

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.

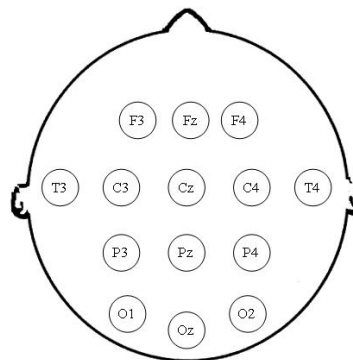


Fig. 2. Distribuzione degli elettrodi sullo scalpo.

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.

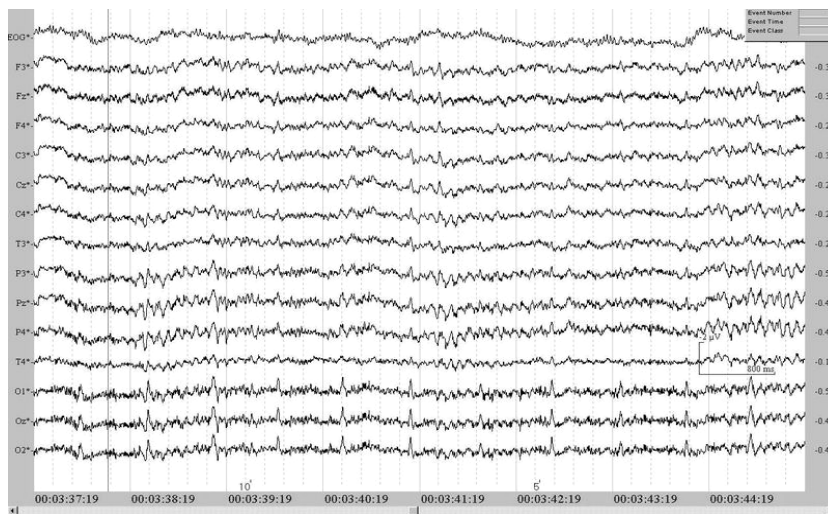
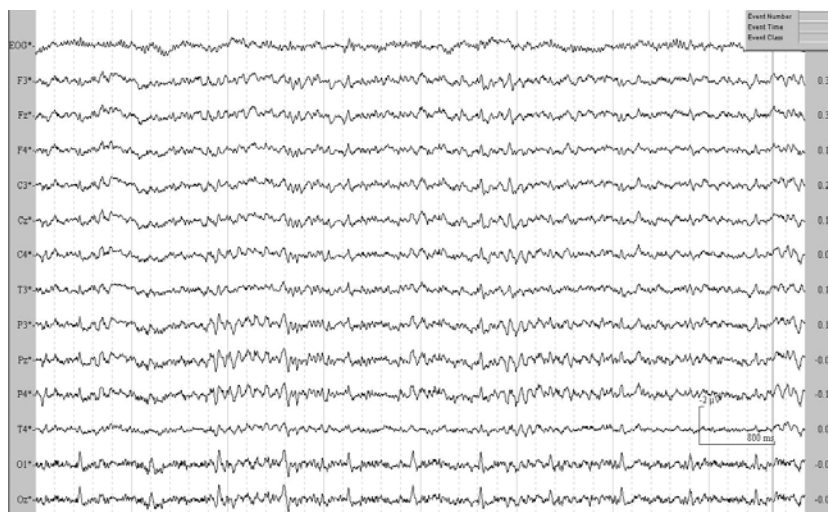


Fig. 3. Output della registrazione prima (a.) e dopo (b.) il filtraggio.

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).

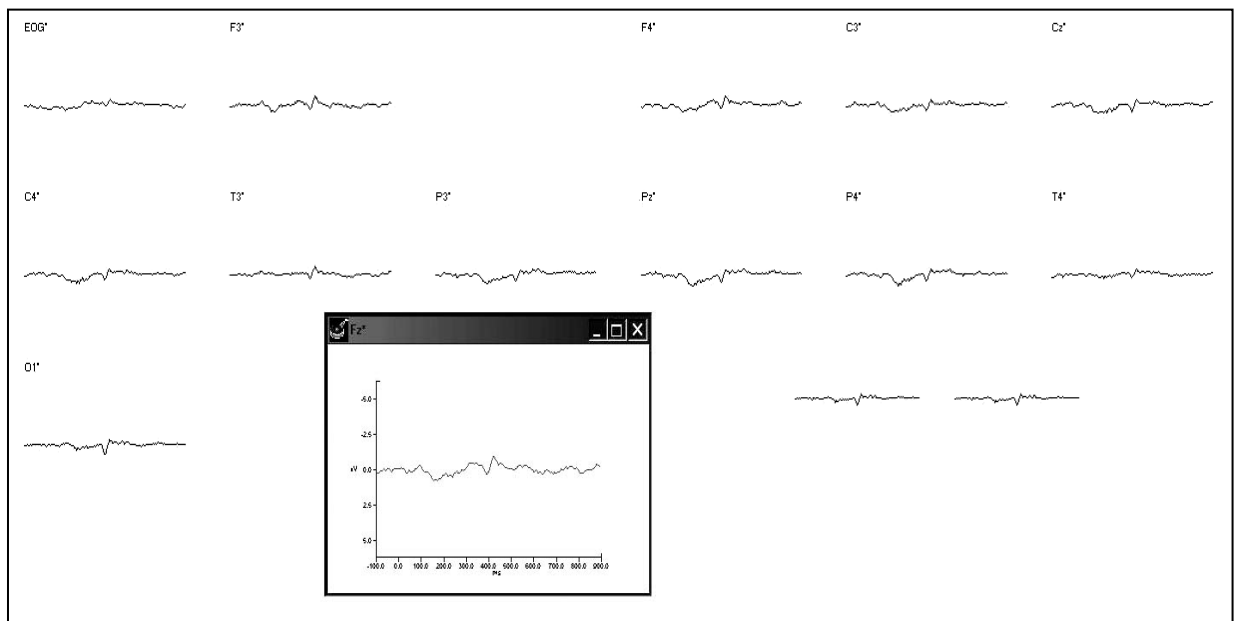


Fig. 4. Esempio di onde epocate.

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).