	p. 50
4. La decodifica audio-visiva dei volti: riconoscere l'identità dal volto e dalla voce	p. 51
5. La convergenza di pattern emotivi	p. 52
6. L'apporto degli studi comportamentali	p. 53
6.1. <i>La funzione dell'integrazione delle informazioni emotive multimodali</i>	p. 55
6.2. <i>L'integrazione come processo precoce ed automatico</i>	p. 56
7. Il contributo della neuropsicologia	p. 57
7.1 <i>La componente MMN (mismatch negativity) come indicatore indiretto della convergenza audio-visiva</i>	p. 58
7.1.1. <i>Il ruolo dell'amigdala</i>	p. 59
7.1.2. <i>Valenza edonica e integrazione cross-modale</i>	p. 60
7.2. <i>Un altro indice indiretto: la componente N1</i>	p. 61
7.3. <i>L'indice di integrazione multimodale P2b</i>	p. 63
8. Gli studi sui casi clinici	p. 64

8.1. <i>Il fenomeno del blindsight</i>	p. 64
8.2. <i>La prosopagnosia</i>	p. 66
8.3. <i>Sistemi multipli di decodifica delle emozioni</i>	p. 68
PARTE II: LA RICERCA	
CAPITOLO 5: LA RICERCA: GLI OBIETTIVI E LA METODOLOGIA	p. 72
1. Introduzione	p. 72
2. Obiettivi e ipotesi	p. 73
3. Metodo	p. 75
3.1. <i>Partecipanti</i>	p. 75
3.2. <i>Stimoli</i>	p. 76
3.3. <i>Procedura sperimentale</i>	p. 78
3.4. <i>Procedure di registrazione degli indici ERP</i>	p. 80
CAPITOLO 6: ANALISI DEI DATI E RISULTATI	p. 84
1. Introduzione	p. 84
2. I tracciati ERP: descrizione qualitativa	p. 84
3. I dati ERP: analisi statistica	p. 87
3.1. <i>Confronto tra stimoli emotivi congruenti e incongruenti</i>	p. 89
3.1.1. <i>Indice N100</i>	p. 89
3.1.2. <i>Indice P100</i>	p. 90
3.1.3. <i>Indice N200</i>	p. 92
3.1.4. <i>Indice P200</i>	p. 95
3.1.5. <i>Indice N300</i>	p. 98
3.1.6. <i>Indice P300</i>	p. 101
3.1.7. <i>Indice N400</i>	p. 103
3.2. <i>L'universo emotivo: le analisi emotion-by-emotion</i>	p. 103
3.2.1. <i>Indice N100</i>	p. 105
3.2.2. <i>Indice P100</i>	p. 108
3.2.3. <i>Indice N200</i>	p. 112
3.2.4. <i>Indice P200</i>	p. 117
3.2.5. <i>Indice N300</i>	p. 127

3.2.6. <i>Indice P300</i>	p. 134
3.2.8. <i>Indice N400</i>	p. 142
3.3. <i>La condizione di incongruenza: ulteriori analisi</i>	p. 147
3.3.1. <i>Indice N200</i>	p. 148
3.3.2. <i>Indice P200</i>	p. 149
4. I dati comportamentali: analisi statistica	p. 151
CAPITOLO 7: DISCUSSIONE	p. 154
1. Per una visione d'insieme	p. 154
2. Le coordinate temporali del fenomeno: componenti percettive e componenti cognitive	p. 154
2.1. <i>I fenomeni percettivi precoci</i>	p. 155
2.2. <i>I fenomeni cognitivi di media latenza</i>	p. 155
2.3. <i>I fenomeni cognitivi di lunga latenza</i>	p. 156
3. La congruenza vs. incongruenza delle informazioni audiovisive	p. 156
3.1. <i>La P200: l'indice di integrazione intersensoriale di stimoli emotivi</i>	p. 157
3.2. <i>La N300: una N400 anticipata?</i>	p. 158
3.3. <i>La N200: il marker del contenuto emotivo</i>	p. 159
3.4. <i>La P300: l'updating cognitivo</i>	p. 160
3.5. <i>La N400: l'elaborazione delle anomalie semantiche</i>	p. 161
4. Emotion-by-emotion: la specificità dei differenti correlati emotivi	p. 162
4.1. <i>N200, arousal e valenza edonica: le coordinate dell'universo emotivo</i>	p. 162
4.2. <i>Indipendenza della P200 rispetto al contenuto emotivo</i>	p. 163
5. La convergenza intersensoriale: processo automatico o processo decisionale?	p. 163
5.1. <i>L'automaticità del processo durante le fasi percettive precoci</i>	p. 164
5.2. <i>L'intervento degli elementi decisionali</i>	p. 165
6. I tempi di risposta: l'efficienza dei fenomeni integrativi emotivi	p. 165
CAPITOLO 8: CONCLUSIONI	p. 167
RIFERIMENTI BIBLIOGRAFICI	p. 170
ALLEGATI	p. 186
Allegato 1: Questionario di validazione	p. 187

Allegato 2: I patterns facciali	p. 192
Allegato 3: Questionario finale	p. 198
Allegato 4: La consegna sperimentale implicita	p. 204
Allegato 5: La consegna sperimentale esplicita	p. 205

INTRODUZIONE

1. Premessa

Nel contesto delle interazioni quotidiane, ci troviamo ogni giorno ad osservare e interpretare il linguaggio del corpo altrui, inferendone pensieri, atteggiamenti, intenzioni, emozioni. Rispetto alle emozioni, sono soprattutto l'espressione del viso e il tono della voce a fornirci gli indizi più rilevanti per operare queste inferenze. La maggior parte delle volte, ciò avviene in modo del tutto spontaneo e automatico, cosicché ci rappresentiamo l'altro come "felice", o "triste" o "arrabbiato" senza neppure essere consapevoli degli indizi corporei su cui abbiamo basato la nostra attribuzione. Questo perchè, normalmente, i diversi canali sensoriali esprimono in modo sinergico la medesima emozione. A chiunque però è capitato di avvertire, di fronte ad un'espressione emotiva altrui, una sensazione di sgomento e di incertezza, che spinge a soffermarsi a riflettere sull'emozione che effettivamente l'altro sta comunicando. "Che gioia vederti qui!", dicono le parole e l'espressione del viso, eppure ci sembra di avvertire che la voce trema per la rabbia o sia incrinata dal pianto. Questa situazione si verifica ogni qualvolta che cogliamo nelle informazioni espresse dai differenti canali sensoriali emozioni contrastanti.

La ricerca descritta in queste pagine nasce proprio dal desiderio di approfondire alcuni particolari aspetti di quei complessi e intricati processi che ci portano ad individuare e a comprendere le emozioni altrui. Nello specifico, essa è stata condotta allo scopo di indagare il modo in cui il nostro sistema cognitivo è in grado di integrare in un'unica rappresentazione le diverse informazioni, permettendoci così di riconoscere le emozioni che le altre persone - spesso in modo congruente e talvolta invece in modo incongruente - esprimono utilizzando in modo congiunto i due principali canali comunicativi emotivi, ovvero la mimica facciale e la componente vocale.

2. Il viaggio è appena iniziato

Per facilitare la lettura, viene fornita una sintetica descrizione delle tematiche trattate nei seguenti capitoli.

Ritengo che, prima di intraprendere una più particolareggiata trattazione di quello che è l'oggetto specifico dello studio condotto, sia utile offrire una più generale e concisa panoramica dello stato attuale della ricerca che indaga il complesso universo delle emozioni. Il primo capitolo risponde a tale esigenza fornendo, d'altro canto, alcune essenziali basi teoriche che potranno essere utili in quanto chiave interpretativa di lettura dei seguenti capitoli. Il medesimo capitolo è inoltre dedicato al canale espressivo delle emozioni che maggiormente ha conquistato l'attenzione dei ricercatori, ovvero *il volto*. Vengono dunque esposti i meccanismi attraverso i quali le espressioni facciali fungono da strumento di comunicazione delle emozioni.

Nel secondo capitolo viene invece dato spazio al ruolo che un altro canale comunicativo, *la voce*, ha nell'espressione e nel riconoscimento degli stati emotivi.

La trattazione di tali tematiche suggerisce tuttavia che studiare singolarmente la funzione della componente visiva o di quella vocale nell'encoding e nel decoding emotivi è limitante, in quanto la realtà è ben più complessa: le emozioni infatti vengono espresse e identificate grazie ad una costellazione di informazioni appartenenti a differenti modalità sensoriali. Per questo motivo, nel terzo capitolo vengono illustrati i meccanismi dell'*integrazione intersensoriale* e nel quarto vengono messe in evidenza le peculiarità di un particolare tipo di integrazione, ovvero la *convergenza di informazioni emotive* di natura uditiva e visiva. Se fin dall'antichità le emozioni costituiscono l'oggetto della curiosità e dello studio dell'uomo, solo recentemente invece i ricercatori hanno rivolto l'attenzione verso i processi di integrazione cross-modale ad esse connessi. Di conseguenza, nonostante il fatto che numerose ricerche siano state condotte a fine di esplorare tale processo, tuttavia esse ancora non sono sufficienti per averne una conoscenza esauriente. Lo studio empirico che viene esposto nei capitoli relativi alla parte sperimentale mira dunque a portare nuova luce su quegli aspetti del fenomeno integrativo intersensoriale che maggiormente risultano essere lacunosi.

Nell'intento di perseguire tale scopo, si è ritenuto opportuno ricorrere alla rilevazione e analisi di indici di natura psicofisiologica. Negli ultimi decenni, infatti, lo studio delle emozioni si è orientato verso l'utilizzo sempre più massiccio di metodi di indagine di natura neuropsicologica e psicofisiologica. D'altro canto, dal momento che le emozioni scaturiscono da funzioni biologiche del sistema nervoso, non è possibile comprenderle appieno se non cercando di chiarire il loro funzionamento a livello cerebrale.

PARTE I

LO STATO ATTUALE DELLA RICERCA

IL DECODING DELL'ESPRESSIONE FACCIALE DELLE EMOZIONI

1. La comunicazione delle emozioni come processo

Che cos'è un'emozione? Oggi, c'è accordo tra la maggior parte degli studiosi nel ritenere che un'emozione sia un *fenomeno multicomponentiale* che permette all'individuo di intrattenere un'interazione dinamica con l'ambiente (Frijda, 2000).

Tuttavia, diverse interpretazioni sono state di volta in volta fornite per rispondere a tale domanda. Ad esempio, alcuni approcci considerano le emozioni come stati, mentre altri le identificano come processi, che hanno inizio con l'appraisal, ovvero la valutazione cognitiva, dello stimolo e terminano con la produzione di una risposta o di un set di risposte comportamentali. Inoltre, le emozioni possono essere viste come stati intrapersonali, mettendo così in primo piano il vissuto soggettivo, o lo stato di attivazione fisiologica, o il reclutamento di determinati pattern muscolari, oppure può essere intesa in primo luogo come un fenomeno interattivo, che mette in comunicazione l'individuo e l'ambiente. La scelta di prendere in esame un livello piuttosto che un altro ha importanti implicazioni dal punto di vista teorico. Ad esempio, la scelta di considerare le emozioni come stati soggettivi interni facilmente porta a considerare l'intenzionalità come un elemento marginale.

Ci si è anche chiesti che cosa caratterizzi le emozioni e permetta di differenziarle rispetto ad altre entità, come il pensiero o la motivazione. Molti teorici hanno posto l'accento sul *feeling* o *affect*, cioè su quell'esperienza che non può essere ridotta ad una sensazione corporea o ad un giudizio cognitivo (Arnold, 1960). Alcuni autori hanno centrato l'attenzione sull'impulso all'azione (Tomkins, 1962). Altri sostengono che l'elemento distintivo delle emozioni è il loro carattere valutativo: esse implicano sempre una valutazione che porta all'accettazione o al rifiuto dello stimolo che le ha elicitate. Si tratta di una valutazione molto differente da quella tipicamente cognitiva: non

sapremmo dare una motivazione razionale del perchè il profumo della rosa è buono e ci invoglia ad annusarlo. Questo tipo di valutazione, che in parte costituisce il risultato pressoché automatico di certe stimolazioni ed in parte scaturisce dalla valutazione cognitiva delle stesse, viene frequentemente definito appraisal (Lazarus, 2000). In particolare, le *teorie dell'appraisal* assumono che le emozioni abbiano origine da una valutazione soggettiva del significato degli eventi in termini di rilevanza rispetto ai propri bisogni e scopi (Scherer, 2003). Un importante stimolo all'indagine delle emozioni è giunto, a partire dagli anni '70, con l'emergere l'approccio dell'*information processing*. Esso ha contribuito largamente a spiegare quella concatenazione di processi mentali che, a partire dalla percezione dello stimolo emotigeno, portano all'insorgere dell'emozione. Ma non solo: è stato grazie all'affermarsi di questo approccio teorico che i ricercatori hanno iniziato a volgere la loro attenzione su un altro tipo di processo, ovvero quello che, partendo dalla percezione di un'emozione espressa da un'altra persona, porta al riconoscimento e alla comprensione della stessa. Frijda (2000) suggerisce che in realtà ciò che caratterizza l'emozione non sia una di queste specifiche componenti, ma la loro combinazione.

2. I correlati emotivi: la prospettiva dimensionale

Un'altra questione ampiamente discussa riguarda i criteri che permettono di distinguere le differenti emozioni. In base a che cosa definiamo le emozioni come gioia, rabbia, tristezza, e le distinguiamo l'una dall'altra?

Per rispondere a tale interrogativo, si è fatto riferimento ad elementi come il livello di prontezza all'azione, il comportamento espressivo, il risultato dell'appraisal, il tipo di evento elicitante. E in effetti, questi elementi, e soprattutto la combinazione tra di essi, permette di differenziare le diverse emozioni, sia che le si consideri come categorie e stati dai confini ben definiti sia che le si consideri come processi.

Tuttavia, diversi ricercatori propongono una differente prospettiva, affermando che queste combinazioni di elementi sono piuttosto confuse e variabili, e che in realtà le etichette emotive che le identificano hanno una natura arbitraria. In altri termini, essi sostengono che le etichette linguistiche che utilizziamo per denominare le diverse

emozioni corrispondono a prototipi o script che hanno un'*origine culturale* o linguistica (Russell, 1991). I prototipi riflettono pattern di componenti che si presentano in modo frequente o che hanno una particolare rilevanza dal punto di vista sociale. La prospettiva culturale è stata massimamente enfatizzata dal costruzionismo sociale, secondo il quale il significato dell'emozione è "prescritto" dall'etichetta emotiva, che riflette degli script che hanno un'origine sociale (Harrè & Parrott, 1996). Il significato dell'emozione è quindi fornito dalle norme sociali relative al provare ed esprimere le diverse emozioni. Secondo il modello della continuità proposto da Russell e Fernández-Dols (1998), le emozioni possono essere rappresentate non mediante categorie discrete ma piuttosto lungo un continuum derivato dall'intersezione dei due assi rappresentazionali della *valenza edonica* e del grado di *attivazione*: tali componenti, insieme alle informazioni contestuali, permettono all'osservatore di interpretare e comprendere la mimica emotiva. Gli autori sostengono che, rispetto al decoding emotivo, nel momento in cui l'osservatore percepisce l'espressione emotiva dell'altro, in modo immediato e automatico egli giudica il livello generale di attivazione e di piacere dell'esprimente e, in seguito a tale valutazione, gli attribuisce una specifica categoria emozionale, sulla base del fatto che le informazioni disponibili corrispondono al prototipo di quell'emozione. Il riconoscimento dell'emozione implica quindi un atto di inferenza che avviene a partire dalla valutazione delle due componenti dimensionali dell'arousal e della valenza edonica e delle informazioni contestuali.

3. Le funzioni delle emozioni

Le emozioni, mediando il rapporto tra l'individuo e l'ambiente, sono tutte finalizzate alla sopravvivenza. Più nello specifico poi, ciascuna ricopre una funzione particolare e può quindi essere descritta come un'unità funzionale autonoma.

Facendo riferimento a tale prospettiva evolutivista è necessario tuttavia operare una distinzione tra la *funzione evolutiva* e la *funzione prossimale* e contingente. Dal momento che il significato funzionale delle emozioni si è evoluto nel corso dello sviluppo socio-culturale dell'uomo, la stessa emozione, espressa da un uomo primitivo nella savana o da un uomo nostro contemporaneo nel corso di una riunione di lavoro,

possono avere funzioni diverse. E' infatti ipotizzabile che lo sviluppo culturale abbia oggi almeno in parte soppiantato la componente evolutiva originaria, pur mantenendo la medesima funzione di conservazione personale, sia in termini di salvaguardia fisica sia in termini di salvaguardia della propria immagine di sé e della propria autostima e di perseguimento dei propri scopi e bisogni personali e sociali.

O'Keefe (1988) ha descritto tre "teorie implicite della comunicazione" o "*message designe logics*" che, secondo l'autore, rendono conto delle origini dei messaggi comunicativi. Nella logica espressiva, la comunicazione è un mezzo per esprimere direttamente pensieri e stati d'animo. Nella logica convenzionale, la comunicazione è impiegata per esprimere ciò che è adeguato al contesto sociale in cui essa si esplica: essa quindi ha una funzione cooperativa, in quanto permette di gestire le relazioni sociali in accordo a determinate regole e procedure convenzionalizzate. Nella logica retorica infine la comunicazione è utilizzata in modo strategico per perseguire e negoziare i reciproci scopi. Planalp e Knie (2002) suggeriscono che tali principi possono essere applicati anche a quella particolare forma di comunicazione che è la comunicazione emotiva: infatti le espressioni delle emozioni possono essere una diretta esternazione dello stato interiore, oppure possono essere consapevolmente adattate alla situazione e alle regole sociali oppure possono essere manifestate al fine di raggiungere uno specifico scopo sociale. O ancora, si può ritenere che nella maggioranza dei casi questi tre livelli, anziché essere indipendenti, data la complessità delle situazioni sociali, operino simultaneamente.

4. Il contributo delle neuroscienze

Esiste uno stretto legame tra le emozioni e i processi cognitivi, che costituiscono il principale fuoco di interesse delle neuroscienze. Lane e colleghi (Lane et al., 2000), riflettendo su questo attuale orientamento metodologico, esplicitano tale legame individuando alcuni punti di stretto contatto:

- L'emozione, anche se in misura variabile in funzione del tipo di processamento dello stimolo, prevede sempre un'appraisal cognitivo, in virtù del quale lo stimolo elicitante è sottoposto ad un'articolata valutazione cognitiva;
- L'emozione, comprendendo una dimensione di esperienza soggettiva, chiama in causa la consapevolezza, che è oggetto di studio delle neuroscienze. In realtà è solo il prodotto del processamento dello stimolo ad essere cosciente, mentre sia la valutazione dello stesso sia l'esecuzione della risposta emotiva avvengono a livello inconscio (LeDoux, 1996);
- E' possibile che i processi di encoding e decoding delle emozioni e i processi cognitivi superiori condividano sistemi di risposta che si sovrappongono;
- Le emozioni possono influenzare pesantemente i processi cognitivi, come nel caso della memoria, dell'attenzione o della percezione.

Come ha saggiamente scritto LeDoux (1998, p. 107), "cercare di immaginare come funzioni una mente è un processo descritto dal linguista Steven Pinker come ingegneria inversa. La macchina c'è e bisogna capire come funziona: perciò smontiamo il cervello nella speranza di vedere a cosa mirasse l'evoluzione quando lo ha assemblato".

La difficoltà che incontrano tutti coloro che si accingono a questa impresa consiste nel fatto che le basi neurali delle emozioni, costituendo esse stesse degli strumenti comunicativi evoluti e complessi, sono a loro volta estremamente complesse. Un risultato comune delle ricerche che hanno indagato le emozioni a livello psicofisiologico è che l'esperienza emotiva di cui noi siamo consapevoli non è altro che una piccola parte del sistema che ha generato l'emozione. Come avviene per i processi di tipo strettamente cognitivo, solo il risultato dell'elaborazione cerebrale raggiunge la soglia della consapevolezza. E tra l'altro, a volte neppure il risultato dell'elaborazione diviene cosciente. Di conseguenza, occorre considerare gli stati affettivi soggettivamente esperiti come il prodotto di un processo che si svolge inconsciamente. Questo processo valuta lo stimolo emotigeno preparando così una base per progettare una condotta comportamentale coerente con la situazione. Un importante corollario di questi risultati è che, una volta appurato che le emozioni, così come i pensieri, sono il

prodotto cosciente di un'elaborazione non conscia di stimoli, la separazione netta che tanto a lungo è stata postulata tra cognizione ed emozione decade, lasciando intravedere non soltanto divergenze ma anche punti di contatto tra questi due mondi apparentemente tanto diversi.

5. Il volto delle emozioni

Con il termine "emozioni" facciamo comunemente riferimento ad una varietà pressoché infinita di fenomeni complessi e multifaccettati. Tali fenomeni, benché prima di tutto percepiti come esperienze interne e fortemente personali, trovano la loro più alta espressione nel momento in cui vengono condivisi all'interno di un contesto sociale. I seguenti paragrafi sono dedicati ad uno dei più potenti mezzi che abbiamo a disposizione per condividere le nostre emozioni, ovvero il volto.

6. Breve storia dello studio delle espressioni facciali delle emozioni

Lo studio delle espressioni facciali delle emozioni ha avuto inizio con la pubblicazione da parte di Charles Darwin del trattato "L'espressione delle emozioni nell'uomo e negli animali" (1872-1998). Nel libro, lo studioso sostiene che ogni emozione viene espressa tramite particolari configurazioni facciali e che tali espressioni sono innate e universali. A partire dagli studi di Darwin, tra i vari canali espressivi delle emozioni, la mimica facciale è quella che maggiormente ha catturato l'attenzione dei ricercatori e che maggiormente ha dato adito a numerose controversie (Keltner & Ekman, 2000).

6.1. Le teorie discrete delle emozioni

Le argomentazioni di Darwin sono rimaste indiscusse per decenni, avvalorate dalle *teorie discrete* delle emozioni (Ekman, 1994). Secondo la *teoria neuroculturale* di Ekman esiste infatti un numero ristretto di emozioni di base, che sono determinate geneticamente, che attivavano differenti percorsi neurali e che vengono espresse

attraverso uno specifico pattern espressivo neuromuscolare. Tali programmi sarebbero innescati da specifici eventi prototipici la cui salienza si è venuta a costituire nel corso dell'evoluzione (Izard 1977; Tomkins, 1962). I sostenitori delle teorie discrete tuttavia non hanno mai negato un ruolo, seppure piuttosto limitato, al contesto socio-culturale e all'apprendimento: essi infatti sostengono che, man mano che acquisisce le regole sociali che caratterizzano la comunità di appartenenza, l'individuo impara a modulare l'espressione spontanea delle sue emozioni, in modo tale da renderla adeguata rispetto alle norme sociali vigenti. Più nello specifico, attraverso quelle che Ekman chiama *regole di esibizione*, l'individuo impara a simulare, a dissimulare, a intensificare o deintensificare la manifestazione dei propri stati emotivi, adattandoli al contesto.

6.2. La prospettiva dell'ecologia comportamentale

I presupposti della teoria neuroculturale sono stati negli anni messi in discussione dai sostenitori dell'*ecologia comportamentale*, i quali hanno portato numerosi dati sperimentali a sostegno dell'ipotesi che le espressioni facciali non sono espressione di stati emotivi interni discreti e neurologicamente determinati, ma sono *segnali sociali* che hanno lo scopo di influenzare il comportamento dell'altro: esse sono "strumenti sociali" che permettono la negoziazione delle interazioni tra le persone (Fridlund, 1994). Le espressioni facciali sono quindi da considerarsi come una manifestazione di intenzioni e atteggiamenti sociali. Ad esempio, quella che per la prospettiva categoriale è l'espressione facciale della rabbia per la prospettiva dell'ecologia comportamentale è espressione della prontezza all'attacco, l'espressione della paura è espressione della prontezza alla sottomissione, il sorriso è prontezza all'affiliazione. In altri termini, le espressioni hanno la caratteristica di essere specifiche rispetto all'intenzione sottesa e rispetto al contesto, piuttosto che essere vincolate ad uno specifico stato emotivo interno.

Ne deriva che, a differenza che nella teoria neuroculturale, non viene fatta nessuna distinzione tra espressioni spontanee ed espressioni "artefatte" per aderire alle norme socio-culturali, poiché tutte le espressioni sono considerate strumento attraverso il quale l'individuo realizza i suoi scopi sociali nel contesto delle interazioni. Inoltre, il

significato delle espressioni deriva dal contesto in cui esse si manifestano, e il contesto è costituito sia dalla situazione fisica e sociale sia dal sostrato comune di conoscenza e dal “campo comune” che si è venuto a costituire a seguito delle precedenti interazioni. A tal proposito, bisogna considerare come fondamentale elemento il contesto privato di cui ciascun individuo è portatore, e che comprende le aspettative e i bisogni che lo spingono verso l'interazione. Per l'ecologia comportamentale, il ricevente dell'espressione facciale e la decodifica della stessa hanno un ruolo di primo piano: la mia espressione acquista un significato nel momento in cui l'altro la utilizza per inferire il mio stato d'animo e le mie intenzioni e per anticipare la mia futura condotta comportamentale.

6.3. I modelli dimensionali

Ad integrazione delle precedenti teorizzazioni e a seguito di decenni di ricerca e di dibattito, oggi si sono largamente affermate le *teorie dimensionali*. Tra queste, sicuramente interessante è il modello elaborato da Russell e Fernández-Dols (1998): esso propone che le emozioni possano essere rappresentate non mediante categorie discrete e indipendenti, di natura innata, ma piuttosto lungo un continuum derivato dall'intersezione di due assi rappresentazionali di base, quali il significato edonico dell'emozione e il grado di attivazione ad essa correlato (arousal), componenti che, assieme alle informazioni contestuali, guidano l'osservatore nell'attribuire un significato complessivo alla mimica emotiva.

In linea con quanto ipotizzato dall'approccio della dimensionalità emotiva (Ellsworth & Scherer, 2003), l'universo delle emozioni si configura per la distribuzione in “famiglie emotive” al loro interno caratterizzate da una consistenza e omogeneità di proprietà strutturali, queste ultime definite principalmente dai due assi categoriali della valenza edonica e dell'arousal. D'altro canto, l'importanza di questi due parametri è stata sottolineata anche da quelle ricerche che hanno indagato il riconoscimento delle espressioni facciali delle emozioni in età evolutiva (Balconi & Carrera, 2006) (vedi par. 3). Ad integrazione di tale modello, utilizzando un recente approccio alla rappresentazione dell'universo emotivo che introduce il costrutto di “sfocatura”

concettuale tra le categorie emotive (categorie *fuzzy*) (Rosch & Mervis, 1981; Violi, 1997), appare fuorviante ipotizzare l'esistenza di confini rigidi tra correlati emotivi, ma piuttosto diviene più utile parlare di somiglianze tra categorie di correlati emotivi, la cui maggiore o minore vicinanza è definita dal possesso di alcuni elementi prototipici e discriminanti oppure no (Bullock & Russell, 1986).

7. L'apporto degli studi sul riconoscimento delle emozioni in età evolutiva

Un importante aiuto nel tentativo di comprendere i meccanismi della decodifica della mimica emotiva giunge da quelle ricerche che ne hanno indagato il riconoscimento in età evolutiva. Per questo motivo, di seguito viene illustrato lo sviluppo di tale fondamentale abilità sociale.

Dal momento che l'abilità nel riconoscere e comprendere la mimica emotiva altrui costituisce una competenza sociale di primaria importanza, già nelle prime fasi dello sviluppo i bambini presentano una discreta competenza nel decoding, che viene progressivamente affinata fino al raggiungimento dell'età adulta. Nello specifico, l'accuratezza nel *face processing* aumenta al crescere dell'età ed inoltre differisce in funzione delle diverse emozioni: tra le emozioni fondamentali, quelle più facilmente e precocemente riconosciute sono la gioia, la tristezza e la rabbia, seguite dalla paura ed infine dal disgusto e dalla sorpresa (Widen & Russell, 2003; 2004). Adottando una prospettiva di più ampio respiro, l'analisi dei processi sottostanti al riconoscimento delle emozioni a partire dalle espressioni facciali fornisce importanti informazioni non solo sulla *lessicalità* della mimica facciale ma anche sulle modalità di *concettualizzazione* della stessa. Bullock e Russell (1986) suggeriscono che i bambini sviluppino un sistema di rappresentazione e classificazione delle emozioni diverso rispetto a quello utilizzato dall'adulto. Tale sistema sarebbe caratterizzato dalla presenza di un esiguo numero di categorie molto ampie, che si costituiscono sulla base di due assi dimensionali: la valenza edonica e l'arousal. Tale ipotesi ha trovato conferma grazie ad alcuni studi sperimentali che hanno rilevato come inizialmente i bambini interpretino le espressioni facciali facendo riferimento alle dimensioni di piacere-dispiacere (valenza edonica bipolare) e di intensità (alta o bassa attivazione): solo successivamente essi

utilizzerebbero etichette verbali che, dapprima esigue e ampie in termini di inclusività, si fanno progressivamente sempre più articolate e specifiche (Schwarzer, Zauner, & Korell, 2003). Alcuni studi hanno indagato un'ulteriore rilevante acquisizione, ovvero la denominazione lessicale delle espressioni facciali: al crescere dell'età aumenta la tendenza a ricorrere alla denominazione lessicale delle emozioni, il lessico emotivo si arricchisce e diminuiscono gli errori commessi nei compiti di etichettamento delle espressioni facciali (Russell & Widen, 2002). Inoltre, le emozioni più precocemente e facilmente riconosciute a partire dall'espressione del volto (gioia, rabbia e tristezza) sono anche quelle le cui etichette verbali compaiono per prime - tra i 2 e i 3 anni - e quelle che, anche successivamente, risultano più accessibili e quindi più ampiamente utilizzate. Le etichette più tardive, che appaiono intorno ai 5 anni, sono quelle relative alle emozioni della sorpresa e del disgusto. Più in generale è possibile sostenere che l'etichetta emotiva costituisce la meta finale di un processo di sviluppo che passa da definizioni di tipo dimensionale (ad esempio le dimensioni della piacevolezza/spiacevolezza del correlato) a definizioni legate alla situazione (la rappresentazione dello script emotivo) per arrivare infine alle etichette verbali, che si collocano ad un livello di maggiore astrazione.

Una ricerca di Widen & Russell (2004) introduce un'ulteriore distinzione nella rappresentazione del correlato emotivo attribuendo un ruolo rilevante alla rappresentazione dei legami di causa-effetto delle emozioni, inteso come insieme di azioni causali da un lato e di conseguenze comportamentali dall'altro che costituiscono elementi esplicativi dell'esperienza emotiva, così come essa è stata esperita dal soggetto. Pertanto, un fattore di rilievo nel riconoscimento della mimica emotiva è costituito dal concetto di *contestualità emotiva*. Come sottolineato da Russell e Widen (2002) è importante considerare che, nell'esperienza quotidiana, le espressioni facciali sulla base delle quali il bambino inferisce i vissuti emotivi sono sempre collocate all'interno di uno specifico contesto interattivo. Alla luce di tale considerazione acquista primaria importanza il concetto di script emotivo: l'individuo giunge al riconoscimento di una specifica emozione verificando la presenza di una serie di elementi prototipici e ordinati secondo specifiche sequenze causali e temporali. Oltre alle espressioni facciali, sono

inclusi in tale rappresentazione le cause attivanti, il contesto fisico e sociale, le azioni e le loro conseguenze, l'esperienza soggettiva, la valutazione cognitiva della situazione. Attraverso un progressivo processo di generalizzazione degli script, inizialmente legati a situazioni specifiche e circoscritte, il bambino giunge ad una comprensione situazionale delle emozioni, condizione che caratterizza il decoding emotivo in età adulta (Fridlund, 1994).

8. La specificità del volto

Per il funzionamento cognitivo umano, il volto è uno *stimolo altamente prioritario*, che viene elaborato in modo differenziato rispetto agli altri oggetti (Balconi, 2004). Normalmente, gli stimoli vengono elaborati attuando strategie analitiche, tali per cui le singole proprietà degli oggetti vengono analizzate una ad una (processamento di tipo *piecemeal*). Il volto costituisce invece, proprio in funzione della sua pregnanza, un caso particolare, in quanto viene percepito come una Gestalt, ovvero come una configurazione unitaria non scomponibile in ulteriori componenti, e viene elaborato applicando una *strategia di tipo olistico* che, quando il volto esprime un'emozione, rende possibile un riconoscimento immediato del significato emotivo. A livello neurofisiologico, l'importanza del volto come canale comunicativo è evidenziata dall'esistenza di un sistema neurale volto-specifico, e cioè di un insieme di neuroni collocati nella corteccia inferotemporale dotati di indipendenza funzionale e deputati al processamento del volto (Allison et al., 1999). Alcune prove della specificità del volto giungono anche dagli studi condotti utilizzando i potenziali evocati corticali, i quali hanno messo in evidenza l'esistenza di componenti specifiche legate alla decodifica dei volti e non evocate invece dalle presentazioni di altri stimoli visivi, che compaiono tra i 140 e i 300 ms dopo la presentazione dello stimolo (Boetzel & Grusser, 1989).

8.1. Il modello gerarchico di Ellis e Young

Secondo il modello gerarchico di riconoscimento dei volti di Ellis e Young (1995) il processo di decodifica dei volti prevede il contributo di molteplici livelli informativi: ad esempio un livello di elaborazione strutturale, un livello di screening del grado di familiarità, un livello di discriminazione dell'espressione emotiva. Per ogni livello informativo, esiste un sistema neurale dedicato, che implica l'attivazione di codici differenziati e specifici (vedi fig. 1).

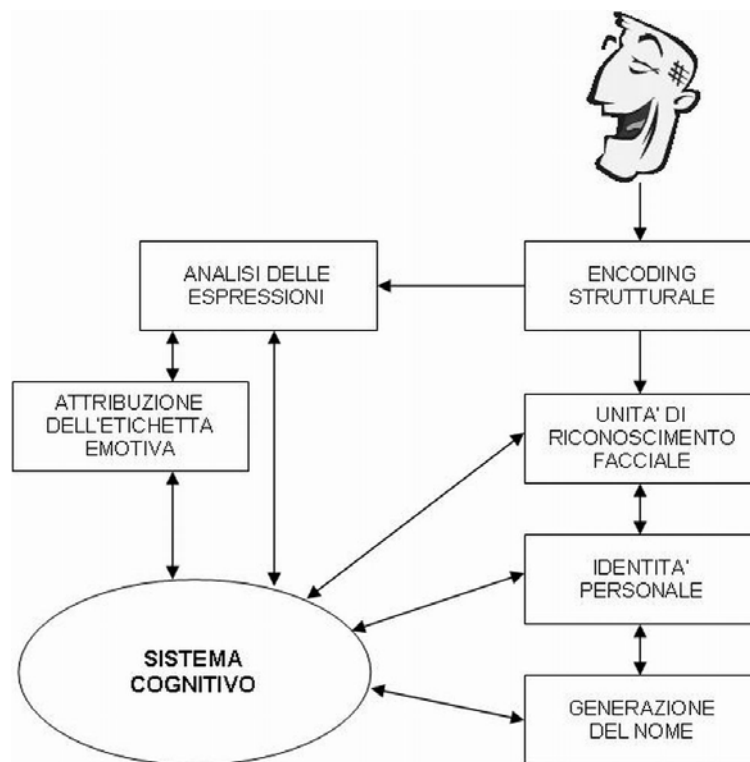


Fig. 1. Processamento gerarchico dei volti, comprendente diversi livelli informativi.

Tra i diversi codici di elaborazione, uno in particolare è dedicato all'identificazione del significato emotivo dello stimolo facciale. Questo complesso sistema di elaborazione dei volti comprende componenti consapevoli e componenti non consapevoli, che vengono regolate da due differenti vie neurali cortico-limbiche. Il decoding esplicito delle informazioni passa attraverso la via ventrale, mentre il processamento implicito impiega la via dorsale. La conferma della presenza di questi due distinti sistemi di elaborazione giunge dagli studi condotti su pazienti prosopagnosici, i quali

presentano difficoltà di riconoscimento dei volti. Nonostante tale deficit, essi infatti conservano quegli specifici aspetti del riconoscimento che avvengono in modo automatico senza raggiungere il livello della coscienza.

L'attendibilità del modello gerarchico è avvalorata da numerosi dati sperimentali. Ad esempio, in un esperimento condotto utilizzando un campione clinico, Etcoff (1986) ha rilevato la presenza di abilità cognitive e processi cerebrali differenziati per il riconoscimento di volti e per il riconoscimento di volti esprimenti un'emozione, sebbene entrambi questi processi fossero mediati dall'emisfero destro.

9. I processi neurali sottesi al riconoscimento delle emozioni

Diversi studi di neuroimaging o basati su casi clinici indicano che il decoding delle espressioni emotive del volto è un processo complesso che, dopo un iniziale stadio percettivo in cui vengono elaborate le caratteristiche strutturali del volto, si differenzia dalla semplice percezione del viso con espressione neutra. Infatti, esiste una differenziazione neurale e funzionale tra gli specifici processi percettivi responsabili del decoding strutturale del volto e un più elevato livello di processamento che permette di associare la rappresentazione di un viso con informazioni di tipo semantico, come ad esempio l'espressione emotiva (Bentin & Deouell, 2000). Nella prima fase di elaborazione del volto, che termina intorno ai 180 ms dopo la presentazione dello stimolo, vengono elaborate le *caratteristiche strutturali*. Questa fase è rivelata dagli indici ERP P120, localizzato nelle aree occipitali posteriori e N170, localizzato nella corteccia anteriore e laterale (Linkenkaer et al., 1998). Tali deflessioni sono sensibili alle caratteristiche strutturali del volto (Balconi & Lucchiari, 2005), mentre non presuppongono alcuna elaborazione del contenuto semantico dello stimolo. Successivamente, lo stimolo ha accesso a stadi di *processamento cognitivo* che coinvolgono la sfera semantica. La decodifica del correlato emotivo espresso dalla mimica si colloca appunto in questa fase tardiva di elaborazione (Junghöfer et al., 2001).

Di seguito, viene presentata una panoramica dei principali indici ERP impiegati nello studio della decodifica delle espressioni facciali delle emozioni.

9.1 L'indice N400: l'elaborazione delle anomalie semantiche

E' nella fase più propriamente cognitiva che alcune ricerche hanno segnalato la comparsa del picco ERP N400, un indice di rielaborazione che si manifesta quando le informazioni sono percepite come anomale dal punto di vista semantico o a seguito della violazione dei vincoli contestuali dello stimolo (Debruille et al., 1996). Recentemente uno studio ERP (Balconi, 2005) condotto allo scopo di operare un confronto tra il processamento di informazioni semantiche di tipo linguistico e l'elaborazione semantica di stimoli facciali esprimenti un'emozione ha evidenziato che il decoding di espressioni facciali anomale (gli stimoli erano stati sottoposti ad un'operazione di *morphing*, da cui risulta un'interpolazione tra due diverse espressioni facciali) rispetto alla decodifica di espressioni congruenti elicitava intorno ai 360 ms dopo l'onset dello stimolo una più ampia deflessione negativa identificabile come una N400. Tale indice ERP, localizzato nella porzione posteriore dello scalpo, viene ritenuto un marker specifico della rilevazione di una *anomalia semantica*. Nonostante il fatto che la N400 sia stata studiata soprattutto impiegando stimoli di tipo linguistico, essa in realtà è presente in ogni caso in cui avviene una violazione delle rappresentazioni semantiche (Balconi, 2002), indicando una indipendenza rispetto al task, che può essere linguistico o non linguistico.

9.2. Gli indici N230 e N270: i marker del decoding della mimica emotiva

Altre ricerche hanno dimostrato la presenza di indici cognitivi che sono strettamente legati al *contenuto emotivo*: la N230 presenta una notevole intensificazione quando vengono percepiti stimoli che esprimono un'emozione rispetto a quando vengono presentati stimoli neutri (Eimer & McCarthy, 1999). Streit (Streit et al., 2000) in un compito di decoding ha rilevato che l'attività puramente percettiva individuata dagli indici ERP P120 e N170 non differisce per gli stimoli neutri e quelli emotivamente connotati. Già dopo 180 ms dalla presentazione dello stimolo, tuttavia, l'inizio del decoding specifico del contenuto emotivo è rivelato da un ampio picco negativo che

raggiunge la massima intensità intorno ai 240 ms (Marinckovic & Halgren, 1998) e che è pressoché assente quando viene percepito uno stimolo neutro. Tale elaborazione specifica dedicata al contenuto emotivo della stimolazione sarebbe soprattutto a carico dell'emisfero destro (Banich, 1997) e principalmente localizzata nella zona postero-occipito-temporale. I risultati delle ricerche consentono tra l'altro di supportare quanto previsto dal modello gerarchico dell'elaborazione dei volti di Ellis e Young (1998) (vedi par. 8.1. del presente capitolo). In particolare, gli indicatori ERP hanno confermato il significato funzionale degli stadi di elaborazione del volto, identificabili tramite "codici" cognitivi distinti, di cui uno direttamente implicato nella decodifica del contenuto emotivo veicolato dal volto.

La pregnanza delle informazioni visive di tipo emotivo emerge anche da un recente studio (Sato et al., 2001) che ha dimostrato come il processamento visivo può essere modulato dal significato emotivo degli stimoli emotivi percepiti. Rispetto ai volti neutri, quelli che esprimono emozioni (nello studio sono state utilizzate le emozioni della gioia e della paura) elicitano infatti un più ampio picco negativo intorno ai 270 ms nelle aree temporali posteriori. L'incremento della N270 indica che la presenza di informazione emotiva nel volto provoca un'intensificazione dei processi di elaborazione visiva. Gli autori suggeriscono che tale effetto sia implementato da proiezioni che arrivano dall'amigdala.

9.3. L'incidenza del contenuto emotivo

Una recente ricerca (Balconi, 2004), che ha confermato la specificità del marker N230 per la decodifica del contenuto semantico della mimica emotiva, ha avuto il merito di portare in luce sull'incidenza delle diverse *categorie emotive*. Infatti, i risultati mostrano che sia l'intensità sia la latenza della deflessione N230 variano in funzione dei due parametri dell'arousal e della valenza edonica. Nello specifico, si osserva un netto incremento in risposta alle emozioni negative ad alto arousal (rabbia, sorpresa e soprattutto paura), mentre in risposta alla gioia (emozione positiva) e alla tristezza (emozione a basso arousal) si rileva una chiara diminuzione dell'attivazione. Inoltre, per le emozioni ad alto arousal, il picco risulta essere anticipato, quasi a testimoniare la

necessità di una risposta repentina di fronte a stimoli emotigeni dotati di un valore funzionale pregnante (Balconi & Pozzoli, 2003).

Calder (Calder et al., 1996) riportano il caso di due pazienti di cui l'uno riusciva a riconoscere piuttosto bene le espressioni della gioia e della tristezza ma aveva notevoli difficoltà nell'identificazione del disgusto, della rabbia e della paura, mentre l'altro era in grado di riconoscere la gioia, la tristezza e il disgusto, manifestava qualche problema con l'emozione della rabbia ma risultava totalmente incapace di decodificare le espressioni di paura. Secondo Calder, il fatto che non tutte le emozioni siano compromesse in egual misura fa supporre l'esistenza di specifici sostrati neurali dedicati per i differenti correlati emotivi (Adolphs et al., 1995). Tale differenziazione sarebbe spiegabile in riferimento al fatto che le diverse emozioni hanno funzioni notevolmente differenziate da un punto di vista evolutivo. Tra queste in particolare la paura, segnalando la presenza di un pericolo per l'organismo, riveste un ruolo chiave ai fini della sopravvivenza.

9.3.1. Il contributo dell'elettromiografia facciale

Una netta differenziazione tra i diversi correlati emotivi emerge anche da quelle ricerche che hanno indagato l'espressione delle emozioni attraverso la mimica facciale. Particolarmente utile si sono rivelate a tale scopo le tecniche elettromiografiche. L'elettromiografia (EMG) facciale misura i pattern di attività dei muscoli facciali. Il vantaggio di questa tecnica sta nel fatto di poter mappare anche i movimenti facciali minimi, non rilevabili all'osservazione. Le ricerche indicano che le emozioni negative sono caratterizzate da una maggiore attività nella zona alta del volto, in particolare nella zona delle sopraciglia, mentre le emozioni positive presentano una maggiore attività dei muscoli della parte bassa del volto (in particolare lo zigomatico superiore, coinvolto nel movimento tipico del sorriso) e perioculare. Tanto maggiore è l'intensità dell'emozione sperimentata, e tanto maggiore è l'attività elettrica dei muscoli. Concludendo, è possibile affermare che, applicando la tecnica dell'EMG, ciò che emerge non è tanto una differenziazione tra le diverse specifiche emozioni, ma tra due costellazioni di emozioni, quelle con valenza edonica positiva e quelle con valenza edonica negativa (Cacioppo et al., 2000).

10. Il ruolo dell'amigdala

Numerose ricerche hanno messo in evidenza l'importanza dell'amigdala nella decodifica delle espressioni facciali delle emozioni, sottolineando il suo prezioso contributo nella sfera del comportamento sociale (Young et al., 1996). Diversi studi, condotti su pazienti che avevano subito un danno bilaterale dell'amigdala dimostrano che, sebbene questa struttura ricopra un ruolo focale soprattutto nella decodifica della paura, più in generale tuttavia essa è coinvolta nel riconoscimento delle diverse emozioni.

L'importanza dell'amigdala nel decoding dell'espressione facciale della paura è stata messa in evidenza da un recente studio (Adolphs et al., 1995) che è stato effettuato su una paziente con lesione all'amigdala, la quale era in grado di riconoscere tutte le espressioni facciali delle emozioni, tranne quella della paura. Questo dato è stato confermato dalle tecniche di neuroimaging, che hanno rilevato che la presentazione del volto della paura, confrontata con le altre emozioni, produce una maggiore attività dell'amigdala (Morris et al., 1999). Ciò avviene anche quando i soggetti non percepiscono consapevolmente il volto (Whalen et al., 1998): l'amigdala agisce, di fronte a stimoli legati alla paura, come un campanello d'allarme che precede la presa di coscienza del pericolo (LeDoux & Phelps, 2000).

10.1. La natura automatica del processamento dell'espressione emotiva

Recentemente, studi che hanno utilizzato tecniche di neuroimaging hanno dimostrato la presenza di un'attivazione dell'amigdala in risposta agli stimoli facciali (Morris et al., 1998). Sapendo che la risposta agli stimoli emotivi avviene in modo automatico e non richiede il contributo della consapevolezza, Whalen (Whalen et al., 1998) ha dimostrato che tale attivazione ha luogo anche in assenza di percezione esplicita degli stimoli. Grazie ad una procedura di tipo *backward masking*, stimoli facciali esprimenti gioia e paura sono stati presentati ai soggetti senza che essi ne fossero consapevoli. La risonanza magnetica funzionale ha rivelato che tali stimoli, se pur non

elaborati a livello della coscienza, tuttavia provocavano un significativo aumento dell'attività dell'amigdala, in particolar modo nel caso degli stimoli esprimenti paura.

Un recente studio fMRI condotto da Liddell (Liddell et al., 2005) indica che il processamento non cosciente di volti esprimenti paura utilizza una via che coinvolge il tronco dell'encefalo, l'amigdala e la corteccia prefrontale, bypassando la corteccia visiva, la cui attivazione viene riscontrata quando gli stimoli sono riconosciuti a livello cosciente. Tale via diretta funge da sistema di allarme che in modo tempestivo, prima dell'intervento dell'appraisal cosciente, attiva l'organismo e ne orienta l'attenzione verso un potenziale pericolo. In risposta a stimoli visivi esprimenti paura e percepiti consapevolmente, si attivano l'amigdala e la corteccia visiva primaria (Adolphs, 2002; LeDoux, 1998). L'amigdala ricopre un ruolo primario nelle risposte dell'organismo a stimoli connessi alla paura (Zald, 2003), orientandolo rapidamente verso una potenziale fonte di pericolo presente nell'ambiente. Questo meccanismo adattivo può avvenire dunque anche in assenza della consapevolezza. Due diversi ordini di dati suggeriscono che questo avvenga coinvolgendo l'amigdala ma escludendo il contributo della corteccia striata. In primo luogo, gli studi condotti su pazienti con blindsight dovuto a lesione della corteccia visiva indicano che, se pur incapaci di percepire consapevolmente stimoli esprimenti paura presentati nel campo cieco, tuttavia in risposta ad essi presentano l'attivazione dell'amigdala e di strutture del tronco encefalico, in particolare il collicolo superiore e il pulvinar (Morris et al., 2001). Questa via diretta subcorticale è ad esempio attivata nel processamento di volti esprimenti paura (de Gelder et al., 1999; Vuilleumier et al., 2002). In secondo luogo, anche gli studi di neuroimaging che impiegano stimoli subliminali, presentati per meno di 30 ms e immediatamente seguiti da uno stimolo neutro (tecnica di *backward masking*) suggeriscono l'attivazione di vie dirette subcorticali, che permettono un'elementare elaborazione percettiva dello stimolo, anche in assenza dell'intervento della corteccia visiva (Morris et al., 1999; Vuilleumier et al., 2003).

Partendo da questi presupposti, lo studio di Liddell ha dimostrato che, quando viene presentato uno stimolo che potenzialmente segnala un pericolo - nello specifico un volto esprime paura -, esso viene rapidamente elaborato senza ricorrere ad un

appraisal cosciente, funzionando come un sistema di allarme che attiva l'organismo. A tal proposito, Halgren & Marinkovic (1995) parlano di un "riflesso di orientamento", in cui la corteccia visiva viene bypassata. Si osserva l'attivazione di strutture del tronco dell'encefalo quali il collicolo superiore, il locus ceruleo e il pulvinar, che ricevono afferenze direttamente dalla retina (Vuilleumier et al., 2003) e che permettono un processamento di basso livello delle caratteristiche sensoriali dello stimolo (Morris et al., 2001). L'elaborazione coinvolge inoltre l'amigdala e, infine, la corteccia prefrontale. Si può ipotizzare che l'amigdala abbia la funzione di modulare la vigilanza prima dell'elaborazione cosciente di stimoli che richiedono una risposta tempestiva da parte dell'organismo. Essa quindi garantisce un monitoraggio costante dell'ambiente, allo scopo di rilevare rapidamente quegli stimoli che potrebbero segnalare pericolo (Davis & Whalen, 2001). Il pattern di attivazione che coinvolge amigdala e corteccia prefrontale osservato in questo studio è coerente con quelle ricerche ERP che hanno evidenziato la presenza del complesso N2-P3 (Liddell et al., 2004; Williams et al., 2004). I volti esprimenti paura elicitano infatti un complesso ERP detto N2-P3 che si manifesta tra i 200 e i 300 ms dopo la presentazione dello stimolo. Quando i volti sono presentati in modalità subliminale, essi elicitano un complesso N2-P3 maggiore rispetto ai volti neutri e ai volti esprimenti paura percepiti consapevolmente (Liddell et al., 2004). Il complesso N2-P3 è associato all'orientamento verso stimoli significativi dal punto di vista adattivo (ad esempio un volto che esprime paura) o stimoli nuovi. Inoltre, esso è legato all'elaborazione di stimoli discrepanti. È interessante osservare che la N2 è legata all'attività dell'amigdala (Krolak-Salmon et al., 2004) e la P3 presenta una localizzazione nelle aree prefrontali (Halgren & Marinkovic, 2005), che sappiamo essere legate al processamento di stimoli nuovi (Ranganath & Rainer, 2003) e all'elaborazione del significato di stimoli emotivi e permettono di riorientare l'attenzione in risposta all'esposizione a stimoli emotivi percepiti non consapevolmente (Killgore & Yungelun-Todd, 2004).

Un'ulteriore conferma della natura automatica del processamento delle espressioni emotive facciali giunge da quegli studi che hanno rilevato come la visione di volti esprimenti emozioni provochi un'involontaria interferenza in un compito di

valutazione lessicale. Ancora una volta, ciò suggerisce che il processamento semantico della mimica emotiva sia involontario e automatico (Stenberg et al., 1998).

LA COMUNICAZIONE DELLE EMOZIONI ATTRAVERSO IL CANALE VOCALE NON VERBALE

1. Introduzione

Il capitolo precedente illustra il ruolo fondamentale dell'espressione del viso nel permetterci di comunicare, recepire e condividere le emozioni, traducendo una sensazione che altrimenti rimarrebbe un fatto privato in un efficace mezzo di comunicazione di scambio interpersonale. Come anticipato tuttavia, le espressioni facciali non sono l'unico mezzo che abbiamo a disposizione. La voce è infatti uno strumento potente, che permette di esprimere e di cogliere l'infinita gamma delle sfumature emotive.

2. Gli indicatori vocali implicati

La voce comunica le emozioni soprattutto attraverso i parametri dell'intonazione, dell'intensità e del ritmo che, nella classificazione delle caratteristiche paralinguistiche, costituiscono il *sistema prosodico o soprasegmentale* (Anolli & Cicceri, 1997). Innanzitutto, ciò che noi percepiamo come *tonalità* è dato dalla frequenza fondamentale (F_0), che è determinata dalla tensione e dalla vibrazione delle corde vocali e dell'intera laringe. Più esse sono tese, più la frequenza è alta e più la tonalità è acuta; al contrario, viene percepito un tono grave quando esse sono allentate. La tonalità viene misurata in Hertz. Più nello specifico, le variabili relative al tono sono le seguenti:

- media di F0 (frequenza fondamentale): è la media aritmetica di F0 in un enunciato;
- gamma di F0: è la differenza tra il tono di voce più alto e quello più basso all'interno di un enunciato;
- variazione di F0: si ottiene contando il numero di variazioni (picchi e deflessioni) dell'intonazione all'interno dell'enunciato;
- profilo di intonazione: andamento dell'enunciato, ottenuto congiungendo tutti i punti indicativi della tonalità.

Rispetto ai *parametri temporali*, vengono distinti la durata, la velocità di emissione e il ritmo. La durata consiste nella lunghezza dell'enunciato misurata in secondi. La velocità di emissione è invece data dal numero di sillabe pronunciate in un secondo. Viene operata un'ulteriore distinzione tra la velocità dell'eloquio (velocità di emissione dell'enunciato comprendente suoni e pause) e velocità di articolazione (velocità di emissione dei soli periodi di suono, escluse le pause). Infine, il ritmo consiste nella combinazione tra durata e accento ed è dato dall'alternarsi di sillabe accentate e sillabe non accentate.

Da ultimo, ciò che noi soggettivamente percepiamo come volume è l'*intensità*, dovuta alla pressione della laringe e alla forza fonoespiatoria. L'intensità viene misurata in decibel (db). Perché un suono venga percepito è necessario che la sorgente sonora emetta vibrazioni di frequenza comprese tra i 20 e i 20.000 cicli al secondo. Al di sotto (infrasuoni) e al di sopra (ultrasuoni) di tali soglie il suono non viene percepito dall'orecchio umano. Rispetto all'intensità vengono considerati i seguenti parametri:

- media: è la media aritmetica dell'intensità in un enunciato;
- gamma: è la differenza tra l'intensità maggiore e quella minore all'interno di un enunciato;
- profilo di intensità: andamento dell'enunciato, ottenuto congiungendo tutti i punti indicativi dell'intensità.

3. Gli studi empirici sulla comunicazione vocale delle emozioni

La comunicazione vocale delle emozioni è stata esplorata focalizzando l'attenzione su due diversi processi:

- le ricerche sul processo di encoding indagano le modalità di codifica vocale delle espressioni emotive;
- gli studi relativi al processo di decoding cercano di capire come avvenga il riconoscimento delle emozioni a partire dalle caratteristiche vocali non verbali.

3.1. Problemi metodologici

Nel loro insieme, le ricerche si sono dovute confrontare con diversi problemi di ordine metodologico (Davidson, Scherer & Goldsmith, 2003). Una prima difficoltà riguarda la *scelta dei parametri* da utilizzare tra i diversi possibili (vedi par. 3.2.). In secondo luogo, i ricercatori devono stabilire se utilizzare *enunciati spontanei*, o *indotti sperimentale* oppure *simulati*. A tale proposito, da tempo sono stati messi in evidenza i vantaggi della simulazione, che permette una maggiore possibilità di controllo e consente di attribuire un carattere di enfasi e ridondanza dell'encoding, riducendo così l'ambiguità. Infatti, l'utilizzo di stimoli raccolti in una situazione in cui realmente le persone esprimono le emozioni attraverso la voce costituirebbe una condizione ideale. La spontaneità ha però come svantaggio la mancanza di controllo. Per questa ragione, anche se ci sono diverse ragioni per sospettare l'esistenza di differenze tra l'*emotional speech* spontaneo e quello recitato, solitamente viene chiesto a dei parlanti - meglio se si tratta di attori professionisti - di simulare le diverse emozioni. Maggiormente naturale, anche se di più difficile realizzazione, è la registrazione di materiale che deriva dall'induzione in laboratorio di una reale emozione nel parlante.

Un ulteriore problema metodologico riguarda la scelta delle *tecniche di mascheramento*, le quali permettono di mantenere inalterate le caratteristiche paralinguistiche eliminando però il contenuto semantico verbale. Esistono diversi tipi di mascheramento:

- utilizzo di suoni, sillabe senza senso o liste di lettere o numeri;
- filtraggio elettronico, che consente di alterare l'informazione verbale, rendendola incomprensibile. Ad esempio, il filtraggio della frequenza fondamentale consente di limitarne la gamma, attraverso filtri passa-basso, che permettono il passaggio delle sole frequenze basse, o attraverso filtri passa-alto, che permettono al contrario il passaggio delle sole frequenze alte. In particolare i filtri passa-basso, eliminando le frequenze alte, che sono indispensabili per il riconoscimento del parlato, rendono inintelligibile il contenuto verbale. Tuttavia, bisogna osservare se l'ascolto viene ripetutamente reiterato, l'ascoltatore può giungere alla comprensione del contenuto verbale;
- capovolgimento o taglio e rimontaggio del nastro;
- utilizzo di una lingua straniera;
- utilizzo di frasi neutre, cioè prive di un contenuto verbale emotivamente connotato. Quest'ultima tecnica di mascheramento presenta evidenti vantaggi, in quanto preserva un profilo paralinguistico che va oltre la singola sillaba o parola, mantenendo inoltre invariati i parametri soprasegmentali.

Recentemente, Douglas-Cowie e colleghi (Douglas-Cowie *et al.*, 2003) hanno passato in rassegna i database esistenti relativi all'*emotional speech* (emozioni espresse attraverso il canale vocale) delineando alcune caratteristiche che dovrebbero essere tenute presenti nella costruzione di un nuovo database. Come essi sottolineano, bisogna innanzitutto tenere in considerazione il fatto che il modo di esprimere vocalmente un'emozione varia considerevolmente da parlante a parlante, da cultura a cultura, tra i generi e tra le condizioni socio-ambientali. Di conseguenza, per garantire che i risultati di una ricerca volta ad indagare l'*emotional speech* siano generalizzabili è necessario, nella costruzione della batteria degli stimoli uditivi, considerare una serie di fattori. In primo luogo, bisogna valutare attentamente il

numero e il tipo di parlanti coinvolto nella costruzione del campione. Gli autori notano che la maggior parte degli studi ha utilizzato un unico soggetto. Per ottenere risultati altamente generalizzabili bisognerebbe coinvolgere diversi parlanti, diversificati sulla base della lingua e del genere. E' auspicabile che i parlanti siano attori professionisti, poiché essi presentano una maggiore competenza nell'utilizzare i parametri vocali in modo intenzionale. Sempre allo scopo di permettere la generalizzazione dei risultati, è bene utilizzare un'ampia gamma di emozioni. La rassegna effettuata indica tuttavia che di norma viene impiegato un numero ridotto di emozioni, per lo più rabbia, tristezza, paura, stress e gioia. Un altro dato che emerge dalla rassegna è che solitamente il materiale è avulso da qualsiasi contesto: si tratta infatti per lo più di registrazioni audio di brevi enunciazioni con contenuto semantico neutro e struttura sintattica prestabilita. A tale proposito bisogna fare alcune considerazioni: nella situazione naturale le persone solitamente utilizzano diversi *indizi contestuali* per integrare le informazioni convogliate dal vocale al fine dell'identificazione dell'emozione. Possiamo indicare almeno quattro diversi tipi di contesto, che dovrebbero essere tenuti in considerazione: un contesto semantico, relativo al contenuto verbale; un contesto strutturale, costituito dalla struttura sintattica della frase; un contesto intermodale, poiché le informazioni acustiche sono integrate da quelle offerte dalle altre modalità sensoriali, principalmente quella visiva; un contesto temporale, dal momento che l'emozione non è espressa in modo puntuale e può essere definita come un processo. Infine, è utile osservare che i parametri più frequentemente utilizzati per analizzare il materiale sono l'intensità, l'intonazione, la struttura formale dello spettro e i parametri temporali (Banse & Scherer, 1996).

3.2. Le ricerche sull'encoding

Le ricerche sulla codifica vocale delle emozioni hanno l'obiettivo di stabilire quali cambiamenti nella voce e nella produzione del parlato - e quali corrispondenti patterns di parametri acustici - si manifestano come conseguenza dell'insorgere di

un'emozione nel parlante. Più nello specifico, tali studi intendono far luce sull'esistenza di differenze nell'espressione vocale delle diverse emozioni. E' possibile operare una rudimentale distinzione tra le differenti emozioni sulla base dei parametri acustici. Tuttavia, il legame tra le caratteristiche prosodiche e la specifica emozione è abbastanza approssimativo, dal momento che esiste un largo margine di variabilità, sia nell'espressione e nel riconoscimento delle emozioni, sia a livello di variabilità intra-individuale e inter-individuale. Aubergé e Cathiard (2003) hanno ad esempio rilevato che le differenze individuali giocano un ruolo non secondario nell'espressione e nel riconoscimento del divertimento, confermando l'intuizione dell'esistenza di diverse strategie per esprimere vocalmente la stessa emozione.

Bisogna inoltre tenere in massima considerazione il fatto che, come precedentemente accennato (vedi par. 3.1), il significato delle caratteristiche acustiche è contesto-dipendente (Cook, 2002).

Come tendenza generale, gli studi che hanno indagato l'encoding delle emozioni a livello vocale non verbale hanno evidenziato l'esistenza di due pattern: l'uno, caratterizzato da alta frequenza fondamentale, ampia gamma e variabilità della frequenza fondamentale, elevata intensità e alta velocità, è tipico delle emozioni ad alto livello di arousal (livello di attivazione fisiologica); l'altro, identificato da bassa frequenza fondamentale, gamma e variabilità di intonazione ristrette, debole intensità e velocità ridotta, è invece tipico delle emozioni a basso arousal (Banse & Scherer, 1996). In effetti, i parametri vocali sono un potente indicatore del livello di attivazione fisiologica, come suggeriscono le cosiddette "teorie dell'arousal" (Pittam & Scherer, 1993). Tuttavia, esistono diverse evidenze, messe in evidenza dalle "teorie dell'appraisal", del fatto che in realtà l'espressione vocale delle emozioni è determinata dall'azione congiunta dell'arousal e della *valenza edonica* : in effetti, le modalità specifiche di encoding di alcune emozioni non possono essere spiegate facendo riferimento esclusivamente al livello di attivazione (Klasmeyer & Sendlmeier, 1997). Di conseguenza, un modello bidimensionale, se non *multidimensionale* dell'encoding emotivo appare oggi più plausibile di un modello unidimensionale. Un simile modello è inoltre supportato da studi recenti che hanno

nesso in evidenza l'esistenza di differenze qualitative nell'espressione delle diverse emozioni (Juslin & Laukka, 2001; Scherer, Johnstone & Klasmeyer, 2003). Nello specifico, Anolli e Ciceri (1997) hanno delineato la "voce" di alcune emozioni. Ad esempio, la gioia è caratterizzata da una media della frequenza fondamentale molto elevata (media di F_0 pari a 228 Hz) e quindi da una voce acuta; elevate sono anche la gamma e la variazione dell'intonazione, con un profilo di intonazione variato e mosso, ma con variazioni sempre gradualmente e lente, e quindi arrotondate. Il profilo discende gradualmente nella parte finale dell'enunciato per poi assestarsi su una tonalità intermedia o per risalire lievemente. Anche l'intensità è elevata e ha un profilo ascendente nella parte conclusiva dell'enunciato. La velocità di eloquio e di articolazione è media, la durata è standard e le pause presentano valori medi. La *tristezza* ha un profilo di intonazione caratterizzato da una frequenza fondamentale bassa (media di F_0 pari a 134 Hz), con variazioni quasi inesistenti e arrotondate e gamma ristretta. Il finale è discendente. Anche il volume e la velocità sono bassi. La tristezza si connota soprattutto per le sue lunghe pause. Nel caso della *paura*, il tono è molto acuto: l'energia si concentra nelle zone di alta frequenza dello spettro (F_0 media pari a 278 Hz), con ampia gamma e numerose variazioni piuttosto brusche, che conferiscono al profilo un andamento spigoloso. L'intensità è elevata e progressivamente crescente: è possibile che l'eloquio giunga addirittura ad essere gridato. Pur avendo una durata della frase standard e una velocità dell'eloquio media, la paura presenta una velocità di articolazione alta. Le pause costituiscono frequenti e rapide "rotture" del parlato, dovute alla respirazione superficiale. La *rabbia* è caratterizzata da una frequenza fondamentale media (F_0 media pari a 178 Hz, inferiore a quella di gioia e paura), con numerose e a volte repentine, spigolose e ampie variazioni del tono. Il profilo ha carattere ascendente. I livelli di intensità sono costantemente alti. Il ritmo è serrato, con una durata della frase contenuta. Si manifesta la tendenza ad "espellere" la frase, in un'unica emissione di respiro e senza pause. Il *disprezzo* si caratterizza per l'articolazione molto lenta delle sillabe e per la durata molto lunga dell'enunciato. Il tono è grave, con una gamma ristretta e variazioni minime e di forma arrotondata. Il profilo presenta un finale discendente.

L'intensità è bassa e costante. Infine, la *tenerezza* presenta un ritmo regolare e costante. Tratto distintivo di questa emozione è la tonalità molto grave, con scarse variazioni di frequenza fondamentale e gamma ristretta. Il volume è basso e decrescente.

Un interessante contributo giunge da uno studio (Gobl & Chasaide, 2003) che ha indagato il ruolo delle *qualità vocali* - ad esempio la voce piena, la voce sottile, la voce ruvida - nel veicolare informazioni circa emozioni, umori e attitudini. A differenza dei parametri di picco, di volume e temporali, che sono relativamente più facili da misurare e da controllare, le qualità vocali sono state finora poco esplorate in relazione alla comunicazione dei correlati emotivi, proprio a causa delle difficoltà metodologiche e concettuali che esse implicano (Scherer, 1986). I risultati indicano che effettivamente le qualità vocali contribuiscono alla comunicazione vocale delle emozioni, non tanto in forza di una corrispondenza biunivoca tra qualità della voce e tipo di emozione espressa ma piuttosto in virtù di un'associazione tra una certa qualità vocale e una costellazione di stati affettivi. Inoltre, la qualità vocale risulta più efficace nel comunicare la valenza edonica rispetto che l'arousal, che viene invece meglio veicolato da intensità, tonalità e durata. Gli autori ipotizzano che la frequenza e la qualità vocale contribuiscano in modo indipendente all'espressione degli stati affettivi: le qualità vocali hanno un ruolo di primo piano nel comunicare gli stati affettivi meno intensi come stati d'animo, umori e attitudini, soprattutto se privi di un correlato fisiologico, mentre la frequenza è fondamentale per l'espressione delle emozioni più intense, accompagnate da una componente a livello fisiologico. La possibilità che la frequenza e le qualità vocali abbiano funzioni differenti e potenzialmente indipendenti nell'espressione degli affetti è stata precedentemente suggerita anche da Scherer (1986), il quale ha ipotizzato che le ampie variazioni dell'intonazione segnalino marcate differenze nei livelli di attivazione, e che invece le qualità vocali comunichino sottili sfumature affettive.

3.2.1. Il modello di Scherer

Scherer (1986) individua una relazione tra i *parametri vocali* delle singole emozioni e le componenti di *valutazione cognitiva* (*appraisal*) delle circostanze attivanti contemplate dal *modello componenziale* da lui stesso proposto. Tale modello individua cinque livelli di valutazione dello stimolo:

- *novità*: il carattere di novità di uno stimolo inatteso provoca una risposta di orientamento, dovuta alla rilevazione di una discrepanza tra le aspettative e lo stimolo. Stimoli poco familiari suscitano reazioni di sorpresa, di curiosità e di avvicinamento, ma anche di paura e diffidenza. Mentre le emozioni positive sono suscitate dalla coincidenza tra le aspettative e i desideri dell'individuo e lo stimolo emotigeno, mentre quelle negative sono il risultato di una discrepanza tra di essi;
- *valenza edonica*: consiste nel carattere di piacevolezza/spiacevolezza dello stimolo. Nel caso in cui esso sia valutato come piacevole, vengono attuate delle risposte di avvicinamento, inclusione e gioia, in caso contrario invece si riscontrano condotte di esitamento, fuga e attacco;
- *rilevanza rispetto ai bisogni e agli scopi del soggetto*: il fatto che uno stimolo costituisca la risposta ad uno scopo, desiderio o bisogno del soggetto, contribuisce ad attribuirgli rilevanza e a suscitare una reazione di avvicinamento, ed in particolar modo di gioia. In caso contrario, la reazione abituale è la collera;
- *capacità di far fronte (coping) allo stimolo*: l'individuo attua due test: tramite il test di controllo, che ha come oggetto le cause elicитanti, e tramite il test di potenza, che ha come oggetto le risorse personali, egli verifica se è possibile controllare o annullare l'evento che suscita la risposta emozionale. La consapevolezza dell'efficacia delle proprie risorse suscita un senso di padronanza, mentre la consapevolezza della propria impotenza genera depressione e ansia;
- *compatibilità con l'immagine di sé e le norme sociali*: l'adesione o la trasgressione delle regole sociali e dei valori personali suscitano differenti reazioni emozionali. A tale proposito, particolare rilievo va attribuito alle cosiddette emozioni sociali

o secondarie, come la vergogna, l'imbarazzo, la colpa e l'orgoglio, che emergono nel processo di socializzazione. Tali emozioni costituiscono un feed-back della propria condotta, ricoprendo così un ruolo di monitoraggio.

Scherer suggerisce che le componenti di valutazione citate producano specifici effetti a livello del sistema nervoso centrale, i quali si traducono a loro volta in modificazioni della tensione della muscolatura dell'apparato fonatorio. In particolare, secondo Scherer, gli *effetti tonici* -più durevoli- non sarebbero controllati volontariamente, a differenza degli *effetti fasici* di rilassamento - a carattere transitorio -, che sarebbero invece il prodotto dei tentativi dell'organismo di controllare l'espressione vocale emotiva. Anche i cambiamenti che avvengono nel sistema nervoso autonomo, quali la modificazione del ritmo respiratorio e della secrezione di saliva, hanno degli effetti sull'emissione vocale.

Scherer infine distingue tra voce piena e voce sottile. La *voce piena*, caratterizzata da una modesta tensione e una respirazione piena, si manifesta quando l'individuo ha fiducia nelle proprie risorse e si prepara all'attacco. La *voce sottile* invece, che risulta da una respirazione poco profonda e una aumento della frequenza fondamentale, contraddistingue quelle situazioni in cui l'individuo, valutando la propria impotenza a contrattaccare, si prepara alla fuga.

3.3. Le ricerche sul decoding

Le ricerche sulla decodifica vocale delle emozioni, che hanno avuto inizio tra gli anni '50 e gli anni '80, sono volte a comprendere quali variazioni dei parametri acustici ci permettano di cogliere la presenza delle emozioni nella voce dell'altro e di discriminarle. In genere, a questo scopo, una serie di stimoli vocali emotivi prodotta da attori professionisti viene sottoposta ad un campione di giudici, cui viene proposto un compito di discriminazione (scelta di una tra le diverse alternative date) o di riconoscimento (*free labelling*, cioè libero attribuzione di un'etichetta emotiva) (Scherer, Johnstone & Klasmeyer, 2003). Alcune ricerche hanno inoltre manipolato i parametri paralinguistici con lo scopo di evidenziare quali caratteristiche abbiano un

maggior peso ai fini del riconoscimento. I principali risultati possono essere così riassunti: le diverse caratteristiche acustiche hanno una diversa incidenza rispetto alle diverse emozioni; in generale, tuttavia, si può affermare che la *frequenza fondamentale* e l'*intensità*, nonché le loro variazioni (principalmente media e gamma), così come anche il *ritmo di articolazione*, comunicano le informazioni essenziali per il riconoscimento delle emozioni.

I dati sperimentali indicano che il riconoscimento delle emozioni sulla base dei segnali vocali non verbali avviene in percentuali cinque o sei volte superiori al caso: è stata infatti evidenziata un'accuratezza media del 60%. Confrontando questo dato con quello relativo all'accuratezza riconoscimento delle emozioni basato sull'espressione facciale, emerge una maggiore difficoltà nel decoding vocale. La differenza è dovuta in particolar modo al riconoscimento della gioia e del disgusto: queste emozioni sono scarsamente discriminate in base ai parametri vocali, mentre invece sulla base della mimica ottengono valori vicini al 100% (Ekman, 1994). La tristezza e la rabbia - ma anche la paura - sono invece meglio riconosciute a livello vocale. Johnstone e Scherer (2000) hanno attribuito a queste differenze un'origine di tipo evolutivo: è utile che un segnale di rabbia o di paura possa essere comunicato anche a grande distanza. Al contrario, un'impressione di disgusto per aver assaggiato del cibo che non è stato apprezzato viene più facilmente espressa ad un cospecifico che si trova nelle vicinanze.

In generale, dalle ricerche emerge che a livello vocale le emozioni negative sono riconosciute più facilmente e con un'accuratezza maggiore rispetto alle emozioni positive. Questo deriva probabilmente dal valore adattivo che il loro riconoscimento ricopre ai fini della salvaguardia dell'individuo (Balconi & Lucchiari, 2005; Balconi & Pozzoli, 2003). Le emozioni più riconosciute sono la rabbia e il dolore, mentre quelle identificate con maggiore difficoltà sono la gioia, il disgusto e il disprezzo. Le emozioni non vengono confuse casualmente, ma in modo sistematico. Ad esempio, tendenzialmente la tristezza viene confusa con l'amore, la vergogna con la tristezza, la paura con la sorpresa. Inoltre, il disprezzo da un lato e il disgusto e la collera dall'altro vengono frequentemente scambiate; lo stesso accade nel caso di gioia e

sorpresa. Il fatto che la sorpresa venga spesso confusa viene talvolta spiegato facendo riferimento alla presenza di un'iniziale elemento di novità/sorpresa dello stimolo, che in un secondo momento della valutazione cognitiva assume una più chiara connotazione di piacevolezza/piacevolezza. La sorpresa è infatti un'emozione ambivalente, che può assumere una valenza positiva o negativa a seconda della specifica natura dello stimolo elicitante. Sono state avanzate due ipotesi, entrambe plausibili e complementari, per spiegare questi errori sistematici. La prima ipotesi fa riferimento alla condivisione del medesimo livello di arousal, la seconda invece, più competente, alla vicinanza semantica (Russell & Fernández-Dols, 1998) tra le emozioni che vengono confuse. Recentemente, una ricerca condotta da Scherer (1994) ha contribuito a spiegare il motivo per cui alcune le caratteristiche soprasegmentali sono un mezzo efficace per esprimere alcune emozioni ma non altre. Scherer introduce il concetto di *affective bursts*, descrivendoli come espressioni non verbali molto brevi e discrete di un'emozione, che si manifestano a livello del volto e della voce e che sono elicitate da eventi chiaramente identificabili. Se si considera l'aspetto funzionale, il concetto di *affective burst* in parte si sovrappone a quello di "interiezione affettiva". Tuttavia i confini dei due costrutti non coincidono: un'esclamazione verbale come "Cielo!", proprio in virtù della sua natura verbale viene classificato come interiezione, a differenza di una risata, che viene invece identificata come *affect burst*.

Scherer offre una classificazione degli *affective bursts* basata sulla sua distinzione push-pull (Scherer, 1988). Gli *effetti push* sono fattori fisiologici, come il dolore, che si traducono in un'espressione. Gli *effetti pull* invece sono regole e aspettative sociali che prevedono espressioni socialmente appropriate per le diverse situazioni, andando così a modulare le espressioni elicitate dagli *effetti push*. Entrambi i fattori quasi sempre contribuiscono alla produzione dell'espressione emozionale, ma uno di essi può prevalere sull'altro in date situazioni. Sulla base di tale distinzione, Scherer propone una differenziazione tra i *raw affect bursts* (*affect bursts* grezzi), che si collocano al polo push del continuum e gli *affect emblems* (emblem affettivi), che si collocano invece all'estremità pull del continuum. Di

conseguenza, i raw effect bursts possono essere definiti come vocalizzazioni grezze tendenzialmente universali e minimamente convenzionalizzate, che presentano forti differenze interindividuali. Al contrario gli affect emblems sono simboli convenzionalizzati, strettamente cultura-dipendenti, che presentano scarse differenze individuali. Ovviamente, tra i due estremi del continuum esistono infinite possibilità di bursts misti, che vanno da suoni non articolati (risata) fino a interazioni con una struttura fonemica se pur non verbali (“wow”) (vedi fig. 1).

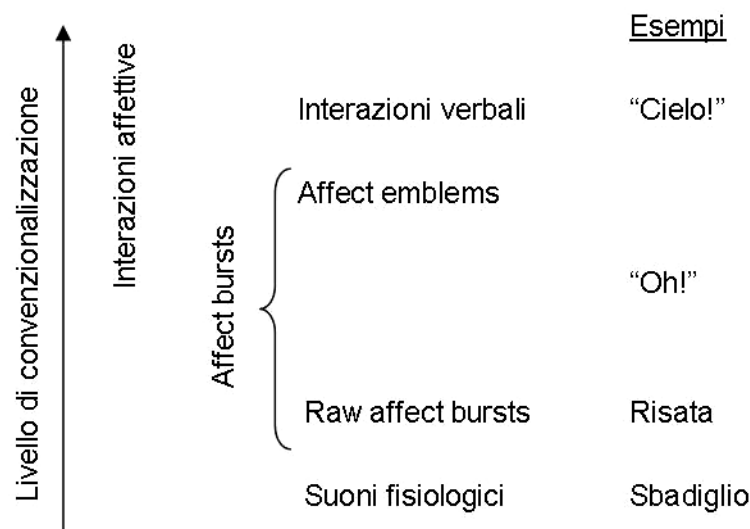


Fig 1. Le interiezioni affettive verbali e gli affective bursts come estremi di un continuum delle espressioni vocali emotivamente connotate (tratto da Schröder, 2003).

In un recente articolo, Schröder (2003) ha dimostrato che gli affective bursts sono in grado, se presentati nella sola modalità uditiva e in assenza di un contesto, di comunicare un significato emotivo chiaramente identificabile. Tutte e 10 le emozioni espresse (ammirazione, paura, disgusto, contentezza, noia, calma, sorpresa, preoccupazione, disprezzo, rabbia) presentano alti indici di riconoscimento (in media 81%). Questo indica che gli affective bursts sono degli efficaci strumenti per l'espressione delle emozioni. Le percentuali di riconoscimento sono notevolmente più alte di quelle rilevate nel caso di decoding basato sulle caratteristiche prosodiche

o sulle qualità vocali del parlato. Ciò è dovuto al fatto che, nel caso degli affective bursts, anche la struttura segmentale stessa, e non solo gli aspetti sovrasegmentali, comunicano un significato. Più nello specifico, si può ipotizzare che alcune emozioni siano più efficacemente espresse dagli affect bursts e altre invece dalla prosodia e dalla qualità della voce. Il disgusto è un chiaro esempio di emozione che viene tipicamente espressa attraverso affect bursts ("bleah"). Banse e Scherer (1996) hanno trovato che gli indici di riconoscimento del disgusto a partire dalla prosodia e dalla qualità vocale del parlato si attestano solo intorno al 15%. Questa emozione infatti viene tipicamente espressa tramite affective bursts anziché attraverso lunghi e articolati enunciati, come supportato dai risultati di questo studio. Al contrario, sembra che altre emozioni, ad esempio la rabbia, che in questo studio ha riportato livelli minori di riconoscimento, siano scarsamente espresse ricorrendo agli affective bursts. In effetti, Johnstone (Johnstone & Scherer, 2000) ha suggerito che alcune emozioni, tra cui la rabbia, vengano espresse tramite una configurazione prototipica di variabili acustiche legate al parlato, e che di conseguenza vengano facilmente prodotte dagli attori e facilmente riconosciute dagli ascoltatori, mentre altre emozioni, come il disgusto, che non presentano questa configurazione tipica, vengano molto più difficilmente prodotte e comprese.

Infine, è emerso che gli affect bursts in generale non sono specifici di una singola emozione ma sono adatti ad esprimere una certa gamma di emozioni, simili se considerate in riferimento alle dimensioni di valenza, arousal e controllabilità.

4. Il contributo della neuropsicologia

La neuropsicologia, ed in particolar modo gli studi sui deficit comunicativi risultanti da lesioni cerebrali, hanno contribuito a chiarire come avviene il processamento delle informazioni emotive espresse dalla voce.

Rispetto all'encoding, nel complesso, i dati disponibili indicano che alcuni specifici nuclei del tronco encefalico, attivati da proiezioni provenienti dal sistema

limbico, coordinano l'attivazione dei muscoli laringei e respiratori, producendo in questo modo specifici pattern vocali (Scherer et al., 2003).

Maggiore attenzione è stata rivolta ai meccanismi cerebrali sottesi al decoding delle emozioni espresse attraverso i parametri vocali. Se da un lato è universalmente riconosciuto che le componenti linguistiche e grammaticali del parlato vengono elaborate in specifici centri dell'emisfero sinistro, d'altro canto non esiste invece un consenso circa la localizzazione dei centri preposti alla decodifica delle componenti emotive non linguistiche del parlato, sempre ammesso che questi esistano. A partire dagli anni '70, diversi studi hanno portato ad ipotizzare una specializzazione dell'emisfero destro per il processamento delle componenti emotive vocali (Tucker, Watson & Heilman, 1977), confermata dall'osservazione di un deficit nella percezione della prosodia emotiva in pazienti con danno cerebrale unilaterale all'emisfero destro ma non in pazienti con danno unilaterale all'emisfero sinistro (Peper & Irle, 1977). Tuttavia, non tutte le ricerche effettuate hanno riscontrato tale lateralizzazione destra. Van Lancker e Sidtis (1992) hanno suggerito che l'emisfero destro sia deputato non tanto all'elaborazione della prosodia emotiva ma più in generale al processamento di alcuni tipi di informazioni acustiche come il livello e la variabilità della frequenza fondamentale, che sono indici fondamentali per il riconoscimento delle emozioni espresse mediante la voce.

Un importante contributo ai fini della localizzazione di questi processi può giungere dalle tecniche di neuroimaging. Tuttavia, gli studi di questo tipo sono poco numerosi. Un interessante studio condotto con la tomografia ad emissione di positroni (PET) ha riportato una notevole attività prefrontale destra durante il processamento del contenuto proposizionale emotivo di alcune frasi (George et al., 1996). Imaizumi (Imaizumi et al., 1998) ha rilevato, utilizzando la magnetoencefalografia (MEG), sia rispetto al processamento linguistico sia rispetto a quello prosodico, l'esistenza di processi specifici di elaborazione delle differenti componenti prosodiche localizzati nell'emisfero destro ma che si avvalgono anche del contributo di alcuni centri dell'emisfero sinistro.

Un apporto giunge anche dagli studi elettroencefalografici. Tra di essi, una ricerca condotta registrando l'elettroencefalogramma ha riportato una specializzazione della corteccia destra nella percezione dei parametri temporali (durata delle sillabe accentate) e intonazionali (gamma di F0). Un recente studio (Alter *et al.*, 2003) ha inoltre messo in evidenza l'esistenza di differenze legate alle caratteristiche della specifica emozione. Alter ha infatti messo in relazione alcuni parametri legati all'encoding del parlato emotivamente connotato (ad esempio il tipo di respirazione o l'asprezza) con le caratteristiche del tracciato ERP registrato su un soggetto mentre ascoltava. Gli autori hanno individuato, in entrambi i casi, una differenziazione basata sulla valenza edonica, positiva, negativa o neutra.

Infine, bisogna osservare una carenza di studi che abbiano utilizzato la tecnica della registrazione dei potenziali evocati, che è invece stata efficacemente utilizzata nello studio del decoding emotivo del volto.

LA CONVERGENZA CROSS-MODALE AUDIO-VISIVA E LA SPECIFICITA' DEI PATTERN EMOTIVI

1. Introduzione

Come anticipato nei precedenti capitoli, la comunicazione degli eventi emotivi è un fenomeno complesso e multicomponentiale, che si esplica attraverso l'azione congiunta di una costellazione di segnali di diversa natura. All'espressione, così come al riconoscimento di un'emozione partecipano infatti numerose informazioni interconnesse: le parole, il tono di voce con cui vengono pronunciate, l'espressione del viso, la postura, un gesto, una risata, un sussulto, che di per sé possono essere ambigui, una volta combinati tra loro in un preciso pattern espressivo, acquistano un notevole valore comunicativo. Tale combinazione viene indicata come *integrazione* o *convergenza cross-modale*. L'integrazione è un fenomeno che si realizza nel momento in cui informazioni appartenenti a diverse modalità sensoriali vengono identificate e processate come appartenenti ad un singolo evento o ad una singola sorgente fisica. Generalmente, la percezione di questa unitarietà ha origine dalla vicinanza spaziale e dalla sincronia con cui le informazioni si presentano. Di seguito, viene inizialmente presentata una panoramica dell'attuale stato della ricerca sui processi integrativi, mentre nella seconda parte del capitolo il focus della trattazione verrà ristretto alle indagini che si sono occupate di comprendere i processi sottesi alla convergenza di informazioni cariche di valenza emotiva.

2. Il contributo degli studi comportamentali

2.1. L'integrazione multimodale come risposta alla complessità ambientale

Oggi si ritiene che i sistemi visivo, uditivo e somatosensoriale abbiano un'organizzazione di tipo gerarchico, tale per cui la stimolazione sensoriale produce una rappresentazione percettiva che passa attraverso una serie di stadi di processamento di complessità crescente. Tale organizzazione gerarchica sembra d'altro canto costituire un principio generale del funzionamento cerebrale.

Per completare il quadro, occorre inoltre soffermare l'attenzione sul fatto che normalmente i nostri organi di senso sono bersagliati contemporaneamente da una molteplicità di stimoli appartenenti a diverse modalità sensoriali. Ed infatti, corrispondentemente, numerosi recenti studi suggeriscono che, fin dalle prime fasi di elaborazione delle informazioni, si verificano, sia a livello delle *aree di convergenza* come la corteccia parietale, temporale (Schroeder & Foxe, 2002) e frontale (Graziano et al., 1997) sia a livello delle *aree specifiche* di elaborazione sensoriale (ad esempio l'area visiva V1 o l'area uditiva A1) fenomeni di convergenza cross-modale.

Ma qual è la funzione di tale processo di integrazione multisensoriale?

L'abilità nell'integrare stimoli ridondanti appartenenti a differenti modalità per formare un unico percolato costituisce una fondamentale componente alla base del comportamento e della cognizione guidati dai sensi. Essa ha una *funzione adattiva*, in quanto è finalizzata a migliorare l'elaborazione dello stimolo, in modo tale da produrre una risposta comportamentale più rapida e accurata. Ad esempio, le ricerche indicano che essa produce una migliore performance nei compiti di orientamento dell'attenzione e di riconoscimento (Schroeder et al., 2003).

2.2. I principi che regolano la percezione multimodale

L'ipotesi che l'integrazione multisensoriale abbia una funzione adattiva di facilitazione della risposta comportamentale è supportata dai risultati ottenuti da

quegli studi che hanno applicato il paradigma classicamente utilizzato per indagare la convergenza multisensoriale, il quale prevede un confronto tra le risposte agli stimoli unimodali con le risposte agli stimoli multimodali che derivano dalla loro combinazione. La letteratura indica che le risposte comportamentali a stimoli multimodali, se confrontate con quelle a stimoli di natura unimodale, sono più rapide in termini di tempi di risposta (TR) e più accurate (Welsch & Warren, 1986). Questo fenomeno viene detto *multisensory enhancement* (miglioramento multisensoriale). Ci si riferisce ad esso anche come *redundant target effect* (effetto di ridondanza del target), poiché l'effetto di miglioramento deriva proprio dal carattere di ridondanza, ripetitività e reciproca conferma degli stimoli. Tale processo, inoltre, è influenzato da alcuni fattori legati alle caratteristiche degli stimoli e alla modalità di presentazione degli stessi. L'azione di tali fattori è descritta da quelli che vengono considerati i tre principi che regolano l'integrazione. Secondo il *principio della vicinanza temporale*, la relazione temporale tra gli stimoli è un fattore critico ai fini della convergenza: gli stimoli separati da un intervallo temporale inferiore ai 100 ms sono quelli che hanno maggiore probabilità di elicitare un miglioramento della risposta. Oltre a ciò, secondo il *principio della vicinanza spaziale*, anche la prossimità spaziale tra gli stimoli ha un'importanza cruciale: quanto più gli stimoli sono ravvicinati e tanto maggiore sarà la possibilità che si verifichi il fenomeno di *multisensory enhancement*. Infine, secondo il principio dell'*inverse effectiveness effect* (effetto dell'efficacia inversa), la forza e l'efficacia dello stimolo unimodale sono inversamente correlate con il miglioramento della risposta multimodale. In altri termini, gli stimoli che di per sé sono poco efficaci producono i massimi livelli di *multisensory enhancement*, mentre gli stimoli che hanno una massima efficacia se presentati singolarmente producono uno scarso miglioramento della risposta se abbinati ad un altro stimolo.

2.3. La natura dell'integrazione: una questione aperta

Nonostante il fatto che l'integrazione multisensoriale sia una caratteristica fondamentale della percezione, tuttavia le nostre conoscenze circa il modo in cui un oggetto caratterizzato da componenti che fanno riferimento a differenti modalità sensoriali venga percepito come un oggetto unitario sono ancora incomplete e molte questioni rimangono aperte. In particolare, il tema di discussione maggiormente controverso riguarda la *natura dei processi implicati nell'integrazione*. Marks, Ben-Artzi e Lakatos (2003) offrono una panoramica degli studi che hanno cercato di dare una risposta a questo quesito, partendo dall'esame delle cosiddette corrispondenze cross-modali. Diversi studi suggeriscono che esistano, nella percezione, delle corrispondenze sensoriali. Tra di esse, quella maggiormente indagata è la sinestesia: nei fenomeni di sinestesia, la stimolazione di una certa modalità sensoriale comporta il coinvolgimento di una seconda modalità. Nella sinestesia audio-visiva gli stimoli acustici evocano risposte sia uditive sia visive, per cui ad esempio la persona può "vedere" colori e forme nei suoni. La sinestesia rivela evidenti corrispondenze cross-modali: ad esempio la brillantezza delle immagini visive aumenta all'aumentare della tonalità dello stimolo acustico. Anche i compiti di confronto e di giudizio sulla similarità cross-modale (Marks, 1989) rivelano le medesime corrispondenze: già in tenera età i bambini individuano delle corrispondenze del tipo brillantezza-intensità o luminosità-tonalità (Marks, 1978). Vi sarebbe poi una corrispondenza diretta tra congruenza e corrispondenza: la corrispondenza si instaura ad esempio tra suoni acuti e immagini molto luminose ma non tra suoni acuti e immagini poco luminose. Queste interazioni legate alla congruenza solitamente hanno natura bidirezionale.

L'interazione cross-modale legata alle corrispondenze cross-modali è stata rilevata anche nel caso dei compiti di discriminazione. I risultati mostrano che, quando si chiede di discriminare tra una luce debole e una brillante che compaiono contemporaneamente ad un suono acuto o grave, i tempi di reazione sono più rapidi

e la performance è più accurata quando vi è congruenza tra i due stimoli (luce brillante e suono acuto; luce debole e suono basso) (Marks, 1987).

In alcuni casi la corrispondenza, e di conseguenza l'integrazione, sembrano derivare da associazioni apprese tra gli stimoli, che potrebbero quindi avere una funzione di facilitazione nell'identificazione degli stimoli stessi. Un esempio di ciò è dato dalla corrispondenza tra colore caldo e temperatura elevata e tra colore freddo e bassa temperatura. L'ipotesi che tali associazioni siano apprese è suggerita dal fatto che esse non sono presenti nei bambini piccoli (Marks, 1987).

Tuttavia, in altri casi, all'origine di una corrispondenza (ad esempio quella tra tonalità del suono e brillantezza o tra intensità del suono e brillantezza) non può essere identificata un'influenza ambientale: queste associazioni infatti sono presenti già nella prima infanzia. In questo caso diventa più difficile individuare la loro natura e la loro funzione nella percezione. L'analisi della letteratura indica l'esistenza di tre diverse possibilità. La prima è che la corrispondenza cross-modale derivi dal fatto che, nei primi stadi di *processamento sensoriale*, si verifichi una qualche forma di dialogo tra l'elaborazione degli stimoli visivi e quella degli stimoli uditivi. Una seconda ipotesi invece attribuisce l'interazione cross-modale non al precoce processamento sensoriale ma ad un *processo decisionale più tardivo* (Marks, 2004). In questo caso la presenza di uno stimolo congruente avrebbe un effetto di facilitazione, abbassando così la soglia del criterio di risposta, senza andare in alcun modo a influenzare il processamento sensoriale dello stimolo. Infine, è plausibile ipotizzare che l'interazione cross-modale coinvolga sia processi di natura sensoriale sia processi di natura decisionale (Odgaard et al., 2003). Una recentissima ricerca (Colin, Radeau & Deltenre, 2005) offre a tal riguardo risultati assai interessanti. La ricerca, che ha sfruttato l'effetto McGurk¹, ha indagato l'*audiovisual speech* (integrazione delle componenti visive e uditive del parlato) modulando alcune variabili di natura sia

¹ L'effetto McGurk (McGurk & McDonald, 1976), indagato in relazione all'*audiovisual speech*, si verifica quando al soggetto vengono presentate due sillabe diverse, una in forma visiva (movimento delle labbra) e una in forma uditiva (parlato). In presenza quindi di una discordanza tra la componente uditiva e quella visiva, i soggetti, cui viene chiesto di riprodurre la sillaba percepita, combinano quanto hanno udito con quanto hanno visto. Ad esempio, le sillabe percepite visivamente "gaga" e le sillabe percepite uditivamente "baba" vengono integrate nelle sillabe percepite "dada".

sensoriale sia cognitiva. Nel corso di due esperimenti, sono state infatti manipolate due variabili sensoriali (l'intensità della voce e la grandezza del volto) e una variabile cognitiva (tipo di task, con risposta a scelta multipla o con risposta libera). I risultati indicano che l'integrazione dipende da entrambi gli ordini di fattori, percettivi e cognitivi.

3. L'apporto della neuropsicologia alla comprensione del processo di integrazione

La questione della natura dell'integrazione cross-sensoriale è stata ulteriormente approfondita affiancando ai dati comportamentali dati di natura neuropsicologica, che si pongono l'obiettivo di esplorarne i processi neurali sottostanti.

Gli studi che hanno utilizzato tecniche elettroencefalografiche e di *neuroimaging* hanno confermato l'esistenza del *redundant target effect*, già rilevato a livello comportamentale (Fort et al., 2002; Teder-Sälejärvi et al., 2002). Gli stimoli bimodali congruenti producono infatti, negli stadi di elaborazione sensoriale, risposte neurali più veloci e di maggiore intensità rispetto a quelle elicitate da stimoli unimodali o da stimoli bimodali incongruenti. L'insieme di tali dati suggerisce che, in presenza di stimoli congruenti, avvenga una qualche forma di facilitazione dovuta all'integrazione cross-modale. A questo proposito, il "*redundant target effect*" ha suggerito diverse interpretazioni circa i processi sottostanti implicati:

- i *race models* affermano che le due componenti vengono processate indipendentemente e che il tempo di risposta coincide con la fine del processamento di quella che, tra le due, termina in tempi più rapidi;
- gli *independent coactivation models* ipotizzano che le due componenti inducano attivazioni indipendenti che vengono sommate per elicitare la risposta;
- infine, gli *interactive coactivation models* affermano che il processamento di uno stimolo in una modalità influenza il processamento di uno stimolo in un'altra modalità, ipotizzando che l'integrazione possa avvenire a diversi livelli: di processamento sensoriale e/o cognitivo di selezione della risposta.

Recentemente alcuni studi neuropsicologici hanno fornito supporto a favore di quest'ultimo tipo di modello (Calvert et al., 1999; Calvert et al., 2000; Hadjikhani & Roland, 1998). In genere, vi è accordo tra i ricercatori sul fatto che l'integrazione abbia inizio ad uno stadio molto precoce del processo di elaborazione sensoriale degli stimoli (Giard & Peronnet, 1999). In un esperimento condotto da Giard & Peronnet, i soggetti furono sottoposti ad un compito di identificazione che utilizzava stimoli unimodali visivi e uditivi e stimoli bimodali congruenti risultanti dalla combinazione delle due componenti. Come previsto, l'identificazione degli stimoli bimodali fu più rapida ed accurata. Un'analisi spaziotemporale degli ERPs ha mostrato che già tra i 40 e i 200 ms dopo la presentazione dello stimolo si manifestano patterns multipli di integrazione cross-modale sia nelle aree corticali specifiche visiva e uditiva sia in aree non specifiche, come nella regione fronto-temporale destra. Gli effetti indotti da stimoli bimodali ridondanti sono stati interpretati come modulazione della risposta unimodale uditiva N1 e della risposta unimodale visiva N185 nelle rispettive cortecce sensoriali, nonché come nuova attività nella corteccia visiva e nelle aree fronto-temporali destre. Anche Teder-Sälejärvi (Teder-Sälejärvi et al., 2002) ha rilevato una prima deflessione che ha inizio attorno ai 130 ms e raggiunge il picco tra i 160 e i 170 ms nelle aree corticali occipito-temporali ventrali. Anche in questo caso, una simile interazione audio-visiva potrebbe essere interpretata come modulazione dell'onda visiva N1. Tale effetto infatti sembra rappresentare un'influenza dell'input uditivo sul processamento che ha luogo in un'area corticale prevalentemente visiva.

Relativamente alla presentazione di stimoli bimodali incongruenti, Fort (Fort et al., 2002) hanno trovato che in generale, come previsto, gli stimoli bimodali non ridondanti non producono un effetto di facilitazione a livello di dati comportamentali, dal momento che per portare a termine il compito di identificazione i soggetti devono processare in modo completo ogni componente dello stimolo. Invece, contrariamente a quanto previsto, sebbene essa sia di minore intensità e più tarda rispetto a quella registrata in presenza degli stimoli ridondanti, è stata rilevata una precoce attività cross-modale in risposta alla presentazione degli

stimoli bimodali non ridondanti, caratterizzata da attivazione sia nelle aree sensoriali specifiche sia nell'area fronto-temporale destra non specifica. Gli autori suggeriscono che l'ipotesi della coattivazione interattiva, che ha ricevuto supporto da recenti studi neuropsicologici, ben si adatta anche ai risultati di questo studio. Tali risultati, inoltre, sono in accordo con i principi neurali di integrazione multisensoriale che si applicano a livello dei singoli neuroni nel collicolo mammale superiore e nella corteccia polisensoriale: secondo tali principi, a questo livello la coincidenza spaziale e temporale degli stimoli è condizione sufficiente per innescare l'integrazione (Bushara et al., 2001; Stein & Wallace, 1996). Si può quindi ipotizzare che, a fronte di una precoce convergenza sensoriale sempre presente, solo in un secondo momento le modalità di elaborazione degli stimoli si differenzino sulla base della natura congruente o incongruente delle informazioni sensoriali. In altre parole, l'attivazione indica un'integrazione cross-modale non tanto nell'identificazione dello stimolo quanto nella sua mera ricezione (*detection*).

I meccanismi fisiologici dell'integrazione sono complessi e molteplici. L'insieme dei risultati dimostra infatti la *flessibilità* dei processi cross-modali, che presentano notevoli possibilità di adattamento in funzione delle caratteristiche dello stimolo. Essi infatti sono influenzati da fattori sia di natura endogena (ad esempio Fort et al. (2002) hanno rilevato un effetto del grado di *expertise* dei soggetti rispetto al task, per cui i soggetti a dominanza visiva e i soggetti a dominanza uditiva presentavano differenti patterns di attivazione) sia di natura esogena (condizioni sperimentali e tipo di compito). La *natura della stimolazione* ha sicuramente un effetto sulle modalità di integrazione (Callan et al., 2001; Calvert et al., 2001). Tale carattere flessibile dei processi di integrazione ha ancora una volta una funzione adattiva, in quanto risponde all'esigenza di produrre una risposta efficiente in presenza di condizioni ambientali variabili.

3.1. Circuiti neurali implicati nel decoding intersensoriale

Relativamente alle aree coinvolte nell'integrazione cross-modale, esistono due differenti scuole di pensiero: alcuni ritengono che ogni specifica combinazione di stimoli sensoriali (ad es. audio-visiva o audio-tattile) venga integrata in una precisa area "associativa" polimodale ad essa dedicata. Per quanto riguarda nello specifico l'integrazione audio-visiva, si ritiene che le aree dedicate siano la corteccia frontale destra inferiore, la corteccia temporale destra, il solco temporale superiore, il giro temporale superiore, l'insula e il lobo parietale, come evidenziato da studi PET (Bushara, Grafman & Hallett, 2001; Hadjikhani & Roland, 1998) e fMRI (Calvert et al., 2000; Downar et al., 2000).

Altri invece sostengono che le aree che processano gli stimoli unimodali processino anche gli stimoli multimodali, ipotizzando che i sensi abbiano accesso l'uno all'altro grazie ad aree di ritrasmissione (*relay*) subcorticali. Questa ipotesi è supportata dall'evidenza che la lesione delle presunte aree polimodali non preclude l'integrazione intersensoriale (Ettlingen & Wilson, 1990 per una rassegna). A tale proposito, con uno studio che ha utilizzato la risonanza magnetica funzionale (fMRI) per indagare il fenomeno del *lip-reading* (comprensione del linguaggio attraverso la lettura dei movimenti labiali), Olson (Olson, Gatenby & Gore, 2002) ha fornito supporto a questa seconda ipotesi dimostrando che le aree unimodali, utilizzando come aree subcorticali di ritrasmissione il claustrum e il putamen, elaborano stimoli appartenenti a diverse modalità sensoriali, mettendoli in comunicazione tra loro.

Recentemente Fort e Giard (2004) hanno suggerito una nuova prospettiva: essi hanno dimostrato che la convergenza ha inizio a livello delle aree sensoriali specifiche in uno stadio veramente molto precoce, intorno ai 40-50 ms dopo la presentazione dello stimolo. Rispetto a questi fenomeni precoci, studi condotti sulle scimmie suggeriscono l'esistenza di proiezioni dirette dalla corteccia uditiva primaria alla corteccia visiva primaria e viceversa (Falchier et al., 2002; Schroeder et al., 2001). Questo dato è difficilmente compatibile con l'ipotesi che siano presenti delle proiezioni dalle aree di convergenza polisensoriali verso le aree specifiche

(Calvert et al., 2001). Gli autori tuttavia ipotizzano che tali proiezioni siano coinvolte in stadi più tardivi di processamento, di natura cognitiva più che percettiva.

3.2. I neuroni multisensoriali

Infine, bisogna specificare che il fenomeno dell'integrazione cross-modale può essere indagato a diversi livelli. Un primo livello di analisi - quello che è stato esposto fino ad ora - pone l'attenzione su specifiche regioni che fungono da aree di convergenza o che partecipano al processo di integrazione. Tuttavia un'ulteriore analisi di tipo strettamente psicofisiologico può essere effettuata anche a livello di singoli neuroni detti *neuroni multimodali*, che hanno la particolarità di elaborare informazioni appartenenti a diverse modalità sensoriali. Esistono infatti neuroni bimodali e trimodali in grado di gestire informazioni sia uditive sia visive sia somatosensoriali.

Mentre della convergenza a livello di area cerebrale abbiamo oggi una discreta conoscenza, i meccanismi sottostanti alla convergenza a livello di singoli neuroni sono ancora poco noti. Ciò di cui siamo a conoscenza è l'esistenza di due tipologie di convergenza multimodale, che vengono differenziate sulla base del loro effetto. Il primo tipo di convergenza neuronale, che è stato studiato a livello del collicolo superiore e a livello della corteccia cerebrale, è detto *convergenza eccitatoria-eccitatoria*. Quando i neuroni multimodali che operano tale tipo di convergenza ricevono informazioni di diverso tipo, le integrano, provocando un miglioramento della risposta. Ad esempio, è possibile che un neurone risponda debolmente ad un certo stimolo uditivo e in modo più accentuato ad un certo stimolo visivo; nel caso in cui esso riceva simultaneamente i due stimoli, la sua risposta sarà nettamente più intensa. Questo tipo di risposta, come nel caso della convergenza a livello delle aree cerebrali, viene detto miglioramento multisensoriale (*multisensory enhancement*). Esso è influenzato da diversi fattori legati alle caratteristiche degli stimoli, alle modalità di presentazione degli stessi e alle caratteristiche del neurone che opera la convergenza e rispetta i principi della vicinanza temporale, della vicinanza spaziale e dell'*inverse*

effectiveness effect. Il secondo tipo di integrazione è denominato *convergenza eccitatoria-inibitoria*. Essa si verifica ad esempio nel caso di soppressione di una risposta in seguito alla presentazione di uno stimolo inatteso (per esempio uno stimolo visivo al posto di uno stimolo uditivo) durante un compito di attenzione selettiva. Si verifica in questo caso un fenomeno di inibizione della risposta. Infatti, a fronte dell'azione eccitatoria di una modalità, è presente un'azione inibitoria esercitata dall'altra modalità. Mentre nel caso della convergenza eccitatoria-eccitatoria l'influenza sulla risposta è molto accentuata, nel caso della convergenza eccitatoria-inibitoria si osserva invece semplicemente una lieve modulazione della risposta (Meredith, 2002).

4. La decodifica audio-visiva dei volti: riconoscere l'identità dal volto e dalla voce

La maggior parte delle ricerche che hanno indagato la cross-modalità hanno impiegato stimoli sensoriali molto semplici, rilevando la centralità delle fasi precoci di elaborazione percettiva. Sostanziali differenze sono state rilevate invece da coloro che hanno utilizzato stimoli che si collocano ad un più elevato livello informativo. Particolarmente interessante è un recentissimo studio di Schweinberger (Schweinberger, in press) che ha dimostrato l'importanza dell'integrazione audiovisiva ai fini del riconoscimento dell'identità delle persone. Da precedenti studi era infatti emerso che sia il volto sia la voce costituiscono delle informazioni importanti quando dobbiamo giudicare il grado di familiarità, ma non era mai stata indagata la convergenza tra i due codici rispetto a tale compito. Un indizio della possibile presenza di fenomeni integrativi era stato fornito da una ricerca che, utilizzando la risonanza magnetica, ha dimostrato che la percezione di una voce familiare attiva la cosiddetta "area fusiforme del volto", che tipicamente viene appunto attivata dalla percezione dei volti (von Kriegsten et al., 2005). Nello studio di Schweinberg, ai soggetti veniva chiesto di giudicare se una frase standardizzata veniva pronunciata da una persona familiare o sconosciuta. Nella condizione unimodale veniva presentata solo la voce, mentre quella audiovisiva era caratterizzata dalla simultanea presentazione di un volto, familiare o sconosciuto,

congruente o incongruente. I risultati dimostrano che, in termini sia di accuratezza sia di tempi di risposta, la simultanea presentazione del volto produce sistematici costi (nella condizione di incongruenza) e benefici (nel caso della congruenza) nella valutazione delle voci familiari, mentre nel caso delle voci non note tali effetti non si verificano. Gli autori suppongono che ciò sia dovuto al fatto che, a seguito della presentazione delle informazioni audiovisive, viene operato un confronto queste e le rappresentazioni multimodali delle persone familiari che sono conservate nella memoria a lungo termine.

5. La convergenza di pattern emotivi

Se il riconoscimento del volto neutro costituisce un processo complesso, ancor più complessi sono i meccanismi che il nostro sistema cognitivo attua quando rileva che le informazioni multimodali sono cariche di significato emotivo.

Quando decodifichiamo un'emozione, utilizziamo molteplici fonti di informazione. Numerose ricerche si sono occupate di capire cosa avviene quando, allo scopo di riconoscere e comprendere un'emozione, l'individuo si trova a dover in qualche modo integrare tali informazioni appartenenti a diversi sistemi sensoriali. Gli studi che sono stati condotti allo scopo di comprendere questo particolare processo di integrazione cross-modale, hanno focalizzato l'attenzione in particolare sulla convergenza tra il *canale visivo* e quello *uditivo* che, come illustrato nei cap. 1 e 2, hanno un ruolo fondamentale nel processo di decoding delle emozioni.

Sembra esistere una stretta interrelazione tra la decodifica della mimica facciale e la decodifica dell'espressione vocale delle emozioni. A tal proposito, van Lancker e Sidtis (1992) hanno trovato che alcuni pazienti con diagnosi di aprosodia presentavano anche un correlato deficit nel riconoscimento dei volti. Parallelamente, Scott et al. (1997) hanno osservato un'incapacità di decodifica delle componenti prosodiche in un paziente con difficoltà nel riconoscimento delle espressioni facciali. Questi dati non permettono tuttavia di stabilire con certezza che le informazioni facciali e prosodiche convergano in una rappresentazione amodale

comune: essi attestano soltanto l'esistenza di una semplice correlazione tra i due ordini di deficit. Tra l'altro, alcuni studi hanno indicato l'esistenza di asimmetrie tra riconoscimento della voce e riconoscimento del volto. Alcune emozioni sono infatti più facilmente riconoscibili sulla base del volto o viceversa: ad esempio la gioia viene facilmente riconosciuta sulla base dell'espressione facciale, ma spesso la voce della gioia viene confusa con l'espressione neutra (Vroomen et al., 1993). L'insieme di queste e simili ricerche ha fornito spunti di riflessione interessanti, che sono stati sviluppati da un filone sperimentale che si è posto l'obiettivo specifico di esplorare il processamento di stimoli emotivi multimodali. Tale corpus di ricerche include sia studi di natura comportamentale sia studi di natura neuropsicologica.

6. L'apporto degli studi comportamentali

Prima di intraprendere un excursus sui risultati delle ricerche che hanno indagato la decodifica cross-modale delle emozioni, è necessario sottolineare che essa presenta delle caratteristiche qualitativamente diverse rispetto alla percezione cross-modale classicamente studiata, che ha impiegato come stimoli lampi di luce e semplici suoni inarticolati. Ciò che differenzia la decodifica cross-modale delle emozioni è la *complessità degli stimoli* implicati. Un parallelo può essere individuato negli studi che si sono occupati di indagare un caso particolare di processamento multimodale: lo *speech reading* o comprensione della lingua parlata attraverso il simultaneo processamento delle informazioni uditive e visive correlate. Normalmente, quando qualcuno ci parla, noi siamo impegnati sia ad ascoltare le sue parole sia a guardare il movimento delle sue labbra. La nostra comprensione è il risultato dell'integrazione tra questi due livelli di informazione. L'effetto McGurk (McGurk & McDonald, 1976) ha dimostrato che tale integrazione ha carattere automatico e obbligato (vedi cap. 2): quando ad un soggetto vengono presentate due diverse sillabe, l'una in forma visiva (movimento delle labbra) e l'altra in forma uditiva (linguaggio parlato) e gli viene chiesto di riferire la sillaba percepita, egli riporta un percelto derivante dalla combinazione delle due. Allo stesso modo, anche

le componenti mimiche e vocali dell'espressione emotiva costituiscono delle informazioni complesse.

Sono stati Beatrice de Gelder e il suo gruppo di ricerca a condurre buona parte degli studi empirici che si sono occupati specificamente di indagare la percezione cross-modale delle emozioni basata su informazioni di natura audio-visiva.

In primo luogo la de Gelder (de Gelder & Vroomen, 2000) si è posta lo scopo di determinare se, in una situazione bimodale in cui le informazioni sullo stato emotivo sono fornite sia attraverso il canale visivo sia attraverso quello vocale, entrambe le modalità contribuiscono al riconoscimento. A tal fine, ai soggetti sono stati mostrati stimoli costituiti da volti e frasi esprimenti tristezza o gioia, in condizione unimodale e bimodale. In una prima fase dell'esperimento, ai soggetti è stato semplicemente chiesto di indicare se la persona cui il volto e/o la voce si riferivano era triste o felice. Nelle due fasi successive è stata invece data loro istruzione di prestare attenzione solo all'espressione del volto o solo al tono di voce. I risultati indicano che, come avviene nei classici esperimenti sulla percezione bimodale, i tempi di latenza sono più veloci quando vengono somministrati due stimoli congruenti (volto e voce esprimenti la stessa emozione) rispetto a quando viene presentato un solo stimolo. Questo fatto indica che, per il sistema di processamento, l'integrazione delle informazioni visive ed uditive costituisce un meccanismo usuale ed efficace.

I tempi più lunghi si registrano invece nel caso di due stimoli incongruenti (volto e voce esprimenti due emozioni diverse), indicando che tale situazione rappresenta un condizione poco naturale e che quindi richiede un maggiore sforzo in termini di decodifica delle informazioni. Per quanto concerne la correttezza del riconoscimento, è stata osservata un'influenza del volto sulla voce e, viceversa, della voce sul volto, come precedentemente riscontrato dallo studio pionieristico di Massaro ed Egan (1996).

6.1. La funzione dell'integrazione delle informazioni emotive multimodali

La de Gelder (de Gelder, 2000) si è chiesta quale sia la funzione della convergenza cross-modale nel decoding delle emozioni. La decodifica simultanea di informazioni acustiche e visive rappresenta infatti un caso di ridondanza. Tre diverse ipotesi sono state formulate per spiegare tale fenomeno:

In primo luogo è possibile che la presenza di due diversi tipi di segnali sia utile quando la ricezione dei segnali provenienti da uno dei due sistemi è povera o assente, ad esempio in presenza di rumore o cecità. Questa ipotesi però non spiega perché, anche nel caso in cui entrambi i sistemi funzionino al meglio, l'organismo processi in modo completo tutte le informazioni disponibili.

Una seconda ipotesi è che l'organismo sia avvantaggiato dalla ridondanza perché i due sistemi sono complementari e che questa condizione gli permetta una maggiore efficienza nella risposta comportamentale. Viene assunto come prova di ciò il fatto che, come già accennato, alcune emozioni vengono meglio espresse dal sistema visivo ed altre da quello uditivo. Tuttavia, bisogna osservare che la convergenza avviene anche quando entrambi gli ordini di informazioni presentano scarsa ambiguità (de Gelder & Vroomen, 2000).

De Gelder (2000) avanza quindi una terza ipotesi, suggerendo che la ridondanza permetta una maggiore *efficienza nella risposta comportamentale* non perché i due sistemi siano complementari ma perché l'organismo già dai primissimi stadi di processamento integra gli stimoli e ciò gli consente di produrre una risposta molto più veloce rispetto al caso in cui processasse gli stimoli separatamente per poi integrare i percetti solo nella fase finale. A conferma di ciò, paragonando il decoding unimodale a quello multimodale, gli studi della de Gelder dimostrano che il processamento degli stimoli integrati precocemente avviene in modo più veloce ed efficiente.

6.2. L'integrazione come processo precoce ed automatico

Come già accennato, a livello di riconoscimento ed etichettamento delle emozioni presentate nella duplice modalità audiovisiva, esiste un reciproco effetto di influenzamento tra le due modalità sensoriali. Il fatto che tale *bias* cross-modale si verifichi anche quando viene esplicitamente richiesto di prestare attenzione ad un'unica modalità sensoriale (quella visiva o quella uditiva) (de Gelder & Vroomen, 2000) rinforza l'ipotesi che l'integrazione avvenga ad uno stadio di processamento molto *precoce* e in modo *automatico ed obbligato*. Ciò induce ad escludere l'ipotesi che il *bias* possa essere il frutto di una valutazione e di un giudizio consapevoli, attuati come risultato della presa di coscienza di un'incongruenza dopo che il processamento separato delle due fonti di informazione è terminato. Anzi, addirittura l'integrazione si verifica nonostante il fatto che i soggetti si dichiarino consapevoli dell'incongruenza. Evidentemente quindi questa impressione fenomenica di incongruenza si colloca ad un livello cosciente e molto differente da quello in cui avviene il processamento cross-modale. Si può quindi ritenere, secondo gli autori, che l'integrazione sia un *fenomeno percettivo*, analogamente a quanto rilevato nel caso dell' *audio-visual speech*. Tale processo percettivo precede ampiamente fenomeni come il riconoscimento e la comprensione delle emozioni, che sono centrati sul significato personale e sociale dell'emozione. A conferma delle proprie affermazioni, de Gelder e colleghi (de Gelder, Vroomen & Bertelson, 1998) hanno replicato gli esperimenti precedentemente condotti introducendo una variante: in alcuni casi, infatti, i volti venivano presentati invertiti. L'inversione del volto comporta una drastica diminuzione della possibilità di identificazione dell'identità e dell'espressione del volto. Ciò è dovuto al fatto che, come precedentemente esposto, l'identificazione del volto si differenzia dall'identificazione di altri tipi di oggetti, in quanto coinvolge la configurazione complessiva e non i singoli attributi. I risultati dello studio mostrano che il decoding dell'espressione facciale influenza il giudizio circa il tono della voce solo quando il volto è presentato dritto ma non quando è capovolto. Questo dato è interessante se

messo in relazione al fatto che il riconoscimento dell'emozione espressa dal volto diventa difficoltoso quando lo stimolo è capovolto. L'effetto cross-modale osservato quando lo stimolo è presentato dritto conferma invece l'ipotesi che tale processo sia un fenomeno percettivo automatico che non può essere ridotto ad un processo post-percettivo volontario di aggiustamento.

7. Il contributo della neuropsicologia

A livello neuropsicologico, diversi studi sono stati effettuati allo scopo di mettere in luce i processi cerebrali implicati nella decodifica multimodale delle emozioni e di spiegare con maggiore chiarezza i dati comportamentali disponibili. Come precedentemente esposto, gli studi behavioural mostrano che, quando vengono presentati simultaneamente due stimoli emotivi congruenti, l'uno vocale e l'altro facciale, la risposta è più accurata e i tempi sono più rapidi rispetto a quando viene presentato uno stimolo unimodale. Ciò suggerisce che l'organismo sfrutti le risorse multiple offerte dall'ambiente ai fini di produrre delle risposte comportamentali più rapide ed efficienti. Tuttavia, i dati disponibili non forniscono alcuna prova certa del fatto che l'integrazione avvenga ad uno stadio precoce del processamento. I tempi più brevi che si presentano in concomitanza con gli stimoli bimodali potrebbero essere spiegati da un *race model*, cioè da un modello secondo il quale i due stimoli vengono processati separatamente, e quello il cui processamento ha termine per primo determina la prestazione. Un'altra ipotesi possibile è che l'integrazione degli stimoli abbia luogo non appena essi si presentano e che il loro processamento congiunto sia il meccanismo che meglio potrebbe sfruttare la ridondanza della stimolazione, come supposto dagli *interactive coactivation models* (vedi cap. 3). Uno degli obiettivi che si pongono le ricerche neuropsicologiche è quindi quello di portare nuove conoscenze che possano disambiguare tale questione.

Un ulteriore principale argomento di discussione nella ricerca sulla convergenza cross-modale degli stimoli di natura emotiva riguarda le sue coordinate temporali. Di conseguenza, lo studio dei *potenziali evocati corticali (ERPs)* si è rivelato,

data la sua alta definizione temporale, particolarmente utile ed efficace (Rugg & Coles, 1997).

Proprio al fine di ampliare le conoscenze relative a tali questioni, pressoché tutti gli studi condotti hanno centrato l'attenzione sugli *stadi percettivi precoci* del processo di elaborazione degli stimoli emotivi. Infatti, gli indici ERP più frequentemente indagati, come la N1 e il MMN, segnalano processi di natura sensoriale.

7.1 La componente MMN (mismatch negativity) come indicatore indiretto della convergenza audio-visiva

I primi studi sull'argomento, effettuati alla fine degli anni '90, hanno utilizzato come indice la *MMN (mismatch negativity)*, un picco negativo che si presenta in concomitanza con stimolazioni di tipo uditivo quando, in una serie di stimoli ripetitivi, appare uno stimolo deviante (Näätänen, 1992). Tale ERP non è sotto controllo attentivo e segnala la ricezione di uno stimolo che tradisce le aspettative (Levänen & Sams, 1997). De Gelder e colleghi (de Gelder et al., 1999) hanno utilizzato la componente MMN al fine di indagare l'influenza dell'espressione facciale sul processamento delle componenti emozionali vocali. Ai soggetti sono state presentate coppie congruenti o incongruenti di stimoli uditivi (parole pronunciate in tono triste o arrabbiato) e stimoli visivi (espressioni facciali di tristezza e rabbia) con la consegna di prestare attenzione al volto e di ignorare la componente uditiva. I risultati indicano che quando, dopo una serie di stimolazioni congruenti, ne viene presentata una incongruente, appare una risposta cerebrale negativa precoce (latenza 178 ms) localizzata nelle aree anteriori della corteccia, in particolare in F3, Cz e soprattutto Fz. Lo stesso avviene quando, dopo una serie di stimoli congruenti, ne compare uno incongruente. I parametri della componente ERP evidenziata dalla de Gelder corrispondono a quelli della MMN, che ha infatti una latenza di 178 ms ed è principalmente localizzato in Fz. Gli autori ipotizzano che l'assenza di una lieve

positività identificabile come P3 o P3a indichi che il processo non avviene sotto controllo attentivo ma è obbligato.

I dati confermano ed estendono i precedenti risultati ottenuti negli studi comportamentali, pur non permettendo, data la tecnica impiegata, di localizzare con precisione le sedi in cui avviene l'integrazione cross-modale. Anche Surakka et al. (1998) hanno utilizzato la MMN per studiare l'integrazione tra stimolazioni visive e stimolazioni uditive, rilevando che gli stimoli visivi hanno un impatto sul processamento degli stimoli uditivi: Surakka ha infatti studiato l'effetto di immagini emotivamente connotate tratte dall'International Affective Picture System su stimoli uditivi (toni standard di 1000 Hz con probabilità pari a 0.85 e toni devianti di 1050 Hz con probabilità pari a 0.15), trovando che l'ampiezza dell'MMN era significativamente attenuata quando l'emozione legata alla figura era a basso arousal e positiva rispetto a quando era negativa o ad alto arousal. Per spiegare questo dato, gli autori suggeriscono che gli stimoli positivi a basso arousal segnalino la presenza di un ambiente non allarmante e non appetitivo, e che questo faccia sì che la tendenza a rilevare automaticamente cambiamenti inaspettati a livello delle stimolazioni uditive sia meno importante da un punto di vista adattivo e di conseguenza presenti un decremento.

7.1.1. Il ruolo dell'amigdala

Surakka inoltre propone un modello secondo il quale l'amigdala costituisce un importante elemento di mediazione nel contesto dell'elaborazione degli stimoli sensoriali. Gli stimoli positivi a basso arousal infatti determinerebbero una diminuzione dell'attività dell'amigdala, che a sua volta causerebbe una diminuzione dell'attività del meccanismo deputato a rilevare automaticamente i cambiamenti inattesi a livello della corteccia uditiva. Il fatto che non sia stato registrato un aumento dell'MMN in presenza di stimoli negativi ad alto arousal viene spiegato ipotizzando che già la detezione degli stimoli incongruenti fosse massimamente attivata. In sintesi quindi, gli autori suggeriscono che il processamento degli stimoli uditivi sia influenzato dall'amigdala e, attraverso quest'ultima, dagli stimoli

emozionali di tipo visivo. Studi precedenti in effetti hanno dimostrato che l'amigdala ha un ruolo di primo piano nel processamento di informazioni connotate emotivamente. Uno studio condotto da Dolan (Dolan et al., 2001) per mezzo della fMRI (risonanza magnetica funzionale) ha mostrato come l'amigdala sia coinvolta nell'integrazione di informazioni visive e uditive legate all'espressione della paura: l'attivazione dell'amigdala e del giro fusiforme aumentano quando un volto esprime paura viene presentato congiuntamente ad un messaggio verbale pronunciato in tono impaurito. Inoltre, l'amigdala riceve proiezioni da tutte le principali aree corticali sensoriali e presenta importanti proiezioni verso le aree visive ed uditive. In base a tutto ciò, è stato suggerito che l'amigdala possa avere una funzione di modulazione sugli stadi relativamente precoci del processamento sensoriale (LeDoux, 1995). In particolare essa, oltre a partecipare all'elaborazione unimodale degli stimoli, è coinvolta nel processamento multimodale degli stimoli che hanno valenza affettiva, sia propriamente in termini emozionali, sia più in generale in termini di valenza edonica: è ritenuta essere una struttura associativa multimodale, perchè riceve afferenze sia dalle diverse aree sensoriali specifiche sia dalle aree polimodali della corteccia temporale (O'Doerty, Rolls & Kringelbach, 2004 per una rassegna).

7.1.2. Valenza edonica e integrazione cross-modale

Relativamente alla *valenza edonica*, la letteratura indica che, nel decoding delle espressioni facciali, le emozioni con valenza negativa vengono processate principalmente nell'emisfero destro, mentre quelle con valenza positiva sono elaborate prevalentemente nell'emisfero sinistro (Davidson & Irwin, 1999). Pourtois e colleghi (Pourtois et al., 2005) hanno voluto indagare con uno studio PET se tale effetto di lateralizzazione è rilevabile anche in presenza di una stimolazione bimodale. Essi hanno utilizzato stimoli visivi (espressioni facciali di gioia e paura), stimoli uditivi (una parola bisillabica pronunciata in tono felice o impaurito) e stimoli bimodali congruenti risultanti dalla combinazione di quelli unimodali. L'originalità dell'esperimento consiste nell'utilizzo di una consegna indiretta o "nascosta" (*covert*):

ai soggetti è stato chiesto di valutare il genere del soggetto che esprimeva l'emozione. Come emerge dai risultati, rispetto agli stimoli unimodali, quelli bimodali attivano maggiormente un'area di convergenza situata nella corteccia temporale sinistra. Tale effetto è descritto anche da Calvert (Calvert et al., 2001). L'attivazione nello specifico coinvolge il giro mediotemporale sinistro (MTG), già precedentemente indicato come area di convergenza multimodale (Mesulam, 1998) e il giro fusiforme sinistro, la cui attivazione era stata rilevata da uno studio fMRI (Dolan, 2001). Inoltre, le analisi condotte separatamente per le due emozioni rivelano la presenza di aree di convergenza supplementari, situate prevalentemente nell'emisfero sinistro per gli stimoli bimodali della gioia e nell'emisfero destro per gli stimoli bimodali della paura. Questo dato indica l'esistenza di sostrati neurali di processamento cross-modale differenziati sulla base della valenza edonica dello stimolo emotivamente connotato. Infine, confermando i dati già presenti in letteratura, i ricercatori hanno evidenziato un'attivazione dell'amigdala per gli stimoli unimodali facciali e per gli stimoli bimodali esprimenti paura. Complessivamente, lo studio condotto dimostra che la presentazione congiunta di stimoli emozionali appartenenti a diverse modalità sensoriali (visiva e uditiva) porta all'attivazione di aree di convergenza eteromodali e che tale processo, data la natura implicita della consegna, ha carattere obbligato.

7.2. Un altro indice indiretto: la componente N1

Come anticipato, la *N1*, componente ERP sensoriale legata al processamento degli stimoli uditivi, è stata utilizzata, al pari della MMN, al fine di indagare le coordinate temporali del fenomeno di integrazione cross-modale che si verifica quando vengono presentati simultaneamente stimoli emotivi visivi e uditivi. In uno studio ERP Pourtois e collaboratori (Pourtois et al., 2000) hanno presentato ai soggetti coppie congruenti e incongruenti di stimoli uditivi (frammenti di 4 sillabe pronunciate in tono triste o arrabbiato) e di stimoli visivi (espressioni facciali tristi o arrabbiate presentate normalmente o capovolte), con la consegna di prestare

attenzione ai volti ignorando la voce. I risultati rivelano che l'informazione visiva influenza il processamento dello stimolo uditivo già dopo 110 ms dopo la stimolazione. Tale influenza si manifesta come un aumento dell'ampiezza di N1, come già rilevato in precedenti studi che hanno utilizzato stimoli non di tipo emotivo (Giard & Peronnet, 1999). Inoltre, l'integrazione avviene solo in presenza di stimoli congruenti, confermando così quanto rilevato a livello comportamentale. Si può dunque ipotizzare che l'elaborazione degli stimoli uditivi venga facilitata dalla presentazione di uno stimolo visivo congruente in termini di contenuto emotivo. L'integrazione non avviene invece quando il volto è capovolto. Questo perché la rotazione di 180° impedisce il normale processo di elaborazione del volto che, come accennato precedentemente, presenta un percorso di processamento specifico e dedicato.

La maggior parte degli studi ha indagato il processamento cross-modale delle emozioni prendendo in esame l'ampiezza delle componenti ERP implicate, rilevando un incremento o decremento delle componenti unimodali precoci, come il picco uditivo N1 o il picco visivo P1, che hanno luogo intorno ai 100 ms di latenza nelle aree sensoriali specifiche (Calvert, Brammer & Iversen, 2000; Giard & Peronnet, 1999; Raij, Uutela & Hari, 2000; Sams et al., 1991). Infatti, l'incremento dell'attività nella corteccia modalità-specifica è considerato un fondamentale correlato elettrofisiologico della cross-modalità (de Gelder, 2000; Driver & Spence, 2000). Ad esempio, è stata segnalata un'attivazione amplificata a livello della corteccia uditiva durante la lettura del labiale (Calvert et al., 1997), del giro fusiforme e dell'amigdala durante la percezione di stimoli emozionali bimodali (Dolan, Morris & de Gelder, 2001) e delle aree tattili durante una stimolazione visuo-tattile (Macaluso, Frith & Driver, 2000). Nel complesso, l'integrazione cross-modale è segnalata da un'amplificazione sia a livello delle aree specifiche sia a livello di quei network corticali, come la corteccia parietale posteriore e il giro temporale mediale, che hanno natura multimodale (Mesulam, 1998).

7.3. L'indice di integrazione multimodale P2b

Solo pochi studi neuropsicologici hanno invece studiato la convergenza audio-visiva a contenuto emotivo tenendo conto del fattore temporale, rappresentato dalla *latenza*.

Tra questi, Pourtois e colleghi (2002) hanno dimostrato empiricamente che l'elaborazione degli stimoli emotivi audio-visivi comporta anche precise implicazioni in termini di latenza. Essi si sono posti infatti l'obiettivo di verificare se la presentazione di uno stimolo facciale può influenzare anche la latenza, oltre che l'ampiezza, dei processi di natura uditiva, prendendo però come oggetto di osservazione un intervallo temporale relativamente meno precoce di quello indagato dagli studi precedentemente citati. In sintesi, le analisi effettuate sulla componente uditiva mostrano l'esistenza di un picco positivo intorno ai 240 ms con una topografia posteriore, che gli autori denominano P2b. Gli autori ritengono che la P2b rappresenti un *indice di integrazione* tra la componente uditiva e quella visiva. Tale picco segue le componenti modalità-specifiche uditive N1 e P2 e precede il complesso amodale N2-P3, che si sa essere deputato all'elaborazione cognitiva ad un più tardivo stadio decisionale. Gli stimoli bimodali congruenti elicitano una P2b più precoce rispetto agli stimoli incongruenti, suggerendo che il processamento uditivo, in presenza di informazioni incongruenti, sia ritardato. Questi risultati sono in accordo con i precedenti dati comportamentali, che dimostrano un accorciamento dei tempi di processamento in presenza di stimoli multimodali congruenti. Inoltre, essi sono in accordo con quelli recentemente ottenuti da studi basati sulla risonanza magnetica funzionale (Calvert, Campbell & Brammer, 2000) o sulla magnetoencefalografia (Raij et al., 2000). Un'analisi di localizzazione della fonte effettuata durante l'intervallo temporale corrispondente alla P2b ha messo in evidenza un'implicazione della corteccia cingolata anteriore, che è implicata nel processamento della congruenza/incongruenza tra stimoli (McLeold & McDonald, 2000). I risultati sono coerenti con un coinvolgimento della corteccia cingolata anteriore nell'integrazione audio-visiva intorno ai 220 ms.

Nel complesso, lo studio ancora una volta dimostra che l'integrazione cross-modale degli stimoli emotivi audio-visivi avviene nel corso del processamento percettivo (intorno ai 220 ms nella zona posteriore) e non ad uno stadio decisionale più avanzato. Pourtois e colleghi suggeriscono che già intorno ai 100 ms, nella fase *percettiva*, avvenga un incremento in termini di ampiezza delle componenti modalità-specifiche (de Gelder et al., 1999; Giard & Peronnet, 1999; Pourtois et al., 2000), successivamente seguito dalla comparsa di altre componenti, come la P2b, sensibili al *contenuto* dello stimolo audio-visivo. Solo in un secondo momento le informazioni avrebbero accesso a stadi cognitivi più avanzati di natura *decisionale*.

8. Gli studi sui casi clinici

Un contributo fondamentale per la comprensione del decoding intersensoriale delle emozioni giunge dagli studi condotti su casi clinici, che ne mettono in risalto l'importante funzione adattiva: ad esempio, in un recente studio condotto su un paziente con grave deficit della localizzazione uditiva dovuto a lesione, Bolognini, Rasi e Ladavas (2005) hanno dimostrato che la contemporanea comparsa di uno stimolo visivo che viene presentato nella medesima posizione spaziale dello stimolo target uditivo migliora fortemente la localizzazione del suono.

Alcuni studi clinici si sono rivelati utili anche per chiarire le funzioni delle strutture corticali implicate nel processamento cross-modale. Tra di essi, uno studio condotto con la risonanza magnetica da Taylor e Brugger (2005) sul caso di un paziente affetto da sclerosi multipla e vittima di allucinazioni audio-visive ha portato gli autori ad ipotizzare che tali allucinazioni fossero legate ad un deficit nella regolazione dell'attività di integrazione cross-modale localizzata nel collicolo superiore e nel solco temporale superiore.

8.1. Il fenomeno del blindsight

Gli studi condotti su casi clinici si sono in particolar modo rivelati utili per indagare il ruolo della consapevolezza nell'integrazione intersensoriale delle

informazioni emotive di natura uditiva e visiva. A tal proposito, de Gelder e colleghi (de Gelder, Pourtois & Weiskrantz, 2002) hanno sottoposto ad uno dei classici esperimenti sul processamento cross-modale delle emozioni due pazienti che presentavano blindsight (emianopia) unilaterale. Essi, a causa di una lesione alla corteccia striata (V1), erano in grado di discriminare le espressioni del volto ma senza essere consapevoli di percepirle. Il processamento cosciente degli stimoli emotivi, che è di tipo cortico-corticale, coinvolge, oltre alla corteccia V1, la corteccia fusiforme e quella orbitofrontale, oltre a provocare un aumento dell'attivazione dell'amigdala destra. La percezione non consapevole coinvolge invece l'amigdala sinistra (Morris, Öhman & Dolan, 1998), il pulvinar e il collicolo superiore (Morris, Öhman & Dolan, 1999), che sono implicati in un circuito sottocorticale di elaborazione delle espressioni facciali. Queste strutture, nei due pazienti che hanno partecipato all'esperimento, erano intatte, permettendo così che venisse conservata l'elaborazione implicita degli stimoli emotivi. Finora il fenomeno del blindsight affettivo era stato studiato soltanto utilizzando come stimolo le espressioni facciali. De Gelder e colleghi hanno introdotto un secondo tipo di stimolo visivo, mostrando ai pazienti anche delle scene a contenuto emotivo. Nel presente esperimento, gli autori si sono chiesti se l'integrazione audiovisiva in soggetti con blindsight emotivo avvenga solo in presenza di abbinamenti naturali (volto della paura - voce della paura), o anche in presenza di abbinamenti semantici (immagine paurosa - voce della paura). Gli autori hanno ipotizzato che, se quest'ultimo caso si verifica, bisogna supporre che i circuiti sottocorticali compensino l'assenza di quelli corticali; nel caso in cui invece l'integrazione non abbia luogo, allora bisogna ipotizzare che la percezione cross-modale delle coppie immagine - voce richieda necessariamente l'intervento di circuiti di ordine superiore, deputati all'elaborazione delle proprietà semantiche che essi condividono. Propendendo per la seconda possibilità, gli autori hanno ipotizzato nello specifico che, nel caso delle coppie naturali, si verificasse un decremento dell'ampiezza di N1 in presenza di coppie incongruenti e che, nel caso delle coppie semantiche, tale decremento si verificasse solo quando gli stimoli erano presentati all'emisfero intatto e quindi processati consapevolmente. Essi quindi, con il loro

esperimento, si aspettavano che la presentazione di un'immagine all'emisfero danneggiato (assenza di consapevolezza) non potesse interferire con il processamento uditivo e che quindi non ci fosse integrazione, per il fatto che il circuito subcorticale in questo caso non sarebbe sufficiente. Dai risultati emerge che, in assenza di percezione consapevole (presentazione nel campo visivo danneggiato), la presentazione dello stimolo visivo influenza il processamento della voce solo nel caso in cui lo stimolo visivo sia costituito da un volto, come suggerito dall'analisi dell'indice ERP N1, che rivela un decremento nelle coppie incongruenti. L'effetto della percezione visiva su N1, indipendentemente dal lato della lesione, presenta una lateralizzazione, essendo maggiormente evidente nell'emisfero destro. Ciò è in linea con il dato che soprattutto tale emisfero è implicato nell'elaborazione delle componenti prosodiche (Ross, 2000). Per spiegare i risultati, viene ipotizzato che quando la corteccia visiva primaria è danneggiata, alcune strutture che ricevono afferenze dirette dalla retina, come il collicolo superiore e il pulvinar, possano compensare fino ad un certo punto la mancata attività di V1, e che tuttavia non possano compensare l'assenza di alcune proiezioni di feedback che mettono in collegamento V1 e aree corticali anteriori garantendo la percezione combinata audio-visiva (Lamme, 2001). La percezione congiunta di stimoli affettivi uditivi e visivi abbinati sulla base del contenuto semantico richiede l'intervento di circuiti corticali deputati all'elaborazione semantica che coinvolgono V1 così come aree corticali anteriori di alto livello. Questo suggerisce che invece, nel caso dell'abbinamento volto-voce, l'intervento dei circuiti corticali non sia del tutto cruciale ai fini dell'integrazione. Gli autori concludono che il riconoscimento delle emozioni a partire dal volto, o dalla voce o dall'integrazione tra i due può avvenire bypassando la coscienza e che ciò, molto probabilmente, è dovuto alla rilevanza che le emozioni rivestono da un punto di vista adattivo.

8.2. La prosopagnosia

Un risultato simile (de Gelder et al., 2000) è stato ottenuto indagando l'integrazione cross-modale di stimoli emotivi in una paziente con prosopagnosia

dovuta a lesione bilaterale dei lobi occipitali, del tutto incapace di riconoscere consapevolmente, sulla base del volto, l'identità e l'espressione emotiva. La paziente non presentava invece problemi nel riconoscere le emozioni espresse attraverso il tono di voce. La prosopagnosia consiste infatti nella compromissione della capacità di riconoscere i volti e di identificarne l'identità e l'espressione (Tranel, Damasio & Damasio, 1995). Tuttavia studi che hanno utilizzato metodi elettrofisiologici come la rilevazione della conduttanza cutanea (Tranel & Damasio, 1987) o come la registrazione dei potenziali evocati (Renault et al., 1989) hanno evidenziato come i pazienti affetti da prosopagnosia siano in grado di riconoscere in modo implicito e latente l'identità, a partire dall'osservazione del volto. Nessuno studio precedente aveva invece studiato l'esistenza di un riconoscimento latente dell'espressione del volto. La ricerca ha indagato il riconoscimento implicito dell'espressione emotiva andando a verificare se, presentando contemporaneamente un'espressione facciale e una parola pronunciata in tono emotivamente connotato, esistesse un'interferenza tra le due modalità sensoriali. Gli stimoli utilizzati a questo scopo esprimevano gioia o tristezza ed erano abbinati in coppie congruenti o incongruenti. In una prima fase, volta ad indagare l'effetto del tono di voce sul riconoscimento del volto, la paziente veniva invitata a ignorare la voce e di identificare il volto come felice o triste. A differenza di quanto rilevato sui soggetti normali, che presentavano un effetto di interazione tra volto e voce, il giudizio della paziente, nonostante la consegna, era interamente basato sul tono di voce. Un risultato molto diverso è stato ottenuto nella seconda fase dell'esperimento, che esplorava l'effetto del volto sul riconoscimento della voce e in cui la paziente veniva invitata a ignorare il volto e ad etichettare la voce come felice o triste. In questo caso è emerso un effetto cross-modale, tale per cui l'espressione del volto aveva un sistematico impatto sulla valutazione del tono di voce. Nel complesso, i risultati indicano l'esistenza di un riconoscimento implicito dell'espressione del volto. Inoltre, contribuiscono a validare l'ipotesi della de Gelder che il processo di integrazione cross-modale abbia carattere obbligato e che avvenga in una fase percettiva precoce: il fatto che la paziente non percepisca consapevolmente il volto esclude infatti che l'integrazione avvenga ad uno stadio

cognitivo decisionale. Diverse ipotesi sono state formulate per spiegare il riconoscimento latente: in primo luogo, esso potrebbe scaturire da una forma di rappresentazione degradata, impoverita e che quindi non ha la possibilità di essere concettualizzata a livello cosciente (Farah, O'Reilly & Vecera, 1993); inoltre, tale riconoscimento latente potrebbe indicare l'esistenza di due sistemi distinti di processamento del volto, l'uno ventrale dedicato alle rappresentazioni manifeste e l'altro dorsale dedicato a quelle latenti (Bauer, 1984). A questo proposito, è significativo il fatto che la paziente presentava una compromissione della via ventrale (occipitotemporale) a fronte della conservazione di quella dorsale. Infine, è stata ipotizzata l'esistenza di due diversi tipi di processamento, qualitativamente differenti e corrispondenti all'elaborazione implicita ed esplicita dei volti. Nella paziente sarebbe conservato solo il primo tipo di elaborazione, che tuttavia non prevede l'accesso alla consapevolezza. In effetti, recenti studi indicano che gran parte del processamento delle emozioni (LeDoux, 1996) e, nello specifico, delle espressioni facciali (Morris, Öhman & Dolan, 1998) avviene al di fuori della consapevolezza. Sarebbero necessarie ulteriori indagini per capire quale di queste ipotesi meglio rende conto del fenomeno osservato.

8.3. Sistemi multipli di decodifica delle emozioni

Recentemente, è stata proposta l'ipotesi che esistano diversi sistemi indipendenti di riconoscimento delle emozioni, che si differenziano sia per il tipo di modalità sensoriale implicata (visiva, uditiva o audio-visiva) sia - nel caso della modalità visiva - per la natura dello stimolo (dinamico o statico). Oggi si tende ad evidenziare il contributo di strutture bilaterali come l'amigdala, il giro cingolato e i gangli basali, oltre alla corteccia prefrontale nel processamento delle emozioni (Adolphs, 2002; Phillips et al., 2003). Diversi lavori hanno tuttavia messo in luce anche l'esistenza di una specializzazione emisferica (Borod, 1993; Tranel et al., 2002). Inoltre, un importante ruolo è ricoperto dalla corteccia somatosensoriale destra, che sembra essere fondamentale nella comprensione delle espressioni emotive facciali, perchè permette al soggetto di accedere alle qualità dell'espressione osservata "come

se" fosse la propria (Adolphs et al., 2003). Nei pazienti neurologici, il deficit nel processamento delle espressioni emotive coinvolge soprattutto specifiche categorie di emozioni, più frequentemente quelle negative, come paura, disgusto e tristezza. Proprio questa osservazione ha suggerito la possibilità che esistano diversi sistemi specializzati di processamento (Adolphs & Tranel, 2004).

Un interessante studio condotto da McDonald e Saunders (2005) su pazienti con severo danno cerebrale traumatico (*traumatic brain injury*, TBI) ha portato supporto a questa ipotesi. I danni cerebrali traumatici consistono in ampie lesioni delle aree frontali e temporali, oltre che delle strutture limbiche e di altre strutture ad esse associate. Possono comportare la disconnessione tra le strutture limbiche e le aree somatosensoriali, disconnessione che spesso è causa di deficit nel riconoscimento delle emozioni altrui (Green et al., 2004). In effetti, la maggior parte dei pazienti con TBI presenta evidenti difficoltà quando viene chiesto di decodificare le emozioni sulla base delle diverse modalità sensoriali (McDonald & Flanagan, 2004). In particolare, la lesione dei lobi frontali e parietali e delle strutture limbiche causa delle difficoltà nel riconoscimento delle espressioni sia facciali sia vocali delle emozioni (Adolphs, 2002). Bisogna poi osservare che spesso questi pazienti presentano deficit maggiori quando viene loro chiesto di riconoscere espressioni facciali statiche anziché dinamiche. Questo dato ha portato Adolphs (Adolphs et al., 2003) ad ipotizzare che esistano due differenti processi e che, in particolare, le espressioni statiche richiedano il contributo del sistema limbico e della corteccia prefrontale associata, a differenza delle espressioni dinamiche che sarebbero invece processate a livello della corteccia parietale.

McDonald e Saunders, sottoponendo i pazienti ad un compito di riconoscimento delle emozioni, hanno rilevato che essi presentavano una competenza deficitaria nel decoding delle emozioni, ma con alcune interessanti specificità. In particolare, gli stimoli facciali dinamici erano normalmente riconosciuti, a differenza degli stimoli facciali statici, degli stimoli emotivi di natura uditiva e, soprattutto degli stimoli audio-visivi, riconosciuti in modo altamente deficitario. Ciò supporterebbe l'ipotesi di sistemi di processamento indipendenti e

qualitativamente distinti. Nello specifico, gli autori suggeriscono che l'elaborazione degli stimoli visivi dinamici sia localizzata principalmente nelle aree parietali (Adolphs et al., 2003), non compromesse dalla lesione, e che invece quella degli stimoli facciali statici avvenga ad opera delle aree fronto-temporali danneggiate e delle strutture limbiche ad esse correlate. Rispetto agli stimoli uditivi, è possibile che il deficit sia dovuto al fatto che i pazienti elaborano il contenuto linguistico, a scapito dell'espressione emotiva. In effetti, tali pazienti normalmente tendono ad interpretare gli enunciati in modo molto letterale, tralasciando di effettuare delle inferenze (McDonald & Flanagan, 2004). Infine, per quanto riguarda gli stimoli audio-visivi, è possibile che, nonostante la compresenza di entrambe le modalità, i pazienti si focalizzino su una sola di esse e che non utilizzino le strategie di processamento normalmente impiegate nel riconoscimento delle informazioni emotive bimodali.

PARTE I

LO STATO ATTUALE DELLA RICERCA

LA RICERCA: GLI OBIETTIVI E LA METODOLOGIA

1. Introduzione

Nel capitolo precedente, la panoramica dello stato attuale della ricerca nel campo della decodifica cross-modale delle emozioni ha evidenziato numerosi punti di interesse; essa, tuttavia, ha anche permesso di rilevare che molti aspetti di questo complesso processo sono tuttora inesplorati o richiedono ulteriori approfondimenti sperimentali. La presente ricerca si propone per l'appunto di portare ulteriore luce su alcuni specifici aspetti dell'integrazione multimodale di stimoli emotivamente connotati che, per diversi motivi, sono stati fino ad oggi ignorati o scarsamente approfonditi.

A tale scopo, si è scelto di affiancare alla raccolta di dati comportamentali la registrazione dei *potenziali evocati corticali*, dal momento che questa metodologia di indagine, garantendo un'alta definizione temporale dei fenomeni, si è rilevata particolarmente efficace nello studio dei fenomeni integrativi. Nello specifico, tale procedura consente di analizzare le differenze funzionali nei processi cognitivi rilevando le variazioni elettriche dello scalpo in risposta a specifici stimoli. Ciò avviene garantendo un'alta definizione temporale e conservando l'ordine sequenziale di elaborazione dello stimolo, per cui è possibile indagare l'evoluzione dell'intero processo. Vengono in questo modo rilevati sia i potenziali corticali percettivi o esogeni, che hanno una latenza breve e che si riferiscono all'elaborazione puramente sensoriale e strutturale dello stimolo, sia i potenziali di lunga latenza detti endogeni o cognitivi che invece rappresentano i successivi e più complessi stadi di processamento cognitivo e semantico (Rugg & Coles, 1995).

Nel presente capitolo, verranno quindi in primo luogo esposti gli obiettivi dello studio, insieme alle relative ipotesi sperimentali. Successivamente, saranno descritti il campione, gli stimoli sperimentali impiegati e la loro preparazione,

nonché la procedura sperimentale e la procedura di registrazione dei potenziali evocati adottata.

2. Obiettivi e ipotesi

Il presente studio è stato condotto allo scopo di analizzare i processi implicati nella decodifica bimodale audio-visiva di stimoli a contenuto emotivo, derivanti dalla combinazione tra la modalità vocale (componenti soprasegmentali di un breve stimolo sintagmatico a contenuto emotivamente neutro) e quella visiva (espressioni facciali). Nello specifico:

1. In primo luogo, al fine di integrare i dati sperimentali forniti da quegli studi che si sono limitati ad indagare la fase meramente percettiva dell'elaborazione multimodale, si è voluta verificare la possibile presenza di effetti di natura *cognitiva a media e lunga latenza* nella integrazione cross-modale delle emozioni. In altri termini, ci si attende di rilevare, oltre alla presenza di indici precoci di natura puramente percettiva (Bentin & Deouell, 2000; Linkenkaer et al., 1998), anche la presenza di variazioni ERP di media e lunga latenza che rendano conto dell'integrazione cognitiva, quale risultato della sintesi di informazioni intermodali di natura emotiva. In particolare, in linea con la letteratura esistente, ipotizziamo che, tra i 150 e i 250 ms il tracciato evidenzi la presenza di un indice specifico legato alla decodifica del *contenuto emotivo* (Sato et al., 2001; Streit et al., 2000). Inoltre, nella medesima finestra temporale, ci attendiamo di riscontrare la presenza di un marker che indichi la presenza di processi specificamente dedicati all'*integrazione intersensoriale audiovisiva emotiva* (Pourtois et al., 2002). E' inoltre plausibile che, in una fase più tardiva dell'elaborazione entrino in gioco, a seguito della rilevazione delle anomalie semantiche, fattori di natura più propriamente *cognitiva e decisionale*, normalmente rivelati dalla presenza della deflessione N400 (Balconi, 2005);
2. In secondo luogo, si è inteso verificare la presenza di differenti profili ERP per le due condizioni di *congruenza* e *incongruenza* emotiva. In particolare, si suppone che anche a livello cognitivo - scarsamente indagato dalle precedenti ricerche -

possano essere rilevate differenze dei “marcatori di integrazione” ERP (in termini di *ampiezza* o di *latenza*) in caso di convergenza delle informazioni emotive o di divergenza delle stesse (Pourtois et al., 2002). Inoltre, ci si è proposti di definire un eventuale *effetto di localizzazione* dei marker suddetti in relazione alla *condizione sperimentale*. Una distribuzione eterogenea delle variazioni di picco in funzione della convergenza o divergenza delle informazioni emotive potrebbe indurre a ipotizzare la presenza di moduli corticali multipli in relazione alle diverse condizioni sperimentali. Più nello specifico, ci attendiamo che soltanto quegli indici ERP che sono direttamente implicati nei processi di convergenza manifestino una sensibilità rispetto alla congruenza/incongruenza audiovisiva, con valori di maggiore ampiezza per la condizione di congruenza (Calvert, Brammer & Iversen, 2000; Raij, Uutela & Hari, 2000).;

3. Utilizzando, rispetto alla maggioranza degli studi condotti fino ad ora, un *range più ampio di componenti emotive* (sei patterns emotivi) è possibile analizzare in modo più approfondito l’incidenza del contenuto delle singole emozioni nel processo di integrazione. La domanda che occorre porsi è relativa ad una possibile differente sensibilità alla congruenza e all’incongruenza intersensoriale in relazione al tipo di emozione espressa. Una seconda fase di analisi dei dati (*emotion-by-emotion*) ha pertanto l’obiettivo di definire l’effetto delle configurazioni emotive specifiche sul processo di integrazione cross-modale. In particolare, è presumibile che venga rilevata una differenziazione delle emozioni che rende conto delle dimensioni di *valenza edonica* ed *arousal* (Pourtois et al., 2005);
4. Inoltre, integrando lo sforzo di quelle ricerche che si sono poste l’obiettivo di comprendere se il processo di integrazione intersensoriale abbia una natura automatica e involontaria o sia il risultato di un processo decisionale che debba essere attivato in modo premeditato, è stata introdotta una variabile relativa alla consegna sperimentale, mettendo a confronto una situazione in cui viene esplicitamente richiesto di elaborare la congruenza/incongruenza tra gli stimoli (*task esplicito*) ed una situazione in cui invece non viene data alcuna indicazione in proposito (*task implicito*). Se in letteratura sono presenti numerose prove del fatto

che a livello percettivo precoce il processo ha carattere *automatico* (de Gelder & Vroomen, 2000), si può al contrario ipotizzare che a livello degli stadi cognitivi più tardivi la consegna possa influenzare il processamento, con un rallentamento dei tempi di latenza nel caso in cui l'attenzione del soggetto sia intenzionalmente indirizzata verso l'elaborazione della convergenza vs. divergenza delle informazioni, implicando un maggiore impiego di *risorse attentive*.

5. Infine, rispetto ai *tempi di risposta comportamentale*, in precedenza è stato rilevato un effetto di facilitazione, con riduzione dei tempi di risposta (TR), per stimoli congruenti rispetto a stimoli incongruenti (Pourtois et al., 2002). Intendiamo approfondire lo studio ponendo come oggetto di analisi l'*incidenza dei singoli patterns emotivi* sulle variazioni dei TR. A tale proposito, ci si attende che anche a livello comportamentale le proprietà edoniche e di arousal dello stimolo abbiano specifici effetti sui processi di decodifica. In particolare, possiamo aspettarci che le emozioni negative ad alto arousal, data la pregnanza del loro significato adattivo, producano le risposte più rapide.

3. Metodo

3.1. Partecipanti

30 soggetti hanno preso parte all'esperimento (età: $M = 24.4$; $DS = 3.4$; 9 maschi e 21 femmine). Essi hanno fornito il proprio consenso a partecipare come volontari ad uno studio volto ad indagare la percezione combinata di stimoli visivi e uditivi e condotto presso il Laboratorio di Psicologia Cognitiva dell'Università Cattolica di Milano. Tutti i soggetti sono studenti universitari iscritti all'Università Cattolica di Milano.

L'intero campione risultava normodotato rispetto alle capacità visive e uditive. Tutti erano destrimani.

3.2. Stimoli

a) Componente vocale

Per la produzione della componente vocale si è optato per la costruzione di un unico *stimolo sintagmatico breve* (composto da una sola sillaba) *a contenuto emotivamente neutro* (“ora”)¹. Gli stimoli sono stati riprodotti nella loro valenza emotiva da un locutore maschio (un attore teatrale professionista), cui è stato esplicitamente richiesto di codificare le diverse espressioni emotive utilizzando le componenti vocali, veicolanti differenti correlati emotivi grazie alle variazioni del profilo di intonazione (rispettivamente *rabbia, paura, sorpresa, gioia, disgusto, tristezza* ed una condizione di controllo rappresentata da un’espressione neutra) (Portois et al., 2002). Di conseguenza, il tono, l’intensità e la velocità dell’eloquio sono stati modificati in funzione delle singole emozioni (Scherer, 1986). Al fine di garantire una corretta modificazione dei parametri vocali per ciascuna emozione, è stata prevista la lettura di un breve script emotivo che consentisse l’immedesimazione del locutore con la situazione elicitante. La registrazione è stata effettuata utilizzando il programma Sonic Foundry Acid Music. Rispetto ai parametri temporali, la durata degli stimoli vocali varia dai 300 agli 850 ms (M = 550 ms; DS = 223 ms; range 300-850 ms).

Il valore semantico e la pertinenza degli stimoli emotivi sono stati verificati da un campione di soggetti (6 studenti universitari) in una fase pre-sperimentale, mediante un apposito *questionario di validazione* (vedi allegato 1). Un primo item verteva sul corretto etichettamento delle singole emozioni (con metodo *free labeling*), che ha fatto rilevare un riconoscimento adeguato per ciascuna di esse (disgusto 80.3%; gioia 100.0%; neutro 83.3%; paura 100.0%; rabbia 83.3%; sorpresa 100.0%; tristezza 76.7%). Successivamente è stata testata la consistenza di giudizio del valore emotivo degli stimoli vocali (scala Likert a 5 passi), chiedendo al soggetto quanto era sicuro che lo stimolo udito esprimesse l’emozione da lui/lei indicata. Non si

¹ Si è scelto un solo sintagma al fine di poter manipolare con maggiore facilità i parametri vocali senza che vi fosse interferenza delle componenti di pronuncia dell’eloquio (profilo di intonazione dell’enunciato, variazioni dei parametri temporali) interpolanti le variazioni vocali indotte dalle emozioni. In secondo luogo, ciò ha consentito di non far convergere l’attenzione dei soggetti principalmente sulla componente vocale rispetto a quella mimica.

osservano differenze significative di giudizio dei soggetti per le singole emozioni, poiché tutti i patterns emotivi sono stati considerati adeguati per esprimere le emozioni (rispettivamente per disgusto $M = 4.5$; $DS = 1.6$; gioia $M = 3.7$; $DS = 1.4$; neutro $M = 3.5$; $DS = 1.4$; paura $M = 4.2$; $DS = 1.3$; rabbia $M = 4.7$; $DS = 0.5$; sorpresa $M = 4.0$; $DS = 1.1$; tristezza $M = 3.8$; $DS = 1.2$).

b) *Componente mimica*

Il materiale-stimolo mimico (*patterns facciali*) è costituito, come si può notare in figura 1, da stimoli visivi in bianco e nero (11 x 15 cm) raffiguranti il volto di un giovane attore di sesso maschile (Ekman & Friesen 1976), per le medesime emozioni prodotte in condizione vocale. Per una migliore visualizzazione degli stimoli si rimanda all'allegato 2.



Fig. 1. Gli stimoli visivi: patterns facciali esprimenti rispettivamente sorpresa, rabbia, gioia, tristezza, paura, disgusto.

Anche in questo secondo caso, la pertinenza emotiva degli stimoli è stata testata in una fase pre-sperimentale da un campione di 10 soggetti (vedi allegato 1). Sempre attraverso il questionario di validazione, la correttezza di giudizio è stata verificata mediante le due scale (etichettamento; consistenza di giudizio), con dati simili a quanto rilevato per gli stimoli visivi. L'etichettamento ha fatto rilevare percentuali di identificazione elevata per ciascun correlato (disgusto 78.3%; gioia 95.1%; neutro 83.9%; paura 98.0%; rabbia 83.4%; sorpresa 98.0%; tristezza 75.7%); parallelamente, i soggetti hanno espresso una elevata consonanza di giudizio per le singole emozioni (rispettivamente per disgusto $M = 4.3$; $DS = 1.4$; gioia $M = 3.8$; $DS =$

1.5; neutro M = 3.9; DS = 1.8; paura M = 4.9; DS = 1.2; rabbia M = 4.4; DS = 0.9; sorpresa M = 4.1; DS = 1.5; tristezza M = 4.2; DS = 1.1).

c) *Coppie bimodali audio-visive*

A partire dalle due batterie di stimoli mimici e vocali, sono state composte coppie congruenti (volto e voce esprimenti la medesima emozione) (*condizione sperimentale di congruenza*) e coppie incongruenti (volto e voce esprimenti due diverse emozioni) (*condizione sperimentale di incongruenza*). Per la condizione di incongruenza è stato predisposto che ciascuna emozione fosse combinata con tutte le altre (per un totale di 5 coppie incongruenti per ogni emozione). Per ciascuna emozione sono state presentate dieci reiterazioni nella condizione di congruenza e dieci nella condizione di incongruenza (per un totale di 60 coppie congruenti e 60 incongruenti). La coppia con stimolo neutro è stata riprodotta anch'essa per 20 volte, sempre in condizione di congruenza. L'ordine delle coppie di stimoli congruenti e incongruenti è stato opportunamente randomizzato e nella costruzione della sequenza si è inoltre proceduto a verificare che non vi fossero coppie identiche poste consecutivamente. Successivamente, si è proceduto alla sincronizzazione degli stimoli visivi con quelli uditivi. Infatti, mediante il sistema di rilevazione temporale della coincidenza dei tempi di onset dei due stimoli di partenza, è stata garantita la perfetta sincronia intersensoriale.

3.3. *Procedura sperimentale*

Fase pre-sperimentale: Per garantire la familiarizzazione dei soggetti con la procedura sperimentale e per chiarire eventuali dubbi, è stata prevista una sessione precedente all'esperimento, che includeva la presentazione di una sequenza di dieci coppie di stimoli (5 congruenti e 5 incongruenti) identiche a quelle presentate nella successiva fase sperimentale.

Fase sperimentale: L'esperimento è stato condotto in una stanza appositamente oscurata e insonorizzata, al fine di garantire una registrazione ottimale del tracciato

EEG. I soggetti sono stati collocati di fronte ad un monitor (90 cm di distanza) collegato con un PC, al centro del cui schermo sono stati proiettati gli stimoli visivi (angolo visivo orizzontale sotteso 4° e verticale 6°) mediante il software di stimolazione STIM 4.2. La sequenza di stimoli uditivi è stata riprodotta con emissione sonora mediante due casse incorporate al PC, direzionate verso il soggetto e poste ad una distanza di c.ca 90 cm dal soggetto medesimo. Per ciascun soggetto è stata prevista una durata complessiva della stimolazione di 10 minuti circa.

Ai soggetti veniva comunicato che avrebbero partecipato ad un esperimento sulla percezione di stimoli da natura visiva e uditiva, anticipando loro che sarebbero stati presentati abbinati in coppie.

Consegna sperimentale: Tutti i soggetti venivano istruiti a prestare attenzione alle coppie di stimoli. Inoltre, la consegna prevedeva che il soggetto, dopo aver prestato attenzione a ciascuna coppia, premesse il pulsante sinistro del mouse. Questo ci ha permesso, con l'ausilio del software STIM 4.2., di registrare i tempi di risposta (TR).

A metà del campione, tuttavia, veniva richiesto soltanto questo compito attentivo, senza ulteriore indicazione relativa alle modalità di codifica dello stimolo, né allo specifico contenuto emotivo degli stimoli presentati (*task implicito*) (vedi allegato 3). All'altra metà dei soggetti, invece, veniva chiesto non soltanto di prestare attenzione agli stimoli, ma anche di cercare di comprendere se, all'interno di ciascuna coppia, essi erano congruenti o incongruenti (*task esplicito*) (vedi allegato 4).

Al soggetto inoltre era preventivamente richiesto di ridurre al minimo i movimenti corporei e quelli oculari, al fine di non alterare la registrazione dell'EEG.

Fase post-sperimentale: Infine, con l'intento di testare la corretta comprensione del contenuto emotivo degli stimoli presentati, al termine della fase di registrazione degli ERPs, ai soggetti sperimentali è stato richiesto di compilare un questionario analogo a quello pre-sperimentale di validazione (vedi allegato). Eccone i risultati:

a) Per gli stimoli visivi, correttezza dell'etichettamento emotivo (disgusto 76.7%; gioia 96.7%; neutro 93.3%; paura 93.1%; rabbia 80.0%; sorpresa 90.0%; tristezza

83.3%); consistenza di giudizio (disgusto M = 4.1; DS = 0.8; gioia M = 4.5; DS = 0.6; neutro M = 4.1; DS = 1.0; paura M = 4.7; DS = 0.7; rabbia M = 4.2; DS = 0.8; sorpresa M = 4.3; DS = 0.7; tristezza M = 3.8; DS = 0.9);

b) Per gli stimoli uditivi, correttezza dell'etichettamento emotivo (disgusto 78.0%; gioia 83.3%; neutro 80.0%; paura 93.3%; rabbia 76.7%; sorpresa 73.3%; tristezza 83.3%); consistenza di giudizio (disgusto M = 3.9; DS = 0.9; gioia M = 4.0; DS = 1.1; neutro M = 4.1; DS = 0.9; paura M = 4.1; DS = 0.8; rabbia M = 3.9; DS = 1.0; sorpresa M = 3.8; DS = 0.9; tristezza M = 4.4; DS = 0.6).

Al termine del questionario, veniva chiesto al soggetto se gli era sembrato che alcune coppie di stimoli fossero incongruenti. Tutti i soggetti (100%) hanno risposto positivamente. Inoltre, tutti (100%) hanno dichiarato di non aver trovato difficile o faticoso il compito.

3.4. Procedure di registrazione degli indici ERP

Montaggio della cuffia. Le variazioni EEG sono state rilevate mediante l'impiego di una cuffia dotata di elettrodi polarizzati (64 canali). Dei canali disponibili, solo 14 sono stati attivati per la registrazione, in funzione delle aree di maggiore interesse, rispettivamente tre frontali (F3, Fz, F4), due temporali (T3, T4), tre parietali (P3, Pz, P4) e tre occipitali (O1, Oz, O2) (vedi fig. 2). Due elettrodi collocati sul lobo dell'orecchio sono stati adottati come riferimento (sistema internazionale, Jasper, 1958). Per la rilevazione degli artefatti oculari, motori o di deriva degli elettrodi sono stati utilizzati due elettrodi collocati rispettivamente sopra e sotto l'occhio destro.

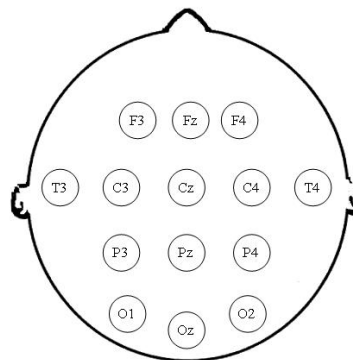


Fig. 2. Distribuzione degli elettrodi sullo scalpo.

Registrazione del segnale elettroencefalografico. Il tracciato EEG è stato campionato mediante il sistema di amplificazione *NeuroScan Synamps 4.2*. Il parametro di impedenza è stato mantenuto al di sotto del valore di 5 Ohm. La registrazione del segnale è compresa tra i 100 ms precedenti lo stimolo (pre-stimolo) e i 1900 ms successivi alla sua comparsa, per un totale di 2000 ms. La figura 3a mostra l'output grezzo della registrazione.

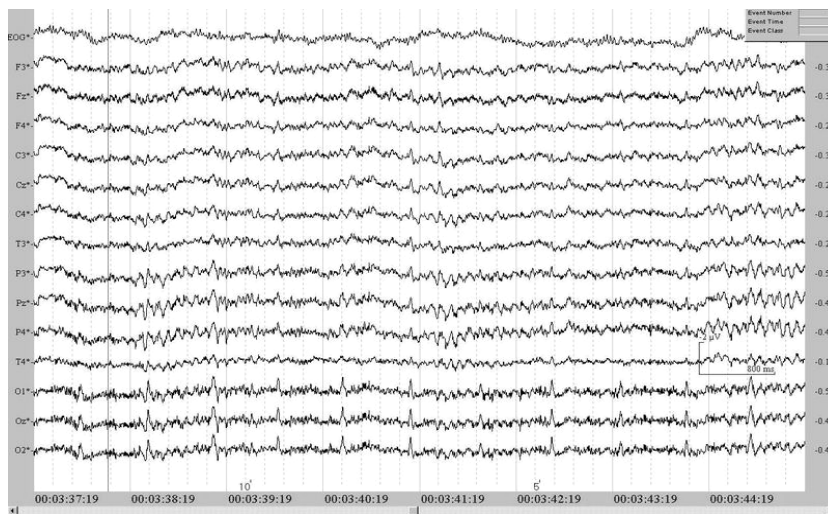
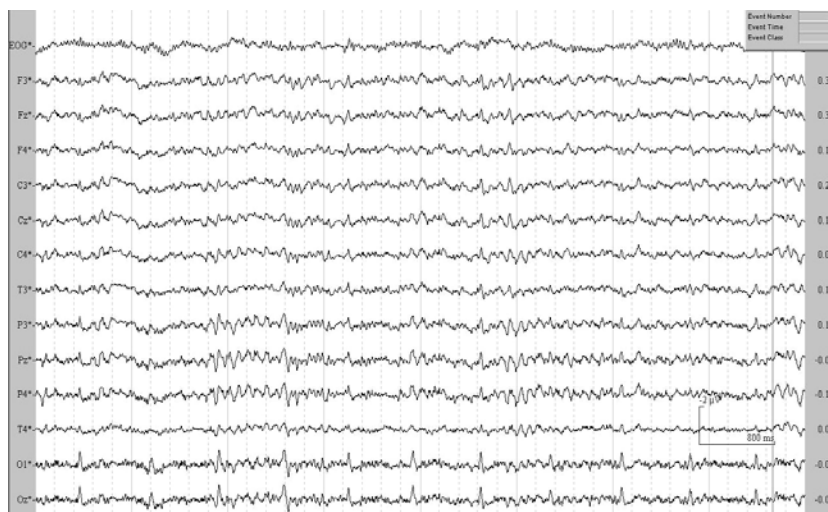


Fig. 3. Output della registrazione prima (a.) e dopo (b.) il filtraggio.

a.



b.

Analisi morfologica preliminare. Prima di procedere nel trattamento di dati, tutte le registrazioni sono state esaminate, allo scopo di eliminare quelle invalidate dalla presenza di artefatti motori o da una registrazione non buona del segnale. A seguito di tale analisi preliminare, sono stati inclusi nelle fasi successive di analisi solo i dati relativi a 30 soggetti (vedi par. 3.1).

Filtraggio. Per una prima reiezione degli artefatti, ci si è avvalsi di un filtraggio passa banda 0.01-60 Hz (frequenza di campionamento 250 Hz). In figura 3b è possibile osservare gli effetti del filtraggio sul tracciato.

Epocaggio. Attraverso un'apposita funzione di NeuroScan, il segnale registrato è stato ridotto in epoche, permettendo così di visualizzare, per ciascun soggetto, un profilo medio dell'onda per ogni elettrodo (vedi fig. 4).

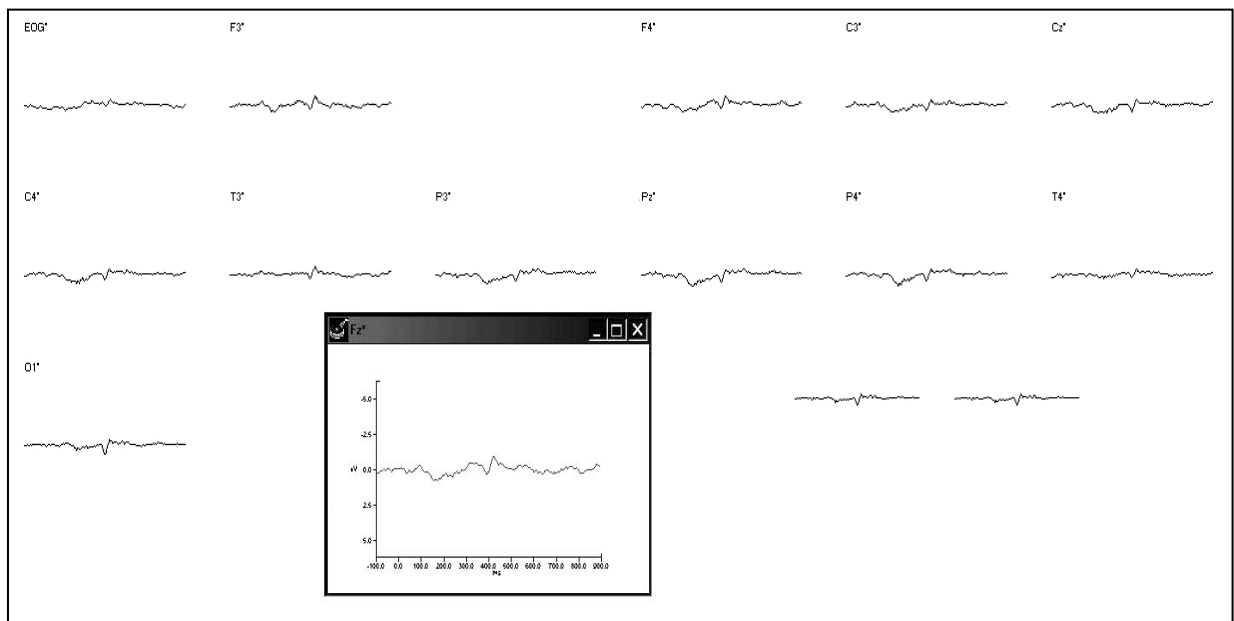


Fig. 4. Esempio di onde epocate.

Reiezione manuale degli artefatti. Ai trials rimanenti è stata applicata una procedura di filtraggio ulteriore mediante ispezione visiva delle singole epoche. Complessivamente, circa il 7% delle epoche è stato giudicato contaminato da artefatti e quindi escluso dall'analisi.

Calcolo delle medie. Per ciascun soggetto, e successivamente per l'intero campione sperimentale, sono stati calcolati i profili medi dell'onda per ciascun fattore sperimentale.

Estrazione e preparazione dei dati numerici. Si è proceduto ad estrarre i dati numerici salvandoli in formato .dat e quindi convertendoli in formato Excel. Tramite l'applicazione di macro sono stati ricavati i database, che sono stati successivamente sottoposti ad analisi statistica (vedi cap. 5).

ANALISI DEI DATI E RISULTATI

1. Introduzione

Nel presente capitolo, dedicato all'analisi dei dati e all'esposizione dei risultati, verrà in primo luogo esposta una descrizione di tipo morfologico dei tracciati ERP ottenuti. Successivamente, verranno presentate le analisi statiche effettuate, illustrandone i risultati. Lo studio ha impiegato due diversi ordini di dati; di conseguenza, una prima sezione sarà dedicata all'analisi dei dati di natura neuropsicologica, ovvero i potenziali evocati corticali (ERPs), e in un secondo momento verrà presentata l'analisi condotta sui dati comportamentali, i tempi di risposta (TR).

2. I tracciati ERP: descrizione morfologica

Osservando i tracciati derivati dalla media tra tutti i 14 elettrodi utilizzati, è possibile notare che la condizione di congruenza e quella di incongruenza presentano un andamento molto simile, caratterizzato dalla presenza di svariati indici ERP descritti in letteratura, nello specifico una P100, una N100, una N200 e una N300. Al di là di tale sostanziale sovrapponibilità, tuttavia i tracciati mostrano alcune differenze sia rispetto all'intensità dei fenomeni sia rispetto alla loro latenza (vedi fig. 1).

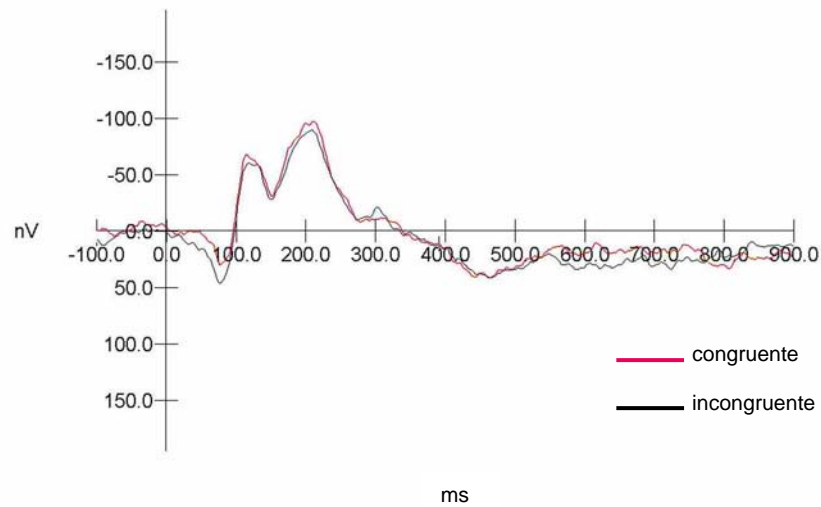


Fig. 1. Tracciati ERP derivati dalla media tra i 14 elettrodi e prodotti dalla presentazione di stimoli emotivi congruenti e incongruenti.

Bisogna inoltre sottolineare che, prendendo in considerazione separatamente i potenziali evocati registrati in corrispondenza delle diverse aree cerebrali, emergono sostanziali differenze.

Nell'area corticale frontale, come osservabile in figura 2, sono presenti una N100, una P100, una N200 e una N300.

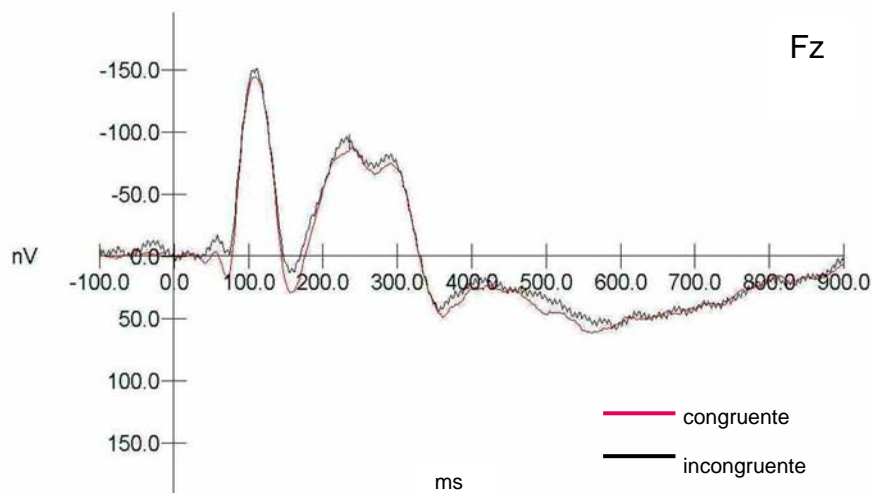


Fig. 2. Confronto tra condizione congruente e condizione incongruente nell'area frontale.

Nella regione centrale, rispetto a quella frontale, il segnale appare meno intenso. I picchi osservabili, se pure appunto meno accentuati, sono i medesimi: N100, una P100, una N200 e una N300 (vedi fig. 3).

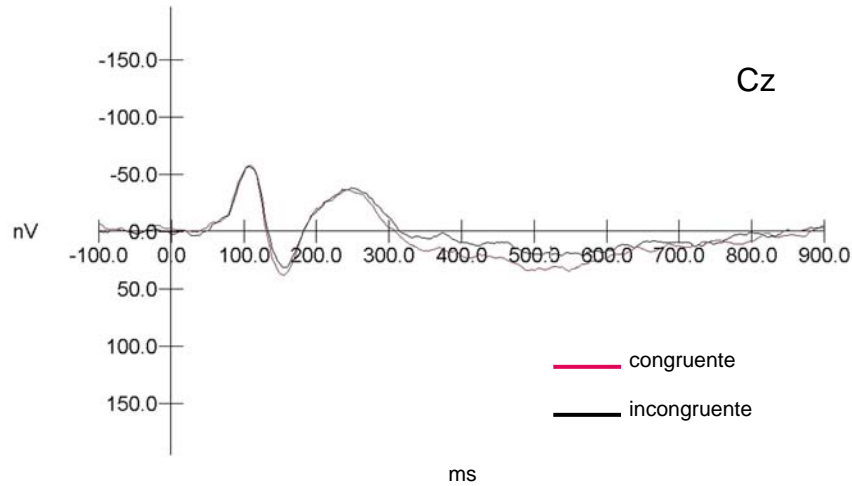


Fig. 3. Tracciati ERP in risposta alla presentazione di stimoli congruenti e incongruenti in corrispondenza della regione centrale.

I tracciati registrati nella parte posteriore della corteccia si differenziano nettamente da quelli rilevati nella zona anteriore.

Nell'area parietale il tracciato mostra infatti i seguenti fenomeni: una N100, una P100, una P200, una P300 e una N400 (vedi fig. 4).

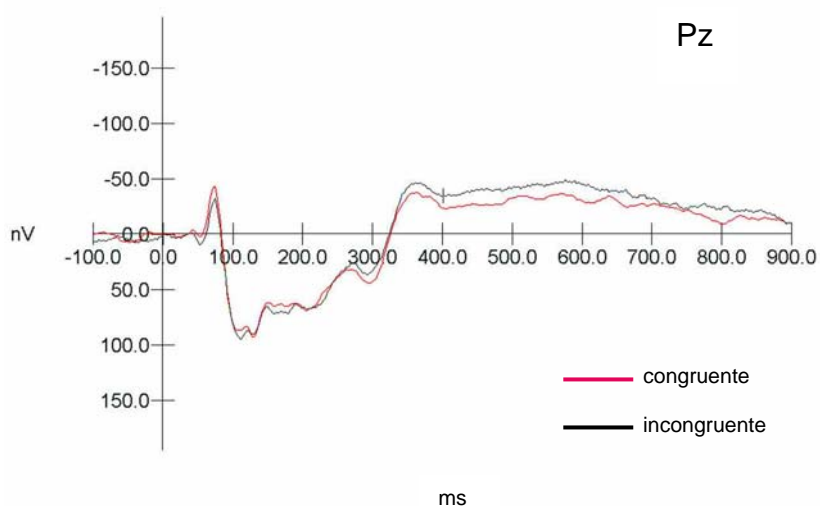


Fig. 4. Aree parietali: tracciati ERP in condizione di congruenza e di incongruenza.

Un andamento simile, con la presenza di una lieve N100, una P100, una P200, una P300 e una N400 si rileva in zona occipitale, come osservabile in figura 5.

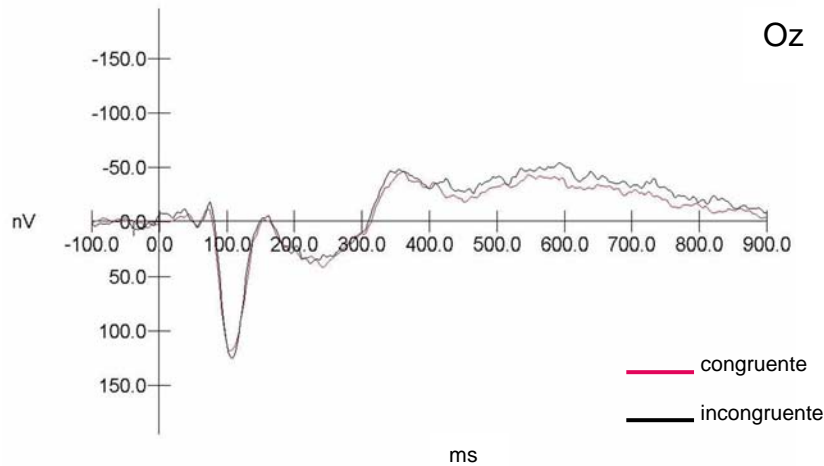


Fig. 5. Tracciati ERP osservabili in zona occipitale in seguito alla presentazione di coppie di stimoli congruenti e incongruenti.

3. I dati ERP: analisi statistica

Una prima fase di analisi dei dati di natura neuropsicologica è stata condotta mettendo a confronto gli effetti della congruenza/incongruenza delle coppie di stimoli, a prescindere dalla specifica emozione espressa da ciascuno di essi. Solo in un secondo momento, sono state poi prese in considerazione le singole emozioni, confrontandole tra loro sia per la condizione di congruenza sia per quella di incongruenza.

Dal momento che una preliminare analisi morfologica dei tracciati aveva fatto rilevare al loro interno la presenza di alcuni indici esplorati da precedenti studi, a partire dall'intero tracciato si è scelto di estrarre tramite l'applicazione di Macro in Excel i dati appartenenti a quattro finestre temporali e di analizzarli, (vedi tab. 1).

Finestra temporale	Indice ERP
0 - 150 ms	N100 P100
150 - 250 ms	N200 P200
250 - 350 ms	N300 P300
350 - 500 ms	N400

Tab. 1. Indici ERP indagati e relative finestre temporali da cui sono stati ricavati per condurre le analisi statistiche.

Per ciascun indice neuropsicologico, sono stati rilevati due parametri, utilizzati nelle analisi come variabili dipendenti:

- *intensità*: la massima intensità di espressione del picco rispetto al baseline, misurata in nanovolts (nV);
- *latenza*: tempo di comparsa del fenomeno, a partire dall'onset dello stimolo, misurata in millisecondi (ms).

La tabella 2 fornisce una descrizione delle variabili impiegate nelle analisi condotte su entrambi i parametri, illustrati nella tabella 3. Le ANOVA a misure ripetute sono state applicate in tutti i casi in cui l'ANOVA a disegno misto non abbia evidenziato alcuna influenza significativa del fattore task. Di conseguenza, nel presente capitolo, i risultati delle ANOVA a disegno misto verranno descritte soltanto nel caso in cui tale fattore sia risultato influente.

Variabile	Tipologia	Livelli
Condizione	within	3: congruente, incongruente, neutro
Elettrodo	within	14: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, T3, T4, O1, Oz, O2
Localizzazione	within	4: frontale (Fz), centrale (Cz), parietale (Pz), occipitale (Oz)
Lateralizzazione	within	2: destra (media di F4, C4, P4, O2), sinistra (media di F3, C3, P3, O1)
Task	between	2: esplicito, implicito

Tab. 2. Variabili indipendenti impiegate nelle analisi condotte sui parametri intensità e latenza.

Anova a disegno misto task x condizione x elettrodo
Anova a misure ripetute condizione x elettrodo
Anova a disegno misto task x condizione x localizzazione
Anova a misure ripetute condizione x localizzazione
Anova a disegno misto task x condizione x lateralizzazione
Anova a misure ripetute condizione x lateralizzazione

Tab. 3 Analisi statistiche condotte in riferimento all'intensità e alla latenza per ciascuna finestra temporale.

Nella descrizione delle analisi che segue verranno presi in considerazione solo i risultati significativi.

3.1. Confronto tra stimoli emotivi congruenti e incongruenti

3.1.1. Indice N100

L'indice N100 è stato ricavato estraendo all'interno della finestra temporale 0 - 150 ms il valore minimo. Dal momento che l'indagine di tale precoce indice di natura percettiva solo parzialmente rientra negli scopi del presente studio, ci si è limitati a rilevarne l'intensità, trascurando invece la latenza.

Intensità

E' stata fatta girare una ANOVA con i fattori condizione ed elettrodo.

Quest'analisi ha evidenziato un effetto significativo della variabile condizione ($F(2, N = 30) = 11,740, p < ,001$). Tuttavia, dall'analisi dei contrasti è emersa un'assenza di differenze significative tra gli stimoli congruenti ($M = -,128; ds = ,010$) e incongruenti ($M = -,121; ds = ,009$). Più interessante è invece il dato relativo agli elettrodi ($F(13, N = 30) = 8,210, p < ,001$), che fa registrare valori notevolmente più elevati nelle aree corticali frontali, come confermato anche dall'analisi dei contrasti.

Al fine di esplorare più nello specifico tale fenomeno, sono state applicate un'ANOVA con i fattori condizione e localizzazione e la relativa analisi dei contrasti.

Ancora una volta, le analisi hanno confermato la prevalente localizzazione frontale dell'indice N100 ($F(3, N = 30) = 6,544, p < ,001$) (vedi fig. 6).

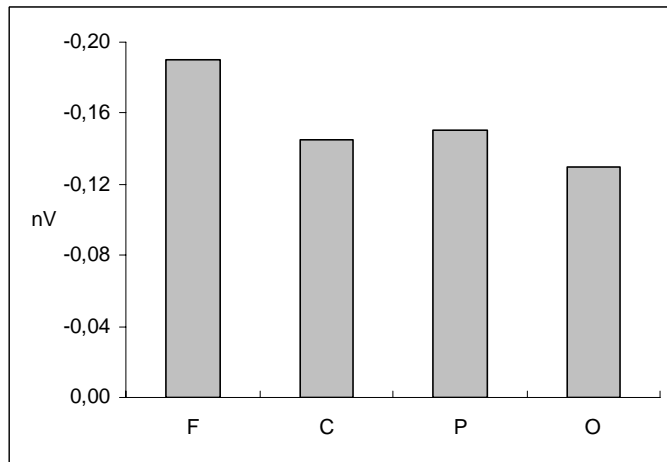


Fig. 6. Intensità di N100 nelle aree corticali frontale, centrale, parietale e occipitale.

Per concludere, abbiamo esplorato la possibile presenza di una lateralizzazione emisferica con un'ANOVA che ha utilizzato i fattori condizione e lateralizzazione, i cui risultati permettono di delineare il picco come un fenomeno prevalentemente lateralizzato nell'emisfero sinistro ($F(1, N = 30) = 4,549, p = ,42$), come osservabile in figura 7.

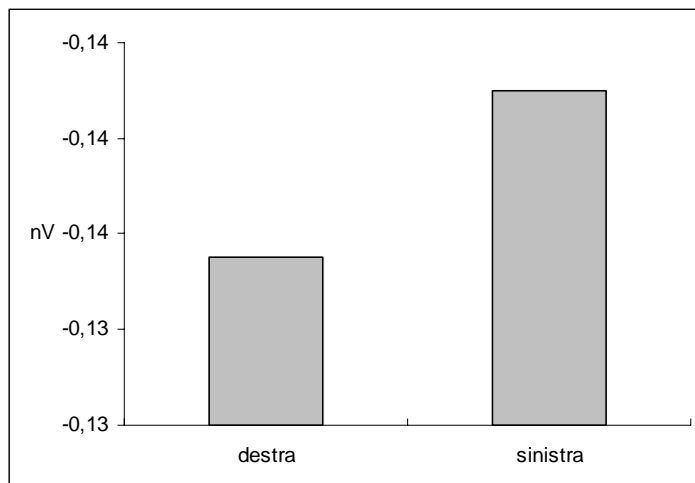


Fig. 7. Lateralizzazione sinistra del picco N100.

3.1.2. Indice P100

Estraendo il valore massimo tra quelli compresi nella finestra temporale 150 - 250 ms è stato ricavato l'indice P100, che è stato esplorato analizzandone

l'intensità. Dato che solo collateralmente tale indice cade nel focus della ricerca, le analisi relative alla latenza sono state tralasciate.

Intensità

Le analisi mostrano la presenza di una differente intensità di attivazione nelle diverse aree corticali ($F(1, N = 30) = 4,549, p = ,42$). Nello specifico, il picco risulta essere più ampio nella porzione posteriore della corteccia, come si rileva anche dai risultati dell'ANOVA che impiega i fattori condizione e localizzazione ($F(3, N = 30) = 15,258, p < ,001$) e dalla successiva analisi dei contrasti (vedi fig. 8).

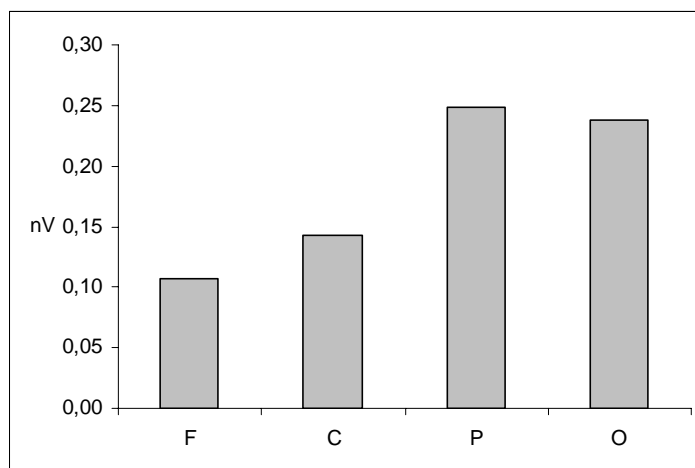


Fig. 8.
Localizzazione
del picco P100
rispetto ai lobi
frontali, centrali,
parietali e
occipitali.

Si rileva inoltre una tendenza alla significatività del fattore condizione ($F(2, N = 30) = 3,066, p = ,054$), sebbene l'analisi dei contrasti indichi che soltanto la condizione neutra presenta valori che si discostano dalle due condizioni sperimentali congruente ($M = ,151; ds = ,14$) e incongruente ($M = ,160; ds = ,11$). Infine, i dati mostrano come il fenomeno P100 sia maggiormente lateralizzato a destra ($F(1, N = 30) = 7,877, p = ,009$), come riportato in figura 9.

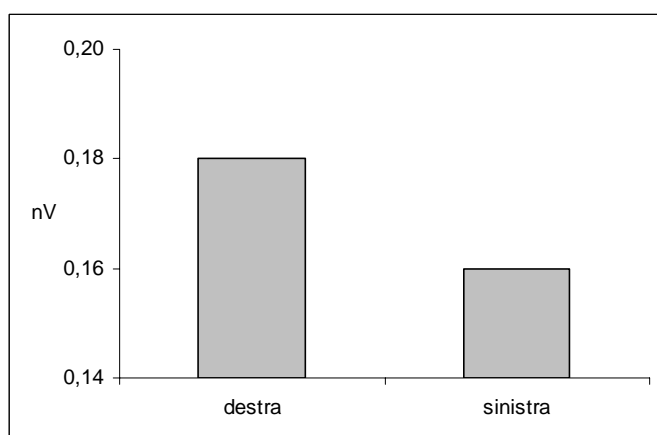


Fig. 9.
Lateralizzazione
destra
dell'indice P100.

3.1.3. Indice N200

L'indice N200 è stato ottenuto estraendo il valore minimo tra quelli compresi nella finestra temporale 150 - 250 ms. Si è quindi proceduto ad esplorarne l'intensità e la latenza.

Intensità

Tutti gli effetti esplorati dall'ANOVA a misure ripetute che ha utilizzato i fattori condizione ed elettrodo sono risultati significativi. In primo luogo, infatti, si è trovato che l'intensità del picco è influenzata dal tipo di stimolo ($F(2, N = 30) = 9,928$, $p < ,001$). Tuttavia, l'analisi dei contrasti non conferma la presenza di differenze significative per le due condizioni sperimentali congruente ($M = -,074$; $ds = ,013$) e incongruente ($M = -,071$; $ds = ,014$), ma solo per il neutro, che si discosta da esse. In secondo luogo, si rilevano delle differenze significative tra i diversi elettrodi ($F(13, N = 30) = 8,256$, $p < ,001$): come confermato dall'analisi dei contrasti, l'attivazione è massima nella zona frontale e, anche se in misura minore, nella zona centrale, per poi decrescere nelle aree posteriori.

Questa tendenza è stata riscontrata anche applicando un'ANOVA con i fattori condizione e localizzazione ($F(3, N = 30) = 11,460$, $p < ,001$), come è possibile osservare nella figura 10.

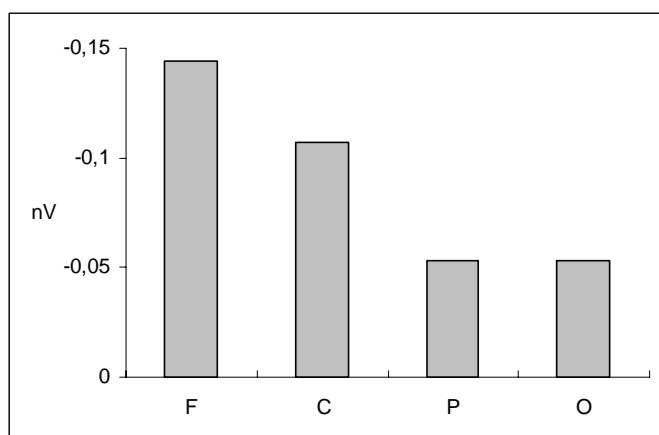


Fig. 10. Intensità del fenomeno N200 nelle aree frontali, centrali, parietali e occipitali.

Una più precisa caratterizzazione del picco è stata ottenuta esplorando l'effetto significativo di interazione condizione*localizzazione ($F(6, N = 30) = 8,106, p < ,001$): nelle aree frontali e centrali, infatti, l'intensità tende ad essere maggiormente accentuata in tutti i casi, ma soprattutto nella condizione di incongruenza; nelle aree posteriori, al contrario, sono gli stimoli incongruenti a produrre un picco maggiormente accentuato (vedi fig. 11).

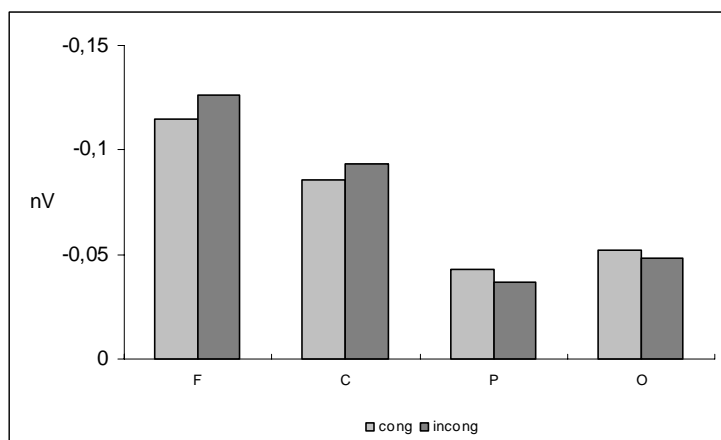


Fig. 11. Indice N200: intensità del picco nei diversi lobi, in risposta a stimoli congruenti e incongruenti.

Latenza

L'ANOVA a disegno misto che ha impiegato i fattori task, condizione ed elettrodo indica che la latenza di picco della N200 varia nelle diverse porzioni di corteccia cui sono stati applicati gli elettrodi ($F(13, N = 30) = 5,602, p < ,001$). In particolare, il segnale risulta essere più ritardato nelle aree frontali, centrali e parietali e più precoce invece in quelle temporali e occipitali. Emerge anche un effetto di interazione task*condizione*elettrodo ($F(26, N = 30) = 1,815, p = ,008$): gli stimoli incongruenti producono un'attivazione più tardiva dei congruenti nella condizione implicita, soprattutto nelle aree anteriori, differenza che non si registra invece nel task esplicito. Inoltre, nella condizione congruente, il task esplicito produce una latenza maggiore nelle zone parietali e occipitali, mentre nella condizione incongruente produce un'attivazione più ritardata a livello della corteccia centrale e parietale (vedi fig. 12).

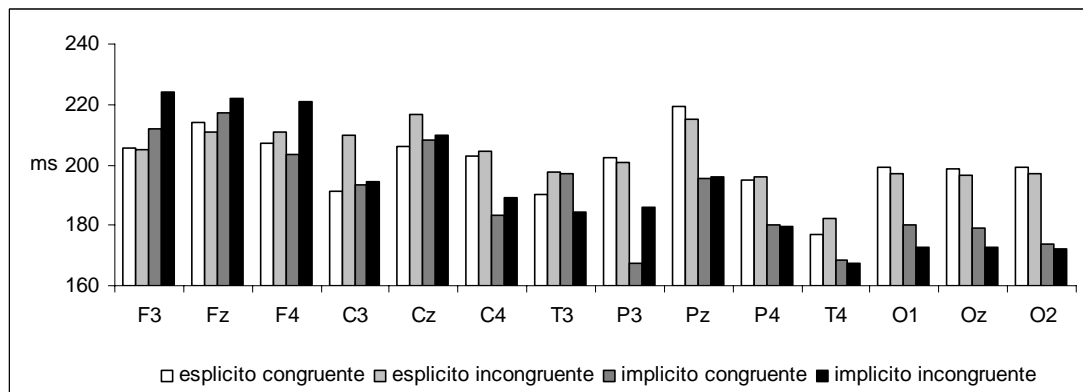


Fig. 12. Indice N200 in funzione dei diversi elettrodi, in condizione di congruenza e di incongruenza e in caso di consegna esplicita e implicita.

L'ANOVA a misure ripetute con i fattori condizione e localizzazione ha ulteriormente evidenziato l'esistenza di una differente latenza di attivazione delle diverse aree corticali, con una sostanziale omogeneità in quelle frontali, centrali e parietali ed invece tempi anticipati nella zona occipitale ($F(3, N = 30) = 5,605, p = ,001$), come confermato dall'analisi dei contrasti e come evidenziato in figura 13, dove si nota il progressivo ridursi dei tempi di attivazione col passaggio dalla porzione anteriore a quella posteriore della corteccia.

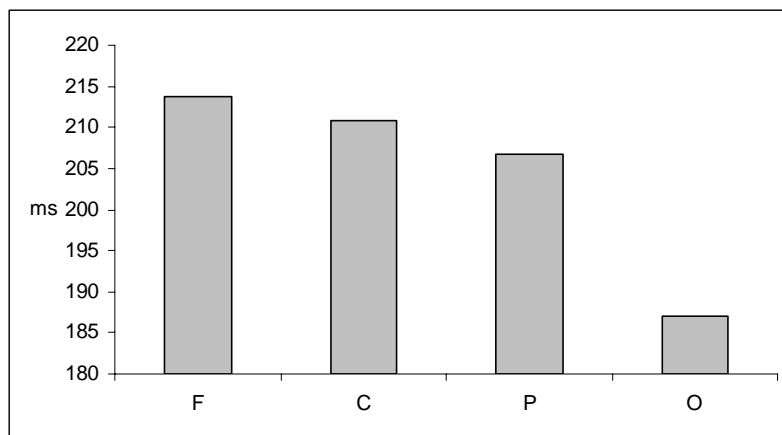


Fig. 13. Tempi di attivazione della N200 nelle aree frontali, centrali, parietali e occipitali.

Un'ultima analisi ha rivelato l'esistenza di un triplice effetto di interazione task*lateralizzazione*condizione ($F(2, N = 30) = 9,623, p < ,001$): quando il task è implicito, rispetto a quando è esplicito, si nota una maggiore differenziazione tra la condizione di congruenza, più anticipata, e di incongruenza, che ha invece tempi più ritardati. Allo stesso modo, sempre nel caso di task implicito, si registrano tempi di

attivazione più brevi nell'emisfero destro rispetto che nel sinistro, differenza non rilevabile invece quando il compito è esplicito (vedi fig. 14).

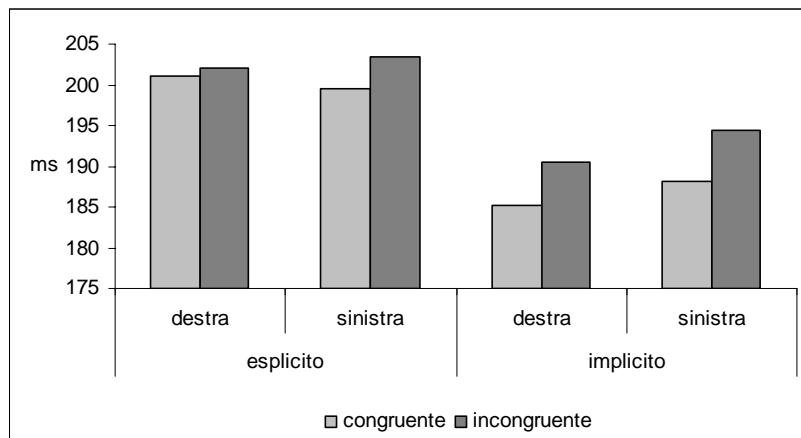


Fig. 14. Tempi di latenza del picco N200 negli emisferi destro e sinistro, in risposta a coppie di stimoli congruenti e incongruenti e con consegna esplicita e implicita.

3.1.4. Indice P200

L'indice P200 è stato ricavato estraendo il valore massimo nella finestra temporale 150 - 250 ms. Di seguito ne verranno esposti i dati relativi a intensità e latenza.

Intensità

La P200 presenta una differente intensità nelle diverse zone corticali considerate ($F(2, N = 30) = 21,608, p = ,039$), con un'amplificazione del picco nelle aree parietali, come emerso dall'ANOVA a misure ripetute e confermato dall'analisi dei contrasti ripetuta. La medesima analisi fa rilevare inoltre la significatività del fattore condizione ($F(2, N = 30) = 3,066, p = ,054$). Tuttavia, l'analisi dei contrasti indica che soltanto la condizione neutra presenta valori che si discostano dalle due condizioni sperimentali congruente ($M = ,178; ds = ,012$) e incongruente ($M = ,178; ds = ,013$), che invece non differiscono tra loro. L'effetto di interazione condizione*elettrodo ($F(26, N = 30) = 2,617, p < ,001$) indica che, tendenzialmente, nelle aree anteriori il picco è più ampio quando il volto e la voce sono congruenti, mentre nelle zone posteriori è più accentuato quando gli stimoli sono incongruenti. Anche in questo caso, però, a produrre delle differenziazioni significative è soltanto la condizione neutra.

Questi effetti sono meglio delineati dall'ANOVA che ha utilizzato i fattori condizione e localizzazione e dalla relativa analisi dei contrasti, che evidenzia l'esistenza di un effetto principale della localizzazione ($F(3, N = 30) = 11,725, p < ,001$), con una massima intensità in zone parietale (vedi fig. 15), e un effetto di duplice interazione condizione x localizzazione ($F(6, N = 30) = 3,580, p = ,002$) che ripropone sinteticamente quanto emerso dalla precedente analisi, e cioè l'assenza di differenze tra la condizione congruente e quella incongruente.

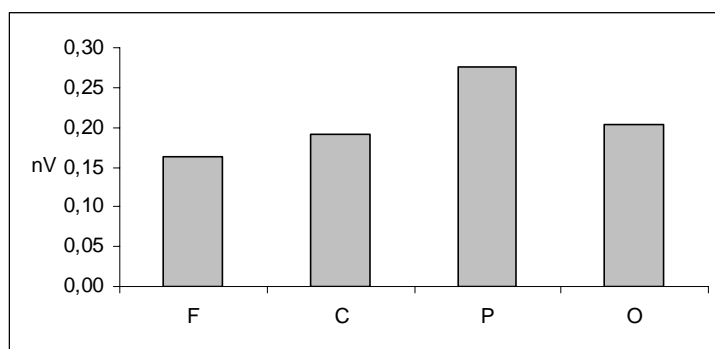


Fig. 15.
Localizzazione
del picco P200.

Latenza

L'ANOVA a disegno misto che ha incrociato i fattori task, condizione ed elettrodo ha messo in evidenza due effetti principali, il task e l'elettrodo, e un effetto di interazione, ovvero condizione*elettrodo.

Per quanto concerne il task ($F(1, N = 30) = 4,696, p = ,039$), quando la consegna è esplicita i tempi di latenza sono più veloci rispetto a quando rimane implicita, come è possibile osservare in figura 16.

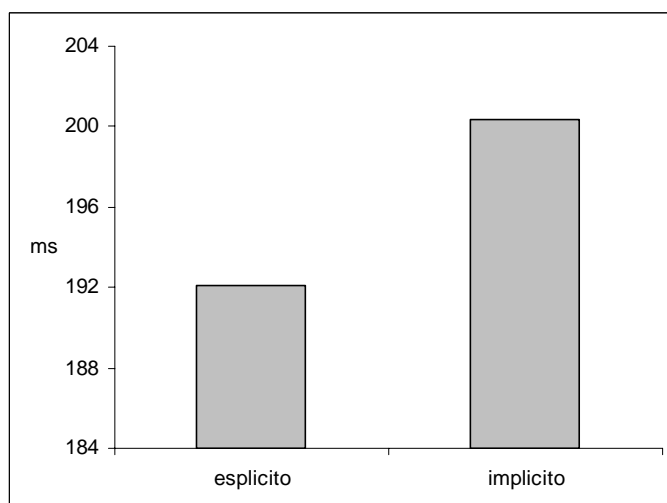


Fig. 16. Tempi
di latenza del
picco P200 in
presenza di
consegna
sperimentale
esplicita e
implicita.

Esaminando la differenziazione tra le diverse porzioni corticali individuate dagli elettrodi ($F(13, N = 30) = 6,898, p < ,001$), si nota come i tempi siano notevolmente più ritardati posteriormente, come del resto delineato anche dai risultati della successiva analisi relativa alla localizzazione ($F(3, N = 30) = 5,147, p < ,003$) riportati in fig. 17. L'analisi dei contrasti ha evidenziato il fatto che sono soprattutto i tempi di latenza dell'area occipitale a discostarsi dai tempi registrati nelle diverse posizioni della corteccia.

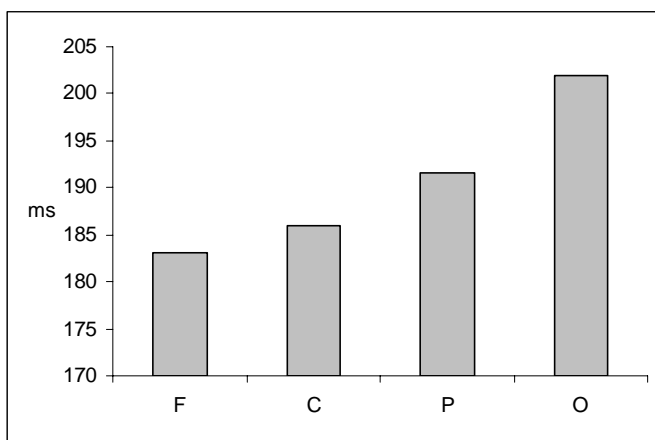


Fig. 17. Tempi di latenza dell'indice P200 nelle diverse aree corticali.

Come precedentemente riportato, i tempi di latenza risentono anche dell'interazione tra gli elettrodi e la condizione ($F(26, N = 30) = 1,508, p = ,051$): nella condizione congruente infatti la latenza è più elevata nelle aree anteriori, mentre in quella incongruente è ritardata posteriormente (vedi fig. 18).

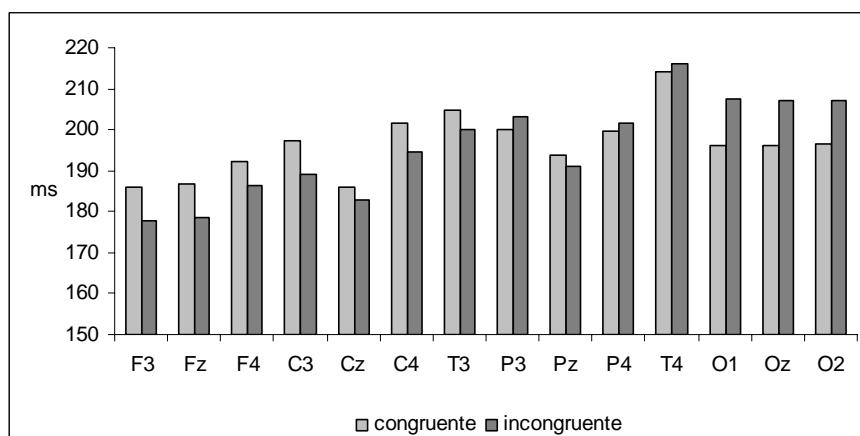


Fig. 18. Latenza della P200 in risposta alla presentazione di coppie di stimoli congruenti e incongruenti.

3.1.5. Indice N300

L'indice N300 deriva dall'estrazione del valore minimo all'interno della finestra temporale 250 - 350 ms. Le analisi hanno evidenziato la presenza di diversi fenomeni di interesse relativi sia all'intensità sia alla latenza.

Intensità

Le analisi indicano che l'intensità del picco è influenzata dal tipo di stimolazione ($F(2, N = 30) = 4,868, p = ,011$). Tuttavia, l'analisi dei contrasti rivela l'assenza di differenze significative per le due condizioni sperimentali congruente ($M = 307,98; ds = 3,62$) e incongruente ($M = 306,83; ds = 4,36$): solo la condizione neutra presenta valori che si discostano da quelli delle altre due condizioni. L'intensità del fenomeno è particolarmente accentuata nella zona frontale e decresce poi posteriormente ($F(13, N = 30) = 8,895 p < ,001$). Più nello specifico, dai risultati emerge che è soprattutto nella condizione di incongruenza che nella zona anteriore i valori di intensità sono più elevati ($F(26, N = 30) = 3,786 p < ,001$) (vedi fig. 19).

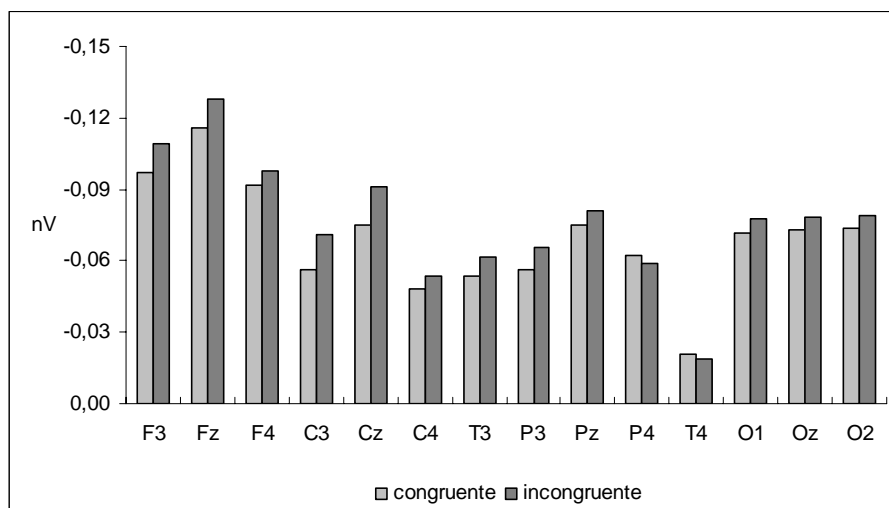


Fig 19.
Distribuzione dell'intensità della N300 rispetto agli elettrodi, in risposta a coppie volto/voce congruenti e incongruenti.

L'ANOVA con i fattori localizzazione e condizione indica la presenza di due effetti significativi, relativi l'uno alla localizzazione ($F(3, N = 30) = 6,833 p < ,001$) e l'altro all'interazione tra localizzazione e condizione ($F(6, N = 30) = 4,870 p < ,001$). Dai test dei contrasti emerge più nel dettaglio quanto illustrato precedentemente,

ovvero l'intensificazione del picco nella corteccia frontale, fenomeno particolarmente accentuato a seguito della presentazione di stimoli incongruenti (vedi fig. 20).

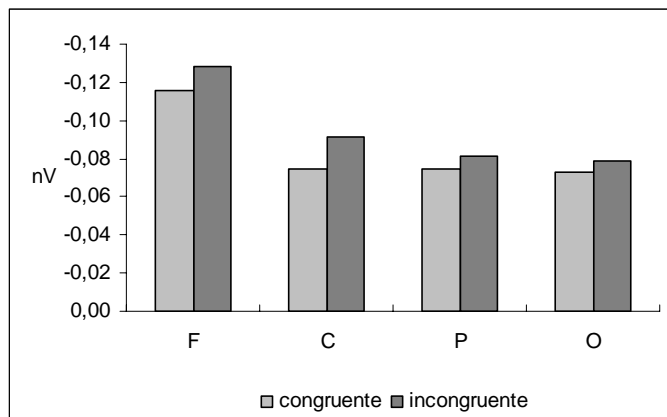


Fig. 20. Intensità del fenomeno N300 nelle aree frontali, centrali, parietali e occipitali in condizione di congruenza e incongruenza.

Latenza

Dall'ANOVA a disegno misto che ha incrociato i fattori task, condizione ed elettrodo emergono numerosi effetti significativi. In primo luogo l'analisi segnala un'influenza della condizione sui tempi di latenza ($F(2, N = 30) = 4,766 p = ,012$); tuttavia, l'analisi dei contrasti indica che gli unici valori che si discostano significativamente dagli altri sono quelli della condizione neutra, con una sostanziale sovrapposizione tra quelli della condizione congruente ($M = 308,29$; $ds = 3,27$) e incongruente ($M = 306,39$; $ds = 3,95$). Più interessante è il fatto che i tempi sono notevolmente più precoci in caso di task esplicito ($F(1, N = 30) = 5,274 p = ,029$), come si può notare in fig. 21.

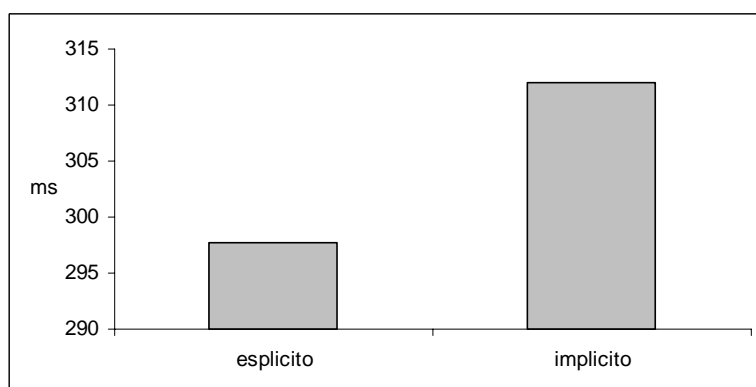


Fig 21. Tempi di latenza della N300 in presenza di consegna esplicita e implicita.

Gli effetti significativi elettrodo ($F(13, N = 30) = 28,395 p < ,001$), condizione x elettrodo ($F(26, N = 30) = 3,046 p < ,001$) ed elettrodo per task ($F(13, N = 30) = 2,587 p$

= ,002) sono efficacemente e sinteticamente illustrati dai risultati della successiva ANOVA che ha impiegato i fattori task, condizione e localizzazione. Quest'ultima analisi infatti indica che la N300 è notevolmente ritardata nelle aree corticali posteriori ($F(3, N = 30) = 3,462 p < ,001$), rispetto a quelle anteriori (vedi fig. 22a), soprattutto in presenza di stimoli incongruenti ($F(6, N = 30) = 3,482 p = ,003$) (vedi fig. 22b). Inoltre, essa è sensibile al tipo di consegna sperimentale, presentando tempi più lunghi quando il task è di natura implicita, soprattutto posteriormente ($F(3, N = 30) = 4,682 p = ,005$) (vedi fig. 22 c).

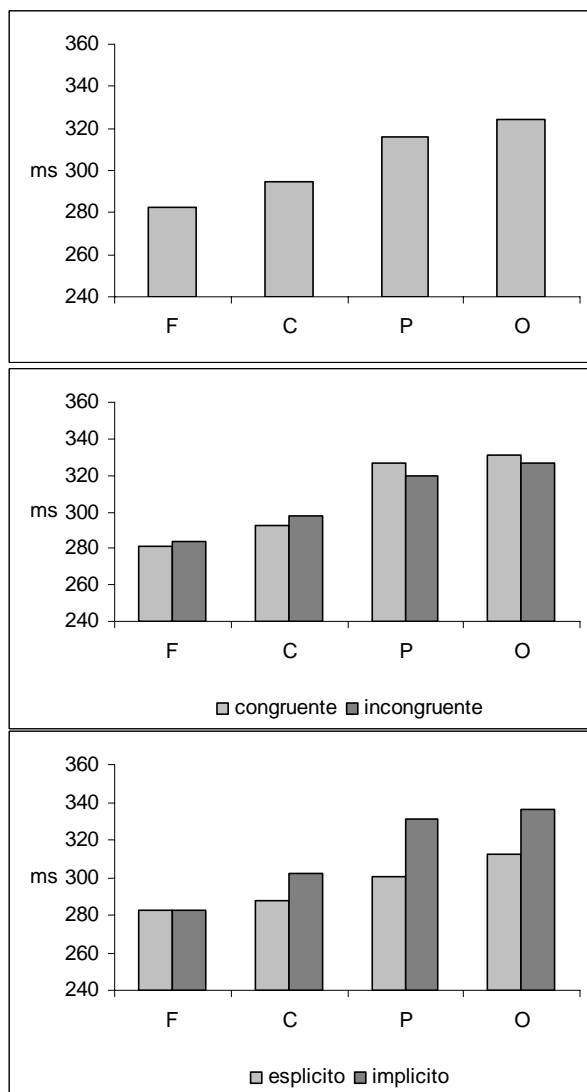


Fig. 22. Latenza di picco della N300 nelle aree frontali, centrali, parietali e occipitali (a), in funzione dell'interazione tra tipo di condizione e localizzazione (b) e tra tipo di compito e localizzazione (c).

Per concludere, le analisi hanno evidenziato un duplice effetto lateralizzazione*condizione sui tempi di latenza ($F(2, N = 30) = 5,742$ $p = ,005$): come osservabile in tabella 4, le medie indicano che nelle condizioni di congruenza e incongruenza i tempi sono più ritardati nell'emisfero sinistro, mentre nella condizione neutra lo sono nell'emisfero destro.

	congruente	incongruente	neutro
destra	314,8	310,2	299,5
sinistra	312,3	307,3	304,6

Tab. 4. Tempi di latenza del picco N300 nei due emisferi, in funzione delle condizioni sperimentali congruente, incongruente e neutra.

3.1.6. Indice P300

L'indice P300 è stato ottenuto estraendo il valore massimo all'interno della finestra temporale 250 - 350 ms. Di seguito, verranno esposti i risultati delle analisi relative all'intensità e alla latenza del picco.

Intensità

Dall'ANOVA a disegno misto con i fattori task, condizione ed elettrodo emergono diversi risultati significativi: l'effetto elettrodo ($F(13, N = 30) = 14,530$ $p < ,001$), l'effetto condizione*elettrodo ($F(26, N = 30) = 2,288$ $p < ,001$) e l'effetto task*condizione*elettrodo ($F(26, N = 30) = 2,285$ $p < ,001$). Questi dati sono più efficacemente rappresentabili grazie ai dati della successiva analisi, che descrivono in primo luogo la P300 come un fenomeno specificamente posteriore ($F(3, N = 30) = 12,484$ $p < ,001$) e soprattutto parietale (vedi fig. 23).

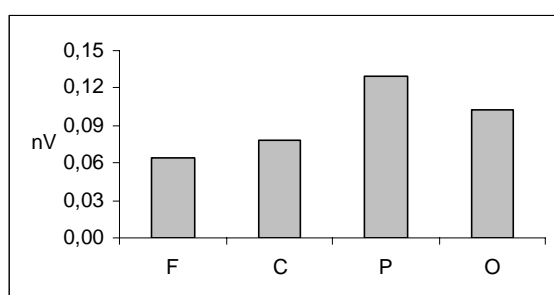


Fig. 23.
Localizzazione
dell'indice P300.

Inoltre, la figura 24 indica che è proprio nell'area posteriore, e soprattutto parietale, che si manifesta una differenziazione relativa al task, con un'accentuazione del picco quando il compito rimane implicito rispetto a quando è esplicito ($F(3, N = 30) = 12,484 p < ,001$). Questa marcata differenziazione non si rileva invece nelle aree anteriori della corteccia.

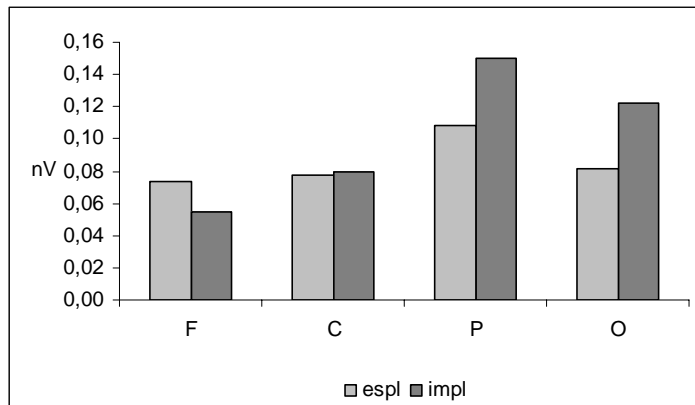


Fig. 24. Rappresentazione dell'indice P300 in condizione di task esplicito e implicito nelle diverse aree corticali.

Infine, l'analisi indica anche la presenza di un'interazione significativa tra la condizione e la localizzazione ($F(6, N = 30) = 2,672 p = ,017$) e tra la condizione, la localizzazione e il task ($F(6, N = 30) = 2,541 p = ,022$). L'analisi dei contrasti tuttavia indica che è in realtà la condizione neutra a discostarsi in modo significativo dalle condizioni sperimentali.

Latenza

Rispetto alla latenza, l'unico dato significativo riguarda gli elettrodi ($F(13, N = 30) = 26,448 p < ,001$) e la localizzazione ($F(3, N = 30) = 34,847 p < ,001$): i tempi infatti sono più ritardati nell'area anteriore, come illustrato in figura 25.

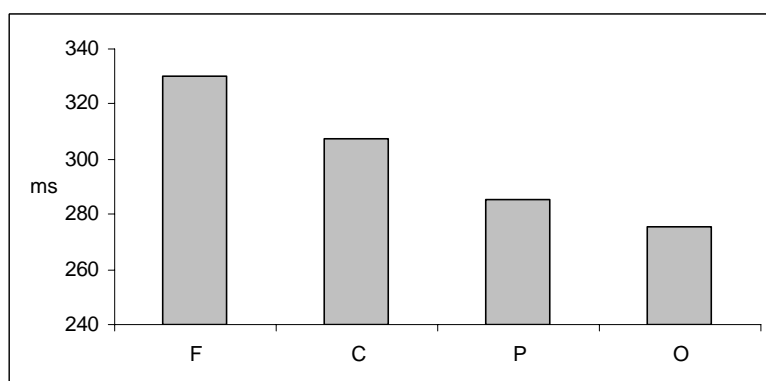


Fig. 25. Latenza della P300 nelle zone frontali, centrali, parietali e occipitali.

3.1.7. Indice N400

L'indice N400 è stato ottenuto estraendo il valore minimo all'interno della finestra temporale 350 - 500 ms. Di seguito, verranno esposti i risultati delle analisi.

Intensità

Il picco N400 si connota come un fenomeno posteriore, come emerge dall'ANOVA con i fattori condizione ed elettrodo ($F(13, N = 30) = 23,538 p < ,001$) e dall'ANOVA con i fattori condizione e localizzazione ($F(3, N = 30) = 18,051 p < ,001$) e come osservabile in figura 26.

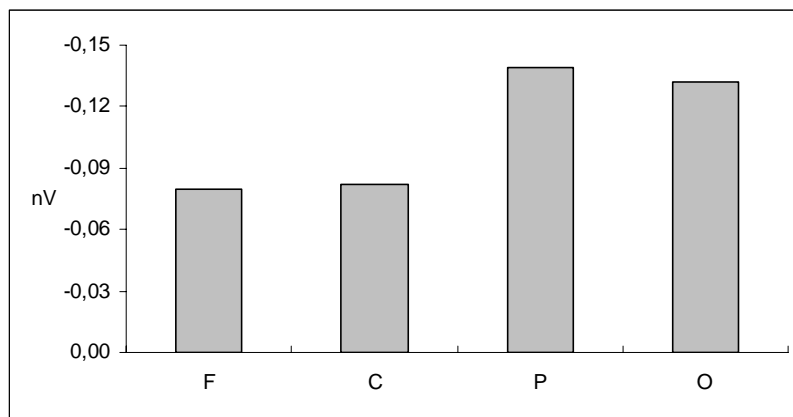


Fig. 26.
Localizzazione
della N400.

Latenza

Le analisi effettuate indicano unicamente la presenza di tempi di latenza più veloci nelle aree anteriori e più ritardati in quelle posteriori ($F(13, N = 30) = 3,698 p < ,001$).

3.2. L'universo emotivo: le analisi emotion-by-emotion

Per esplorare la possibile presenza di differenziazioni all'interno dell'universo delle emozioni congruenti e per metterle a confronto con la condizione sperimentale di incongruenza, sono state applicate a ciascuna delle medesime finestre temporali precedentemente descritte le seguenti analisi accompagnate, dove necessario, da un'analisi dei contrasti:

- Anova a disegno misto localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x emozione (var. within a 6 livelli: disgusto, gioia, paura, rabbia, sorpresa, tristezza) x task (var. between a 2 livelli: esplicito, implicito)
- Anova a misure ripetute localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x emozione (var. within a 6 livelli: disgusto, gioia, paura, rabbia, sorpresa, tristezza)
- Anova a disegno misto condizione (var. within a 2 livelli: disgusto, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x task (var. between a 2 livelli: esplicito, implicito)
- Anova a misure ripetute condizione (var. within a 2 livelli: disgusto, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale)
- Anova a disegno misto condizione (var. within a 2 livelli: gioia, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x task (var. between a 2 livelli: esplicito, implicito)
- Anova a misure ripetute condizione (var. within a 2 livelli: gioia, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale)
- Anova a disegno misto condizione (var. within a 2 livelli: paura, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x task (var. between a 2 livelli: esplicito, implicito)
- Anova a misure ripetute condizione (var. within a 2 livelli: paura, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale)
- Anova a disegno misto condizione (var. within a 2 livelli: rabbia, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x task (var. between a 2 livelli: esplicito, implicito)
- Anova a misure ripetute condizione (var. within a 2 livelli: rabbia, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale)
- Anova a disegno misto condizione (var. within a 2 livelli: sorpresa, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x task (var. between a 2 livelli: esplicito, implicito)

- Anova a misure ripetute condizione (var. within a 2 livelli: sorpresa, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale)
- Anova a disegno misto condizione (var. within a 2 livelli: tristezza, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x task (var. between a 2 livelli: esplicito, implicito)
- Anova a misure ripetute condizione (var. within a 2 livelli: tristezza, incongruente) x localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale)

3.2.1. Indice N100

Intensità

Confronto tra emozioni

Le analisi condotte evidenziano l'assenza di differenze nell'intensità del picco dovute al tipo di emozione o al task. Si rileva invece una differenziazione del segnale rispetto alla localizzazione ($F(3, N = 30) = 4,877 p = ,004$): le emozioni congruenti infatti producono un picco più accentuato in area frontale e parietale (vedi tab. 5).

	M	DS	
F	-,226	,016	Tab 5. Intensità della N100 elicitata dalle emozioni congruenti.
C	-,187	,015	
P	-,205	,019	
O	-,179	,014	

Confronto tra le singole emozioni congruenti e la condizione incongruente

Disgusto/incongruente

L'ANOVA indica che la N100 è più intensa in presenza di stimoli congruenti esprimenti il disgusto rispetto che in presenza di stimoli incongruenti ($F(1, N = 30) = 21,236 p < ,001$), come osservabile in fig. 28; inoltre, il fenomeno è prevalentemente frontale ($F(3, N = 30) = 6,133 p = ,001$) (vedi tab. 6).

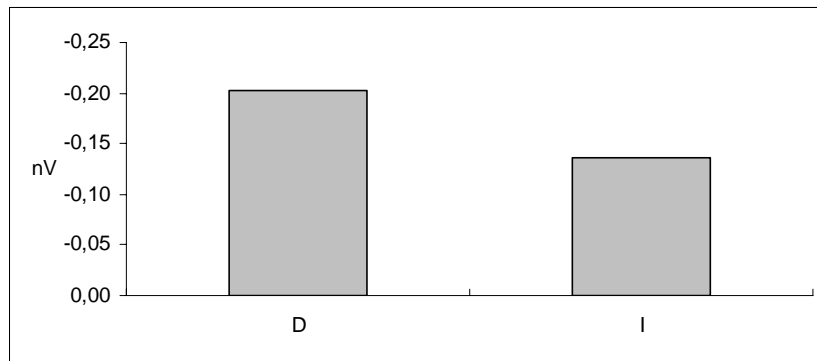


Fig. 28. Intensità del picco N100 in risposta a stimoli congruenti esprimenti disgusto e a stimoli incongruenti.

Gioia/incongruente

Il picco N100 è più intenso in seguito alla presentazione di stimoli congruenti esprimenti gioia rispetto che in presenza di stimoli incongruenti ($F(1, N = 30) = 7,005$ $p = ,013$) (vedi fig. 29); inoltre, il fenomeno è soprattutto frontale ($F(3, N = 30) = 3,612$ $p = ,016$) (vedi tab. 6).

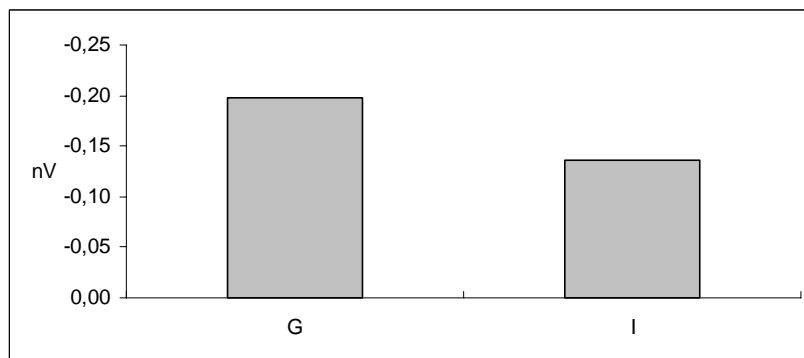


Fig. 29. Intensità del picco N100 in risposta a stimoli congruenti esprimenti gioia e a stimoli incongruenti.

Paura/incongruente

Gli stimoli congruenti esprimenti paura, rispetto agli stimoli incongruenti, producono valori di N100 più elevati ($F(1, N = 30) = 5,741$ $p = ,023$) (vedi fig. 30); inoltre, il fenomeno è soprattutto frontale ($F(3, N = 30) = 6,012$ $p = ,001$) (vedi tab. 6).

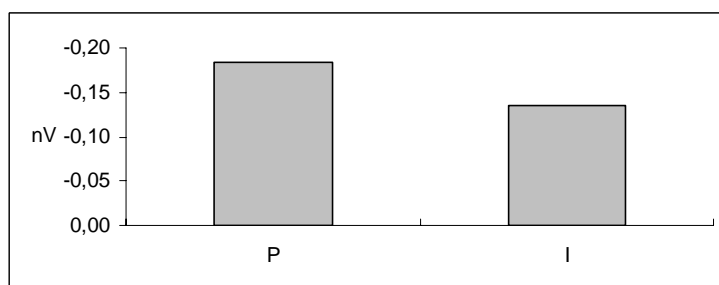


Fig. 30. Intensità del picco N100 in risposta a stimoli congruenti esprimenti paura e a stimoli incongruenti.

Rabbia/incongruente

Gli stimoli congruenti esprimenti rabbia, rispetto agli stimoli incongruenti, producono un picco N100 più accentuato ($F(1, N = 30) = 8,242 p = ,008$) (vedi fig. 31); inoltre, il fenomeno è soprattutto frontale ($F(3, N = 30) = 4,549 p = ,005$) (vedi tab. 6).

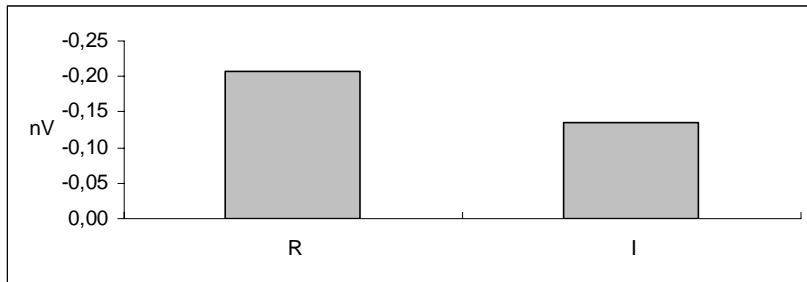


Fig. 31. Intensità del picco N100 in risposta a stimoli congruenti esprimenti rabbia e a stimoli incongruenti.

Sorpresa/incongruente

Il picco N100 è più intenso in seguito alla presentazione di stimoli congruenti esprimenti sorpresa rispetto che in presenza di stimoli incongruenti ($F(1, N = 30) = 24,293 p < ,001$) (vedi fig. 32); inoltre, ancora una volta, il fenomeno è soprattutto frontale ($F(3, N = 30) = 3,612 p = ,016$) (vedi tab. 6).

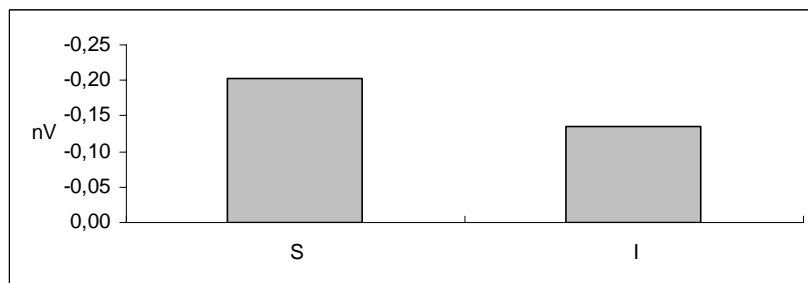


Fig. 32. Intensità del picco N100 in risposta a stimoli congruenti esprimenti sorpresa e a stimoli incongruenti.

Inoltre, si rileva un'interazione tra il tipo di stimolo e la localizzazione ($F(3, N = 30) = 6,271 p = ,001$), che indica come sia soprattutto in area parietale che si verifica una differenziazione del picco in funzione della condizione sperimentale, con valori molto più elevati in caso di stimoli congruenti esprimenti sorpresa. Gli stimoli incongruenti invece producono i valori più elevati in zona frontale (vedi fig. 33).

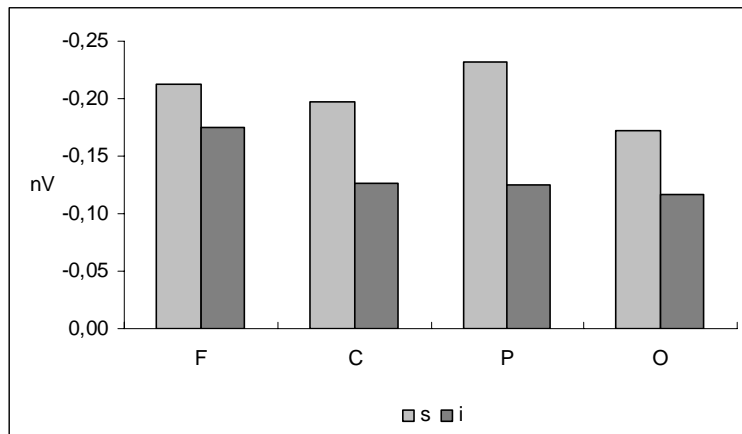


Fig. 33. Intensità del picco N100 in risposta a stimoli congruenti esprimenti sorpresa e a stimoli incongruenti nelle diverse aree corticali.

Tristezza/incongruente

Gli stimoli congruenti esprimenti tristezza, rispetto agli stimoli incongruenti, producono un picco N100 più accentuato ($F(1, N = 30) = 24,820 p < ,001$) (vedi fig. 34); inoltre, il fenomeno è soprattutto frontale ($F(3, N = 30) = 4,578 p = ,005$) (vedi tab. 6).

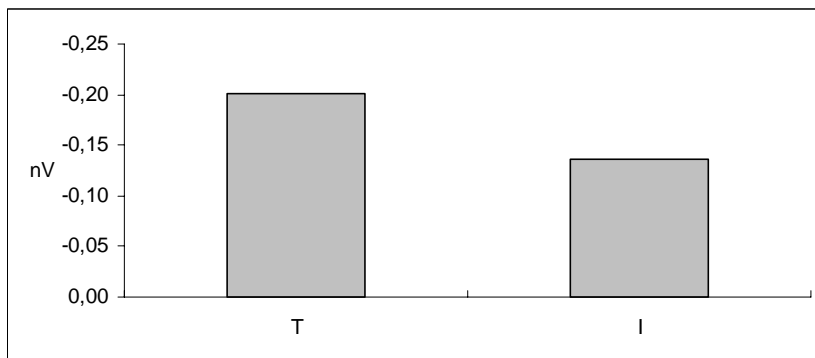


Fig. 34. Intensità del picco N100 in risposta a stimoli congruenti esprimenti tristezza e a stimoli incongruenti.

	D-I		G-I		P-I		R-I		S-I		T-I	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	-,20	,01	-,19	,01	-,20	,01	-,20	,01	-,19	,01	-,20	,01
C	-,15	,01	-,15	,01	-,14	,01	-,15	,01	-,16	,01	-,16	,01
P	-,16	,01	-,15	,01	-,14	,02	-,17	,02	-,17	,01	-,16	,01
O	-,14	,01	-,15	,01	-,14	,01	-,15	,01	-,14	,01	-,14	,01

Tab 6. Media e deviazione standard dei valori della N100 nelle diverse aree corticali, per ciascuna emozione congruente confrontata con la condizione incongruente.

3.2.2. *Indice P100*

Intensità

Confronto tra emozioni

Le analisi evidenziano una differenziazione del segnale rispetto alla localizzazione ($F(3, N = 30) = 15,610 p < ,001$): le emozioni congruenti, infatti, producono un picco P100 che si presenta come un fenomeno posteriore (vedi tab. 7).

	M	DS
F	,164	,015
C	,187	,013
P	,291	,023
O	,276	,022

Tab 7. Localizzazione della P100 elicitata dalle emozioni congruenti.

*Confronto tra le singole emozioni congruenti e la condizione incongruente**Disgusto/incongruente*

La P100 è maggiormente accentuata in caso di presentazione di stimoli esprimenti disgusto rispetto che di stimoli incongruenti ($F(1, N = 30) = 4,622 p = ,040$) (vedi fig. 35). Inoltre, il picco è un fenomeno spiccatamente posteriore ($F(3, N = 30) = 13,552 p < ,001$) (vedi tab. 8).

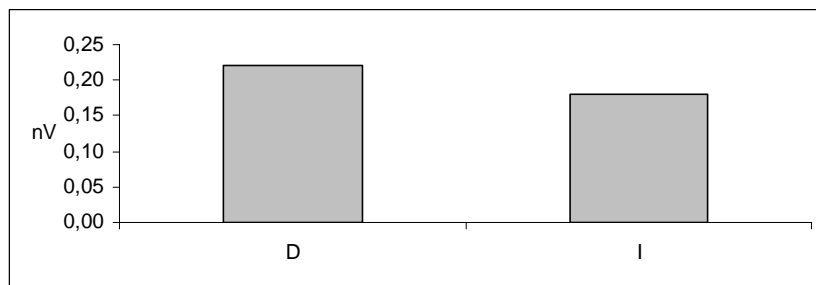


Fig. 35. Intensità del picco P100 in seguito alla presentazione di volti e voci esprimenti disgusto e di volti e voci incongruenti.

Gioia/incongruente

Il picco prodotto dagli stimoli esprimenti gioia mostra un'intensità maggiore rispetto a quello prodotto dagli stimoli incongruenti ($F(1, N = 30) = 11,457 p < ,001$) (vedi fig. 36). Esso, inoltre, si connota come un fenomeno posteriore, come osservabile in tabella 8 ($F(3, N = 30) = 15,978 p < ,001$).

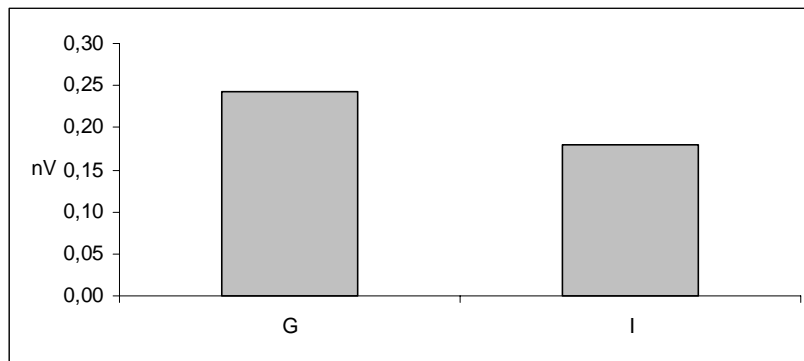


Fig. 36. Intensità del picco P100 in seguito alla presentazione di stimoli esprimenti gioia e di stimoli incongruenti.

Paura/incongruente

Il picco P100, quando elicitato da stimoli esprimenti paura, è sensibile al fattore localizzazione ($F(3, N = 30) = 15,310 p < ,001$), con una prevalenza posteriore e soprattutto parietale (vedi tab. 8), e al tipo di task ($F(1, N = 30) = 5,290 p = ,029$), con valori più elevati in caso di consegna esplicita (vedi fig. 37).

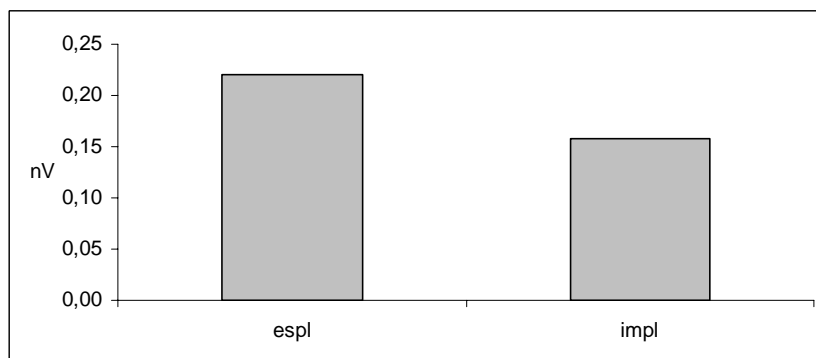


Fig. 37. Valori del picco P100 in presenza di consegna esplicita ed implicita, in seguito alla presentazione di stimoli esprimenti paura e di stimoli incongruenti.

Rabbia/incongruente

Gli stimoli congruenti esprimenti paura provocano un picco P100 più intenso rispetto agli stimoli incongruenti ($F(1, N = 30) = 18,840 p < ,001$), come osservabile in figura 38. In generale, il picco si presenta come un fenomeno particolarmente accentuato in area posteriore ($F(3, N = 30) = 13,856 p < ,001$) (vedi tab. 8).

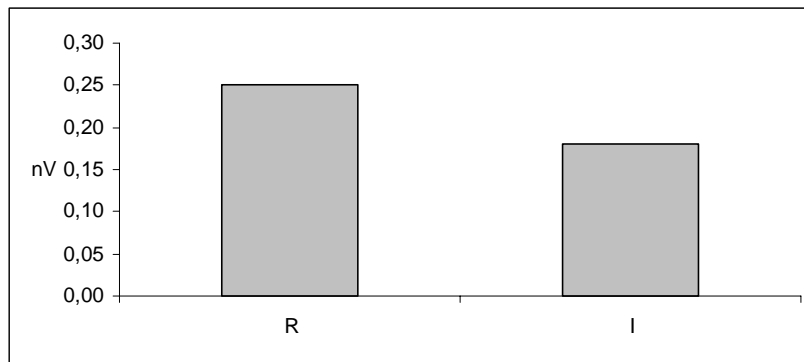


Fig. 38. La P100 risulta essere più accentuata quando vengono presentati stimoli esprimenti rabbia rispetto a quando vengono presentati stimoli incongruenti.

Sorpresa/incongruente

Gli stimoli esprimenti sorpresa elicitano un picco più intenso rispetto agli stimoli incongruenti ($F(1, N = 30) = 16,145 p < ,001$) (vedi fig. 39). Inoltre, tale picco è presente soprattutto nella porzione posteriore della corteccia ($F(3, N = 30) = 14,373 p < ,001$) (vedi tab. 8).

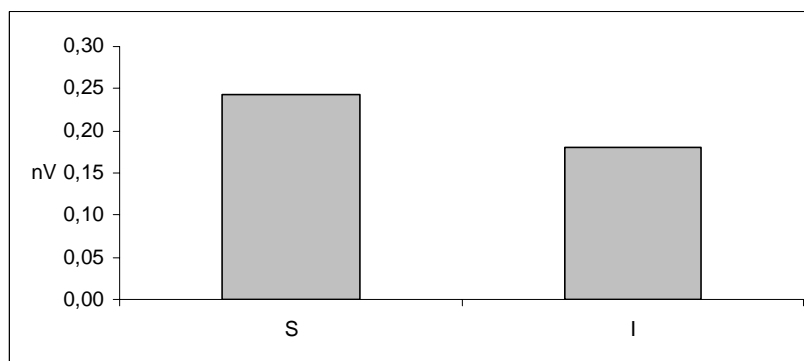


Fig. 39. La P100 presenta valori più elevati quando vengono presentati stimoli esprimenti sorpresa rispetto che nella condizione incongruente.

Tristezza/incongruente

In generale, la P100 si presenta come un fenomeno specificamente posteriore ($F(3, N = 30) = 14,485 p < ,001$) (vedi tab. 8). Inoltre, risulta essere più intensa in presenza di stimoli esprimenti gioia rispetto che in presenza di stimoli incongruenti ($F(1, N = 30) = 11,495 p = ,002$), come osservabile in figura 40.

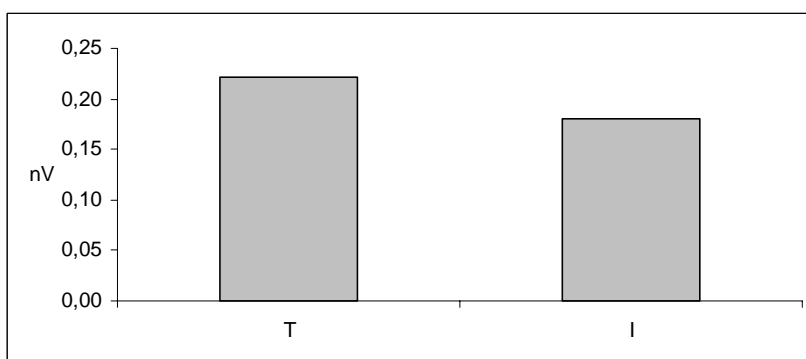


Fig. 40. Valori della P100 in seguito alla stimolazione con stimoli esprimenti tristezza e con stimoli incongruenti.

	D-I		G-I		P-I		R-I		S-I		T-I	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	,130	,016	,130	,016	,121	,015	,142	,019	,140	,019	,128	,017
C	,161	,015	,175	,015	,144	,016	,172	,018	,164	,017	,157	,016
P	,261	,026	,281	,026	,248	,025	,282	,022	,272	,022	,264	,023
O	,251	,022	,263	,025	,240	,024	,264	,022	,270	,022	,257	,026

Tab. 8. Media e deviazione standard dei valori del picco P100 nelle diverse aree corticali, per ciascuna emozione congruente confrontata con la condizione incongruente.

3.2.3. Indice N200

Intensità

Confronto tra emozioni

Dal confronto tra loro le sei emozioni congruenti e dall'analisi dei contrasti emerge, in generale, che la N200 è un fenomeno tipicamente anteriore ($F(3, N = 30) = 7,757 p < ,001$), soprattutto frontale ma anche centrale (vedi fig. 41) e che, più nello specifico, la sua entità differisce in funzione dell'effetto congiunto del tipo di emozione e della localizzazione ($F(15, N = 30) = 2,234 p = ,005$): in figura 42 si può notare come la localizzazione frontale sia accentuata soprattutto per l'emozione della rabbia.

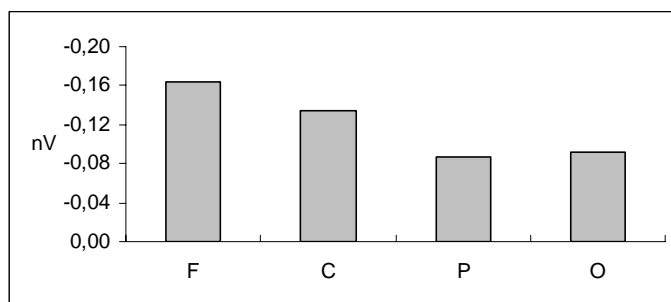


Fig. 41. Localizzazione dell'indice N200, considerando le emozioni congruenti.

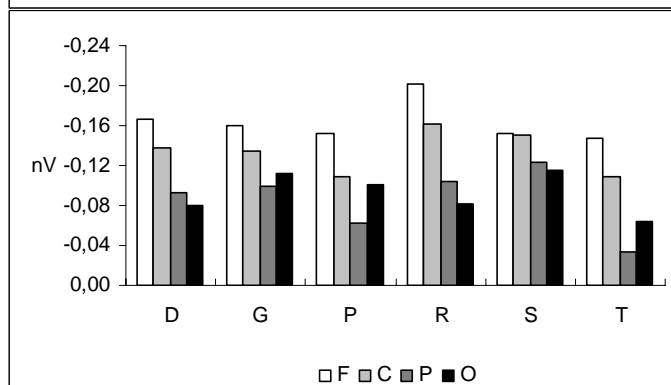


Fig. 42. Localizzazione del picco P100 rispetto a ciascuna delle emozioni congruenti.

Confronto tra le singole emozioni congruenti e la condizione incongruente

Disgusto/incongruente

Confrontando il disgusto e la condizione incongruente, emerge una caratterizzazione della N200 come picco prevalentemente frontale ($F(3, N = 30) = 11,319 p < ,001$) (vedi tab. 9) e maggiormente accentuato nel caso di stimoli esprimenti disgusto ($F(1, N = 30) = 5,029 p = ,033$) (vedi fig. 43).

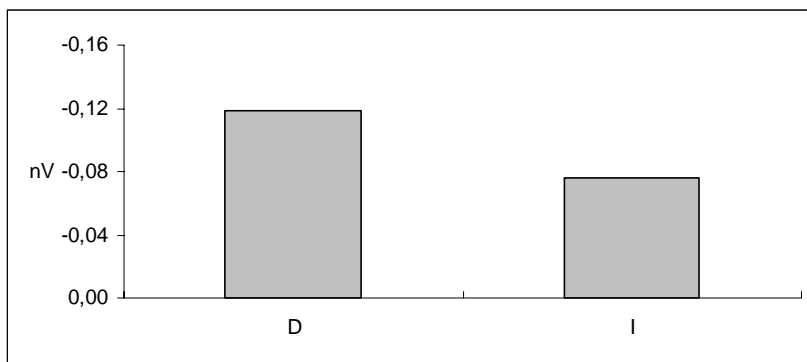


Fig. 43. Intensità della N200 in seguito alla presentazione di stimoli esprimenti disgusto e stimoli incongruenti.

Gioia/incongruente

Anche nel caso della gioia, il picco è localizzato frontalmente ($F(3, N = 30) = 6,253 p < ,001$) (vedi tab. 9) e maggiormente accentuato quando gli stimoli sono congruenti ($F(1, N = 30) = 7,872 p = ,009$), come si può osservare in figura 44.

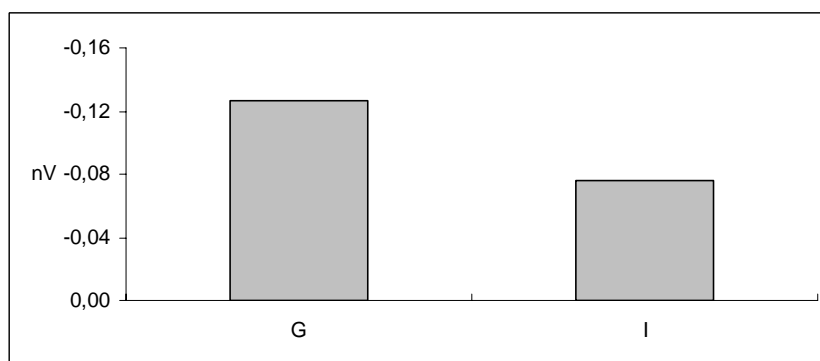


Fig. 44. Intensità della N200 in presenza di stimoli esprimenti gioia e stimoli incongruenti.

Paura/incongruente

Rispetto al confronto tra paura e incongruenza, l'unico dato rilevato è una localizzazione frontale del picco ($F(3, N = 30) = 15,310 p < ,001$) (vedi tab. 9).

Rabbia/incongruente

Le analisi fanno rilevare una localizzazione specificamente frontale del picco N200 ($F(3, N = 30) = 10,803 p < ,001$) (vedi tab. 9) e una sua maggiore ampiezza nel caso di stimoli che esprimono rabbia ($F(1, N = 30) = 12,416 p < ,001$) (vedi fig. 45).

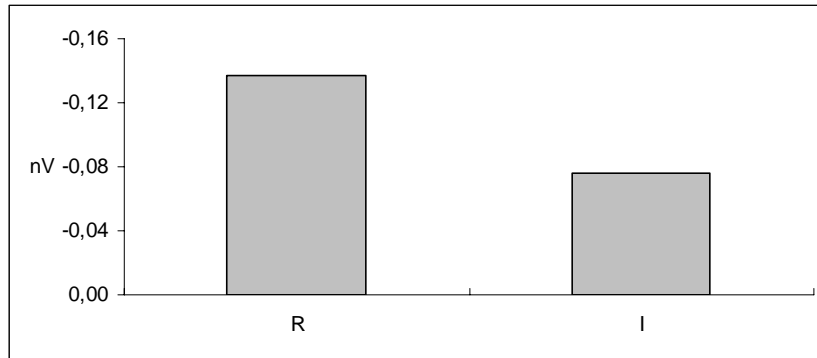


Fig. 45. Ampiezza della N200 elicitata da stimoli esprimenti rabbia e da stimoli incongruenti.

Sorpresa/incongruente

Ancora una volta, la N200 si conferma come un picco con localizzazione prevalentemente frontale ($F(3, N = 30) = 4,109 p = ,009$) (vedi tab. 9) e più accentuata in caso di congruenza ($F(1, N = 30) = 23,556 p < ,001$) (vedi fig. 46). Inoltre, si rileva un effetto di interazione condizione*localizzazione ($F(3, N = 30) = 3,854 p = ,012$) che mette in evidenza come, soprattutto nel caso degli stimoli incongruenti, si verifichi una drastica deintensificazione del picco nella porzione posteriore della corteccia (vedi fig. 47).

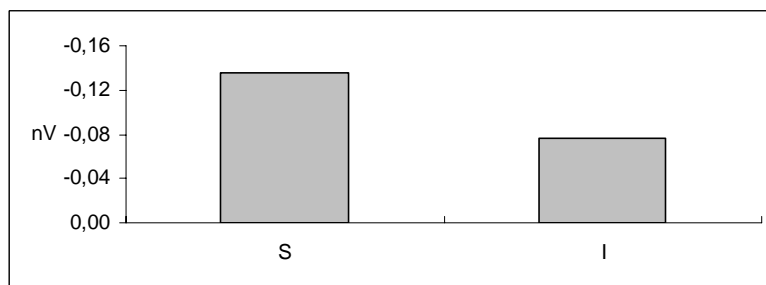


Fig. 46. Ampiezza della N200 elicitata da stimoli esprimenti sorpresa e da stimoli incongruenti.

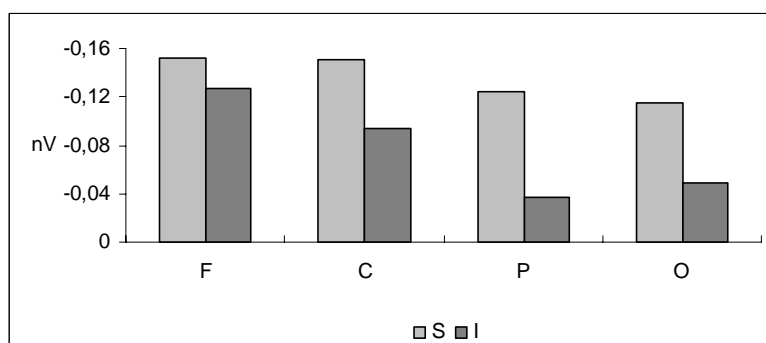


Fig. 47. Ampiezza della N200 in funzione della localizzazione e del tipo di stimolo (rabbia congruente/incongruenza).

Tristezza/incongruente

Rispetto al confronto tra tristezza e incongruenza, si rileva unicamente una localizzazione prevalentemente frontale ($F(3, N = 30) = 11,153$ $p < ,009$), come osservabile in tabella 9.

	D - I		G - I		P - I		R - I		S - I		T - I	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	-,146	,015	-,144	,018	-,139	,017	-,164	,021	-,139	,020	-,136	,018
C	-,115	,020	-,114	,018	-,101	,017	-,128	,023	-,122	,024	-,101	,021
P	-,065	,023	-,068	,022	-,050	,023	-,071	,024	-,081	,027	-,036	,026
O	-,064	,020	-,080	,018	-,074	,020	-,065	,019	-,082	,018	-,056	,019

Tab. 9. Media e deviazione standard dei valori del picco N200 nelle diverse aree corticali, per ciascuna emozione congruente confrontata con la condizione incongruente.

*Latenza**Confronto tra emozioni*

Dal confronto tra le diverse coppie di emozioni congruenti, si rileva un ritardo dei tempi di latenza in area anteriore, a fronte di tempi più rapidi in area parietale e soprattutto occipitale (vedi tab. 10).

	M	DS
F	208,6	4,3
C	207,3	4,6
P	197,2	5,1
O	188,3	5,1

Tab. 10. Tempi di latenza della N200 rispetto alle diverse aree corticali considerate.

*Confronto tra le singole emozioni congruenti e la condizione incongruente**Disgusto/incongruente*

Dalle analisi emerge un effetto di interazione condizione*localizzazione ($F(3, N = 30) = 3,217$ $p = ,027$). Esso indica che, per il disgusto, i tempi di latenza non differiscono grandemente nelle diverse aree corticali, mentre, per la condizione di incongruenza, si nota un rilevante ritardo in area frontale e, anche se in misura minore, parietale (vedi fig. 48).

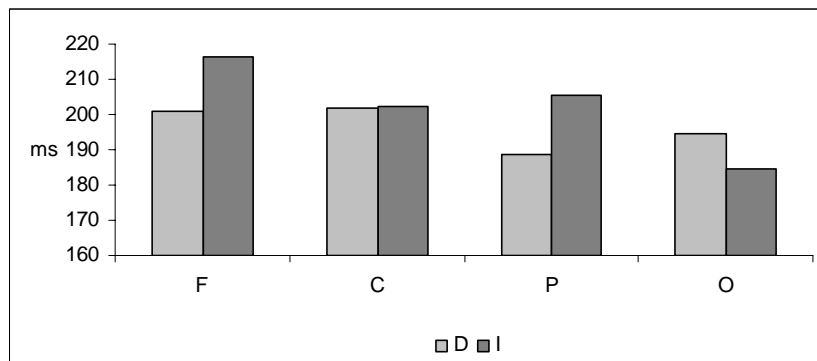


Fig. 48. Tempi di latenza della N200 in funzione della localizzazione e del tipo di stimolo (disgusto congruente/incongruenza).

Gioia/incongruente

L'analisi indica, ancora una volta, il ritardo in area frontale dell'indice N200, come indicato in tab. 11 ($F(3, N = 30) = 5,171 p = ,002$).

Paura/incongruente

Mettendo a confronto paura e condizione incongruente, si nota un allungamento dei tempi di latenza in zona frontale ($F(3, N = 30) = 4,281 p = ,004$) (vedi tab. 11) così come in presenza di un task di tipo esplicito ($F(1, N = 30) = 4,294 p = ,048$) (vedi fig. 49).

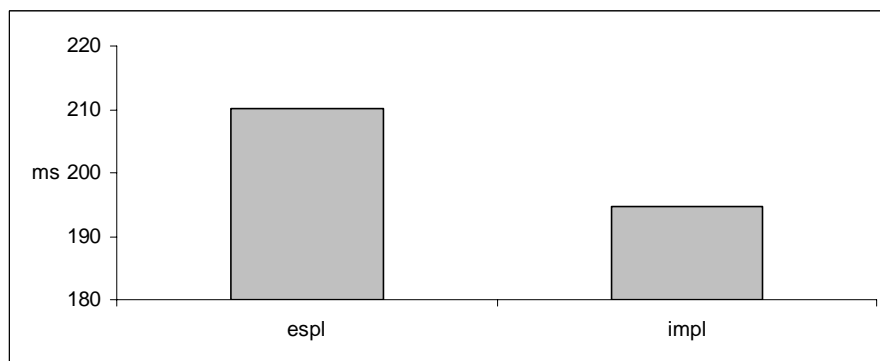


Fig. 49. Tempi di latenza del picco N200 in funzione del tipo di task.

Rabbia/incongruente

Come precedentemente rilevato, il picco presenta tempi più rapidi nelle zone posteriori ($F(3, N = 30) = 6,615 p < ,001$) (vedi tab. 11).

Sorpresa/incongruente

Confrontando sorpresa e incongruenza, le analisi indicano non soltanto la riduzione dei tempi di latenza nelle aree posteriori ($F(3, N = 30) = 5,413 p = ,002$) (vedi tab. 11), ma anche un marcato ritardo della condizione incongruente rispetto a quella congruente ($F(3, N = 30) = 7,019 p = ,013$) (vedi fig. 50).

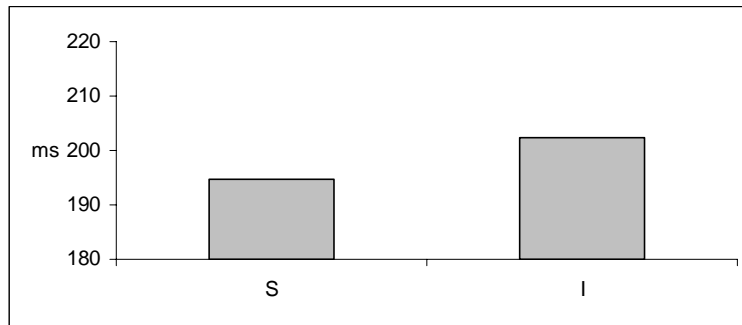


Fig. 50. Tempi di latenza del picco N200 in seguito alla presentazione di stimoli esprimenti sorpresa e stimoli incongruenti.

Tristezza/incongruente

Rispetto al confronto tristezza/incongruente si rileva soltanto, ancora una volta, il ritardo del picco in area anteriore ($F(3, N = 30) = 5,434 p = ,002$) (vedi tab. 11).

	G - I		P - I		R - I		S - I		T - I	
	M	ds	M	ds	M	ds	M	ds	M	ds
F	213,2	5,8	212,3	5,0	215,3	5,4	209,8	6,0	216,0	5,8
C	205,2	6,5	207,9	5,9	205,9	6,1	201,2	6,2	206,5	6,0
P	202,4	7,2	203,2	6,9	200,8	6,2	198,2	6,4	206,4	6,3
O	185,8	6,3	186,5	6,4	185,6	6,0	184,4	5,9	186,9	6,7

Tab. 11. Media e deviazione standard della latenza del picco N200 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

3.2.4. *Indice P200**Intensità**Confronto tra emozioni*

L'indice P200 presenta valori maggiormente accentuati nelle aree posteriori, e in particolar modo in zona parietale, come osservabile in tabella 12 ($F(3, N = 30) = 9,573 p < ,001$).

	M	ds
F	0,213	0,018
C	0,232	0,017
P	0,308	0,021
O	0,242	0,014

Tab. 12. Intensità dell'indice P200 nelle aree frontale, centrale, parietale e occipitale.

Questo dato è ancora più accentuato quando il task sperimentale viene lasciato implicito; in caso di task esplicito, invece, il picco risulta essere più ampio anteriormente ($F(3, N = 30) = 2,811 p = ,044$) (vedi fig. 51).

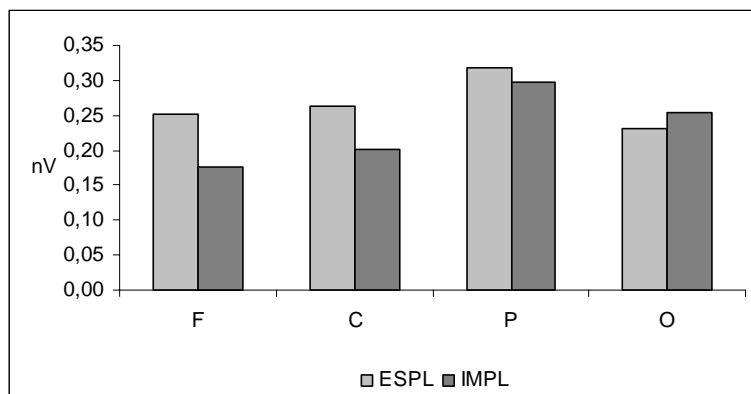


Fig. 51. Intensità del picco P200 nelle diverse aree corticali in presenza di consegna sperimentale esplicita e implicita.

Confronto tra le singole emozioni congruenti e la condizione incongruente

Disgusto/incongruente

Il picco P200 elicito dagli stimoli esprimenti disgusto e dagli stimoli incongruenti è più accentuato in area parietale ($F(3, N = 30) = 8,463 p < ,001$) (vedi tab. 13). Inoltre, l'interazione significativa condizione*localizzazione*task ($F(3, N = 30) = 3,205 p = ,027$) indica che l'intensità del picco diminuisce nelle aree anteriori soprattutto quando la consegna è implicita; sempre in caso di consegna implicita, in area posteriore si ha un'intensificazione del picco in seguito alla presentazione di stimoli incongruenti (vedi fig. 52).

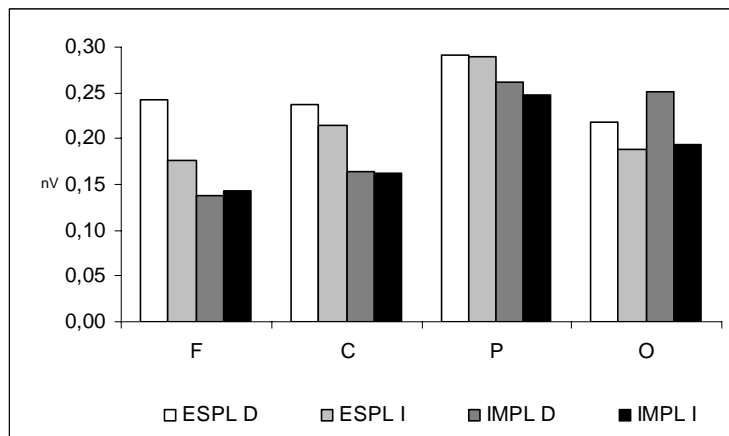


Fig. 51. Intensità del picco P200 nelle diverse aree corticali in presenza di consegna sperimentale esplicita e implicita, per gli stimoli esprimenti disgusto e per gli stimoli incongruenti.

Gioia/incongruente

Relativamente al confronto tra gioia e incongruenza, il picco si conferma prevalentemente parietale ($F(3, N = 30) = 6,638 p < ,001$) (vedi fig. 13), con valori più alti nel caso di stimoli esprimenti gioia ($F(1, N = 30) = 5,907 p = ,022$) (vedi fig. 52).

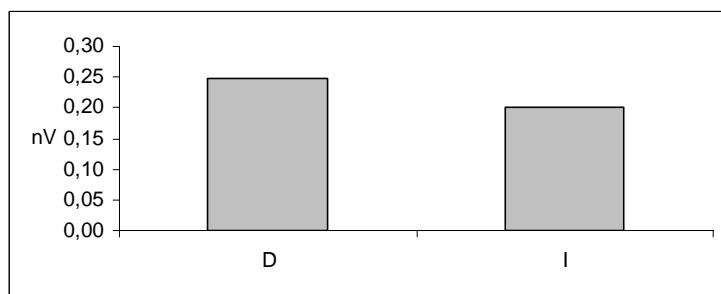


Fig. 52. Intensità del picco P200 in seguito alla presentazione di stimoli esprimenti gioia e coppie di stimoli incongruenti.

Come osservabile in figura 53, la triplice interazione condizione*localizzazione*task ($F(3, N = 30) = 2,399 p = ,055$) indica che questa intensificazione del picco in area parietale si verifica sia in caso di task esplicito sia in caso di task implicito; inoltre, in area frontale, il picco presenta un notevole aumento quando gli stimoli esprimono disgusto e il task è esplicito.

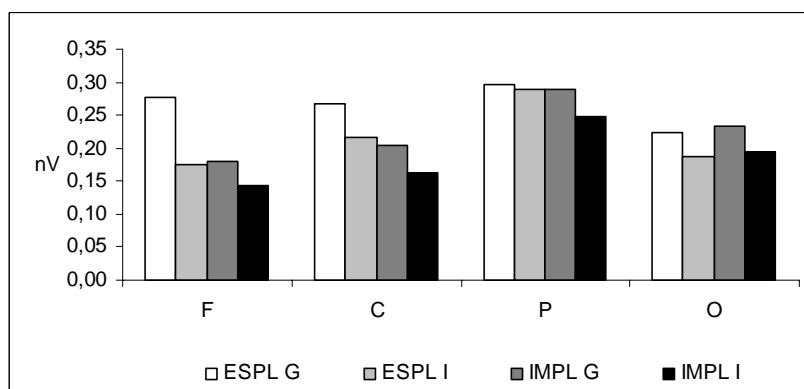


Fig. 53. Intensità della P200 in caso di stimoli esprimenti gioia e di stimoli incongruenti, nelle diverse aree corticali e con task esplicito e implicito.

Paura/incongruente

Il picco risulta essere maggiormente accentuato in caso di stimoli esprimenti paura rispetto che in caso di stimoli incongruenti ($F(1, N = 30) = 27,519 p < ,001$) (vedi fig. 54), oltre ad avere valori più elevati in area parietale ($F(3, N = 30) = 9,971 p < ,001$) (vedi tab. 13).

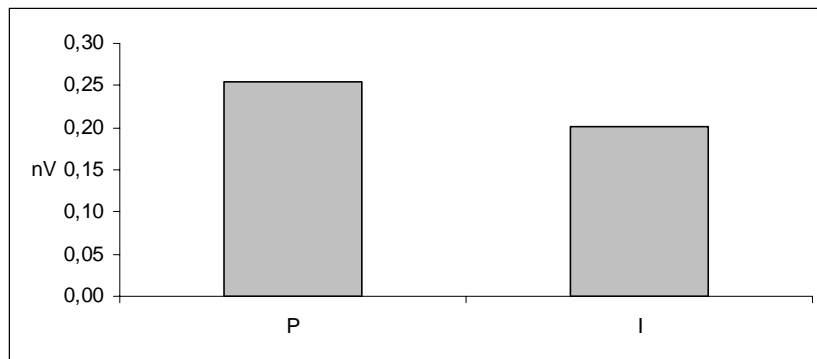


Fig. 54. Intensità del picco P200 in seguito alla presentazione di stimoli esprimenti paura e stimoli incongruenti.

Rabbia/incongruente

La P200 è più marcata in area parietale ($F(3, N = 30) = 11,307 p < ,001$) (vedi tab. 13) così come quando le coppie di stimoli esprimono rabbia ($F(1, N = 30) = 6,222 p = ,019$) (vedi fig. 55).

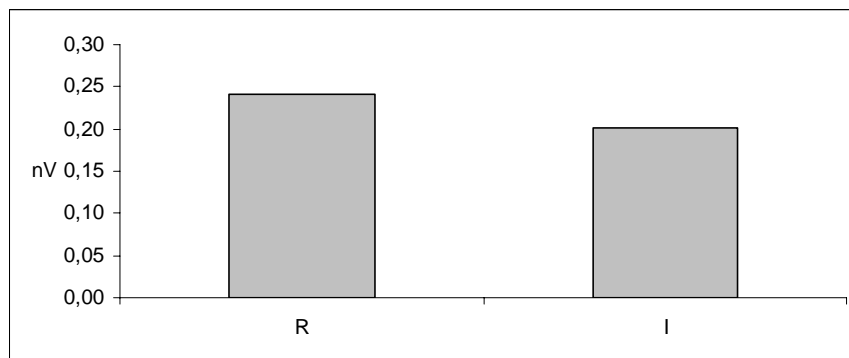


Fig. 55. Intensità del picco P200 elicitato da stimoli esprimenti rabbia e da stimoli incongruenti.

Sorpresa/incongruente

Gli stimoli esprimenti sorpresa provocano un picco più elevato rispetto agli stimoli incongruenti ($F(1, N = 30) = 19,112 p < ,001$) (vedi fig. 56). Il picco, in generale, è principalmente parietale ($F(3, N = 30) = 7,040 p < ,001$) (vedi tab. 13).

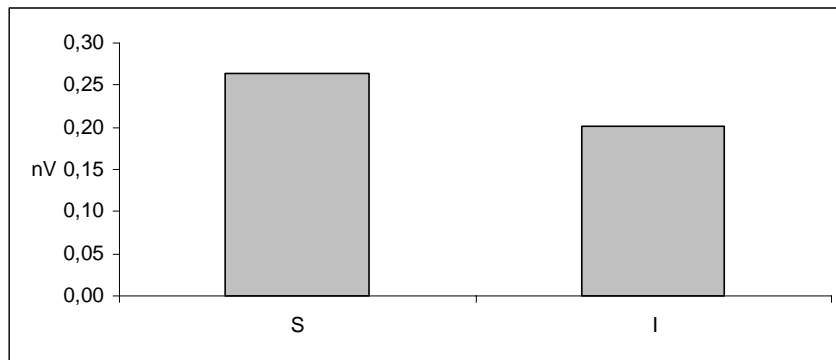


Fig. 56. Intensità del picco P200 per gli stimoli esprimenti sorpresa e gli stimoli incongruenti.

Infine, la triplice interazione condizione*localizzazione*task ($F(1, N = 30) = 4,339$ $p = ,007$) indica che, in area frontale e centrale, il picco presenta un notevole aumento quando gli stimoli esprimono sorpresa e il task è esplicito (vedi fig. 57).

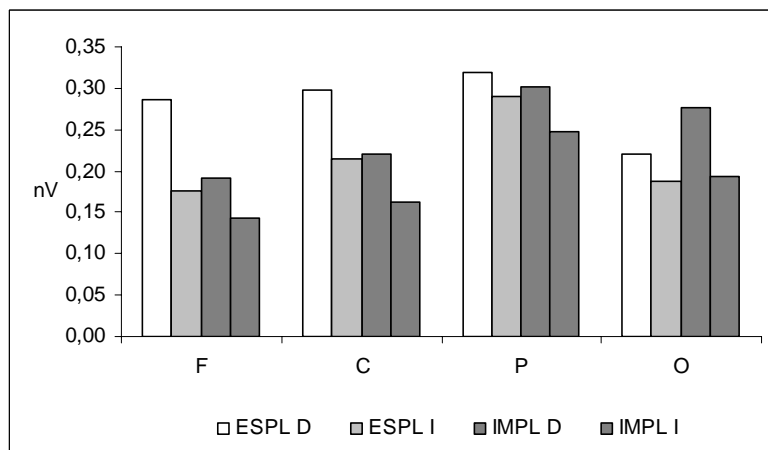


Fig. 57. Intensità del picco P200 in seguito alla presentazione di stimoli esprimenti sorpresa e stimoli incongruenti, in funzione della localizzazione e del tipo di task.

Tristezza/incongruente

Le coppie esprimenti tristezza elicitano un picco più ampio rispetto alle coppie incongruenti ($F(1, N = 30) = 24,650$ $p < ,001$) (vedi fig. 58). Il picco, inoltre, ha localizzazione parietale ($F(3, N = 30) = 12,415$ $p < ,001$) (vedi tab. 13).

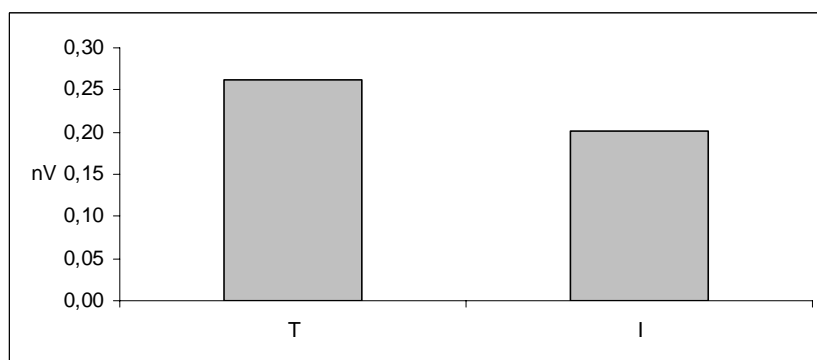


Fig. 58. Ampiezza della P200 elicitata da stimoli che esprimono tristezza e stimoli incongruenti.

	D		G		P		R		S		T	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	,175	,020	,194	,020	,189	,020	,178	,021	,199	,022	,184	,018
C	,195	,016	,213	,021	,207	,020	,206	,018	,224	,021	,216	,019
P	,273	,022	,281	,024	,296	,023	,288	,022	,289	,021	,303	,023
O	,212	,016	,210	,015	,219	,015	,213	,015	,220	,014	,224	,016

Tab. 13. Media e deviazione standard dei valori del picco P200 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

Latenza

Confronto tra emozioni

Dal confronto tra le emozioni congruenti emerge un ritardo dei tempi di latenza nelle aree posteriori rispetto a quelle anteriori ($F(3, N = 30) = 8,022$ $p < ,001$) (vedi tab. 14). L'effetto di interazione emozione*localizzazione ($F(3, N = 30) = 2,118$ $p = ,008$), come illustrato in figura 59, indica che tale ritardo è maggiormente accentuato nel caso di paura, rabbia, sorpresa e tristezza.

	M	ds
F	196,622	3,594
C	197,711	3,611
P	205,767	4,068
O	214,678	3,284

Tab 14. Media e deviazione standard della latenza del picco P200 nelle diverse aree corticali.

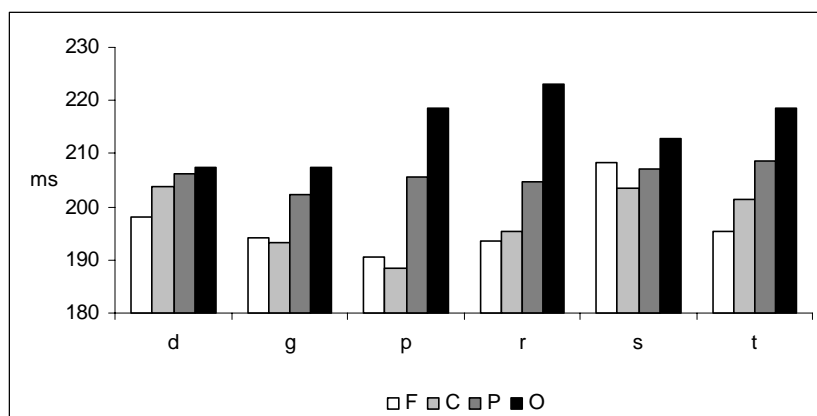


Fig. 59. Ampiezza della P200 elicitata da stimoli che esprimono tristezza e stimoli incongruenti.

Confronto tra le singole emozioni congruenti e la condizione incongruente

Disgusto/incongruente

Rispetto al confronto tra disgusto e incongruenza si riscontrano numerosi fenomeni. In primo luogo, il picco P200 risulta essere ritardato posteriormente (vedi tab. 15); inoltre, il disgusto ha tempi di latenza più ritardati rispetto all'incongruente ($F(1, N = 30) = 13,756 p < ,001$) (vedi fig. 60a). Dall'interazione tra condizione e localizzazione emerge che tale divario risulta pressoché annullato in area occipitale, dove anche nei casi di incongruenza i tempi si allungano ($F(3, N = 30) = 3,719 p = ,014$) (vedi fig. 60b). Il quadro si completa con la triplice interazione condizione*localizzazione*task ($F(3, N = 30) = 3,757 p = ,014$), che indica che i tempi più brevi si registrano in caso di stimoli incongruenti e task esplicito e che invece il picco più ritardato si manifesta in area occipitale per gli stimoli incongruenti in condizione di task implicito (vedi fig. 60 c).

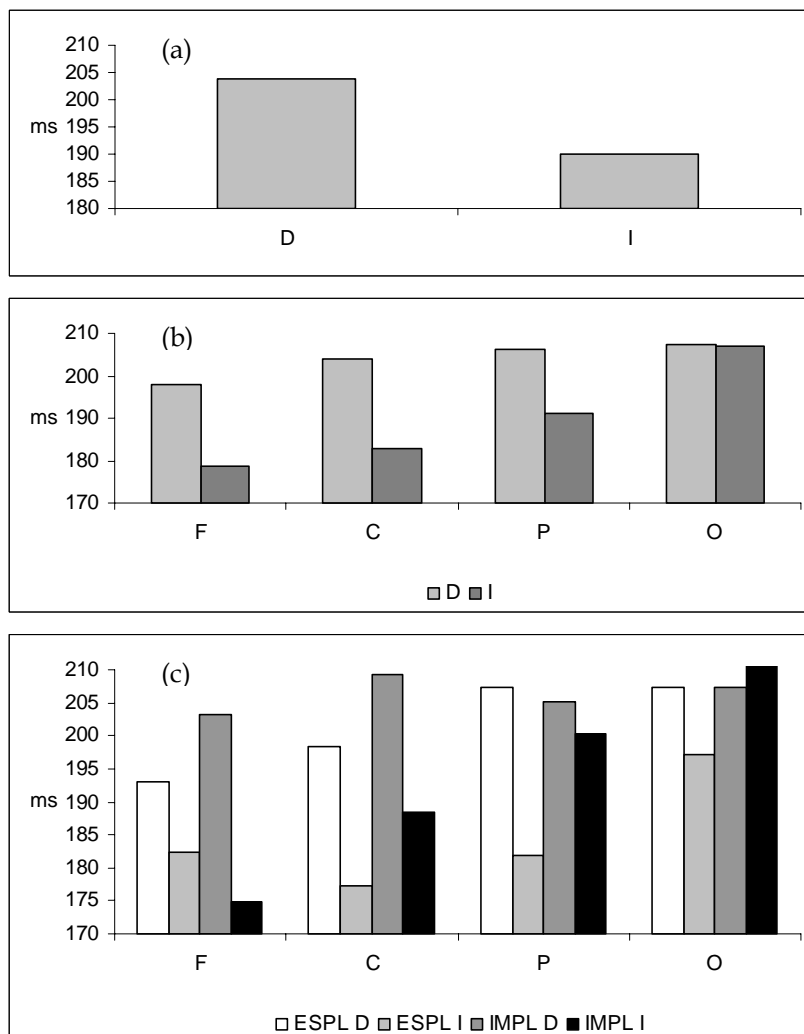


Fig. 60. Tempi di latenza della P200 elicitata da stimoli che esprimono disgusto e stimoli incongruenti (a), nelle diverse aree corticali (b) e in condizione di task esplicito e implicito (c).

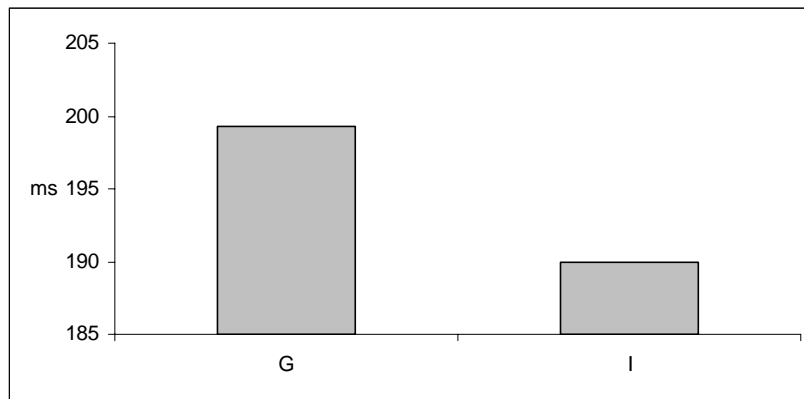


Fig. 61. Latenza della P200 elicitata da stimoli che esprimono gioia e da stimoli incongruenti.

Gioia/incongruente

La latenza dell'indice P200 è maggiore nelle aree posteriori ($F(3, N = 30) = 6,059 p < ,001$) (vedi tab. 15) e per gli stimoli esprimenti gioia rispetto a quelli incongruenti ($F(1, N = 30) = 11,749 p = ,002$) (vedi fig. 61).

Paura/incongruente

I tempi di latenza del picco P200 sono più rapidi anteriormente ($F(3, N = 30) = 12,243 p < ,001$) (vedi tab. 15) e in caso di stimoli incongruenti, se confrontati con quelli esprimenti paura ($F(1, N = 30) = 10,701 p = ,003$) (vedi fig. 62).

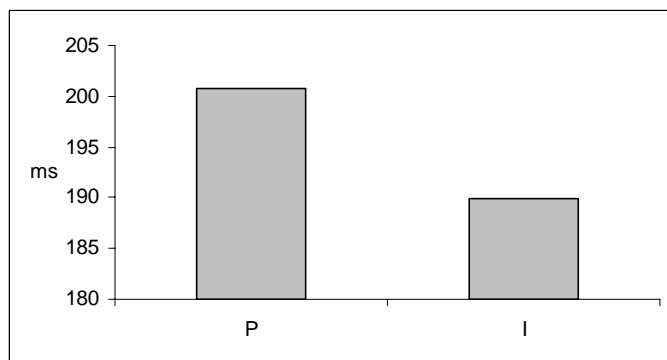


Fig. 62. Latenza del picco P200 elicitato da stimoli esprimenti paura e da stimoli incongruenti.

Rabbia/incongruente

Confrontando l'effetto degli stimoli esprimenti rabbia e degli stimoli incongruenti emerge che i tempi di latenza sono più ridotti nelle aree anteriori ($F(3, N = 30) = 12,438 p < ,001$) (vedi tab. 15), nella condizione di incongruenza ($F(1, N = 30) = 17,700 p = ,003$) (vedi fig. 63a) ed in caso di task esplicito ($F(1, N = 30) = 6,506 p = ,017$) (vedi fig. 63b).

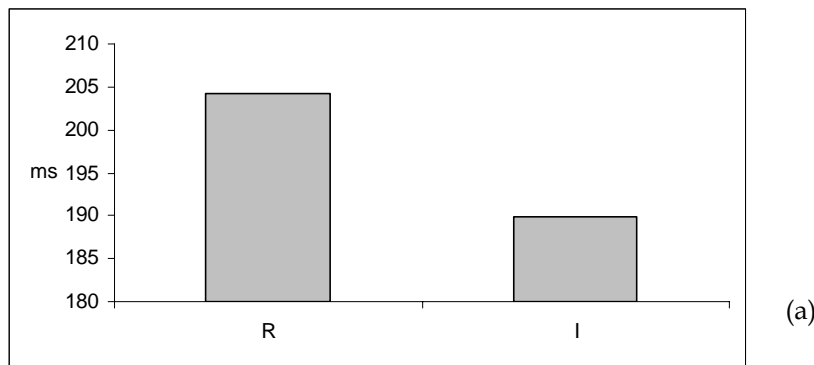
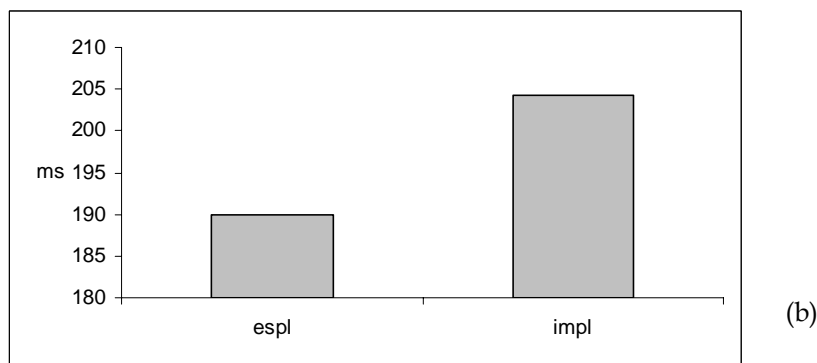


Fig. 63. Latenza del picco P200 con stimoli esprimenti rabbia e stimoli incongruenti (a) e in funzione del tipo di task (b).



Sorpresa/incongruente

L'indice P200 presenta tempi di latenza più ritardati nelle aree posteriori ($F(3, N = 30) = 4,746$ $p = ,004$) (vedi tab. 15) e in caso di stimoli congruenti esprimenti sorpresa anziché stimoli incongruenti ($F(1, N = 30) = 29,026$ $p < ,001$). Inoltre, come si può notare in figura 64a, è soprattutto nel caso degli stimoli incongruenti che si nota un progressivo aumento dei tempi di latenza con il passaggio dalle aree anteriori a quelle posteriori ($F(3, N = 30) = 3,557$ $p = ,018$) (vedi fig. 64b).

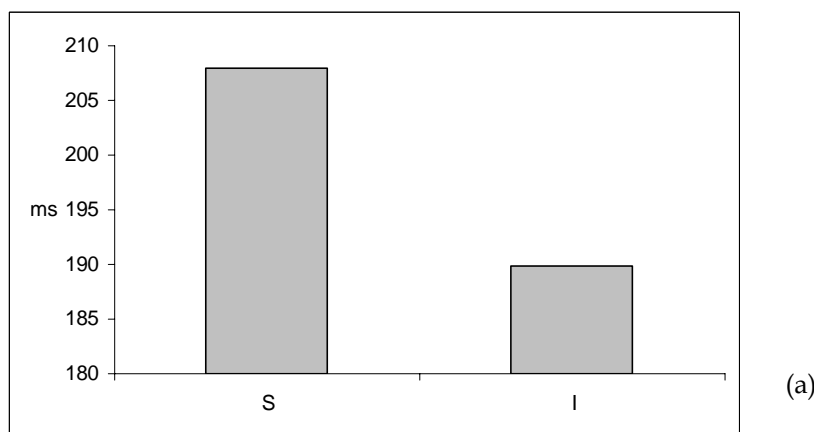
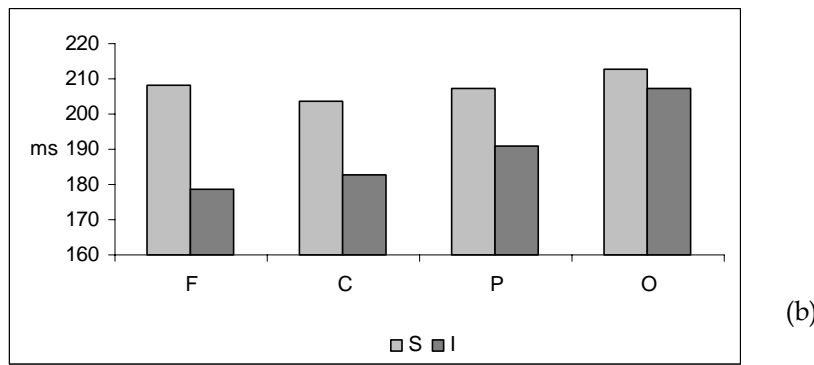


Fig. 64. Latenza del picco P200 in seguito alla presentazione di stimoli esprimenti rabbia e stimoli incongruenti (a) e in funzione dell'effetto congiunto del tipo di stimolo e della localizzazione (b).



Tristezza/incongruente

La P200 elicitata da stimoli esprimenti tristezza e stimoli incongruenti presenta tempi di latenza maggiori posteriormente ($F(3, N = 30) = 8,656 p = ,004$) (vedi tab. 15). Questi tempi, inoltre, sono più ritardati nel caso di presentazione di stimoli esprimenti tristezza ($F(1, N = 30) = 24,150 p < ,001$) (vedi fig. 65a) e quando il compito sperimentale è esplicito ($F(3, N = 30) = 4,451 p = ,004$) (vedi fig. 65b).

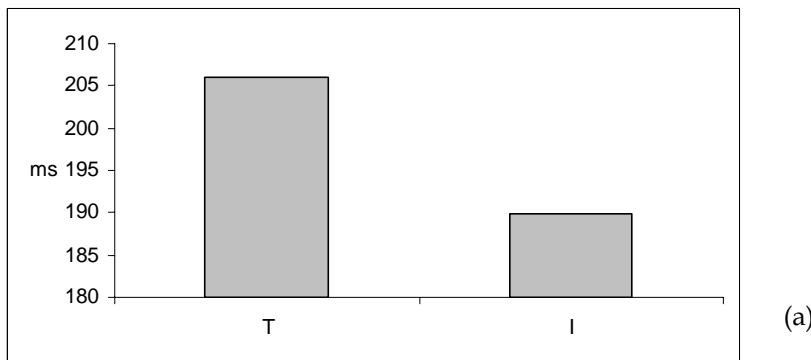
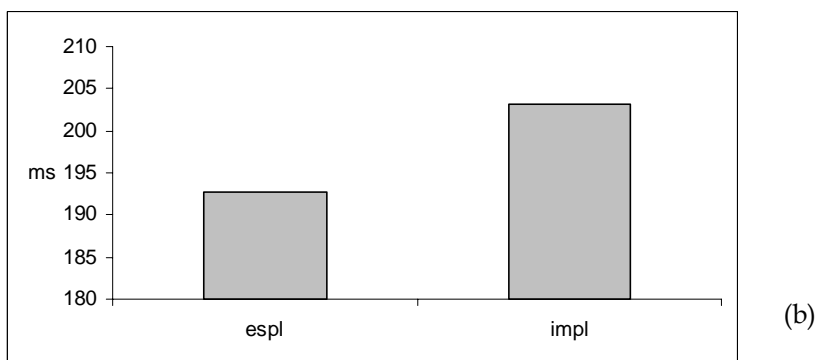


Fig. 65. Tempi di latenza del picco P200 in seguito alla presentazione di stimoli esprimenti tristezza e stimoli incongruenti (a) e in funzione del tipo di task (b).



	D - I		G - I		P - I		R - I		S - I		T - I	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	188,3	3,9	186,3	4,2	184,4	4,5	186,1	3,7	193,3	4,3	186,9	4,3
C	193,3	4,3	188,1	4,4	185,7	3,8	189,1	4,6	193,2	4,5	192,1	4,3
P	198,6	4,5	196,6	4,7	198,4	4,3	197,8	4,9	199,1	5,0	199,9	3,8
O	207,2	4,2	207,3	4,1	212,9	4,1	215,1	4,3	209,9	4,8	212,8	4,2

Tab. 15. Media e deviazione standard della latenza del picco P200 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

3.2.5. Indice N300

Intensità

Confronto tra emozioni

Confrontando tra loro le coppie di emozioni congruenti, il picco N300 si presenta come un fenomeno particolarmente accentuato nell'area frontale ($F(3, N = 30) = 4,004$ $p = ,010$) (vedi tab. 16).

	M	ds
F	-,170	,019
C	-,127	,014
P	-,131	,016
O	-,127	,012

Tab 16. Media e deviazione standard dell'intensità del picco N300 nelle diverse aree corticali.

Confronto tra le singole emozioni congruenti e la condizione incongruente

Disgusto/incongruente

Come già anticipato, la N300 si conferma come un fenomeno frontale ($F(3, N = 30) = 4,023$ $p = ,010$) (vedi tab. 17), oltre che più accentuato in seguito alla presentazione di stimoli esprimenti disgusto, se paragonati agli stimoli incongruenti ($F(1, N = 30) = 7,915$ $p = ,009$) (vedi fig. 66).

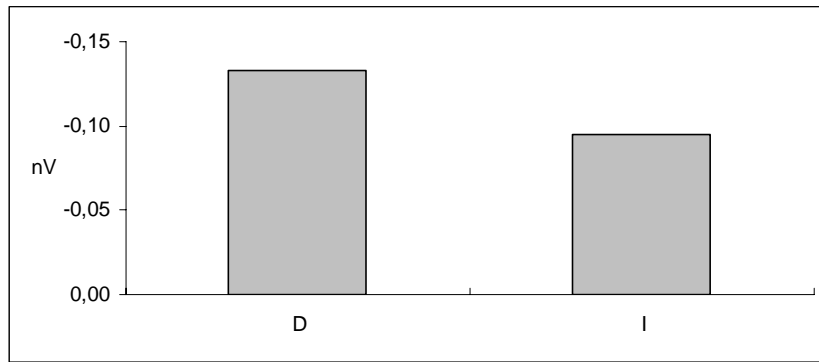
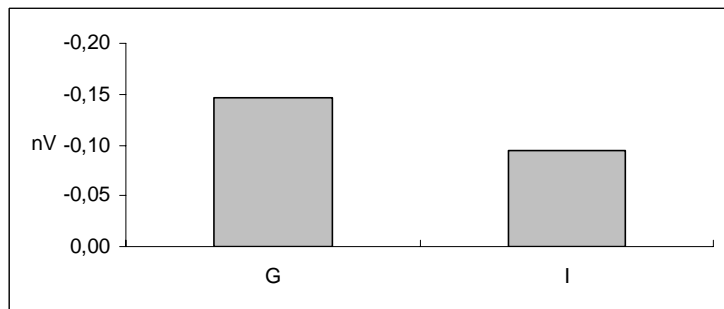


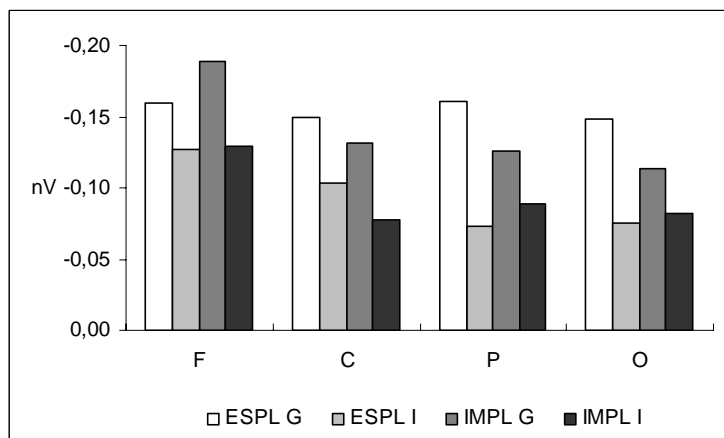
Fig. 66. Intensità del picco N300 in seguito alla presentazione di stimoli esprimenti disgusto e stimoli incongruenti.

Gioia/incongruente

La N300 è più intensa in area frontale ($F(3, N = 30) = 3,303 p = ,024$) (vedi tab. 17) e in presenza di stimoli esprimenti gioia anziché stimoli incongruenti ($F(1, N = 30) = 8,370 p = ,007$) (vedi fig. 67a). Inoltre l'effetto congiunto della condizione, della localizzazione e del task ($F(3, N = 30) = 2,904 p = ,040$) indica che, nelle diverse aree, gli stimoli che esprimono gioia in condizione di task esplicito presentano valori molto elevati; inoltre, in area frontale, sono soprattutto gli stimoli esprimenti gioia ma in condizione di task implicito a provocare un notevole incremento dell'intensità del picco (vedi fig. 67b).



(a)



(b)

Fig. 67. Intensità del picco N300 in seguito alla presentazione di stimoli esprimenti gioia e stimoli incongruenti (a) e in funzione del tipo di stimolo, della localizzazione e

Paura/incongruente

Il picco N300, ancora una volta, appare essere un fenomeno frontale ($F(3, N = 30) = 6,366 p = ,001$), come riportato in tabella 17. Inoltre, i valori aumentano per gli stimoli esprimenti paura rispetto che per gli stimoli incongruenti ($F(1, N = 30) = 13,655 p = ,001$) (vedi fig. 68).

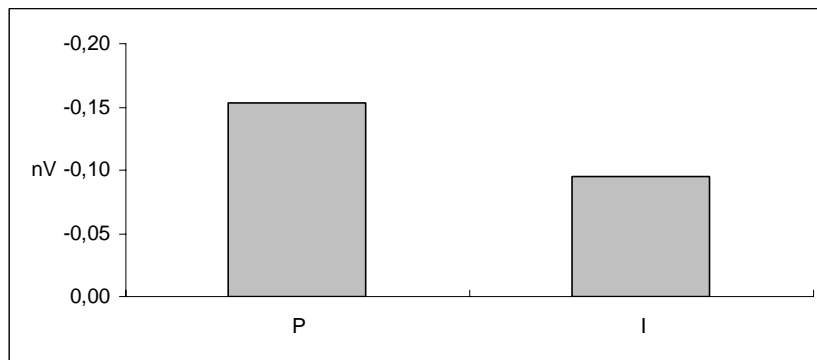


Fig. 68. Valori del picco N300 elicitato dalla presentazione di stimoli esprimenti paura e stimoli incongruenti.

Rabbia/incongruente

Il picco N300 si conferma come un fenomeno frontale ($F(3, N = 30) = 3,987 p = ,010$), come riportato in tabella 17. Inoltre, l'intensità aumenta per gli stimoli esprimenti rabbia rispetto che per gli stimoli incongruenti ($F(1, N = 30) = 7,098 p = ,012$) (vedi fig. 69).

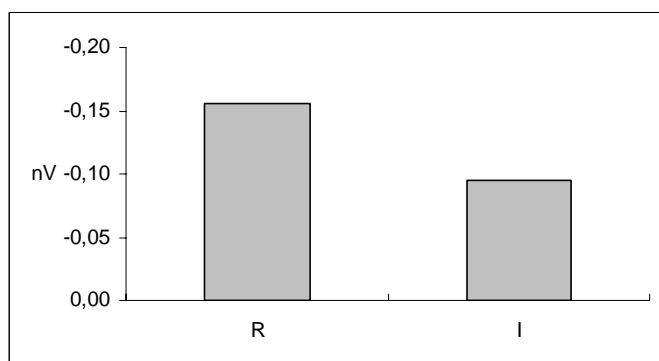


Fig. 69. Intensità della N300 elicitata da stimoli esprimenti rabbia e stimoli incongruenti.

Sorpresa/incongruente

L'effetto condizione*localizzazione ($F(1, N = 30) = 3,496 p = ,019$) rivela che gli stimoli incongruenti producono un picco maggiore nell'area frontale, con un progressivo decremento nel passaggio alle aree posteriori; nel caso degli stimoli esprimenti

sorpresa invece il picco presenta valori elevati sia in area frontale sia in area occipitale (vedi fig. 70).

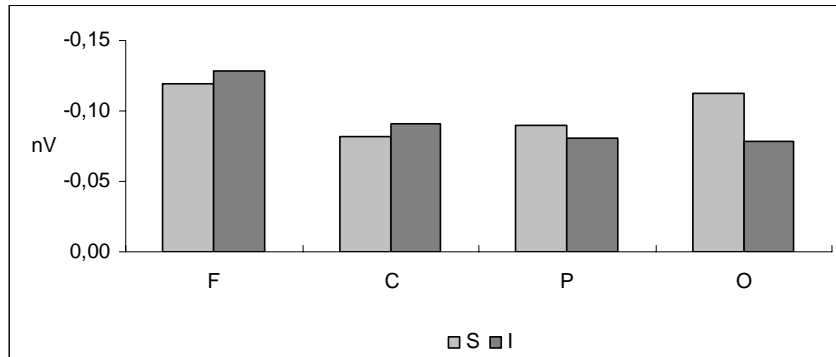


Fig. 70. Intensità della N300 elicitata da stimoli esprimenti sorpresa e stimoli incongruenti nelle diverse aree corticali.

Tristezza/incongruente

Ancora una volta, le analisi indicano la N300 come un fenomeno frontale (F(1, N = 30) = 3,266 p = ,026) (vedi tab. 17) e più accentuato nel caso degli stimoli congruenti (in questo caso esprimenti tristezza) che degli stimoli incongruenti (F(1, N = 30) = 11,953 p = ,002) (vedi fig. 71).

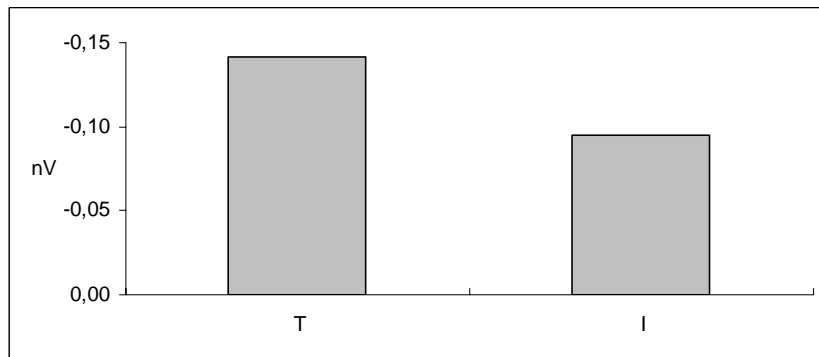


Fig. 71. Intensità dell'indice N300 in presenza di stimoli esprimenti tristezza e stimoli incongruenti.

	D - I		G - I		P - I		R - I		S - I	
	M	ds	M	ds	M	ds	M	ds	M	ds
F	-,144	,016	-,151	,017	-,165	,015	-,160	,022	-,151	,019
C	-,106	,013	-,116	,015	-,119	,015	-,116	,019	-,110	,015
P	-,104	,016	-,112	,016	-,107	,016	-,115	,018	-,110	,017
O	-,101	,012	-,105	,012	-,105	,013	-,111	,012	-,102	,013

Tab. 17. Media e deviazione standard del picco N300 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

*Latenza**Confronto tra emozioni*

Il confronto tra le emozioni congruenti indica che i tempi di latenza della N300, in generale, sono più anticipati frontalmente ($F(3, N = 30) = 38,719 p < ,001$), come rappresentato in tabella 18.

	M	DS
F	288,5	3,105
C	296,6	3,569
P	315,3	4,147
O	320,7	3,306

Tab 18. Media e deviazione standard della latenza del picco N300 nelle diverse aree corticali.

Disgusto/incongruente

Le analisi condotte confermano l'allungamento dei tempi di latenza in zona posteriore ($F(3, N = 30) = 17,388 p < ,001$) (vedi tab. 19). Oltre a ciò, va rilevato che i tempi più lunghi si registrano in caso di consegna sperimentale implicita, sempre posteriormente ($F(3, N = 30) = 17,388 p = ,051$) (vedi fig. 72).

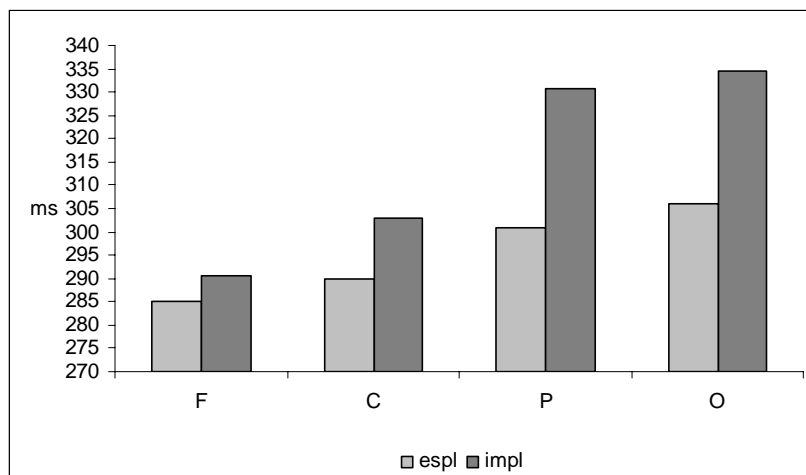


Fig. 72. Latenza dell'indice N300 in presenza di stimoli esprimenti disgusto e stimoli incongruenti.

Gioia/incongruente

Oltre a indicare il ritardo dei tempi di latenza nelle zone corticali posteriori ($F(3, N = 30) = 22,374 p < ,001$) (vedi tab. 19), l'ANOVA condotta mostra un ritardo per gli stimoli incongruenti quando il task è implicito ($F(3, N = 30) = 4,883 p = ,035$), come illustrato in figura 73.

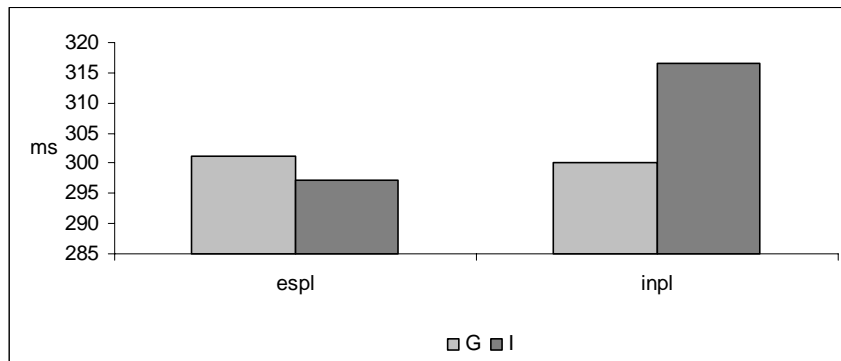


Fig. 73. Tempi di latenza dell'indice N300 per gli stimoli esprimenti disgusto e gli stimoli incongruenti, in condizione di task esplicito e implicito.

Paura/incongruente

Anche in questo caso i tempi più rapidi sono localizzati anteriormente ($F(3, N = 30) = 22,656$ $p < ,001$) (vedi tab. 19). Rispetto alla tipologia di task, la latenza aumenta quando esso è implicito ($F(1, N = 30) = 8,855$ $p = ,006$) (esplicito: $M = 294,217$; $ds = 5,200$; implicito: $M = 316,100$; $ds = 5,200$). Si registra inoltre una triplice interazione condizione*localizzazione*task ($F(3, N = 30) = 3,285$ $p = ,025$), rappresentata in figura 74: il dato più interessante è il fatto che i tempi più lunghi sono prodotti nelle aree parietale e occipitale quando la consegna sperimentale rimane implicita.

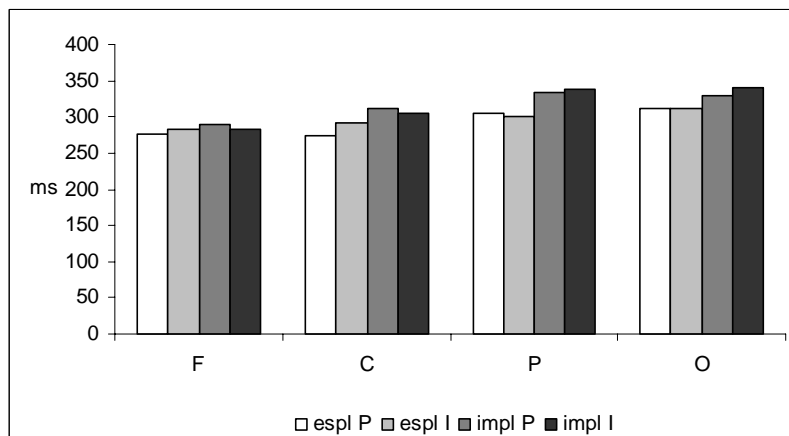
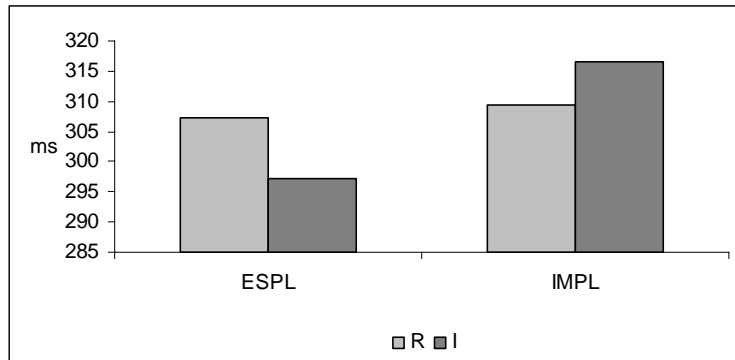


Fig. 74. Latenza dell'indice N300 in risposta a stimoli esprimenti paura e stimoli incongruenti, nelle diverse aree corticali e con consegna sperimentale esplicita e implicita.

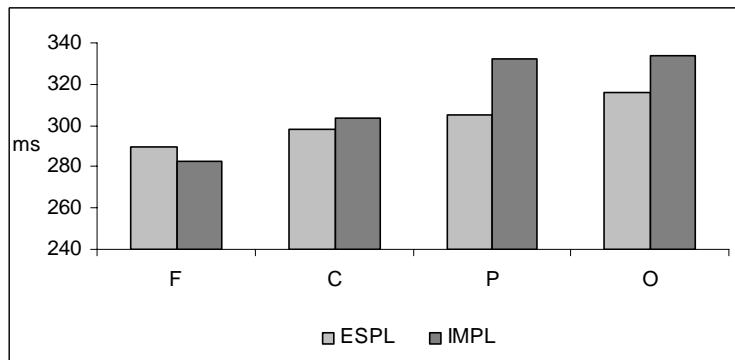
Rabbia/incongruente

In generale, i tempi di latenza sono più rapidi anteriormente ($F(3, N = 30) = 23,301$ $p < ,001$) (tab. 19). Più nello specifico, si registra in primo luogo un'interazione tra la condizione e il task ($F(3, N = 30) = 4,211$ $p = ,050$): a fronte di una certa stabilità della latenza in caso di stimoli esprimenti rabbia, gli stimoli incongruenti producono tempi

molto più ritardati nella condizione implicita rispetto a quella esplicita (fig. 75a). L'effetto localizzazione*task ($F(3, N = 30) = 4,663$ $p = ,005$) infine indica che, pur verificandosi in generale un ritardo della latenza nelle aree posteriori, tale ritardo risulta essere molto più marcato quando il task sperimentale è implicito (fig. 75b).



(a)



(b)

Fig. 75. Tempi di latenza del picco N300 in presenza di stimoli che esprimono rabbia e stimoli incongruenti, con consegna sperimentale esplicita e implicita (a) e nelle diverse aree corticali, in funzione del tipo di task (b).

Sorpresa/incongruente

Rispetto al confronto sorpresa/incongruente si rileva soltanto il ritardo in area posteriore ($F(3, N = 30) = 25,027$ $p < ,001$) (vedi tab. 1^o) (b)

Tristezza/incongruente

Ancora una volta, le analisi indicano che la latenza è più rapida in area anteriore ($F(3, N = 30) = 32,846$ $p < ,001$) (vedi fig. 19).

	D - I		G - I		P - I		R - I		S - I		T - I	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	287,9	4,4	284,5	4,1	283,1	4,0	286,0	3,8	288,9	4,5	285,8	2,9
C	296,4	4,8	295,1	5,1	295,0	5,8	300,8	5,4	299,6	5,7	296,0	5,2
P	315,7	5,5	314,5	6,0	318,9	6,3	318,8	5,7	317,0	5,5	319,1	5,3
O	320,3	5,3	320,8	4,9	323,5	5,2	324,7	4,1	326,1	4,2	326,8	3,8

Tab. 19. Media e deviazione standard della latenza picco N300 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

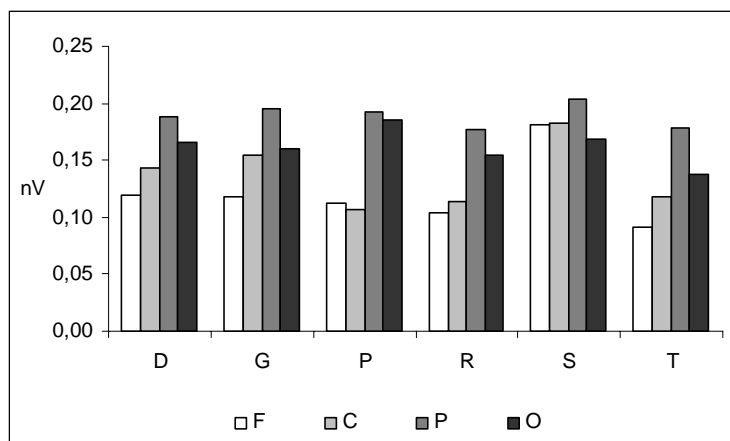
3.2.6. Indice P300

Intensità

Confronto tra emozioni

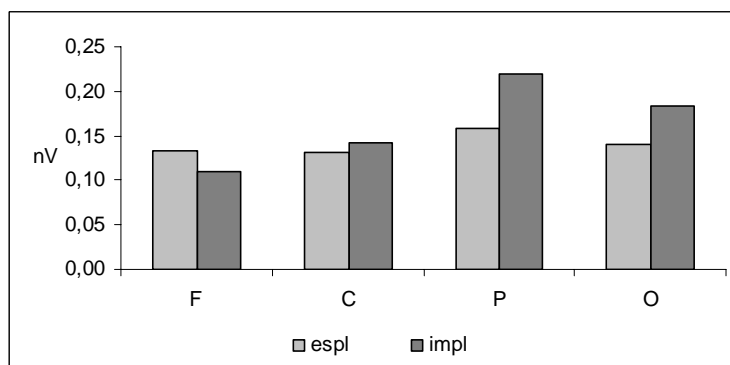
La P300 è un fenomeno prevalentemente posteriore, e più nello specifico parietale ($F(3, N = 30) = 11,369 p < ,001$) (vedi tab. 20). Nella figura 76a è possibile osservare l'effetto condizione*localizzazione ($F(3, N = 30) = 2,778 p < ,001$): per tutte le emozioni il picco presenta valori elevati in area parietale e in misura minore occipitale, mentre nelle zone frontali e centrali si rileva una maggiore differenziazione tra le sei emozioni, con un decremento dell'intensità di picco nel caso di tristezza, paura e rabbia. Infine, l'effetto localizzazione*task ($F(3, N = 30) = 4,871 p = ,004$) indica che, in caso di task implicito, nelle aree parietali e occipitali il fenomeno P300 risulta molto più accentuato in termini di intensità (vedi fig. 76b).

	M	ds	
F	,121	,014	Tab 20. Media e deviazione standard dell'ampiezza del picco P300 nelle diverse aree corticali.
C	,137	,014	
P	,189	,016	
O	,162	,010	



(a)

Fig. 76. Intensità del fenomeno P300 come effetto all'interazione dei fattori localizzazione e condizione (a) e localizzazione e task (b).



(b)

Confronto tra le singole emozioni congruenti e la condizione incongruente

Disgusto/incongruente

La localizzazione del picco P300 è parietale ($F(3, N = 30) = 8,964 p < ,001$) (vedi tab. 21); inoltre, il fenomeno è più intenso in seguito alla presentazione di stimoli congruenti esprimenti disgusto che di stimoli incongruenti ($F(1, N = 30) = 20,642 p < ,001$), come rappresentato in fig. 77.

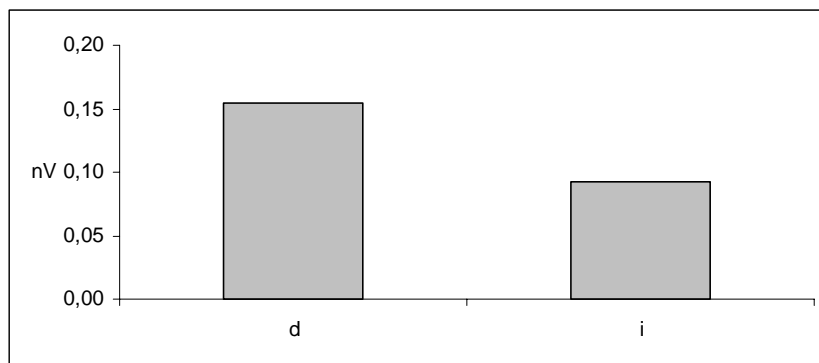


Fig. 77. Intensità del fenomeno P300 come effetto della presentazione di stimoli esprimenti disgusto e di stimoli incongruenti.

Gioia/incongruente

Anche in questo caso, la localizzazione del picco P300 è parietale ($F(3, N = 30) = 7,940 p < ,001$) (vedi tab. 21), con una più elevata intensità in caso di stimoli congruenti esprimenti gioia che di stimoli incongruenti ($F(1, N = 30) = 20,642 p < ,001$) ($F(1, N = 30) = 12,558 p = ,001$) (vedi fig. 78a). Si nota inoltre in figura 78a che, a prescindere dalla tipologia dello stimolo, i valori di picco più elevati si registrano in caso di task implicito in area parietale e occipitale; in area frontale invece il task esplicito produce un picco più ampio di quello prodotto dalla consegna implicita ($F(3, N = 30) = 3,666 p = ,015$).

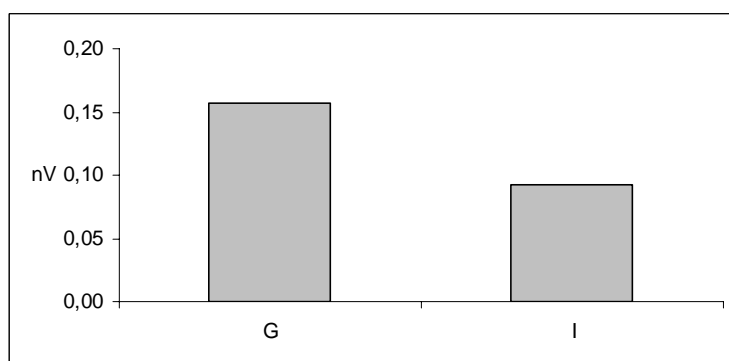
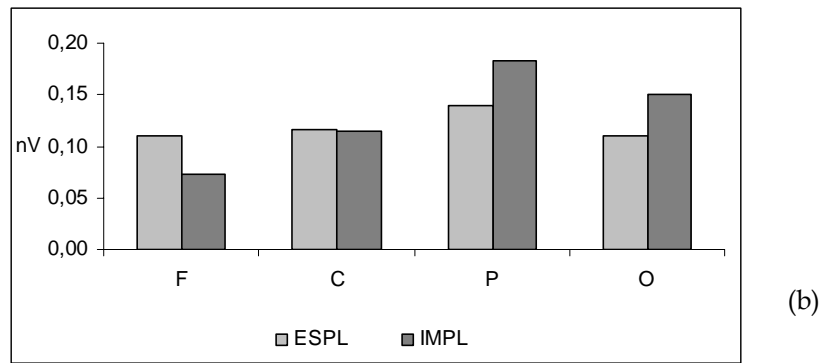


Fig. 78. Intensità del fenomeno P300 elicitato da stimoli esprimenti gioia e stimoli incongruenti (a) e nelle diverse aree corticali in condizione di task esplicito e implicito (b).



Paura/incongruente

In generale, i valori di picco più elevati sono localizzati in area parietale ($F(3, N = 30) = 11,655$ $p < ,001$) (vedi tab. 21) e sono elicitati dalle coppie di stimoli esprimenti paura, rispetto a quelle incongruenti ($F(1, N = 30) = 14,933$ $p = ,001$) (vedi fig. 79a). Ciò avviene in tutte le aree, ma soprattutto in quelle parietali e occipitali ($F(1, N = 30) = 4,354$ $p = ,007$) (vedi fig. 79b). Sempre posteriormente, è il task implicito a produrre i valori di picco più elevati, mentre nelle aree anteriori ciò avviene, anche se in misura minore, per il task esplicito (vedi fig. 79c). Infine, l'interazione condizione*localizzazione*task ($F(1, N = 30) = 3,368$ $p = ,022$) rivela che, se in generale sono gli stimoli esprimenti paura a produrre i picchi più ampi, più nello specifico questo avviene nelle aree anteriori quando il task è esplicito e nelle aree posteriori quando è implicito (vedi fig. 79d).

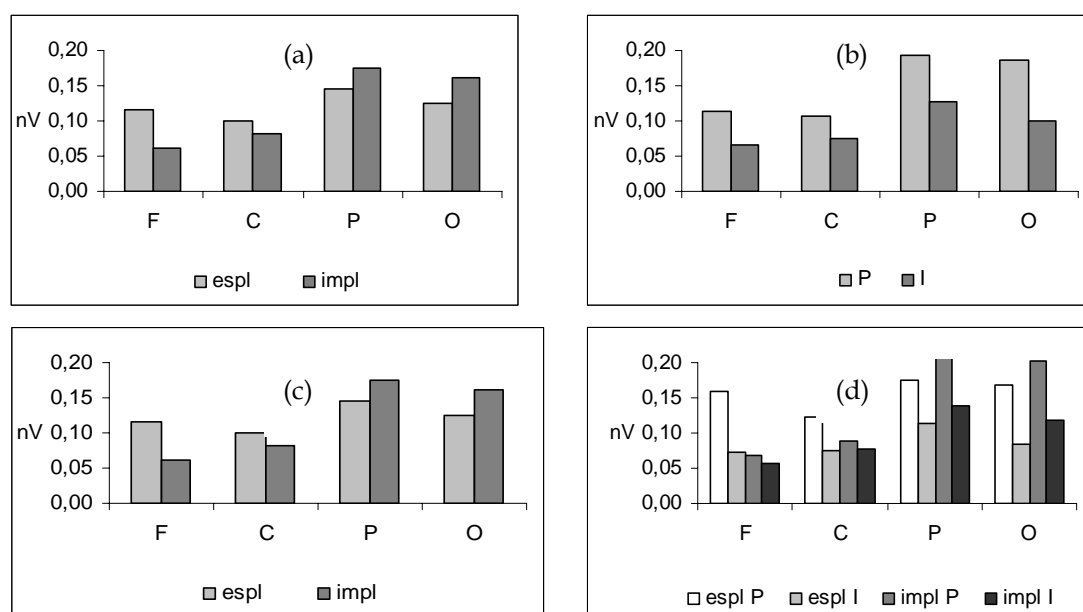


Fig. 79. Confronto paura/incongruente: i grafici illustrano gli effetti condizione (a), condizione*localizzazione (b), localizzazione*task (c) e condizione*localizzazione*task.

Rabbia/incongruente

Confrontando la P300 prodotta da stimoli che esprimono rabbia e stimoli incongruenti emergono la localizzazione parietale del picco ($F(3, N = 30) = 11,941 p < ,001$) (vedi tab. 21) e la maggiore intensità quando le coppie sono congruenti ($F(1, N = 30) = 17,608 p < ,001$), come rappresentato in fig. 80.

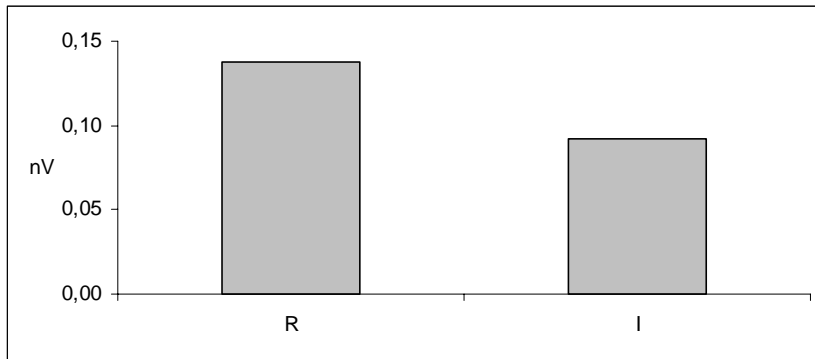
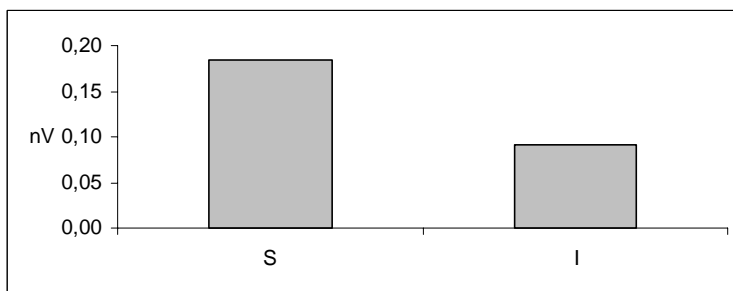


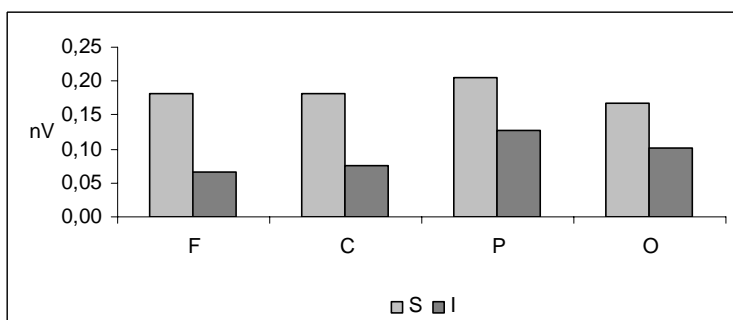
Fig. 80. Intensità del picco P300 elicitato da stimoli esprimenti rabbia e stimoli incongruenti.

Sorpresa/incongruente

Ancora una volta, il picco si conferma prevalentemente parietale ($F(3, N = 30) = 3,358 p = ,022$) (vedi tab. 21). Inoltre, gli stimoli esprimenti sorpresa producono un picco più elevato degli stimoli incongruenti ($F(1, N = 30) = 20,956 p < ,001$) (vedi fig. 81a). Tale divario risulta essere particolarmente accentuato nella aree anteriori ($F(3, N = 30) = 5,220 p = ,002$) (vedi fig. 81b).



(a) Fig. 81. Intensità del picco P300 in seguito alla presentazione di stimoli esprimenti sorpresa e stimoli incongruenti (a) e come effetto dell'interazione condizione*task (b).



(b)

Tristezza/incongruente

La localizzazione del picco è parietale ($F(3, N = 30) = 10,581 p <,022$) (vedi tab. 21) e l'intensità è maggiormente accentuata quando gli stimoli esprimono tristezza ($F(1, N = 30) = 6,914 p =,014$), come rappresentato in fig. 82.

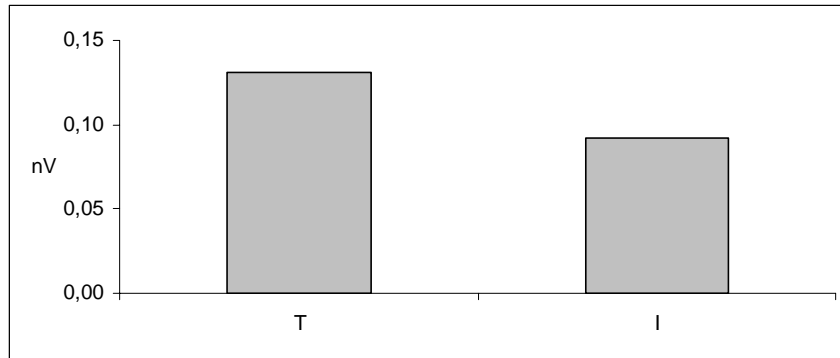


Fig. 82. Ampiezza del picco P300 in seguito alla presentazione di stimoli esprimenti tristezza e stimoli incongruenti.

	D-I		G-I		P-I		R-I		S-I		T-I	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	,092	,013	,092	,016	,089	,014	,084	,009	,123	,016	,078	,013
C	,109	,014	,115	,017	,091	,015	,095	,014	,129	,016	,096	,014
P	,157	,018	,161	,022	,159	,019	,152	,018	,165	,020	,153	,018
O	,134	,013	,130	,014	,143	,012	,128	,011	,135	,013	,120	,011

Tab. 21. Media e deviazione standard dell'intensità picco N300 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

Latenza

Confronto tra emozioni

Rispetto al picco P300, dal confronto tra le diverse emozioni congruenti emerge unicamente la presenza di tempi più brevi nella aree anteriori, a fronte di un ritardo posteriore della comparsa del picco ($F(3, N = 30) = 33,013 p <,001$) (vedi tab. 22).

	M	ds
F	317,27	3,83
C	305,86	3,61
P	290,58	3,59
O	282,80	2,91

Tab 22. Media e deviazione standard dei tempi di latenza del picco P300 nelle diverse aree corticali.

Confronto tra le singole emozioni congruenti e la condizione incongruente

Disgusto/incongruente

Confrontando l'effetto degli stimoli esprimenti disgusto e degli stimoli incongruenti, emerge un ritardo del picco P300 nelle aree anteriori ($F(3, N = 30) = 30,867 p < ,001$) (vedi tab. 23). Più nello specifico, questo ritardo avviene soprattutto quando gli stimoli sono incongruenti; nelle aree posteriori, invece, sono gli stimoli esprimenti disgusto a produrre tempi più lenti ($F(3, N = 30) = 2,817 p < ,044$) (vedi fig. 83).

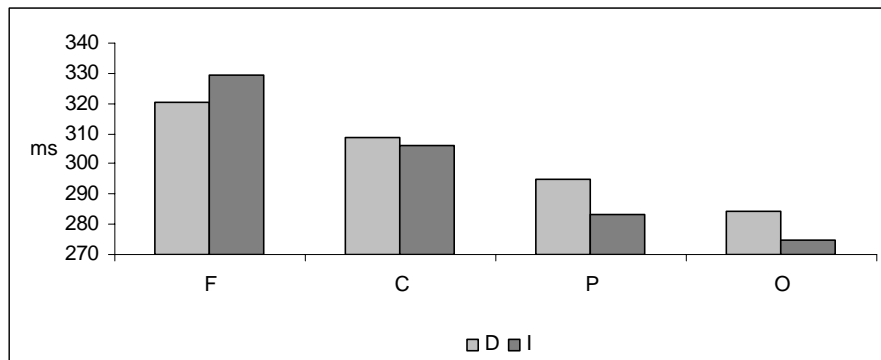


Fig. 83. Latenza del picco P300 in seguito alla presentazione di stimoli esprimenti disgusto e stimoli incongruenti.

Gioia/incongruente

Anche da questa analisi emerge l'allungamento dei tempi in area anteriore ($F(3, N = 30) = 24,629 p < ,001$) (vedi tab. 23) e in caso di congruenza degli stimoli ($F(1, N = 30) = 4,752 p = ,038$) (vedi fig. 84a). L'interazione condizione**localizzazione ($F(3, N = 30) = 5,007 p = ,003$) indica tuttavia che questo trend risulta invertito nella zona frontale, dove i tempi più lunghi sono prodotti dagli stimoli incongruenti (vedi fig. 84b). Infine, in figura 84c si può osservare che sempre in area frontale è il task implicito a portare ad un allungamento della latenza, al contrario di quanto accade nella corteccia centrale, parietale e occipitale ($F(3, N = 30) = 3,012 p = ,035$).

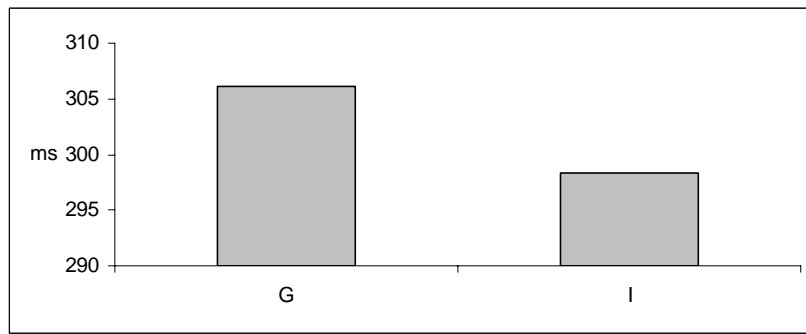
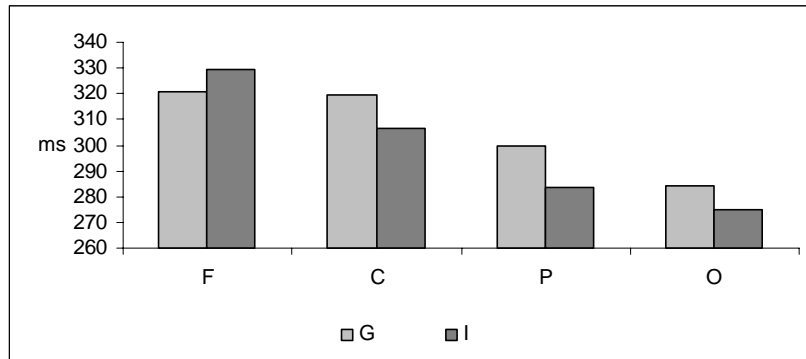
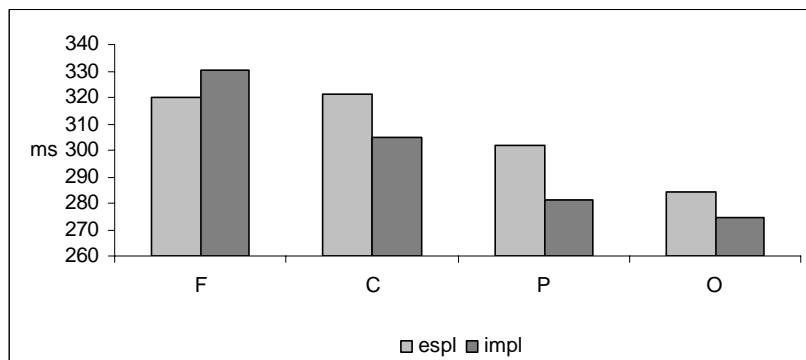


Fig. 84. Latenza del picco P300 rispetto agli effetti condizione (a), localizzazione (b) e localizzazione*task (c).



(b)



(c)

Paura/incongruente

Ancora una volta, i tempi di latenza più rapidi sono localizzati posteriormente ($F(3, N = 30) = 23,546 p < ,001$) (vedi tab. 23). In area frontale è il task implicito a portare ad un allungamento della latenza, al contrario di quanto accade nella corteccia centrale, parietale e occipitale ($F(3, N = 30) = 3,403 p = ,021$) (vedi fig. 85).

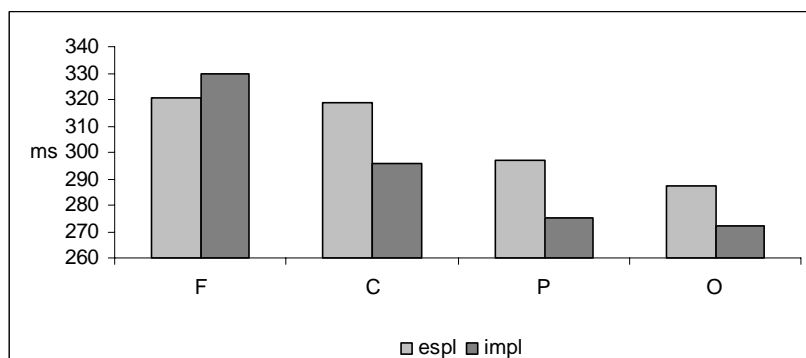


Fig. 85. Latenza della P300 rispetto alla localizzazione e al task.

Rabbia/incongruente

Anche dal confronto tra rabbia e incongruente si evince che i tempi di latenza più rapidi sono localizzati posteriormente ($F(3, N = 30) = 20,431 p < ,001$) (vedi tab. 23). In area frontale e, anche se in misura minore, centrale, i tempi più lunghi sono prodotti dagli stimoli incongruenti ($F(3, N = 30) = 4,896 p = ,003$), come si può osservare in figura 86.

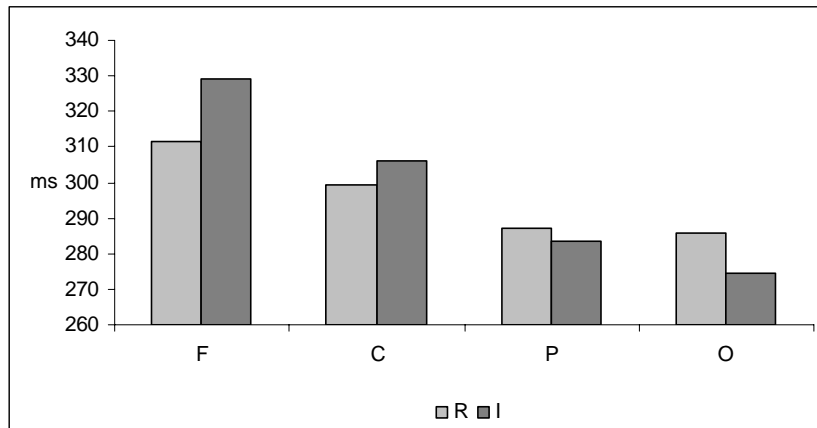


Fig. 86. Latenza della P300 rispetto alla localizzazione e al tipo di stimolo per il confronto rabbia/incongruente.

Sorpresa/incongruente

I tempi di latenza più rapidi sono localizzati posteriormente ($F(3, N = 30) = 19,893 p < ,001$) (vedi tab. 23). I tempi più lunghi sono prodotti dagli stimoli incongruenti in area frontale e, anche se in misura minore, centrale, ($F(3, N = 30) = 4,650 p = ,005$) (vedi fig. 87).

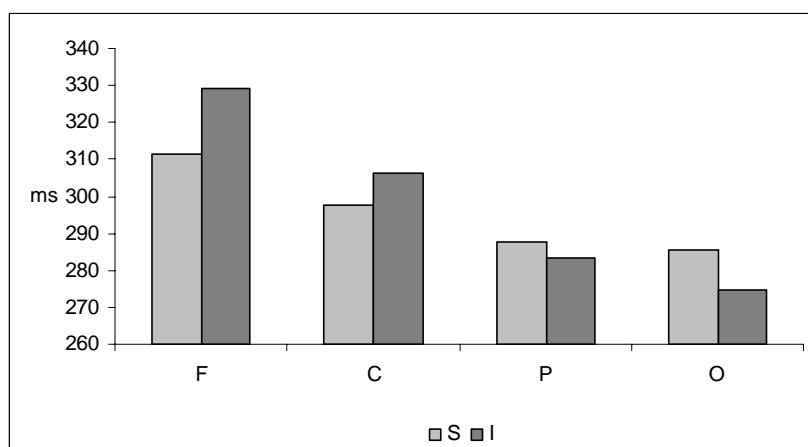


Fig. 87. Latenza della P300 rispetto alla localizzazione e al tipo di stimolo per il confronto sorpresa/incongruente.

Tristezza/incongruente

Dal confronto tra tristezza e condizione incongruente emerge soltanto che i tempi di latenza più rapidi sono localizzati posteriormente ($F(3, N = 30) = 39,539 p < ,001$) (vedi tab. 23).

	D - I		G - I		P - I		R - I		S - I		T - I		Tab. 23. Media e deviazione standard dei tempi di latenza del picco
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds	
F	324,8	5,1	325,2	5,4	325,0	5,0	320,5	5,4	320,4	5,6	323,7	5,1	
C	307,4	5,5	312,9	5,4	307,1	5,9	302,7	5,6	302,0	6,1	304,1	5,8	
P	289,1	4,3	291,4	5,0	286,1	5,2	285,2	4,8	285,5	5,0	284,6	5,7	
O	279,3	4,3	279,5	4,3	279,8	3,9	280,1	4,0	280,1	4,1	273,2	3,5	

N300 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

*3.2.7. Indice N400**Intensità**Confronto tra emozioni*

Dal confronto tra le emozioni emerge la prevalente localizzazione posteriore, ed in particolare parietale, del picco N400 ($F(3, N = 30) = 16,488 p < ,001$) (vedi tab. 24).

	M	ds	Tab 24. Media e deviazione standard dell'intensità del picco N400 nelle diverse aree corticali.
F	-,140	,014	
C	-,141	,011	
P	-,204	,014	
O	-,196	,011	

*Confronto tra le singole emozioni congruenti e la condizione incongruente**Disgusto/incongruente*

Dall'analisi emerge la localizzazione posteriore del picco ($F(3, N = 30) = 17,497 p < ,001$) (vedi tab. 25), come precedentemente esposto, così come la sua accentazione nel caso in cui gli stimoli esprimano disgusto anziché essere incongruenti ($F(1, N = 30) = 54,999 p < ,001$) (vedi fig. 88).

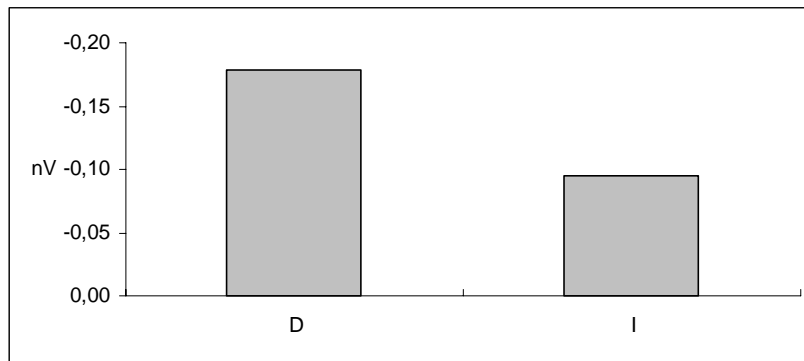


Fig. 88. Intensità del picco N400 in risposta a stimoli esprimenti disgusto e stimoli incongruenti.

Gioia/incongruente

Gli stimoli esprimenti gioia presentano una situazione analoga, con la localizzazione posteriore del fenomeno ($F(3, N = 30) = 11,089$ $p < ,001$) (vedi tab. 25) e valori più elevati quando gli stimoli sono congruenti ($F(1, N = 30) = 37,989$ $p < ,001$) (vedi fig. 89).

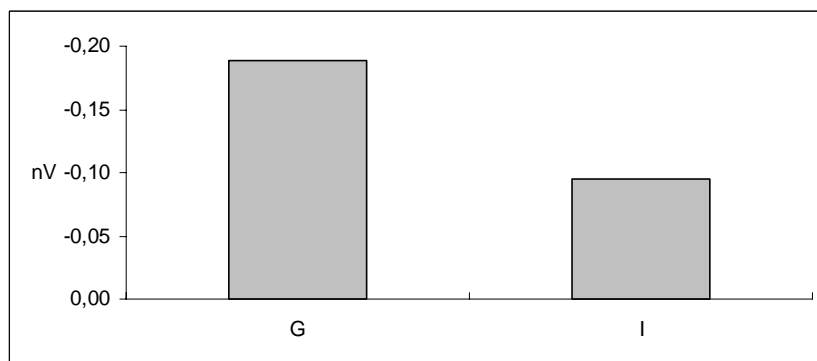


Fig. 89. Intensità del picco N400 in risposta a stimoli esprimenti gioia e stimoli incongruenti.

Paura/incongruente

Anche nel caso della paura, la localizzazione è prevalentemente parietale ($F(3, N = 30) = 12,423$ $p < ,001$) (vedi tab. 25), con un picco più intenso nella condizione congruente rispetto a quella incongruente ($F(1, N = 30) = 24,948$ $p < ,001$), come illustrato in figura 90.

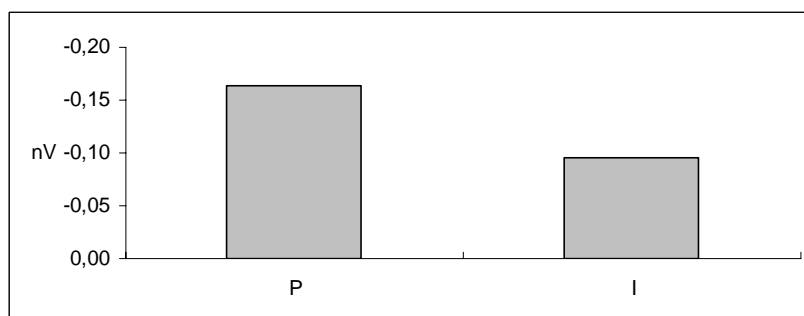


Fig. 90. Ampiezza del picco N400 elicitato da stimoli esprimenti paura e stimoli incongruenti.

Rabbia/incongruente

Ancora una volta, la localizzazione è prevalentemente parietale ($F(3, N = 30) = 12,270$ $p < ,001$) (vedi tab. 25), con un picco più intenso per gli stimoli esprimenti rabbia che per gli stimoli incongruenti ($F(1, N = 30) = 15,108$ $p = ,001$), come illustrato in figura 91a. Bisogna inoltre segnalare una duplice interazione localizzazione*task ($F(3, N = 30) = 2,770$ $p = ,047$) (vedi fig. 91b): nelle aree anteriori, ed in particolar modo frontali, il task esplicito produce un picco molto più intenso di quello prodotto dalla consegna implicita, mentre nelle aree posteriori la situazione risulta essere invertita, se pure con un minore scarto tra le due condizioni.

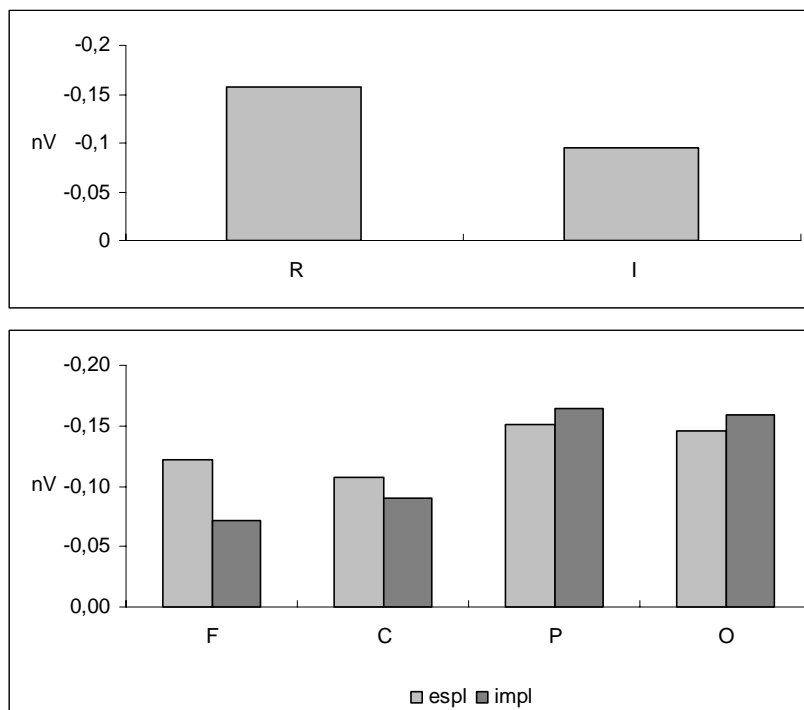


Fig. 91. Ampiezza del picco N400 elicitato da stimoli esprimenti rabbia e stimoli incongruenti (a), e rispetto all'interazione tra il fattore localizzazione e il fattore task (b).

Sorpresa/incongruente

Anche nel caso della sorpresa, la localizzazione è soprattutto parietale ($F(3, N = 30) = 22,047$ $p < ,001$) (vedi tab. 25), con valori più intensi nella condizione congruente rispetto a quella incongruente ($F(1, N = 30) = 13,064$ $p = ,001$) (vedi fig. 92).

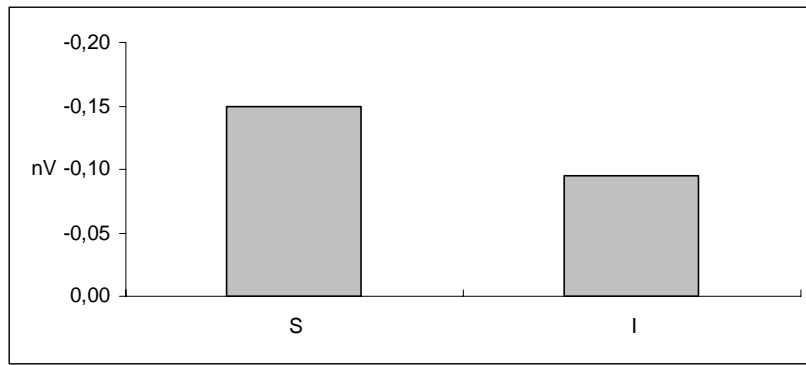


Fig. 92. Intensità del picco N400 elicitato da stimoli esprimenti sorpresa e da stimoli incongruenti.

Tristezza/incongruente

Infine, un'analogha situazione contraddistingue il confronto tristezza/incongruente, con una localizzazione posteriore, in particolare parietale ($F(3, N = 30) = 20,755 p < ,001$) (vedi tab. 25), e con una maggiore intensità di picco nel caso degli stimoli congruenti esprimenti tristezza ($F(1, N = 30) = 40,068 p < ,001$) (vedi fig. 93).

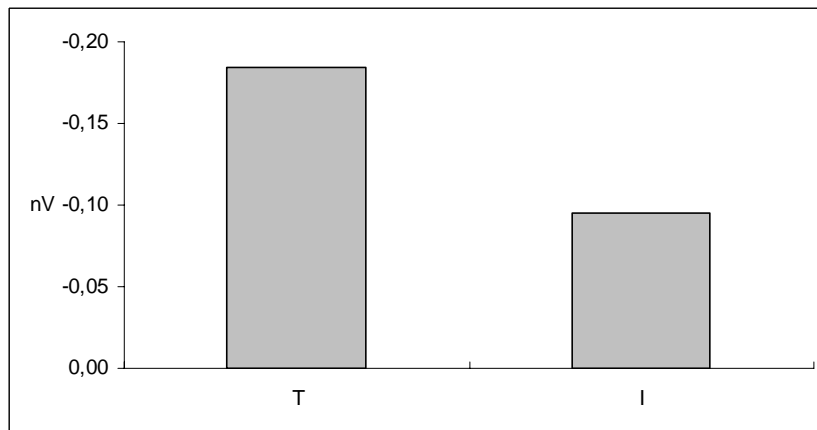


Fig. 93. Intensità del picco N400 in seguito alla presentazione di stimoli congruenti esprimenti tristezza e stimoli incongruenti.

	D - I		G - I		P - I		R - I		S - I		T - I	
	M	ds	M	ds	M	ds	M	ds	M	ds	M	ds
F	-,103	,012	,092	,016	,089	,014	,084	,009	-,085	,012	-,103	,012
C	-,108	,011	,115	,017	,091	,015	,095	,014	-,096	,011	-,113	,010
P	-,173	,015	,161	,022	,159	,019	,152	,018	-,157	,015	-,176	,014
O	-,164	,010	,130	,014	,143	,012	,128	,011	-,151	,012	-,166	,011

Tab. 25. Media e deviazione standard dell'intensità del picco N400 nelle diverse aree corticali, come emerge dal confronto tra le emozioni congruenti e la condizione incongruente.

*Latenza**Confronto tra emozioni*

Confrontando tra loro le diverse emozioni congruenti rispetto alla latenza del fenomeno N400 si rileva un'interazione condizione*task ($F(1, N = 30) = 2,817 p < ,019$): in figura 94 si nota che per la paura e la sorpresa i valori di picco sono nettamente più elevati quando il task è implicito. Per la tristezza invece è il task esplicito a produrre valori molto più elevati. Per il disgusto, la gioia e la rabbia, non si rilevano invece differenze rilevanti tra i due tipi di consegna sperimentale.

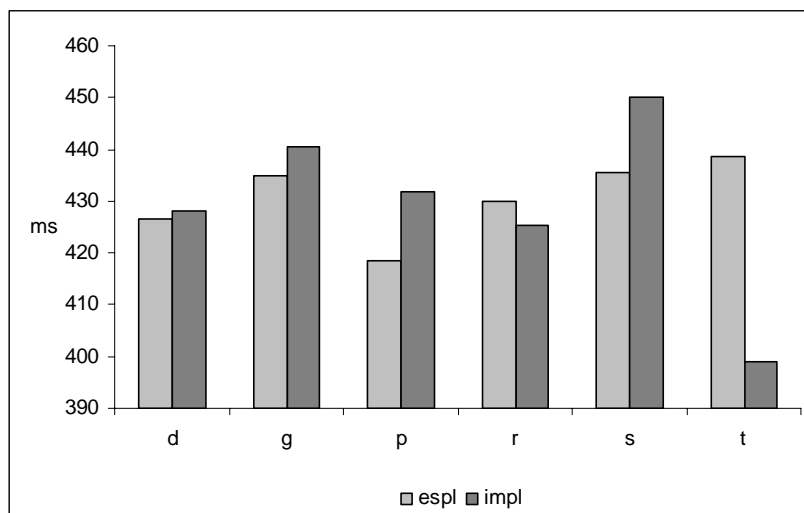


Fig. 94. Intensità del picco N400 in funzione del tipo di emozione e della consegna sperimentale esplicita e implicita.

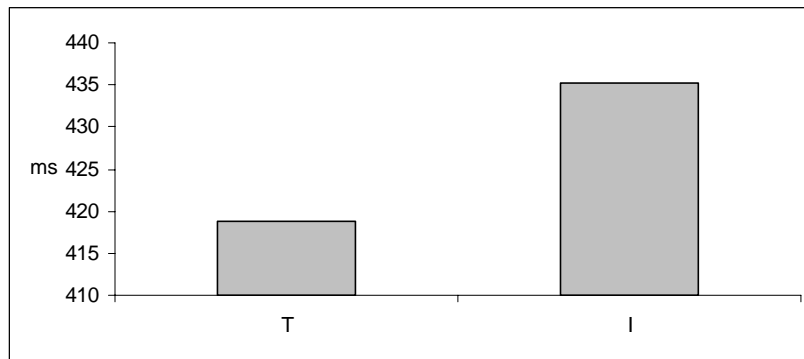
Confronto tra le singole emozioni congruenti e la condizione incongruente

Dal confronto tra ciascuna emozione congruente e la condizione incongruente non emerge nessun risultato significativo per il disgusto, la gioia, la paura, la rabbia e la sorpresa. Vengono invece esposti i dati relativi alla tristezza, la cui analisi hanno messo in evidenza alcune significatività.

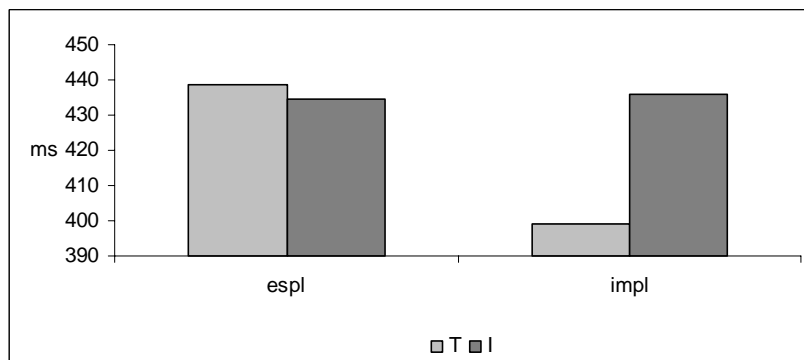
Tristezza/incongruente

Rispetto al confronto tra tristezza e condizione incongruente, emerge un notevole allungamento dei tempi di latenza per la condizione incongruente ($F(1, N = 30) = 4,558 p < ,041$) (vedi fig. 95a). Questa tendenza risulta essere notevolmente accentuata

quando la consegna sperimentale è implicita ($F(1, N = 30) = 8,873 p < ,006$), come rappresentato in figura 95b.



(a) Fig. 95. Latenza del picco N400 per gli stimoli esprimenti tristezza e gli stimoli incongruenti (a) e in funzione della condizione e del task esplicito e implicito (b).



(b)

3.3. La condizione di incongruenza: ulteriori analisi

Al fine di approfondire l'esplorazione di possibili differenziazioni legate alla condizione di incongruenza all'interno dell'universo emotivo, alla finestra temporale 150-250 ms sono state applicate le seguenti analisi accompagnate, dove necessario, dai successivi confronti appaiati:

- Anova a misure ripetute localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x emozione (var. within a 6 livelli: disgusto, gioia, paura, rabbia, sorpresa, tristezza incongruenti)
- Per ciascuna emozione: anova a misure ripetute localizzazione (var. within a 4 livelli: frontale, centrale, parietale, occipitale) x condizione (var. within a 2 livelli: emozione congruente/emozione incongruente)

3.3.1. Indice N200

Intensità

Confronto tra le sei emozioni incongruenti

L'anova evidenzia un effetto significativo emozione*localizzazione ($F(15, N = 30) = 2,571, p = ,001$) (vedi fig. 102). A fronte di una generale prevalente localizzazione frontale, rispetto alle altre aree corticali si nota una distribuzione disomogenea in funzione delle diverse emozioni.

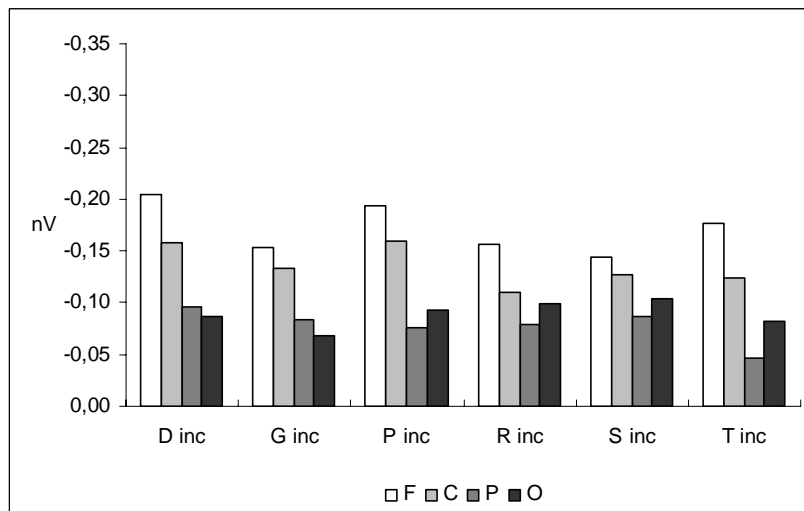


Fig. 102. Intensità del picco N200 per le diverse emozioni incongruenti in funzione della localizzazione corticale.

Analisi condotte sulle singole emozioni

Soltanto le emozioni della paura e della rabbia mostrano differenze significative. Per entrambe, infatti, le analisi hanno fatto emergere un effetto di interazione condizione*localizzazione (rispettivamente per la paura ($F(3, N = 30) = 3,432, p = ,020$) e per la rabbia ($F(3, N = 30) = 3,674, p = ,015$)) (vedi fig. 103). Rispetto alla paura, la condizione incongruente fa registrare valori più elevati in tutte le aree dello scalpo, tranne che in quella occipitale; nel caso della rabbia, si verifica una situazione inversa, con valori più accentati per la condizione di congruenza in tutte le aree tranne che in quella occipitale.

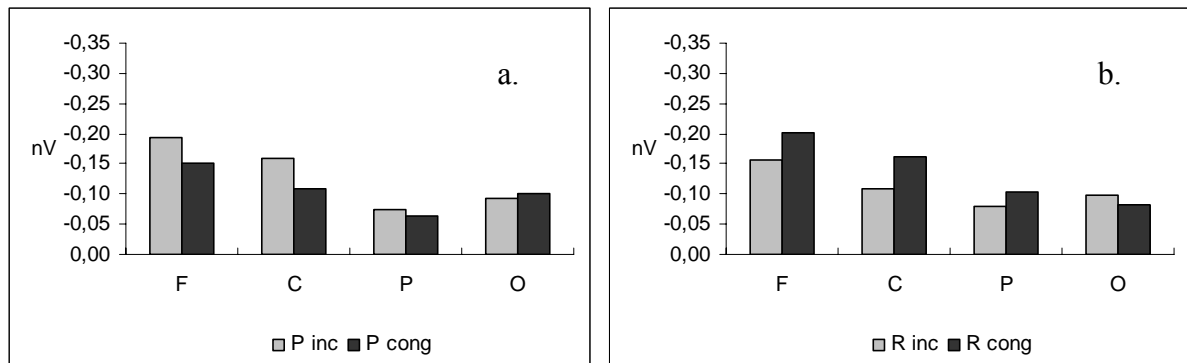


Fig. 103. Ampiezza del picco N200 per paura (a.) e rabbia (b.) in funzione dell'interazione tra condizione e localizzazione.

Latenza

Rispetto alla latenza, dalle analisi non risulta alcun effetto significativo.

3.3.2. Indice P200

Intensità

Confronto tra le sei emozioni incongruenti

Confrontando tra loro le sei emozioni nella condizione di incongruenza, emerge un effetto di interazione tra le variabili emozione e localizzazione ($F(15, N = 30) = 3,960, p < ,001$). In particolare, come si può osservare in figura 104, per tutte le emozioni la localizzazione è prevalentemente parietale, mentre si osserva una distribuzione maggiormente variegata nelle altre zone della corteccia.

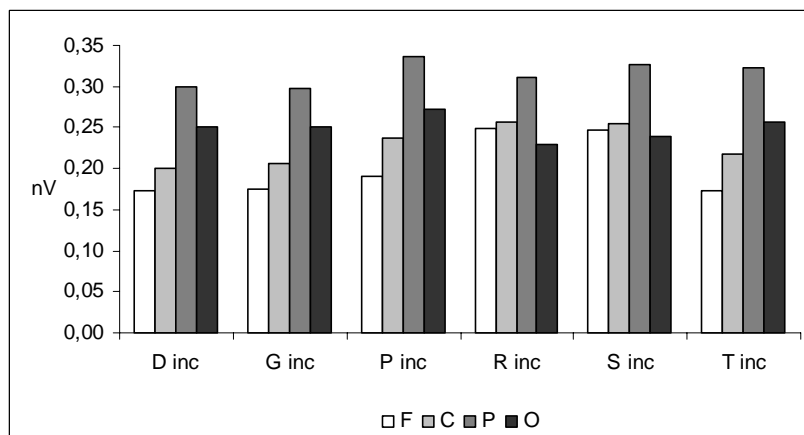


Fig. 104. Intensità del picco P200 per le diverse emozioni incongruenti in funzione della localizzazione corticale.

Analisi condotte sulle singole emozioni

Le emozioni di gioia ($F(3, N = 30) = 5,256, p = ,002$) e paura ($F(3, N = 30) = 3,644, p = ,016$) presentano un effetto significativo condizione*localizzazione (vedi fig. 105). Per l'emozione della gioia, nelle aree anteriori i valori più elevati si manifestano nella condizione di congruenza, mentre nelle aree posteriori è la condizione di incongruenza a elicitare un picco più accentuato. Nel caso della paura, invece, è sempre la condizione di incongruenza a presentare valori più elevati, eccetto che nella zona frontale.

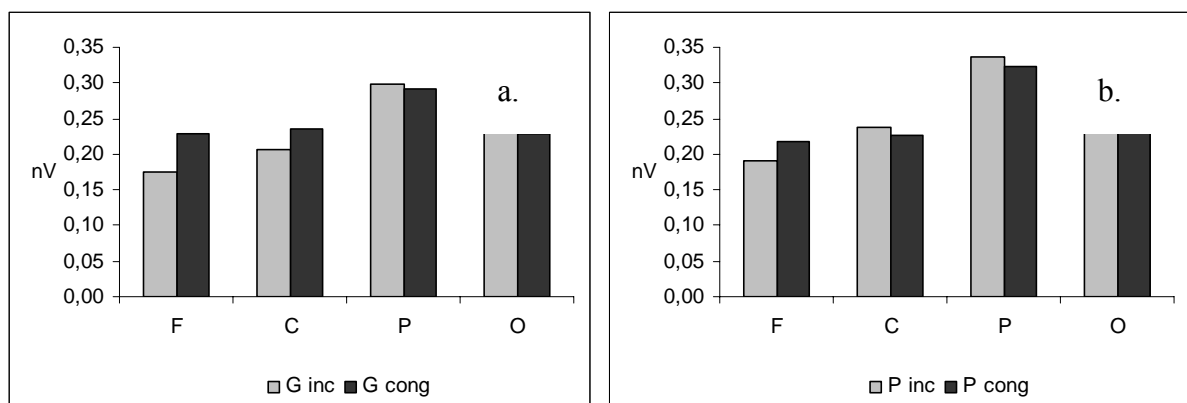


Fig. 105. Ampiezza del picco N200 per gioia (a.) e paura (b.) in funzione dell'interazione tra condizione e localizzazione.

Latenza

Rispetto alla latenza, le analisi indicano una differenziazione dei tempi di comparsa del picco per l'emozione della sorpresa, con un effetto significativo condizione*localizzazione ($F(3, N = 30) = 2,676, p = ,052$) (vedi fig. 106): nelle aree anteriori si rilevano tempi ritardati per la congruenza, mentre in area occipitale i tempi subiscono invece un ritardo nella condizione di incongruenza.

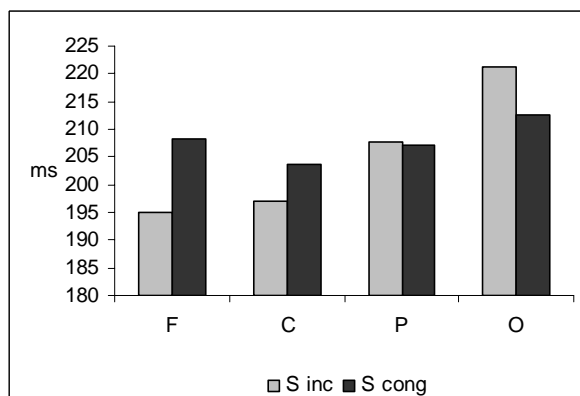


Fig. 106. Tempi di latenza del picco P200 per l'emozione della paura, nella condizione di congruenza/incongruenza e nelle diverse aree corticali.

4. I dati comportamentali: analisi statistica

I dati di natura comportamentale (tempi di reazione) sono stati esplorati utilizzando come prima analisi un'ANOVA a misure ripetute utilizzando il fattore condizione (congruente/incongruente). L'analisi ha mostrato che i tempi di risposta alle coppie di stimoli congruenti e i tempi di risposta alle coppie di stimoli incongruenti non presentano differenze significative ($F(1, N = 30) = ,627 p = ,435$). Anche il fattore task, inserito in un secondo momento nell'analisi, non è risultato essere significativo ($F(1, N = 30) = ,731 p = ,400$).

Successivamente, è stata fatta girare un'ANOVA a misure ripetute impiegando il fattore emozione (variabile within a 6 livelli: disgusto, gioia, paura, rabbia, sorpresa, tristezza congruenti): l'analisi ha rivelato un'incidenza di tale fattore sui tempi di risposta ($F(5, N = 30) = 32,909, p < ,001$) (vedi fig. 107). Rispetto a tale effetto, è stata applicata un'analisi dei contrasti, dalla quale si evince che i TR più rapidi si registrano per le emozioni della tristezza e del disgusto, seguite dalla gioia. I tempi più lunghi sono invece prodotti da paura, rabbia e sorpresa.

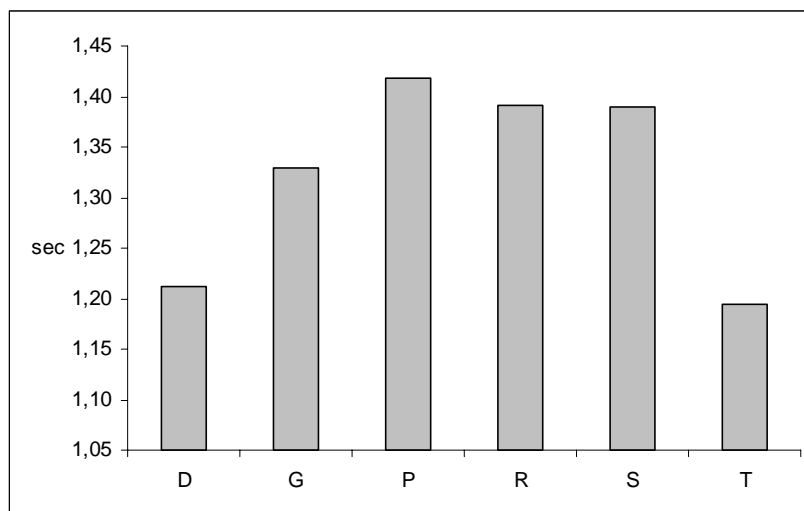


Fig. 102. Tempi di risposta alle diverse coppie di emozioni congruenti.

La medesima analisi è stata condotta mettendo a confronto le diverse emozioni nella condizione di incongruenza: i confronti a coppie indicano che sono in particolare i tempi dell'emozione della rabbia (molto rapidi) a differenziarsi da quelli delle emozioni positive (gioia) e a basso arousal (disgusto e tristezza); esiste

una differenza significativa anche tra tristezza (che ha tempi di risposta piuttosto lunghi) e sorpresa ($F(5, N = 30) = 2,547, p = ,031$) (vedi fig. 103).

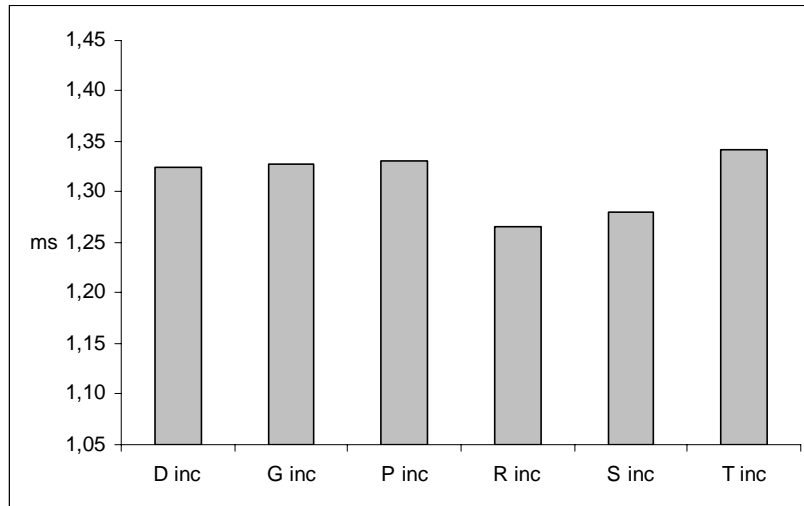


Fig. 103. Tempi di risposta alle diverse coppie di emozioni incongruenti.

In un'ultima fase di analisi, grazie ad una serie di ANOVA a misure ripetute, i tempi di risposta di ciascuna emozione nella condizione di congruenza sono stati confrontati con i tempi di risposta della stessa emozione nella condizione di incongruenza. Eccone i risultati, che mostrano come per tutte le emozioni, tranne che per la gioia, la condizione sperimentale incide sui tempi di risposta (vedi fig. 103):

- disgusto: ($F(1, N = 30) = 12,946, p = ,001$)
- gioia: non significativo
- paura: ($F(1, N = 30) = 10,754, p = ,003$)
- rabbia: ($F(1, N = 30) = 16,527, p < ,001$)
- sorpresa: ($F(1, N = 30) = 16,491, p < ,001$)
- tristezza: ($F(1, N = 30) = 31,653, p < ,001$)

In particolare, si nota che disgusto e tristezza hanno tempi più rapidi per la condizione di congruenza, mentre per le emozioni negative ad alto arousal e cioè rabbia, sorpresa e paura, i tempi sono più brevi per la condizione di congruenza.

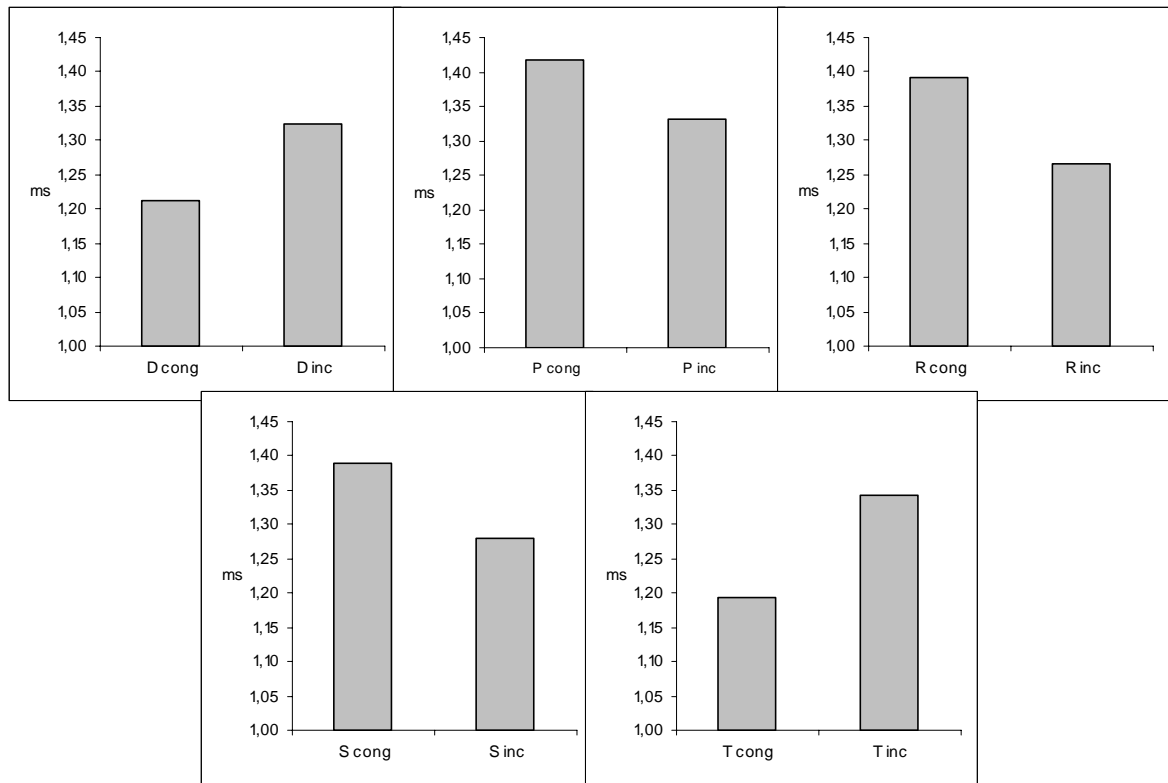


Fig. 103. Tempi di risposta alle diverse emozioni , per la condizione di congruenza e di incongruenza.

DISCUSSIONE

1. Per una visione d'insieme

Dai risultati sperimentali emergono alcuni aspetti principali di rilievo. Gli effetti significativi rilevati riguardano in particolare: 1) La presenza di *indicatori ERP di media latenza*, presumibilmente direttamente correlati alle componenti cognitive di integrazione intersensoriale, con una netta differenziazione del profilo ERP in relazione alla *localizzazione corticale*; 2) La presenza di variazioni specifiche, in termini di ampiezza, latenza e localizzazione, che consentono di differenziare opportunamente i profili per la *congruenza* e *l'incongruenza* dei due canali visivo e uditivo; 3) La rilevanza delle dimensioni di *arousal* e *valenza edonica* nel differenziare la risposta al contenuto dei *singoli patterns emotivi*; 4) La compresenza di *processi automatici* (a breve latenza) e *decisionali* (a media e lunga latenza); 5) Una differenziazione dei *tempi di risposta* in funzione della congruenza/incongruenza degli stimoli e del contenuto emotivo, a fronte di una generale *efficienza* dei fenomeni integrativi emotivi.

2. Le coordinate temporali del fenomeno: componenti percettive e componenti cognitive

Come sottolineato dall'analisi morfologica del profilo medio dell'ERP, la simultanea presentazione di stimoli emotivi visivi e uditivi produce numerose deflessioni che è possibile identificare come fenomeni di precoce, media e lunga latenza.

2.1. I fenomeni percettivi precoci

All'interno della prima finestra temporale indagata (0 - 150 ms), che è dedicata all'elaborazione delle caratteristiche percettive e strutturali della stimolazione (Balconi & Lucchiari, 2005), sono stati rilevati l'indice positivo P100 e l'indice negativo N100. Dal momento che tali indici sono stati già ampiamente esplorati dalle precedenti ricerche (Calvert, Brammer & Iversen, 2000; Giard & Peronnet, 1999; Raij, Uutela & Hari, 2000; Sams et al., 1991) e che, inoltre, l'indagine di tali precoci deflessioni di natura percettiva solo parzialmente rientra negli scopi del presente studio, ci siamo limitati ad analizzarne l'intensità. Questa scelta è motivata dal fatto che, come la letteratura indica, l'aumento dell'ampiezza delle componenti unimodali precoci è un fenomeno tipico della condizione di stimolazione bimodale, se confrontata con quella unimodale (Giard & Peronnet, 1999; Pourtois et al., 2000). La presenza, intorno ai 100 ms dopo la presentazione dello stimolo, di un netto picco negativo con prevalente localizzazione frontale e di un chiaro picco positivo prevalentemente localizzato nelle aree posteriori porta ulteriore conferma a questo dato.

2.2. I fenomeni cognitivi di media latenza

All'interno della finestra temporale di media latenza, si rileva innanzitutto la presenza di due differenti deflessioni, una di natura positiva (P200) ed una di natura negativa (N200). Esse appaiono caratterizzate da una comune disomogeneità della distribuzione della deflessione sullo scalpo. Più specificamente, nel caso dell'indice P200 si registra una localizzazione prevalente del picco nelle aree *parietali* dello scalpo, dal momento che la deflessione raggiunge una intensità massima posteriormente (Pz). Un'analoga sensibilità per la localizzazione si rileva per il picco N200, che fa registrare una distribuzione differenziata tra aree anteriori e posteriori, con prevalenza delle prime, in particolare in zona *frontale* (Fz) rispetto alle seconde (Pz e Oz).

Anche intorno ai 300 ms successivi alla stimolazione si osservano un picco negativo (N300) maggiormente localizzato nella porzione *anteriore* dello scalpo e una deflessione positiva (P300) con prevalente localizzazione *parietale*.

La rilevazione di questi fenomeni costituisce un dato di notevole interesse, poiché avvalorata l'ipotesi che l'integrazione audiovisiva di stimoli emotivi coinvolga, a seguito di una precoce fase di natura percettiva, anche stadi di elaborazione di *natura cognitiva* più tardivi. Pressoché tutti gli studi condotti hanno centrato l'attenzione sugli stadi percettivi precoci del processo di elaborazione, esplorando infatti indici come la N100 e il MMN, che segnalano processi di natura sensoriale (vedi cap. 4). Questi studi hanno dimostrato che la convergenza è un fenomeno che ha inizio in una fase molto precoce del processamento. Partendo da questo dato, numerosi ricercatori hanno tuttavia anche concluso che l'integrazione avvenga senza il contributo di una successiva elaborazione che coinvolge la sfera semantica (de Gelder et al., 1999; Pourtois et al., 2000). I risultati della presente ricerca smentiscono tale possibilità, suggerendo invece che l'integrazione emotiva intersensoriale sia un processo che inizia precocemente ma non si esaurisce nella fase puramente percettiva. Tale ipotesi trova ulteriore conferma nella rilevazione di indici di lunga latenza.

2.3. I fenomeni cognitivi di lunga latenza

Anche la finestra temporale più tardiva (350 - 500 ms) rivela la presenza di alcuni fenomeni, ovvero un picco negativo che si connota come un fenomeno *posteriore* (N400) e un picco positivo marcatamente *anteriore* (P400), confermando la presenza di stadi di elaborazione tardivi, di natura semantica e decisionale.

3. La congruenza vs. incongruenza delle informazioni audiovisive

Considerando le analisi che hanno contrapposto in generale le due condizioni di congruenza e incongruenza senza differenziare le specifiche emozioni, è possibile individuare due categorie di fenomeni, di cui l'uno, che comprende gli indici P200 e N300, mostra di essere sensibile alla condizione di congruenza e incongruenza della stimolazione, mentre l'altro, che raggruppa gli indici N200, P300 e N400, non risente dell'influenza della condizione sperimentale. Di seguito, questi risultati e le loro implicazioni vengono dettagliatamente discussi: riteniamo infatti che le differenti

risposte alla stimolazione emotiva intersensoriale rendano conto della presenza di *fenomeni cognitivi differenti*.

Le analisi indicano che gli indici P200 e N300, in riferimento ai parametri di intensità e latenza, subiscono l'effetto della congruenza o incongruenza tra i due canali visivo e uditivo. Questo ci porta a ritenere che entrambi questi indici ERP siano direttamente implicati nel processo di integrazione intersensoriale di informazioni a contenuto emotivo.

E' necessario tuttavia anticipare che le successive analisi *emotion-by-emotion* hanno evidenziato la presenza di un effetto della variabile condizione a livello di tutti i fenomeni indagati (vedi par. 4.).

3.1. La P200: l'indice di integrazione intersensoriale di stimoli emotivi

E' stata osservata una *variabilità dell'intensità* di picco in funzione della condizione per la P200, con valori di maggiore ampiezza per la condizione di congruenza. Anche rispetto alla *localizzazione* si rileva un dato interessante, ovvero l'accentuazione - nella condizione di congruenza rispetto a quella di incongruenza - della differenziazione anteriore/posteriore in funzione della condizione, con una più netta localizzazione frontale rispetto a quella parietale.

Data la sensibilità dell'effetto P200 per la condizione di congruenza, è possibile ipotizzare che esso rappresenti un *indice di media latenza dei processi integrativi intersensoriali* in risposta a stimoli *a contenuto emotivo*. Nello specifico, esso mostrerebbe una prevalenza per condizioni di ridondanza e convergenza dell'informazione emotiva proveniente da differenti canali (visivo e uditivo). Al contrario, una sua riduzione renderebbe conto della mancata convergenza delle informazioni veicolate dal volto e dalla voce. In precedenti ricerche è stato rilevato un analogo effetto di amplificazione di una deflessione positiva in relazione alla congruenza dei canali, definita come variazione P2b (Pourtois et al., 2002), generalmente successiva ad un complesso di effetti percettivi di breve latenza (N100 -P100) e precedente una deflessione positiva più tardiva (P300). Gli autori ritengono che la P2b rappresenti un indice di integrazione tra la componente uditiva e quella visiva.

Rispetto al parametro della *latenza*, alcune ricerche precedenti hanno rilevato un ritardo di comparsa del picco per l'incongruenza tra canali. Tuttavia, questo dato si riferisce alle componenti percettive e non a quelle di maggiore latenza (cognitive), scarsamente considerate sino ad ora. Contrariamente a quanto rilevato in tali precedenti studi, la latenza del picco P200 non appare essere significativamente differenziata in relazione alla condizione sperimentale, ma unicamente in funzione della localizzazione. Ciò fa supporre che la condizione di incongruenza, nella fase cognitiva dell'elaborazione, non richieda tempi di processamento maggiori.

E' possibile paragonare il significato cognitivo dell'indice P200 a quanto rilevato per un indice di maggiore latenza, la variazione N400 (Debruille et al., 1996). Tale componente è stata in precedenza osservata in risposta a condizioni di anomalia semantica o di generale necessità di riconfigurazione del contesto a fronte di elementi di non consonanza stimolo-background, indipendentemente dalla modalità percettiva (è stata rilevata sia per immagini che per parole) (Balconi & Lucchiari, 2005). E' pertanto possibile stabilire un piano di analogicità tra le due variazioni ERP, in quanto l'una sarebbe indicativa di una condizione di incongruenza semantica dello stimolo di tipo amodale (N400), l'altra della mancata convergenza semantica tra più informazioni modalità-specifiche (P200). Nella fattispecie, nella presente ricerca, essa risulta essere sensibile all'incongruenza tra canale mimico e vocale, indipendentemente dallo specifico contenuto emotivo espresso.

3.2. La N300: una N400 anticipata?

Anche l'indice N300 mostra, in riferimento alla *localizzazione*, un andamento sensibile in funzione della congruenza/incongruenza, con una intensificazione dell'ampiezza del picco nelle aree frontali quando gli stimoli sono incongruenti e con un'accentuazione della differenziazione anteriore/posteriore in funzione della condizione di incongruenza, che rivela una più netta localizzazione posteriore rispetto alla congruenza. Rispetto alla *latenza*, il fenomeno risulta essere ritardato posteriormente soprattutto per le coppie di stimoli incongruenti.

Il comportamento di questo indice, soprattutto in riferimento alla sua localizzazione frontale e alla sua sensibilità alla rilevazione di incongruenze, potrebbe indurre a considerare la sua funzione come analoga a quella che è caratteristica della N400 (Debruille et al., 1996) e a considerare quindi il fenomeno da noi rilevato come un'anticipazione di tale deflessione.

3.3. La N200: il marker del contenuto emotivo

Ciò che caratterizza la deflessione N200 è una sostanziale stabilità dell'intensità del picco per entrambe le condizioni sperimentali. Un secondo dato di rilievo riguarda la *latenza di comparsa* della variazione ERP. La latenza del picco, infatti, rimane invariata nelle due condizioni di congruenza e incongruenza. Il picco N200, in altri termini, non presenta specifici effetti correlati alla convergenza o divergenza delle informazioni emotive cross-modali. Piuttosto, esso appare mantenere un analogo profilo tra le due condizioni sperimentali. Sulla base di quanto rilevato, è possibile concludere che tale deflessione non rappresenti una risposta cognitiva specifica alle stimolazioni emotive intersensoriali. Seppure rilevante ai fini della comprensione dei processi cognitivi sottesi all'integrazione cross-modale, tale variazione funge più specificamente da indicatore emozione-correlato (Eimer & McCarthy, 1999; Streit et al., 2000), come suggerito dai risultati delle analisi condotte sulle singole emozioni (vedi par. 4.1.). Studi precedenti hanno infatti rilevato la sua sensibilità rispetto allo specifico contenuto emotivo dello stimolo (tipo di emozione), in funzione di alcuni parametri, tra cui le componenti di arousal e valenza edonica del pattern (Balconi & Pozzoli, 2003; Balconi, 2004; Sato et al., 2001). Tale sensibilità ha indotto i ricercatori a considerare il picco negativo che si manifesta intorno ai 200 ms di latenza come un *marker del decoding della mimica emotiva*.

3.4. La P300: l'updating cognitivo

La localizzazione e la latenza del picco permettono di identificarlo chiaramente come un fenomeno P300 (Meinhardt & Pekrun, 2003). Esso non appare influenzato dalla variabile condizione. La sua presenza suggerisce che non soltanto nel riconoscimento unimodale delle emozioni (Schupp et al., 2004), ma anche in quello bimodale, esso giochi un ruolo importante, senza tuttavia essere direttamente legato ai processi di integrazione.

La P300 rappresenta un *fenomeno attentivo* (Meinhardt & Pekrun, 2003). Non a caso, tale picco è stato utilizzato al fine di verificare l'effetto della presenza di un'emozione nei compiti di performance. In accordo con il modello dell'allocazione delle risorse (Ellis & Ashbrook, 1988), che si basa sull'assunzione che le capacità di processamento umane sono limitate (Kahneman, 1973), le emozioni e le loro conseguenze cognitive (per esempio i pensieri intrusivi da esse evocati) assorbono risorse attentive che altrimenti verrebbero dedicate al compito. Tali studi che hanno utilizzato il paradigma del *dual task* hanno dimostrato che le emozioni positive, ma soprattutto quelle negative, riducono le risorse indirizzate al compito. Tale effetto delle emozioni è appunto rivelato da una variazione di intensità del picco P300. Normalmente, il fenomeno P300 si manifesta in risposta alla presentazione di stimoli nuovi (Knight & Scabini, 2003). Ed infatti, il paradigma tradizionalmente utilizzato per studiarlo prevede la presentazione di uno stimolo distrattore all'interno di una sequenza di stimoli target. Recentemente, Katayama e Polich (1998) hanno messo in risalto il ruolo che tale componente svolge nel ridirezionamento dell'attenzione verso lo stimolo che, all'interno di un determinato contesto, risulta essere saliente.

Per cogliere appieno la funzione di questo picco, bisogna fare riferimento alla distinzione che Luu e Tucker (2003) operano tra la componente P3a, che ha una distribuzione fronto-centrale e compare intorno ai 250-350 ms, e la componente P3b, che ha invece una localizzazione parietale e una latenza più ritardata. Le componenti contestuali sono fondamentali nell'elicitarla P3a. Essa è implicata nella risposta rapida a segnali discrepanti significativi. Complessivamente, è possibile affermare che presumibilmente essa segnala una risposta a violazioni del controllo predittivo. I

risultati della presente ricerca suggeriscono che la deflessione rilevata corrisponda invece ad una P3b. E' stato suggerito che essa rappresenti un processo di *updating cognitivo* che ha luogo quando la nostra rappresentazione interna del contesto viene violata e deve quindi essere modificata (Donchin & Coles, 1988) e che, inoltre, essa sia coinvolta nei conseguenti processi di pianificazione strategica dell'azione. E' possibile che sia parte di un sistema di regolazione dell'azione sensibile al contesto e implicato nei processi di apprendimento (Knight & Scabini, 2003).

Un importante contributo giunge da Schupp (Schupp et al., 2004), il quale ha proposto un modello che presuppone l'intervento di un'*attenzione motivata* nel decoding emotivo. Egli ha osservato che non solo le variazioni della P300 si manifestano durante il processamento di stimoli emotivi, ma nello specifico l'entità di tali variazioni è strettamente legata al significato motivazionale della stimolazione. Infatti, l'entità delle variazioni è maggiore quando gli stimoli emotivi sono caratterizzati da un alto livello di arousal (Schupp et al., 2000). Questa tendenza, non a caso, è massima in risposta alla presentazione scene erotiche, di minaccia o mutilazione, che a causa del loro elevato significato adattivo, attivano fortemente i sistemi motivazionali appetitivi o difensivi.

Le ricerche hanno rilevato una deflessione P300 legata a fenomeni di updating e decisionali durante il decoding emotivo (Balconi & Lucchiari, 2006). Se ciò avviene nella condizione unimodale, è chiaro che a maggior ragione la P300 ricopra un ruolo importante nel caso in cui il sistema cognitivo debba riconoscere un'emozione che viene presentata attraverso fonti di informazione multiple.

3.5. La N400: l'elaborazione delle anomalie semantiche

L'indice N400 è un indice di rielaborazione che si manifesta quando le informazioni sono percepite come anomale dal punto di vista semantico o a seguito della violazione dei vincoli contestuali dello stimolo (Debrulle et al., 1996). Esso è stato rilevato da precedenti studi che hanno indagato il riconoscimento delle emozioni espresse attraverso il volto, in corrispondenza di espressioni facciali anomale. Tale indice ERP, localizzato nella porzione posteriore dello scalpo, viene ritenuto un marker specifico della rilevazione di una *anomalia semantica*. Essa si manifesta infatti in ogni

caso in cui avviene una violazione delle rappresentazioni semantiche (Balconi, 2002), indicando una indipendenza rispetto al tipo di stimolazione. Non stupisce quindi il fatto che tale picco venga elicitato anche nel caso in cui le emozioni vengano presentate attraverso il duplice canale audiovisivo.

4. Emotion-by-emotion: la specificità dei differenti correlati emotivi

Appare a questo punto necessario comprendere meglio quale ruolo svolga lo specifico contenuto emotivo, analizzando l'effetto dei singoli patterns sul tracciato ERP. L'impiego di una vasta gamma di correlati emotivi ha consentito di fornire informazioni aggiuntive circa lo specifico significato cognitivo delle deflessioni rilevate. Essa, inoltre, ha permesso di rilevare alcuni effetti specifici della condizione sperimentale che non erano emersi dalle analisi generali. Di seguito, vengono discussi i risultati maggiormente significativi.

4.1. N200, arousal e valenza edonica: le coordinate dell'universo emotivo

In particolare, l'effetto N200 mostra una diversa rispondenza alle singole emozioni, poiché esso appare differenziato tra le emozioni indipendentemente dalla condizione sperimentale. Si rileva una sostanziale differenziazione tra un primo gruppo di emozioni, ovvero *rabbia, paura, sorpresa e disgusto* e un secondo gruppo rappresentato da *gioia e tristezza*. Nel primo caso la deflessione negativa mostra una maggiore intensità di picco. In linea con quanto rilevato in precedenti studi, i dati indicano che la N200 potrebbe essere un buon indice della dimensione di *arousal* e, in secondo luogo, della *valenza edonica* dello stimolo (Balconi, 2004; Balconi & Pozzoli, 2003). E' stato ipotizzato che esso possa costituire un marker del contenuto emotivo, variabile in funzione del maggior investimento emotivo dei soggetti in risposta a stimoli maggiormente attivanti (alto arousal, per emozioni come la rabbia, la paura e la sorpresa) e determinati da situazioni "aversive" o di possibile minaccia per il soggetto percipiente (per emozioni negative, come il disgusto).

4.2. Indipendenza della P200 rispetto al contenuto emotivo

Un profilo contrapposto è emerso per gli altri fenomeni, che non rivelano una particolare sensibilità allo specifico contenuto emotivo della stimolazione bimodale.

In particolare, per quanto concerne la *deflessione P200*, anche per le singole emozioni essa appare localmente circoscritta nelle aree posteriori (Pz) dello scalpo rispetto a quelle anteriori (Fz). In secondo luogo, complessivamente, le analisi indicano ancora una volta che essa è sensibile alla condizione ma non allo specifico contenuto emotivo. Infatti, a fronte di una mancanza di differenze significative tra le emozioni, esso appare circoscrivere una particolare risposta alla condizione di incongruenza tra canali, mostrando una sostanziale riduzione del picco in questa condizione e un incremento invece nel caso di congruenza. Come precedentemente esposto, in precedenti ricerche è stato rilevato un analogo effetto di amplificazione di una deflessione positiva in relazione alla congruenza dei canali, definita come variazione P2b (Pourtois et al., 2002), generalmente successiva ad un complesso di effetti percettivi di breve latenza (N1-P1) e precedente una deflessione positiva più tardiva (P3). Tuttavia, come già sottolineato, gli studi precedenti hanno esteso la propria analisi solo ad alcuni dei possibili correlati emotivi (generalmente due emozioni contrapposte come gioia e rabbia o gioia e tristezza), non potendo pertanto escludere un rapporto diretto dell'effetto P200 con lo specifico contenuto emotivo veicolato. Nel presente contributo, grazie all'ampliamento del *range* delle emozioni, possiamo al contrario escludere con maggiore sicurezza un effetto determinato dal contenuto emotivo a vantaggio del solo effetto della convergenza/divergenza delle informazioni.

5. La convergenza intersensoriale: processo automatico o processo decisionale?

Il paradigma sperimentale della presente ricerca prevede una duplice consegna sperimentale: nella condizione di *task esplicito*, ai soggetti è stato espressamente richiesto di elaborare la congruenza/incongruenza tra i canali visivo e uditivo, mentre nella

condizione di *task implicito*, essi sono stati semplicemente invitati a prestare attenzione agli stimoli, senza che venisse fatto alcun riferimento ad essa. Questa scelta metodologica è stata dettata dalla volontà di indagare la natura dei processi integrativi.

Diversi ricercatori, infatti, a fronte di un gran numero di studi empirici, sostengono che i processi di integrazione cross-modale delle emozioni siano fenomeni automatici che avvengono in modo obbligato e senza che sia necessario l'intervento della coscienza (vedi cap. 2, par. 2.2.). Si tratterebbe di un processo percettivo di integrazione obbligata di informazioni di diversa natura sensoriale e non di un processo post-percettivo decisionale operante sotto controllo attentivo (de Gelder & Vroomen, 2000; De Gelder et al., 2004). Tale processo precederebbe ampiamente fenomeni frequentemente studiati come il riconoscimento e la comprensione delle emozioni, che sono centrati sul significato personale e sociale dell'emozione.

E' d'obbligo tuttavia osservare che gli studi che hanno dato origine a tale affermazione si sono limitati all'indagine dei processi percettivi di breve latenza, trascurando le successive fasi di elaborazione.

Il presente studio, considerando l'intero intervallo temporale compreso tra gli 0 e i 500 ms successivi alla presentazione dello stimolo, ha permesso di sottoporre tale ipotesi ad una più completa verifica empirica, portando alla luce un quadro caratterizzato da maggiore complessità.

5.1 L'automaticità del processo durante le fasi percettive precoci

Una prima considerazione che emerge dall'osservazione dei risultati concerne le prime fasi di elaborazione, nelle quali lo stimolo viene sottoposto ad un processamento di tipo percettivo e strutturale (Linkenkaer et al., 1998). In linea con quanto rilevato dalle precedenti ricerche, in generale in questo primo stadio non è stata riportata alcuna rilevante differenza tra la condizione esplicita e quella implicita. Il fatto che l'*integrazione percettiva* delle informazioni emotive abbia luogo anche quando non viene richiesto, indica che essa avviene obbligatoriamente, in modo *automatico*. Ciò trova una chiara spiegazione se si considera la funzione del decoding intersensoriale delle emozioni da un punto di vista adattivo. I dati sperimentali confermano tale ipotesi. In

alcuni specifici casi, infatti, è stata rilevata una differenziazione tra i due task, che si manifesta prevalentemente come un ritardo dei tempi di latenza per la condizione di consegna implicita. Questo ritardo, tuttavia, è stato riscontrato quasi esclusivamente per l'emozione della gioia (emozioni positiva a basso arousal) e in alcuni casi per il disgusto e la tristezza ma non per le emozioni negative ad alto arousal (rabbia, paura, sorpresa), ovvero per quelle emozioni la cui esatta o errata discriminazione può essere determinante ai fini della sopravvivenza.

5.2. *L'intervento degli elementi decisionali*

Se effettivamente, a livello percettivo, l'integrazione si configura come un processo automatico che si verifica indipendentemente dal fatto che venga esplicitamente richiesta, a *livello cognitivo più avanzato* al contrario essa è sensibile alla natura della consegna sperimentale. Infatti, i dati indicano un allungamento dei tempi di latenza in caso di task implicito proprio in riferimento a quei due indici ERP, la P200 e la N300, che risultano essere sensibili anche alla condizione di congruenza/incongruenza e che quindi abbiamo ritenuto essere direttamente implicati nei processi integrativi intersensoriali di natura emotiva. Ciò suggerisce che, con l'avvento di processi più propriamente cognitivi, intervengano elementi di natura *decisionale*. Si può ipotizzare che, nel caso in cui l'individuo intenzionalmente elabori la congruenza/incongruenza, il sistema cognitivo metta in atto un processo di updating e di verifica del risultato dell'elaborazione sensoriale precoce, il che provoca un rallentamento dei tempi di processamento.

6. I tempi di risposta: l'efficienza dei fenomeni integrativi emotivi

Un dato interessante e consonante con quanto rilevato dal profilo ERP riguarda le componenti comportamentali in termini di tempi di risposta (TR). I soggetti mostrano infatti una sostanziale *abilità di decoding* della congruenza così come dell'incongruenza, poiché complessivamente si rileva un'abilità nel riconoscere la condizione di incongruenza in tempi pressoché identici a quelli impiegati per identificare la

congruenza. Ciò probabilmente risponde alla capacità di base di identificare situazioni di non consonanza delle informazioni emotive provenienti da differenti canali comunicativi in modo *automatico e repentino*. Occorre di converso sottolineare che la ridondanza di informazioni prodotta nel caso di congruenza non costituisce per tutte le emozioni un fattore di facilitazione del processo di integrazione.

E' interessante osservare la presenza di un differente profilo delle risposte in relazione alle diverse emozioni, seppure con una tendenza inversa rispetto alle nostre attese. L'effetto di ridondanza prodotto dalla congruenza tra canali produce infatti un concomitante "*effetto di disturbo*" per alcune emozioni (paura, rabbia e sorpresa), mentre ha un "*effetto di facilitazione*" per altre emozioni (soprattutto la tristezza e in secondo luogo il disgusto). In quest'ultimo caso i soggetti impiegano un tempo di risposta inferiore nella condizione di congruenza rispetto all'incongruenza. Possiamo pertanto complessivamente ipotizzare la presenza di un effetto di facilitazione del processo di elaborazione delle informazioni in condizione di congruenza per le emozioni con minore potere di attivazione (a basso arousal), mentre un effetto opposto, distraente, per le emozioni di maggiore impatto emotivo (a elevato arousal). Paradossalmente i soggetti in questo secondo caso impiegano tempi maggiori nel produrre una risposta se vengono fornite più informazioni aggiuntive, seppure esse vadano nella medesima direzione. E' possibile supporre che i soggetti appaiono in grado di rilevare rapidamente l'incongruenza mentre sono ostacolati nel proprio processo di comprensione se vengono fornite più informazioni di quanto cognitivamente necessario. In altri termini, tali emozioni funzionerebbero bene con livelli informativi minimi e in modo peggiore con livelli informativi superiori ed eccedenti. Un andamento opposto può essere ipotizzato per le emozioni "secondarie" (apprese successivamente nel corso dello sviluppo), come tristezza e disgusto, che potrebbero beneficiare di un "rinforzo informativo", garantito dalla ridondanza tra canali.

CONCLUSIONI

La panoramica della letteratura attuale che ha come oggetto il riconoscimento intersensoriale audiovisivo delle emozioni rivela che numerosi studi hanno aperto la strada verso la comprensione di questo processo, ma che al tempo stesso, trattandosi di un settore di indagine che soltanto negli ultimi decenni ha ricevuto l'attenzione dei ricercatori, siamo ben lungi dall'afferrare appieno i meccanismi neurali e cognitivi implicati in questo complesso fenomeno.

Il presente studio è stato condotto nell'intento di fare luce su quegli aspetti del processo di integrazione degli stimoli emotivi che fino ad ora erano rimasti al margine dell'attenzione dei ricercatori. La gran parte degli studi che hanno esplorato la percezione cross-modale ha utilizzato stimoli estremamente semplici, come flash di luce e brevi suoni inarticolati. Considerando la natura di tali stimolazioni, è ben comprensibile che la ricerca abbia investigato i fenomeni integrativi scandagliando le fasi precoci di elaborazione strutturale-percettiva. Quando ci si appresta ad indagare la convergenza cross-modale di stimoli come il volto e la voce emotivamente connotati, è tuttavia necessario considerare che si tratta di stimolazioni molto complesse e ricche di significato sociale. Di conseguenza, non sarà più sufficiente limitare il campo d'indagine al processamento strutturale, ma si renderà necessario includere nell'analisi anche le fasi di elaborazione più propriamente di natura cognitiva che sono successive al processamento puramente percettivo. Il limite della maggior parte delle ricerche consiste proprio nell'aver trascurato di compiere questo passo. E la prova che un tale ampliamento dell'ambito di ricerca è sensato giunge dai risultati ottenuti con la presente ricerca che, sfruttando l'alta definizione temporale permessa dalla registrazione e analisi dei potenziali evocati corticali, oltre a confermare il coinvolgimento dei processi percettivi di breve latenza, ha appunto rilevato la presenza di *fenomeni cognitivi* più tardivi. Particolarmente interessante da questo punto di vista è un picco ERP positivo rilevato intorno ai 200 ms dopo la presentazione degli stimoli

bimodali in zona parietale. La sua sensibilità alla congruenza intermodale e, al contrario, la sua indipendenza rispetto allo specifico contenuto emotivo fanno supporre che si tratti di un indice di integrazione intersensoriale di stimoli emotivi di media latenza.

Inoltre, mettendo a confronto una situazione in cui i soggetti erano esplicitamente istruiti ad elaborare la congruenza o discrepanza del contenuto emotivo veicolato dal volto e dalla voce con una condizione in cui invece non ricevevano alcuna specifica direttiva, è stato possibile verificare che, contrariamente a quanto avviene negli stadi percettivi di breve latenza, in cui l'integrazione ha luogo in modo automatico senza raggiungere la soglia della coscienza, nella fase cognitiva successiva intervengono processi intenzionali di *natura decisionale*. Ciò è oltretutto attestato dalla presenza di indici come il picco P300 che rendono conto dell'impiego di *risorse attentive* nella valutazione della congruenza/incongruenza delle informazioni emotive e come la deflessione N400, che viene ritenuta un marker specifico della rilevazione di una *anomalia semantica*.

Nondimeno, il dato di maggior interesse emerso dallo studio condotto riguarda la *salianza*, nell'ambito del funzionamento cognitivo dell'uomo, del processo attraverso il quale, integrando informazioni appartenenti a differenti sistemi sensoriali, riusciamo a riconoscere e a discriminare le emozioni. Indipendentemente dal livello di elaborazione indagato, infatti, ogni qualvolta il sistema cognitivo rivela la presenza di informazioni emotive multiple, le integra allo scopo di attribuire loro un significato univoco. A tale proposito, tra l'altro, i dati comportamentali rilevati indicano che le situazioni di consonanza o non consonanza delle informazioni emotive provenienti da differenti canali comunicativi vengono rilevate in modo automatico e repentino. Tale salienza è spiegabile in termini adattivi: identificare correttamente le emozioni è un'abilità che, a livello evolutivo, può essere fondamentale per la sopravvivenza e, a livello prossimale e contingente, costituisce un'indispensabile competenza sociale. Non a caso infatti la ricerca, che ha il merito di aver ampliato il range di emozioni indagate, ha messo in evidenza una differenziazione dei processi integrativi in funzione delle componenti dimensionali dell'arousal e della valenza edonica. E' infatti possibile, sulla

base delle specifiche modalità integrative rilevate, delineare due gruppi di emozioni che comprendono l'uno le emozioni negative ad alto arousal (rabbia, paura, sorpresa e disgusto) - che sono indicative di una possibile minaccia - e l'altro le emozioni positive (gioia) o a basso arousal.

RIFERIMENTI BIBLIOGRAFICI

- Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinions in Neurobiology*, 12, 169-177.
- Adolphs, R., & Tranel, D. (2004). Impaired judgements of sadness but not happiness following bilateral amygdala damage. *Journal of Cognitive Neuroscience*, 16, 453-462.
- Adolphs, R., Tranel, D., & Damasio, H. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition*, 52, 61-69.
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1995). Impaired recognition of emotion in facial expression after bilateral damage to the human amygdala. *Nature*, 372, 669-672.
- Allison, T., Puce, A., Spencer, D.D., & McCarthy, G. (1999). Electrophysiological studies of human face perception : Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex*, 9, 415-430.
- Alter, K., Rank, E., Kotz, S.A., Toepel, U., Besson, M. Schirmer, A., & Friederici, A.D. (2003) Affective encoding in the speech signal and in event-related potentials. *Speech Communication*, 40, 61-70.
- Anolli, L., & Ciceri, R. (1997). *La voce delle emozioni. Verso una seriosità della comunicazione vocale non-verbale delle emozioni*. Milano: Franco Angeli.
- Arnold, M.B. (1960). *Emotion and personality*. New York: Columbia University Press.
- Aubergé, V., & Cathiard, M. (2003). Can we hear the prosody of smile?. *Speech Communication*, 40, 87-97.
- Balconi, M. (2002). Neuropsicologia della comunicazione. In L. Anolli (Ed.), *Manuale di psicologia della comunicazione* (pp. 85-125). Bologna: Il mulino.

-
- Balconi, M. (2004). Decoding dell'espressione facciale delle emozioni, arousal e valenza edonica. Analisi mediante potenziali evocati corticali (ERPs). *Ricerche di Psicologia*, 2, 7-34.
- Balconi, M. (2005b). An ERP study on facial expression of emotion: comparison of linguistic and visual semantic decoding. *Perceptual and Motor Skills*, 100, 129-134.
- Balconi M. & Carrera A. (2006). Lessicalizzazione e concettualizzazione nella comprensione della mimica emotiva. Analisi mediante script emotivi. *Età Evolutiva*, 83, 18-29.
- Balconi M., Lucchiari C. (2005). Consciousness, emotion and face: An event-related potentials (ERP) study. In R.D. Ellis & N. Newton (Eds.), *Consciousness and emotion. Agency, conscious choice, and selective perception* (pp. 121-135). John Benjamins Publishing Company: Philadelphia.
- Balconi, M., & Lucchiari, C. (2006). EEG correlates (event-related desynchronization) of emotional face elaboration: A temporal analysis. *Neuroscience Letters*, 392, 118-123.
- Balconi M., Pozzoli U. (2003). Face-selective processing and the effect of pleasant and unpleasant emotional expressions of ERP correlates. *International Journal of Psychophysiology*, 49, pp. 67-74.
- Banich, M.T. (1997). *Neuropsychology. The neural bases of mental function*. Boston: Houghton Mifflin Company.
- Banse, R., & Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70, 614-636.
- Bauer, R.M. (1984). Autonomic recognition of names and faces in prosopagnosia: A neuropsychological application of the knowledge test. *Neuropsychologia*, 22, 457-469.
- Bentin, S., & Deouell, L.Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology*, 17, 35-54.

- Boetzel, K., & Grusser, O.J. (1989). Electric brain potentials evoked by pictures of faces and non-faces: a search for "face-specific" EEG-potentials. *Experimental Brain Research*, 77, 349-360.
- Bolognini, N., Rasi, F., & Ladavas, E. (2005). Visual localization of sounds. *Neuropsychologia*, 43, 1655-1661.
- Bullock M., Russell J. A. (1986). Concepts of emotion in developmental psychology. In C. E. Izard & P. B. Read, *Measuring emotions in infants and children* (pp. 203-237). Cambridge: Cambridge University Press.
- Bushara, K.O., Grafman, J., Hallett, & M. (2001). Neural correlates of auditory-visual stimulus onset asynchrony detection. *Journal of Neuroscience*, 21, 300-304.
- Cacioppo, J.T., Berntson, G.G., Larsen, J.T., Poehlmann, K.M., & Ito T.A. (2000). The psychophysiology of emotion. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 173-191). New York: Guilford Press.
- Calder, A.J., Young, A.W., Rowland, D., Perrett, D., Hodges, J.R., & Ectoff, N.L. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology*, 13, 699-745.
- Callan, D.E., Callan, A.M., Kroos, C., & Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. *Cognitive Brain Research*, 10, 349-353.
- Calvert, G.A. (2001). Cross-modal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex*, 11, 1110-1123.
- Calvert, G.A., Brammer, M.J., Bullmore, E.T., Campbell, T., Iversen, S.D., & David, A.S. (1999). Response amplification in sensory-specific cortices during cross-modal binding. *Neuroreport*, 10, 2619-2623.
- Calvert, G.A., Brammer, M.J. & Iversen, S.D. (1998). Cross-modal identification. *Trends in Cognitive Sciences*, 7, 247-253.
- Calvert, G.A., Bullmore, E.T., Brammer, M.J., Campbell, M., & Williams, S.C.R. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276, 593-596.

-
- Calvert, G.A., Campbell, M., & Brammer, R.J. (2000). Evidence from functional magnetic resonance imaging in cross-modal binding in the human heteromodal cortex. *Current Biology*, 10, 649-657.
- Calvert, G.A., Hansen, P.C., Iversen, S.D., & Brammer, M.J. (2001). Detection of audio-visual integration sites in human by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, 14, 427-238.
- Colin, C., Radeau, M., & Deltenre, P. (2005). Top-down and bottom-up modulation of audiovisual integration in speech. *European Journal of Cognitive Psychology*, 17, 541-560.
- Cook, N.D. (2002). *Tone of voice and mind. The connections between intonation, emotion, cognition and consciousness*. Amsterdam/Philadelphia: John Benjamins Publishing company.
- Darwin, C. (1872/1998). *The expression of emotions in man and animals*. New York: Oxford University Press.
- Davidson, R.J., & Irwin, W. (1999) The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Science*, 3,11-21.
- Davidson, R.J., Scherer, K.R., & Goldsmith, H.H. (2003). *Handbook of affective sciences*. Oxford/New York : Oxford University Press.
- Davis, M., & Whalen, P.J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry*, 6, 13-34.
- Debruille, J.B., Pineda, J., & Renault, B. (1996). N400-like potentials elicited by faces and knowledge inhibition. *Cognitive Brain Research*, 4, 133-144.
- De Gelder, B. (2000). Recognizing emotions by ear and by eye. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 84-105). New York: Oxford University Press.
- De Gelder, B., Böcker, K. B. E., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: Early interaction revealed by human electric brain responses. *Neuroscience Letters*, 260, 133-136.

- De Gelder, B., Pourtois, G., & Weiskrantz, L. (2002). Fear recognition in the voice is modulated by unconsciously recognized affective pictures. *Proceedings of the National Academy of Science, USA*, 99, 4121-4126.
- De Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14, 289-311.
- De Gelder, B., Vroomen, J., & Bertelson, P. (1998). Upright but not inverted faces modify the perception of emotion in the voice. *Cahiers de Psychologies Cognitive/Current Psychology of Cognition*, 17, 1021-1031.
- De Gelder, B., Vroomen, J., & Pourtois, G. (2004). Multisensory perception of emotion, its time course, and its neural basis. In G. Calvert, C. Spence & B.E. Stein (Eds.), *The handbook of multisensory processes* (pp. 581-596). Cambridge: The MIT Press.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Cross-modal binding of fear in voice and face. *Proceedings of the National Academy of Science*, 98, 10006-10010.
- Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of contextual updating ?. *Behavioral and Brain Sciences*, 11, 357-374.
- Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003) Emotional speech: Towards a new generation of database. *Speech Communication*, 40, 33-59.
- Downar, J., Crawley, A.P., Mikulis, D.J., & Davis, K.D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, 3, 277-283.
- Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology*, 10, 731-735.
- Eimer, M., & McCarthy, R.A. (1999). Prosopagnosia and structural encoding of face: evidence from event-related potentials. *Neuroreport*, 10, 255-259.
- Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, 115, 268-287.
- Ekman P., Friesen W. V. (1976). *Pictures of facial affects*, Consulting Psychologist Press: Palo Alto.

-
- Ellis, H.C., & Ashbrook, P.W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition and social behaviour* (pp. 25-43). Toronto: Hogrefe.
- Ellis, A.W., & Young, A.W. (1995). *Human cognitive neuropsychology: A textbook with readings*. Hove: Psychology Press.
- Ellsworth P.C., Scherer, K.R. (2003). Appraisal processes in emotion. In R.J. Davidson, K.R. Scherer, & H.H. Goldsmith, *Handbook of affective sciences* (pp. 572-595). Oxford/New York : Oxford University Press.
- Etcoff, N.L. (1986). The neuropsychology of emotional expression. In G. Goldstein & R.E. Tarter (Eds.), *Advances in clinical neuropsychology* (pp. 49-76). New York: Plenum Press.
- Ettlingen, G., & Wilson, W.A. (1990). Cross-modal performance: Behavioural processes, phylogenetic considerations and neural mechanisms. *Behavioural Brain Research*, 40, 169-192.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience*, 22, 5749-5759.
- Farah, M.J., O'Reilly, R.C., & Vecera, S.P. (1993). Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review*, 100, 571-588.
- Fort, A., Delpuech, C., Pernier, J., & Giard, M.-H. (2002). Early auditory-visual interactions in human cortex during nonredundant target identification. *Cognitive Brain Research*, 14, 20-30.
- Fort, A., & Giard, M.-H. (2004). Multiple electrophysiological mechanisms of audiovisual integration in human perception. In G. Calvert, C. Spence, & B.E. Stein (Eds.), *The handbook of multisensory processes* (pp.503-513). Cambridge: The MIT Press.
- Fridlund A.J. (1994). *Human facial expression*. San Diego, Academic Press.

- Frijda, N.H. (2000). The psychologists' point of view. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 59-74). New York: Guilford Press.
- George, M.S., Parekh, P.I., Rosinsky, N., Ketter, T.A., Kimbrell, T.A., Heilman, K.M., Herscovitch, P. & Post, R.M. (1996). Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology*, *53*, 665-670.
- Giard, M.H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*, 473-490.
- Gobl, C., & Chasaide, A.N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, *40*, 189-212.
- Graziano, M.S., Hu, X.T., & Gross, C.G. (1997). Visuospatial properties of ventral premotor cortex. *Journal of Neurophysiology*, *77*, 2268-2292.
- Green, R.E.A., Turner, G.R., & Thompson, W.F. (2004). Deficits in facial emotion perception in adults with recent traumatic brain injury. *Neuropsychologia*, *42*, 133-141.
- Hadjikhani, N., & Roland, P.E. (1998). Cross-modal transfer of information between the tactile and the visual representations in the human brain: A positron emission tomographic study. *Journal of Neuroscience*, *18*, 1072-1084.
- Halgrean, E., & Marinkovic, K. (2005). Neurophysiological networks integrating human emotions. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1137-1152). Cambridge, The MIT Press.
- Harrè, R., & Parrott, W.G. (1996). *The emotions: Social, cultural, and biological dimensions*. Thousand Oaks: Sage.
- Imaizumi, S., Mori, K., Kiritani, S., Hosoi, H., & Tonoike, M. (1998). Task-dependent laterality for cue decoding during spoken language processing. *Neuroreport*, *9*, 899-903.
- Izard, C.E. (1977). *Human emotions*. New York: Plenum.
- Jasper, H.H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371-375.

- Johnstone, T., & Scherer, K.R. (2000). Vocal communication of emotion. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 220-235). New York: Guilford Press.
- Junghöfer, M., Bradley, M.M., Elbert, T.R., & Lang, P.J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology*, *38*, 175-178.
- Juslin, P.N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, *1*, 381-412.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice Hall.
- Katayama, J., & Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology*, *35*, 23-33.
- Keltner, D., & Ekman, P. (2000). Facial expression of emotion. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of emotions*, 2(nd) edition, (pp. 236-249). New York: Guilford Publications, Inc..
- Killgore, W.D., & Yungelun-Todd, D.A. (2004). Activation of the amygdala and the anterior cingulate during the processing of nonconscious sad versus happy faces. *Neuroimage*, *21*, 1215-1223.
- Klasmeyer, G., & Sendlmeier, W.F. (1997). The classification of different phonation types in emotional and neutral speech. *Forensic Linguistics*, *4*, 104-124.
- Knight, R.T., & Scabini, D. (2003). Funzioni di controllo e lobi frontali del cervello. In Zani A. & Mado Proverbio A. (Eds.), *Elettrofisiologia della mente. Il cervello e le funzioni cognitive* (pp. 215-242). Roma, Carocci.
- Krolak-Salmon, P., Henaff, M.-A., Vighetto, A., Bertrand, O., & Manguiere, F. (2004). Early amygdala reaction to fear spreading in occipital, temporal and frontal cortex: A depth electrode ERP study in human. *Neuron*, *42*, 665-676.
- Lamme, V.A. (2001). Blindsight: The role of feedforward and feedback corticocortical connections. *Acta Psychologica*, *107*, 209-228.
- Lane, R.D., Nadel, L., Allen, J.J.B., & Kaszniak, A.W. (2000). The study of emotion from the perspective of cognitive neuroscience. In R.D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 3-11). New York: Oxford University Press.

- Lazarus, R.S. (2000). Appraisal, relational meaning, stress, and emotion. In K.R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal process in emotion: Theory, methods, research*. New York: Oxford University Press.
- LeDoux, J.E. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. London: Weidenfeld & Nicolson. Trad. It.: *Il cervello emotivo: Alle origini delle emozioni*. Milano: Baldini e Castoldi, 2003.
- LeDoux, J.E., & Phelps, E.A. (2000). Emotional networks in the brain. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 157-172). New York: Guilford Press.
- Levänen, S., & Sams, M. (1997). Disrupting human auditory change detection: Chopin is superior to white noise. *Psychophysiology*, 34, 258-265.
- Liddell, B.J., Brown, K.J., Kemp, A.H., Barton, M.J., Das, P., Peduto, A., Gordon, E., & Williams, L.M. (2005). A direct brainstem-amygdala-cortical "alarm" system for subliminal signals of fear. *Neuroimage*, 24, 235-243.
- Liddell, B.J., Williams, L.M., Rathjen, J., Shevrin, H. & Gordon, E. (2004). A temporal dissociation of subliminal versus supraliminal fear perception: An event-related potential study. *Journal of Cognitive Neuroscience*, 16, 479-486.
- Linkenkaer, K., Palva, J.M., Sams, M., Hietanen, J.K., Aronen, H.J., & Ilmoniemi, R.J. (1998). Face-selective processing in extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neuroscience Letters*, 253, 147-150.
- Luu, P., & Tucker, D.M. (2003). Self-regulation and the executive functions: Electrophysiological clues. In A. Zani & A. Mado Proverbio (Eds.), *In the cognitive electrophysiology of mind and brain* (pp. 199-223). California, Elsevier.
- Macaluso, E., Frith, C.D., & Driver, J. (2000). Modulation of human visual cortex by cross-modal spatial attention. *Science*, 289, 1206-1208.
- Marinckovic, K., & Halgren, E. (1998). Human brain potentials related to the emotional expression, repetition, and gender of faces. *Psychobiology*, 26, 348-356.

-
- Marks, L. E. (1978). *The unity of the senses. Interrelations among the modalities*. New York: Academic Press.
- Marks, L. E. (1987). On cross-modal similarity: audio-visual interactions in speeded discrimination. *Journal of Experimental Psychology. Human Perceptual Performances*, 13, 384-394.
- Marks, L. E. (1989). On cross-modal similarity: the perceptual structure of peach, loudness and brightness. *Journal of Experimental Psychology. Human Perceptual Performances*, 15, 586-602.
- Marks, L. E., Ben-Artzi, E., E, & Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination. *International Journal of Psychophysiology*, 50, 125-145.
- Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review*, 3, 215-221.
- McDonald, S., & Flanagan, S. (2004). Social perception deficits after traumatic brain injury: The interaction between emotion recognition, mentalising ability and social communication. *Neuropsychology*, 18, 572-579.
- McDonald, S., & Saunders, J.C. (2005). Differential impairment in recognition of emotion across different media in people with severe traumatic brain injury. *Journal of the International Neuropsychological Society*, 11, 392-399.
- McGurk, H., & McDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- McLeod, C.M., & McDonald, P.A.(2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, 4, 383-391.
- Meinhardt, J, & Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. *Cognition and Emotion*, 17, 477-500.
- Meredith, M. A. (2002). On the neuronal basis of multisensory convergence: A brief overview. *Cognitive Brain Research*, 14, 31-40.
- Mesulam, M.M. (1998). From sensation to cognition. *Brain*, 121, 1013-1052.

- Morris, J.S., de Gelder, B., Weiskrantz, L., & Dolan, R.J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, 124, 1241-1252.
- Morris, J.S., Öhman, A., & Dolan, R.J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393, 467-470.
- Morris, J.S., Öhman, A., & Dolan, R.J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Science of the United States of America*, 96, 1680-1685.
- Näätänen, R. (1992). *Attention and brain function*. Mahwah: Lawrence Erlbaum Associates.
- Odgaard, E.C., Arieh, Y., & Marks, L.E. (2003). Cross-modal enhancement of perceived brightness: intersensory interaction vs. response bias. *Perceptual Psychophysiology*, 65, 123-132.
- O'Doerty, J., Rolls, E.T., & Kringelbach, M. (2004). Neuroimaging studies of cross-modal integration for emotion. In G. Calvert, C. Spence & B.E. Stein (Eds.), *The handbook of multisensory processes* (pp. 563-579). Cambridge: The MIT Press.
- O'Keefe, B.J. (1988). The logic of message design: Individual differences in reasoning about communication. *Communication Monographs*, 55, 80-103.
- Olson, I.R., Gatenby, J.C., & Gore, J. C. (2002). A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. *Cognitive Brain research*, 14, 129-138.
- Peper, M., & Irle, E. (1977). Categorical and dimensional decoding of emotional intonations in patients with focal brain lesion. *Brain and Language*, 57, 233-264.
- Phillips, M.L., Young, A.W., Scott, S.K., Calder, A.J., Andrew, et al. (2003). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, 265, 1809-1817.
- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Pittam, J., & Scherer, K.R. (1993). Vocal expression and communication of emotion. In M. Levis, J.M. Haviland (Eds.), *Handbook of emotions* (pp. 185-198). New York: Guilford Press.

- Planalp, S., & Knie, K. (2002). Integrating verbal and nonverbal emotion(al) messages. In S.R. Fussell, (Ed.), *The verbal communication of emotions: Interdisciplinary perspectives* (pp. 55-77). Mahwah: Erlbaum.
- Pourtois, G., Debatisse, D., Despland, P.-A., & de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Cognitive Brain Research*, 14, 99-105.
- Pourtois, G., de Gelder, B., Bol, A., & Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex*, 41, 49-59.
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11, 1329-1333.
- Raij, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron*, 28, 617-625.
- Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Natural Review, Neuroscience*, 4, 193-202.
- Renault, B., Signoret, J.L., DeBruille, B., Breton, F. & Bolgert, F. (1989). Brain potentials reveal covert facial recognition in prosopagnosia. *Neuropsychologia*, 27, 905-912.
- Rosch E., Mervis C. (1981). Categorisation of natural objects. In M. Posenszweig & L. Porter (Eds.). *Annual Review of Psychology*, 32.
- Ross, E.D. (2000). Affective prosody and the aprosodias. In M.M. Mesulam (Ed.), *Principle of behavioral and cognitive neurology* (pp. 316-331). Oxford: Oxford University Press.
- Rugg, M.D., & Coles, M.G.H. (1995). *Electrophysiology of mind*. Oxford: Oxford University Press.
- Russell, J.A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, 110, 426-450.
- Russell J. A., Fernández-Dols J. M. (1998). *Psicologia delle espressioni facciali*. Trento: Erikson. Ed. originale: Cambridge: Cambridge University Press, 1997.

- Russell J. A., & Widen S. C. (2002). Words versus faces in evoking preschool children's knowledge of the causes of emotions. *International Journal of Behavioral Development*, 26, 97-103.
- Sams, M., Aulanko, R., Hamalainen, M., Hari, R., Lounasmaa, O.V., Lu, S.T., & Simola, J. (1991). Seeing speech: Visual information from lip movements modify activity in the human brain cortex. *Neuroscience Letters*, 127, 141-145.
- Sato, W., Kochiyama, T., Yoshikawa, S., & Matsumura, M. (2001). Emotional expression boots early visual processing of faces: ERP recording and its decomposition by independent component analysis. *Neuroreport*, 12, 709-714.
- Scherer, K.R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143-165.
- Scherer, K.R. (1988). Criteria for emotion-antecedent appraisal: A review. In V. Hamilton, G.H. Bower & N.H. Frijda (Eds.), *Cognitive perspectives on emotion and motivation* (pp. 89-126). Dordrecht: Kluwer.
- Scherer, K.R. (1994). *Affect bursts*. In S.H.M. van Goozen & Sergeant (Eds), *Emotions: Essays on motion theory* (pp 161-193). Hillsdale: Lawrence Erlbaum.
- Scherer, K.R. (2003). Introduction: Cognitive components of emotion. In R.J. Davidson, K.R. Scherer, H.H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 563-595). Oxford/New York: Oxford University Press.
- Scherer, K.R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of the Affective Sciences* (pp. 433-456). New York: Oxford University Press.
- Schröder, M. (2003). Experimental study of affect burst. *Speech Communication*, 40, 99-116.
- Schroeder, C.E., & Foxe, J.J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Cognitive Brain Researches*, 14, 187-198.
- Schroeder, C.E., Lindsley, R.W., Specht, C. Marcovici, A., Smiley, J.F., & Javuit, D.C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *Journal of Neurophysiology*, 85, 1322-1327.

-
- Schroeder, C. E., Smiley, J., Fu, K. J., McGinnis, T., O'Connell, M. N., & Hackett, T. A. (2003). Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. *International Journal of Psychophysiology*, *50*, 5-17.
- Schupp, H.T., Cuthbert, B.N., Bradley, M.M., Hillman, C.H., Hamm, A.O., & Lang, P.J. (2004). Brain processes in emotion perception: Motivated attention. *Cognition and emotion*, *18*, 593-611.
- Schupp, H.T., Cuthbert, B.N., Bradley, Cacioppo, J.T., Ito, T., & Lang, P.J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*, 257-261.
- Schwarzer G., Zauner N., & Korell M. (2003). Face processing during the first decade of life. In G. Schwarzer & H. Leder (Eds.), *The development of face processing* (pp. 55-80). Göttingen: Hogrefe & Huber.
- Schweinberger, S.R., Robertson, D., & Kaufmann, J.M. Hearing facial identities. *In press*.
- Scott, S.K., Young, A.W., Calder, A.J., & Hallowell, D.J. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature*, *385*, 254-275.
- Stein, B.E., & Wallace, M.T. (1996). Comparisons of cross-modality integration in midbrain and cortex. *Progress in Brain Research*, *112*, 289-299.
- Stenberg, G., Wiking, S., & Dahl, M. (1998). Judging words at face value: Interference in a word processing task reveals automatic processing of affective facial expressions. *Cognition and Emotion*, *12*, 755-782.
- Streit, M., Brinkmeyer, J., Ihl, R., & Gaebel, W. (2000). Electrophysiological correlates of emotional and structural face processing in humans. *Neuroscience Letters*, *278*, 13-16.
- Surakka, V., Tenhunen-Eskelinen, M., Hietanen, J. K., & Sams, M. (1998). Modulation of human auditory information processing by emotional visual stimuli. *Cognitive Brain Research*, *7*, 159-163.

- Taylor, K.I., & Brugger, P.U. (2005). Audiovisual peduncular hallucinations: A release of cross-modal integration sites? *Cognitive & Behavioral Neurology*, 18, 135-136.
- Teder-Sälejärvi, W.A., McDonald, J.J., Di Russo, F., & Hillyard, S.A. (2002). An analysis of audio-visual cross-modal integration by means of event-related potential (ERP) recordings. *Cognitive Brain Research*, 14, 106-114.
- Tomkins, S.S. (1962). *Affect, imagery and consciousness*. New York: Springer.
- Tranel, D., Bechara, A., & Dunburg, N.L. (2002). Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision making and emotional processing. *Cortex*, 38, 589-612.
- Tranel, D., & Damasio, A.R. (1987). Evidence for covert recognition of faces in global amnesic. *Journal of Clinical and Experimental Neuropsychologia*, 9, 15.
- Tranel, D., Damasio, H. & Damasio, A.R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience*, 7, 425-432.
- Tucker, D.M., Watson, R.T., & Heilman, K.M. (1977). Discrimination and evocation of affectively intoned speech in patients with right parietal disease. *Neurology*, 27, 947-950.
- Van Lancker, D., & Sidtis, J.J. (1992). The identification of affective-prosodic stimuli by left- and right. Hemisphere-damaged subjects: All errors are not created equal. *Journal of Speech and Hearing Research*, 35, 963-970.
- Violi P. (1997). *Significato ed esperienza*. Milano: Bompiani.
- von Kriegsten, K., Kleinschmidt, A., Sterzer, P., & Giraud, A.L. (2005). Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience*, 12, 367-376.
- Vroomenn, J., Collier, R., & Mozziconacci, S. (1993). Duration and intonation in emotional speech. In *Proceedings of the Third European Conference on Speech Communication and Technology*, Berlin, pp. 577-580.
- Vuilleumier, P., Armony, J.L., Clarke, K., Husain, M., Driver, J., & Dolan, R.J. (2002). Neural response to emotional faces with and without awareness:

- Event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia*, 40, 2156-2166.
- Vuilleumier, P., Armony, J.L., Driver, J., & Dolan, R.J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6, 524-631.
- Welsch, R.B., & Warren, D.H. (1986). Intersensory interactions. In K.R. Kaufman & J.P. Thomas, *Handbook of human perception and human performance*, Vol. 1 (pp. 1-36). New York: Wiley.
- Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., & Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18, 411-418.
- Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., & Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience*, 18, 411-418.
- Widen S. C., & Russell J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, 39, 114-128.
- Widen, S. C., & Russell, J. A. (2004). The relative power of an emotion's facial expression, label and behavioral consequence to evoke preschoolers' knowledge of its causes. *Cognitive Development*, 19, 111-125.
- Williams, L.M., Liddell, B., Rathjen, J., Shevrin, H., Gray, J.A. et al. (2004). Mapping the time course of nonconscious and conscious perception of fear: An integration of central and peripheral measures. *Human Brain Mapping*, 21, 64-74.
- Young, A.W., Hellawell, D.J., Van de Wal, C., & Johnson, M. (1996). Facial expression processing after amygdectomy. *Neuropsychologia*, 34, 31-39.
- Zald, D.H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research, Brain Research Review*, 41, 83-123.

ALLEGATI

ALLEGATO 1

QUESTIONARIO DI VALIDAZIONE

Il Laboratorio di psicologia cognitiva dell'Università Cattolica del Sacro Cuore di Milano sta conducendo una ricerca sulla percezione delle emozioni espresse mediante il volto e mediante la voce.

Le chiediamo una sua gentile collaborazione.

DATI ANAGRAFICI

Età

Genere

Corso di Laurea

1. Osservi i volti qui di seguito riprodotti e, per ognuno, indichi quale emozione, secondo lei, esso esprime. Inoltre segni su una scala da 1 a 5, per ciascuno dei volti, quanto è sicuro che essa esprima l'emozione da lei indicata.

Ad esempio, se è sicuro che il volto esprima l'emozione da lei indicata, indichi in questo modo la sua risposta:

Per nulla 1 2 3 4 ~~5~~ Del tutto

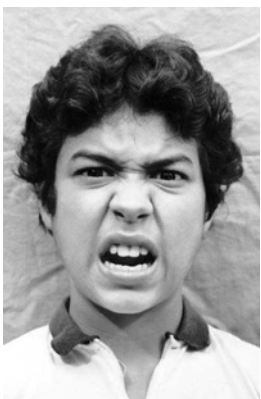
Se invece è poco sicuro della sua risposta:

Per nulla 1 ~~2~~ 3 4 5 Del tutto



1.

Per nulla 1 2 3 4 5 Del tutto



2.

Per nulla 1 2 3 4 5 Del tutto



3.

.....

Per nulla 1 2 3 4 5 Del tutto



4.

.....

Per nulla 1 2 3 4 5 Del tutto



5.

.....

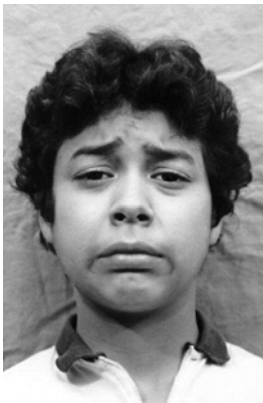
Per nulla 1 2 3 4 5 Del tutto



6.

.....

Per nulla 1 2 3 4 5 Del tutto



7.

.....

Per nulla 1 2 3 4 5 Del tutto

2. Le verranno fatte ascoltare alcune serie di voci. Per ognuna, indichi quale emozione, secondo lei, essa esprime. Inoltre segni su una scala da 1 a 5, per ciascuna delle voci, quanto è sicuro che essa esprima l'emozione da lei indicata. Ad esempio, se è sicuro che la voce esprima l'emozione da lei indicata, indichi in questo modo la sua risposta:

Per nulla 1 2 3 4 ~~5~~ Del tutto

Se invece è poco sicuro della sua risposta:

Per nulla 1 ~~2~~ 3 4 5 Del tutto

Voce 1.....

Per nulla 1 2 3 4 5 Del tutto

Voce 2.....

Per nulla 1 2 3 4 5 Del tutto

Voce 3.....

Per nulla 1 2 3 4 5 Del tutto

Voce 4.....

Per nulla 1 2 3 4 5 Del tutto

Voce 5.....

Per nulla 1 2 3 4 5 Del tutto

Voce 6.....

Per nulla 1 2 3 4 5 Del tutto

Voce 7.....

Per nulla 1 2 3 4 5 Del tutto

ALLEGATO 2
I PATTERNS FACCIALI



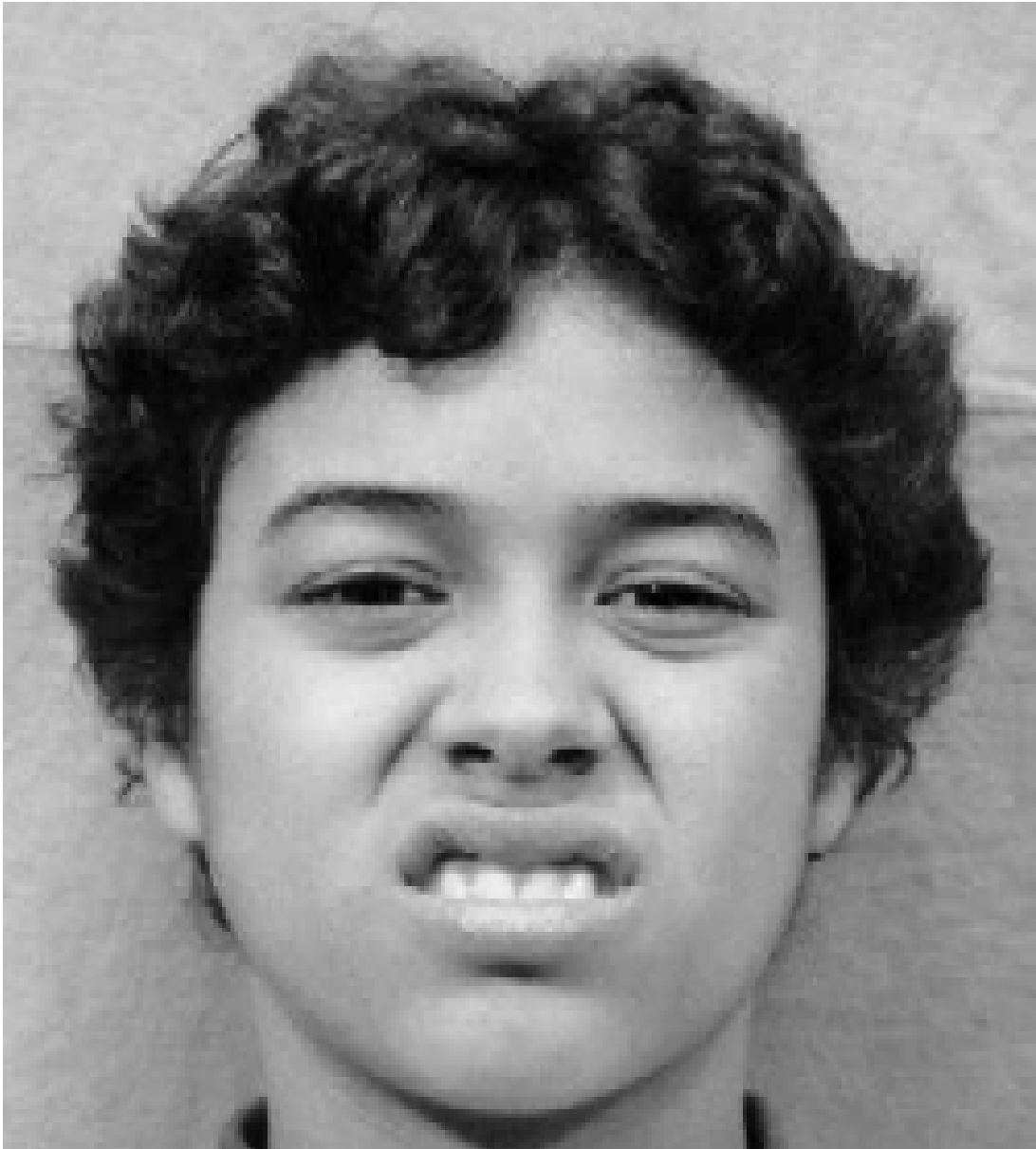
- GIOIA -



- PAURA -



- RABBIA -



- DISGUSTO -



- SORPRESA -



- TRISTEZZA -

ALLEGATO 3
QUESTIONARIO FINALE

Ora che ha terminato l'esperimento, le chiediamo di compilare gentilmente il presente questionario in tutte le sue parti.

DATI ANAGRAFICI

Età

Genere

Corso di Laurea

1. Osservi i volti qui di seguito riprodotti e, per ognuno, indichi quale emozione, secondo lei, esso esprime. Inoltre segni su una scala da 1 a 5, per ciascuno dei volti, quanto è sicuro che essa esprima l'emozione da lei indicata.

Ad esempio, se è sicuro che il volto esprima l'emozione da lei indicata, indichi in questo modo la sua risposta:

Per nulla 1 2 3 4 ~~5~~ Del tutto

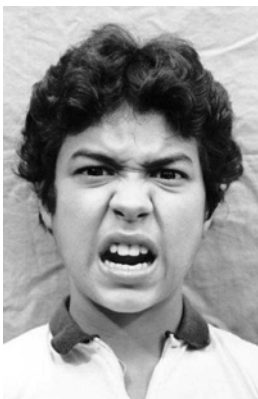
Se invece è poco sicuro della sua risposta:

Per nulla 1 ~~2~~ 3 4 5 Del tutto



1.

Per nulla 1 2 3 4 5 Del tutto



2.

Per nulla 1 2 3 4 5 Del tutto



3.

.....

Per nulla 1 2 3 4 5 Del tutto



4.

.....

Per nulla 1 2 3 4 5 Del tutto



5.

.....

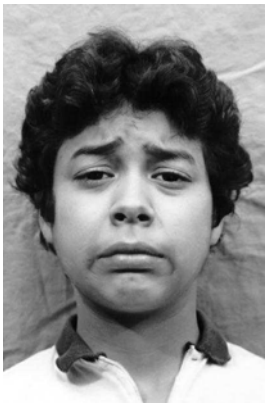
Per nulla 1 2 3 4 5 Del tutto



6.

.....

Per nulla 1 2 3 4 5 Del tutto



7.

.....

Per nulla 1 2 3 4 5 Del tutto

2. Le verranno fatte ascoltare alcune serie di voci. Per ognuna, indichi quale emozione, secondo lei, essa esprime. Inoltre segni su una scala da 1 a 5, per ciascuna delle voci, quanto è sicuro che essa esprima l'emozione da lei indicata. Ad esempio, se è sicuro che la voce esprima l'emozione da lei indicata, indichi in questo modo la sua risposta:

Per nulla 1 2 3 4 ~~5~~ Del tutto

Se invece è poco sicuro della sua risposta:

Per nulla 1 ~~2~~ 3 4 5 Del tutto

Voce 1.....

Per nulla 1 2 3 4 5 Del tutto

Voce 2.....

Per nulla 1 2 3 4 5 Del tutto

Voce 3.....

Per nulla 1 2 3 4 5 Del tutto

Voce 4.....

Per nulla 1 2 3 4 5 Del tutto

Voce 5.....

Per nulla 1 2 3 4 5 Del tutto

Voce 6.....

Per nulla 1 2 3 4 5 Del tutto

Voce 7.....

Per nulla 1 2 3 4 5 Del tutto

3 4 5 Del tutto

3. Le sembra che in alcuni casi durante l'esperimento i due stimoli, visivo ed uditivo, comunicassero emozioni diverse?

.....
.....
.....
.....

4. Infine, questo esperimento le è sembrato difficile e - in caso di risposta affermativa - in quali parti?

.....
.....
.....
.....

5. Infine, l'esperimento le è sembrato faticoso e - in caso di risposta affermativa - in quali parti?

.....
.....
.....
.....

ALLEGATO 4

LA CONSEGNA SPERIMENTALE IMPLICITA

Università Cattolica del Sacro Cuore di Milano

Laboratorio di Psicologia Cognitiva

Dipartimento di Psicologia

La Facoltà di Psicologia, Università Cattolica del Sacro Cuore di Milano, sta conducendo una ricerca presso il Laboratorio di Psicologia Cognitiva.

Stiamo svolgendo un esperimento inerente alla percezione combinata di stimoli visivi e uditivi. Le verranno presentate delle coppie di stimoli visivi (volti) e uditivi (voci). Le chiediamo di prestare attenzione a ciascuna coppia e di premere il tasto sinistro del mouse quando ha terminato questo compito.

Al termine dell'esperimento, le verrà somministrato un breve questionario.

Nel caso avesse qualche dubbio, può tranquillamente richiedere chiarimenti allo sperimentatore che segue l'esperimento.

Grazie per la gentile collaborazione.

ALLEGATO 5

LA CONSEGNA SPERIMENTALE ESPLICITA

Università Cattolica del Sacro Cuore di Milano

Laboratorio di Psicologia Cognitiva

Dipartimento di Psicologia

La Facoltà di Psicologia, Università Cattolica del Sacro Cuore di Milano, sta conducendo una ricerca presso il Laboratorio di Psicologia Cognitiva.

Stiamo svolgendo un esperimento inerente alla percezione combinata di stimoli visivi e uditivi. Le verranno presentate delle coppie di stimoli visivi (volti) e uditivi (voci). Le chiediamo di prestare attenzione a ciascuna coppia e di valutare se gli stimoli sono congruenti o incongruenti. Terminata la valutazione, preme il tasto sinistro del mouse.

Al termine dell'esperimento, le verrà somministrato un breve questionario.

Nel caso avesse qualche dubbio, può tranquillamente richiedere chiarimenti allo sperimentatore che segue l'esperimento.

Grazie per la gentile collaborazione.