

# Verso una nuova frenologia? Considerazioni sull'uso dei metodi di brain imaging e di strategie sottrattive per lo studio della cognizione e delle sue basi neurali.

Dario Taraborelli

► **To cite this version:**

Dario Taraborelli. Verso una nuova frenologia? Considerazioni sull'uso dei metodi di brain imaging e di strategie sottrattive per lo studio della cognizione e delle sue basi neurali.. Logic and Philosophy of Science: An Electronic Journal (LPS), 2003, 1 (1). <ijn\_00000313>

**HAL Id: ijn\_00000313**

**[https://jeannicod.ccsd.cnrs.fr/ijn\\_00000313](https://jeannicod.ccsd.cnrs.fr/ijn_00000313)**

Submitted on 7 Jan 2003

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Verso una nuova frenologia ?

### **Considerazioni sull'uso dei metodi di brain imaging e di strategie sottrattive per lo studio della cognizione e delle sue basi neurali**

DARIO TARABORELLI\*

*Institut Jean-Nicod, Paris*

*Abstract* - Metodi non invasivi per la visualizzazione dell'attività cerebrale (*brain imaging*) sono impiegati da ormai due decenni per lo studio della cognizione e delle sue basi neurali. Il loro impianto metodologico - basato su strategie sottrattive importate da studi comportamentali - si fonda su una serie complessa di assunzioni spesso non tematizzate nella letteratura. Nel loro insieme, queste assunzioni disegnano un quadro teorico in cui il cervello umano è concepito come in larga misura costituito di "centri di elaborazione" compatti, altamente specializzati e funzionalmente indipendenti. Questo articolo intende esplorare le assunzioni implicite di tali metodi e le strategie inferenziali che essi adoperano per definire ipotesi di correlazione fra struttura cerebrale e funzione. Un'attenzione particolare sarà dedicata al cambiamento nella nozione di 'sistema isolabile' che l'introduzione di tali metodi ha comportato e al ruolo che l'idea di localizzazione va guadagnando nella spiegazione dei fenomeni cognitivi.

***Towards a new phrenology? Considerations on the use of brain imaging methods and subtractive strategies for the study of cognition and its neural bases.***

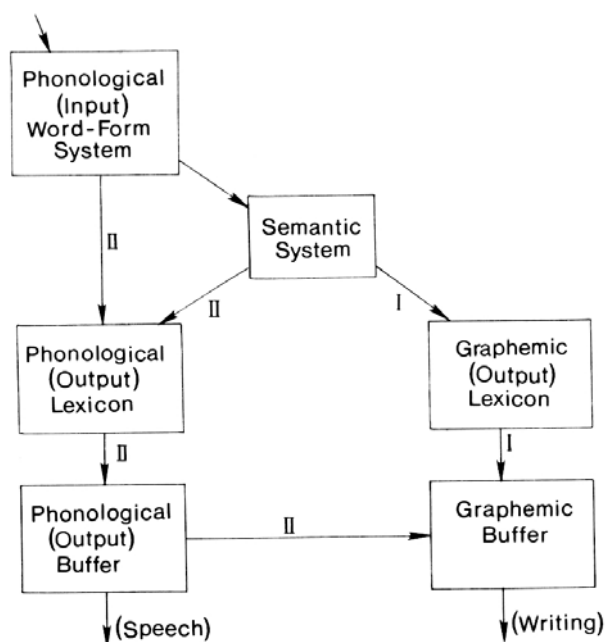
*Abstract* - Non-invasive techniques for the visualization of neural activity in the human brain have been used in the past two decades for studying the neural bases of cognition. Their methodology – based on subtractive strategies that are drawn from behavioral studies – relies on a complex set of assumptions that are seldom made explicit in the literature. Taken together, these assumptions sketch a theoretical framework according to which the human brain is seen as mainly constituted of clustered, highly specialized and functionally independent “processing centers”. The aim of this paper is to explore the implicit assumptions of these methods and the inferential strategies they use to build hypotheses of correlation between structure and function. Special attention will be devoted to the change in the idea of 'isolable subsystem' that these methods have brought in and the new role of localization hypotheses in the explanation of cognitive phenomena.

*Keywords* - brain imaging; PET; fMRI; cognitive subtraction; additive factors; modularity.

\* [dario.taraborelli@ens.fr](mailto:dario.taraborelli@ens.fr)- Institut Jean-Nicod, 1bis avenue de Lowendal, 75007, Paris, France

## 1. Dall'architettura funzionale alle mappe anatomo-funzionali

‘Spiegare i fenomeni della cognizione’ è stato sinonimo per diversi decenni di ‘fornire un modello analitico dei sottosistemi che ne permettono il corretto funzionamento’, così come ‘spiegarne le disfunzioni’ ha significato - e in larga misura significa tuttora - ‘ridescrivere le disfunzioni in termini di deficit o conflitti reciproci fra tali sottosistemi’<sup>1</sup>. L’individuazione di tali componenti (o sottosistemi funzionalmente isolabili<sup>2</sup>), lo studio dei loro deficit e della loro sensibilità a fattori contestuali – in una parola, lo studio dell’architettura funzionale dei processi cognitivi – è stato oggetto di ricerca per intere generazioni di neuropsicologi e psicologi sperimentali.



[Architettura dei processi di decodifica fonologica coinvolti nella scrittura, da SHALLICE (1988)]

Gli ultimi dieci anni hanno visto incrinarsi quest'idea di spiegazione sotto la spinta di diverse tendenze: il crescente interesse per lo studio della cognizione ‘incorporata’, propria di un essere vivente situato nel proprio ambiente, una sempre maggiore

attenzione per le basi biologiche del comportamento e per la storia filogenetica dell'essere vivente e, non ultimo, lo sviluppo di metodologie per lo studio non invasivo dell'attività cerebrale nell'uomo hanno ridimensionato l'attenzione per gli aspetti *esclusivamente* funzionali dei fenomeni cognitivi.

L'indebolimento di un'idea di spiegazione come semplice modellazione in termini di *information processing* (tendenza dominante nella scienza cognitiva dalle sue origini fino alla fine agli anni '80<sup>3</sup>) è andato di pari passo con l'affermarsi delle neuroscienze come disciplina dominante fra i settori di ricerca coinvolti in questo paradigma scientifico.

È forse questione oziosa chiedersi se tale svolta metodologica sia stata effetto di particolari progressi nelle tecniche di rilevamento dell'attività cerebrale nell'uomo o se di questi stessi progressi essa sia piuttosto la causa<sup>4</sup>. È comunque innegabile che la nascita delle "neuroscienze cognitive" sia indissolubilmente legata alla comparsa di nuove strategie sperimentali per l'indagine dei rapporti fra mente e cervello in soggetti umani.

L'impiego di metodi non invasivi per la visualizzazione dell'attività cerebrale nell'uomo, quali PET (*Positron Emission Tomography*) e fMRI (*Functional Magnetic Resonance Imaging*) – metodi che d'ora innanzi indicheremo genericamente come metodi di *brain imaging* – ha permesso di estendere al cervello umano quegli studi che da almeno un secolo a questa parte investigavano la geografia anatomo-funzionale del cervello di diverse specie di mammiferi e di confrontarli con i risultati della neuropsicologia<sup>5</sup>. I metodi di *brain imaging*, sfruttando proprietà note del metabolismo cerebrale (come la correlazione fra attività neurale e afflusso sanguigno) e tecniche avanzate di rilevamento, hanno cioè reso possibile la visualizzazione delle modifiche dell'attività cerebrale sottese allo svolgersi dei diverse 'routines' mentali (siano esse percettive, cognitive, attenzionali, emotive e via dicendo). In altre parole, hanno permesso per la prima volta di testare sull'uomo in modo non invasivo l'ipotesi secondo la quale *specifiche* strutture cerebrali sarebbero deputate a *specifiche* forme di trattamento dell'informazione.

L'ipotesi che il cervello sia in larga misura costituito da popolazioni neurali altamente specializzate per dominio o tipo di trattamento – d'ora innanzi *ipotesi modulare* – insieme con la nascita di metodi per testare empiricamente tale ipotesi hanno influenzato radicalmente il modo in cui la scienza cognitiva *spiega* i suoi fenomeni. Tali metodi hanno cioè contribuito a ridefinire i criteri in virtù dei quali i processi cognitivi sono scomposti e analizzati in sottosistemi.

Con l'avvento del *brain imaging*, i modelli “architetonici” della psicologia cognitiva classica hanno avuto accesso a un confronto con le mappe anatomico-funzionali del cervello umano: i *sottosistemi* postulati dalla psicologia sono cioè divenuti suscettibili di possibili raffronti, conferme o confutazioni sperimentali con l'organizzazione del cervello in *aree corticali*, con la loro struttura anatomica e la loro specializzazione funzionale.

L'idea è colta con particolare efficacia da POSNER, uno dei padri dell'uso del *brain imaging* per lo studio della cognizione.

*To move from elementary operations and their effects upon performance to the level of neural systems, it is important to have methods of localization. Although theories of localization have been present for decades, methods for examining the relation between cognitive function and brain activity have been indirect. (...) The heart of the problem of constructing neural models of cognition is to move from the level of performance to underlying neural systems or microcircuits. Neuroscience approaches have placed somewhat greater emphasis on spatial methods that give hope for studying localization. Cognitive approaches have tended to place emphasis on the temporal organization of information flow in the nervous system. Cognitive neuroscience requires the integration of methods that trace the time dynamics of information processing with those that provide information on the location of neural systems activated* <sup>6</sup>

Intendo partire da questa affermazione programmatica per sottoporre a critica l'impianto teorico e le conseguenze di questo nuovo paradigma metodologico.

Come siamo in grado di stabilire l'esistenza di popolazioni neurali specializzate nel trattamento di determinati tipi di informazione tramite la misura di variazioni del metabolismo cerebrale rilevate in particolari contesti sperimentali ? Come possiamo *isolare* l'attività neurale correlata a un preciso aspetto del comportamento ? Di quali aspetti del comportamento è verosimile cercare basi neurali ? Perché sapere *dove* hanno luogo determinati fenomeni nella corteccia cerebrale può rivestire interesse per capire *come* funzionano i processi mentali ? E prima ancora: siamo in grado di stabilire in modo attendibile *dove* abbiano luogo i processi che ci interessa visualizzare ? Siamo in grado di stabilire *a che cosa* corrisponda l'attività cerebrale misurata ?

Dopo una caratterizzazione dell'*explanandum* e della strategia inferenziale introdotta dai metodi di *brain imaging* (paragrafo 2), i paragrafi successivi del presente articolo - dedicati a un'analisi delle assunzioni che l'uso di tali metodi comporta - cercheranno di strutturare e dare una risposta a tali questioni. I paragrafi 3 e 4 tratteranno delle assunzioni richieste per stabilire correlazioni fra misure dell'attività cerebrale locale e organizzazione anatomo-funzionale del cervello. I paragrafi 5 e 6 approfondiranno a quali condizioni i modelli funzionali della psicologia cognitiva possono essere applicati allo studio delle mappe di attivazione neurale. Il paragrafo conclusivo tenterà un bilancio provvisorio sul ruolo che l'idea di 'localizzazione' è venuta ad assumere nella spiegazione dei fenomeni cognitivi rispetto ai modelli tradizionali.

## 2. Metodo ed *explanandum*: strategie inferenziali negli studi di *brain imaging*.

Gli studi "pilota" che hanno aperto la strada all'impiego di tecnologie di *brain imaging* per investigare i correlati neurali dei processi mentali risalgono agli anni '80: scopo di uno dei primi esperimenti di *brain imaging*<sup>7</sup> era quello di individuare modifiche nel flusso sanguigno locale di particolari aree corticali in corrispondenza della lettura, della ripetizione ad alta voce e della produzione di singole parole. L'aspirazione dichiarata di questo esperimento era quella di riuscire a isolare e dissociare le basi neurali dei vari sottosistemi coinvolti in tali attività: dai meccanismi di riconoscimento visivo degli stimoli, a quelli di trattamento fonologico, a processi genericamente classificati come "operazioni semantiche".

A partire da queste prime esplorazioni, gli studi di *brain imaging* hanno conosciuto una fortuna straordinaria. Una rassegna dei contributi dell'ultimo decennio nelle riviste specializzate mostra a quanti problemi tradizionali della psicologia cognitiva classica sia stata applicata la ricerca di correlati neurali: dalle basi neurali della *visione* – percezione del movimento, del colore, riconoscimento visivo di forme, oggetti, volti... – della *percezione uditiva* – timbro, altezza tonale, localizzazione spaziale degli stimoli sonori... – così come delle altre *modalità sensoriali*, allo studio delle basi del *comportamento motorio* o dei *processi attenzionali*, fino alle ricerche più audaci volte a studiare i correlati neurali della *coscienza*, della *concettualizzazione*, del *comportamento emotivo* – paura, ansietà, eccitazione sessuale – dei meccanismi alla base delle *interazioni sociali* – ‘mind reading’ e capacità di mentalizzazione – o di determinate *sindromi psicopatologiche* – quali autismo e schizofrenia).

Diversi studiosi hanno deplorato, a margine di questa fioritura di indagini sperimentali, la mancanza di un'adeguata riflessione sull'impianto metodologico degli studi di *brain imaging*: PET e fMRI si sono rivelate tecniche di indagine talmente versatili da favorire in certi casi un abuso della loro applicazione: la realizzazione di esperimenti in cui al rilevamento dell'attività neurale non sarebbe stata affiancata una adeguata caratterizzazione dell'*explanandum*<sup>8</sup>.

La prima distinzione che ritengo necessaria per una valida caratterizzazione epistemologica di tali studi è quella che riguarda i tipi di questioni cui l'applicazione del *brain imaging* può rispondere. Tali questioni e le rispettive strategie esplicative possono essere raggruppate in due grandi classi: la prima (A) parte da ipotesi circa l'architettura funzionale di determinate capacità per interpretare le mappe di attivazione neurale; la seconda (B) parte, inversamente, da ipotesi circa la specializzazione di aree corticali per inferire l'intervento o l'assenza di particolari operazioni nello svolgimento di un processo.

A. La prima classe di domande concerne il modo in cui il trattamento dell'informazione è implementato nel cervello. Partendo da un modello che descriva le operazioni richieste per la realizzazione di un determinato processo, si tratta di investigare se esistano aree cerebrali correlate a ciascuna di queste operazioni. Consideriamo un esempio elementare, quello della lettura ad alta voce di parole presentate visivamente. Un modello dei meccanismi coinvolti dalla lettura ad alta voce dovrà rendere conto almeno dei seguenti processi: la decodifica visiva della parola scritta, la sua eventuale associazione ad un *item* lessicale corretto, il recupero di una forma fonologica adeguata e la coordinazione e programmazione motoria richiesta dall'articolazione della parola. Definito l'insieme di queste operazioni, è possibile porre la domanda se esistano circuiti neurali distinti nella corteccia cerebrale sottesi a ciascuna di esse. L'*explanandum* di questa classe di domande riguarda, in sintesi, la localizzazione cerebrale di determinati componenti postulati a livello funzionale.

B. Una diversa classe di domande riguarda la ricerca dei *contesti* in cui ricorre l'attivazione di determinate aree corticali. Ovverosia, a partire da ipotesi circa la specializzazione funzionale di determinate regioni del cervello, è possibile chiedersi quali processi le coinvolgano e quali invece non le coinvolgano. Supponiamo, per esempio, di disporre di un'ipotesi empiricamente consolidata circa l'esistenza di un'area specifica della corteccia linguistica deputata al trattamento fonologico della parola scritta. Possiamo chiederci quali tipi di compiti richiedano l'intervento di tale trattamento, ad esempio se il compito di associare un'immagine al significato di una



parola presentata visivamente coinvolga o meno tali meccanismi di decodifica fonologica: il rilevamento di un'attività cerebrale localizzata in quest'area potrà essere interpretato come indice di un trattamento fonologico in corso. L'*explanandum* di questa seconda classe di domande, a differenza di quelle della classe precedente, concerne l'*architettura* di determinate capacità cognitive e la sensibilità al contesto di determinati sottosistemi.

Strategie del secondo tipo (B) sono evidentemente subordinate alla disponibilità di specifiche ipotesi sulla specializzazione funzionale di particolari strutture, ipotesi che possono provenire a loro volta da studi di *brain imaging* del primo tipo (A) o da evidenza sperimentale di altra natura (studi di lesione e considerazioni di anatomia comparata, dati elettrofisiologici etc.). Questioni di tipo (B) richiedono cioè che disponiamo già di un criterio autonomo per stabilire '*come isolare un'area cerebrale*'<sup>9</sup>.

Le ricerche improntate al primo tipo di strategia inferenziale risultano quindi in un certo senso prioritarie rispetto alle ricerche del secondo tipo e non è un caso che queste ultime siano di gran lunga più giovani<sup>10</sup>

Riassumendo, ho distinto due classi di domande alle quali corrispondono due distinte strategie inferenziali: possiamo infatti formulare ipotesi sull'anatomia funzionale del cervello a partire da modelli dell'organizzazione funzionale di date capacità (A: '*quale popolazione neurale svolge una data funzione?*') o, viceversa, costruire ipotesi sull'architettura funzionale di tali capacità a partire da osservazioni sull'attivazione di particolari circuiti neurali e da ipotesi sul ruolo funzionale di tali circuiti (B: '*quali funzioni svolge una data popolazione neurale?*').

In entrambi i casi, qualsiasi tipo di inferenza fra l'anatomia funzionale del cervello e l'architettura funzionale dei processi cognitivi si fonda su una serie articolata di nessi epistemologici, la validità dei quali è legata a specifiche ipotesi. L'affidabilità di queste strategie inferenziali richiede cioè che disponiamo di vincoli metodologici per interpretare il rilevamento di specifici *pattern* di attività cerebrale come *indice* dell'intervento di specifiche operazioni o, viceversa, per interpretare l'intervento di specifiche operazioni come *causa* di particolari *pattern* di attività neurale.

Il presente articolo si concentrerà sulle varie garanzie metodologiche richieste dall'interpretazione 'funzionale' delle mappe di attivazione neurale – garanzie che possiamo classificare in tre livelli distinti, dei quali almeno il secondo e soprattutto il terzo hanno un carattere genuinamente epistemologico.

- Il primo livello, legato ad aspetti puramente tecnici del metodo di rilevamento, riguarda l'accuratezza e i limiti di risoluzione dei dati acquisiti.
- Il secondo livello di problemi riguarda l'affidabilità della misura dei parametri fisiologici (volume del flusso sanguigno, tasso di ossigenazione, concentrazione glucidica) come indice dell'attività cerebrale.
- Il terzo livello, infine, riguarda l'affidabilità delle strategie di scomposizione funzionale impiegate per 'selezionare' l'attività cerebrale pertinente e rappresenta il cuore dell'impianto teorico del *brain imaging*.

Ai primi due livelli sono dedicati i due paragrafi successivi, il terzo livello merita invece una discussione più articolata cui toccheranno i paragrafi seguenti.

### 3. I limiti della misura: risoluzione spaziale e temporale

I vincoli meno interessanti dal punto di vista teorico, nondimeno cruciali per stabilire l'affidabilità del metodo riguardano la precisione e i limiti della misura. Da un punto di vista tecnico, i tempi di acquisizione delle immagini e la risoluzione spaziale dei dati acquisiti vanificano il sogno di una visualizzazione dettagliata e in tempo reale dei cambiamenti globali che intervengono nel metabolismo cerebrale nel corso dello svolgimento di una data attività.

Sotto il profilo spaziale, le tecniche di *brain imaging* coniugano una risoluzione spaziale sufficientemente fine (dell'ordine di alcuni millimetri) ad un volume di campionamento adeguato alla misurazione dell'attività di intere aree corticali. Ciò che va perso a questo livello di misura è la possibilità di confrontare le mappe del metabolismo cerebrale con quell'insieme considerevole di informazioni di carattere anatomo-fisiologico sull'organizzazione e la connettività dei tessuti neurali che provengono dagli studi microstrutturali delle neuroscienze tradizionali. A quali precisi neuroni corrisponda l'attivazione rilevata dalle scansioni, è difficile stabilire con precisione. La conseguenza più grave di questo limite emerge soprattutto nella notevole variabilità anatomica che si riscontra da individuo a individuo e nella conseguente necessità di ricorrere a metodi di normalizzazione della conformazione anatomica (riduzione ad un cervello standard descritto tramite un sistema di coordinate tridimensionali) e - problema più dibattuto - al confronto statistico di scansioni effettuate su diversi soggetti per generalizzare i risultati ottenuti<sup>11</sup>. Senza entrare nel merito della discussione, mi limiterò qui ad osservare che i problemi relativi alla normalizzazione anatomica nascondono assunzioni fondamentali relative all'omogeneità interindividuale delle strutture corticali e della loro funzione. Anche in presenza di precise omogeneità anatomiche, morfologiche e citoarchitettoniche fra aree cerebrali corrispondenti di diversi individui, è lecito sollevare il dubbio se strutture cerebrali *corrispondenti* in diversi individui sottendano allo *stesso* tipo di trattamento. Su questo punto è possibile misurare la distanza che intercorre fra l'affermare, ad esempio, che "il lobo anteriore dell'ipofisi è responsabile della secrezione dell'ormone STH" e affermare invece che "l'area FFA localizzata nel

giro fusiforme destro è associata al riconoscimento dei volti”. In mancanza di livelli descrittivi intermedi tra il livello fisiologico e quello comportamentale, non esistono garanzie a tutela del fatto che la stessa popolazione neurale svolga la stessa funzione in individui diversi.

Una seconda classe di problemi riguarda la *risoluzione temporale* dei rilevamenti. La versatilità degli strumenti di cui le neuroscienze cognitive dispongono sta nella possibilità di effettuare misure ripetute, se non in tempo reale, dell’attività cerebrale: si distinguono infatti tecniche in grado di eseguire misurazioni indipendenti ripetute, misurazioni temporalmente continue, singole misurazioni di lunga durata o singole misurazioni simultanee di diverse variabili fisiologiche. Mentre tecniche meno precise sotto il profilo spaziale (come l’EEG) permettono misurazioni temporali continue (tali cioè da rilevare cambiamenti che avvengono anche all’ordine del centinaio di millisecondi – l’ordine di grandezza della maggior parte dei fenomeni di interesse per la scienza cognitiva), le tecniche di *brain imaging* in senso stretto (PET e fMRI) permettono esclusivamente scansioni reiterate e discontinue. I parametri in gioco, in questo caso, sono la *risoluzione temporale* propriamente detta e la *frequenza di scansione*. Dal momento che i dati ottenuti attraverso *brain imaging* sono immagini “statiche” di un’attività cerebrale che si sviluppa nel tempo, è chiaro che entrambi questi parametri sono cruciali per la rilevanza dei dati rispetto ai fenomeni che si intendono studiare. Più in particolare, quando si parla di risoluzione temporale per tecniche di *brain mapping* è necessario considerare due diversi aspetti: la risoluzione temporale dello strumento di misura e l’ordine di grandezza temporale del processo che si intende misurare. La condizione minima che va rispettata è ovviamente che la risoluzione dello strumento sia commisurata alla grana temporale del processo. Nel caso della PET tale condizione viene palesemente violata: le immagini ottenute attraverso questa tecnica richiedono infatti un tempo di acquisizione che varia dai 30 secondi ai 2 minuti, durante i quali al soggetto viene richiesto di reiterare o protrarre un preciso compito. L’*fMRI*, al contrario, permette di ottenere e confrontare risposte evocate a partire da singoli stimoli o eventi con una frequenza che può giungere fino a circa 20 scansioni al secondo, con il vantaggio di liberare le misure da quelle interferenze che nella *PET* possono derivare da

modulazioni attenzionali richieste per la reiterazione o la protrazione del compito effettuato dal soggetto<sup>12</sup>. Va però osservato che neppure l'*fMRI* può considerarsi una tecnica di misurazione *continua* sotto il profilo temporale. Se infatti da un lato i tempi teorici di misura sono notevolmente ridotti (per quanto riguarda singoli “voxel” o unità spaziali di rilevamento del segnale), dall'altro, non vanno trascurati due fattori limitanti: in primo luogo, la scansione globale di intere sezioni cerebrali richiede tempi dell'ordine del secondo, ovvero tempi di almeno un ordine di grandezza superiore a quello dei fenomeni che si intende misurare; in secondo luogo, il metabolismo sanguigno dei tessuti (che, come vedremo nel paragrafo successivo, è utilizzato come indice dell'attività neurale locale) presenta tempi di latenza superiori a quelli dell'attività neurale. L'insieme di questi fattori fa sì che sotto il profilo della dinamica temporale - uno dei parametri principali della psicologia sperimentale per lo studio dei fenomeni cognitivi - i metodi di *brain imaging* siano estremamente poveri e richiedano quindi l'integrazione di altre metodologie più precise da questo punto di vista.

#### 4. L'affidabilità degli "indici": dal metabolismo all'attività neurale

I primi problemi più propriamente epistemologici si riscontrano al livello dell'affidabilità della misura. È nota la differenza che intercorre tra metodi di misurazione *diretta* dell'attività neurale (tecniche elettrofisiologiche, potenziali evocati, magnetoencefalografia etc... - tecniche che misurano il segnale elettrico o il campo magnetico da esso indotto) e metodi di misurazione *indiretta* della stessa, metodi che cioè inferiscono la variazione di tale attività elettrica in funzione della variazione di altri parametri fisiologici. È questo il caso di tutte le tecniche appartenenti alla categoria del *brain imaging* propriamente detto: tanto la *PET* quanto l'*fMRI* misurano indirettamente l'attività del cervello sotto l'ipotesi che essa sia direttamente legata alle variazioni dell'irrorazione sanguigna locale dei tessuti (*cerebral blood flow, CBF*) o di parametri fisiologici più o meno direttamente correlati a quest'ultima (come il tasso di ossigenazione o la concentrazione glucidica). L'affidabilità di questa misurazione indiretta dell'attività neurale è vincolata da una nutrita serie di assunzioni sull'uso di tali parametri in qualità di "indici"<sup>13</sup>.

- a. variazioni nel CBF riflettono variazioni nell'attività neurale locale.
- b. i profili di distribuzione di CBF rilevati forniscono una base affidabile per predire proprietà permanenti delle aree osservate (la loro attivazione in corrispondenza di tipi particolari di trattamento), anche se essi non manifestano altro che processi fisiologici temporanei.
- c. aumenti di CBF corrispondono normalmente a processi eccitatori.
- d. Il CBF aumenta monotonicamente rispetto al grado di attività neurale, ovvero più un'area cerebrale è attiva, più aumenterà l'ossigenazione dei suoi tessuti.
- e. maggiore è l'incremento del CBF, maggiore il contributo dell'area cerebrale alla realizzazione del processo interessato.
- f. aree coinvolte nella stessa misura da un processo mostrano un pari aumento di CBF.
- g. assenza di variazioni di CBF è indice di assenza di variazioni nell'attività neurale dei tessuti corrispondenti.

Alcune di queste assunzioni sono talmente consolidate da non essere generalmente messe in discussione (è il caso di a. e d.)<sup>14</sup>. Le altre invocano, allo stato attuale delle conoscenze, un'ulteriore convalida empirica e sono di conseguenza più esposte delle prime a possibili revisioni.

L'assunzione b., intesa in senso debole, esclude che la misura dell'attività cerebrale attraverso le variazioni del flusso sanguigno sia suscettibile di importanti variazioni dovute a fattori quali l'affaticamento, l'ansietà o particolari condizioni del soggetto – fattori che, pur senza interessare l'anatomia, possano temporaneamente alterare, a livello fisiologico, la regolarità dei processi misurati. Interpretata in senso forte, tale assunzione richiede che le *stesse* popolazioni neurali realizzino lo *stesso* tipo di trattamento e producano le *stesse* variazioni metaboliche, indipendentemente dal contesto in cui tale trattamento è richiesto: è questo un problema di primaria importanza, le cui conseguenze esulano da quelle strettamente relative all'affidabilità della misura, e sulle quali torneremo di conseguenza nei paragrafi successivi.

Per quanto riguarda l'assunzione c. va rilevato che la “polarità” dell'attivazione cerebrale (eccitatoria o inibitoria) non è discernibile attraverso i parametri di misura adottati dal *brain imaging*. Particolari aumenti di flusso sanguigno in alcune aree potrebbero cioè indicare dispendiosi processi inibitori richiesti dal compito e suggerire dunque un aumento di attività in aree la cui funzione deve, a rigore, essere mantenuta inibita. Per quanto anche una stessa attività inibitoria possa essere di potenziale interesse, le conseguenze di questa indiscernibilità sul piano dell'interpretazione dei dati rilevati sono cruciali<sup>15</sup>.

Dall'assunzione e. dipende in parte l'interpretabilità dei dati ricavati dal *brain imaging* funzionale. Non esiste garanzia del fatto che un incremento dell'attività neurale locale sia legato in modo sistematico al grado di partecipazione di determinate popolazioni neurali al processo in questione. È noto come il *priming* e l'apprendimento possano ridurre notevolmente le risorse impiegate per realizzare uno stesso tipo di operazione, con la conseguenza che tra l'incremento di irrorazione sanguigna locale di una popolazione neurale e la sua partecipazione al tipo di trattamento studiato non esiste necessariamente una relazione diretta.

L'assunzione f. rappresenta un criterio minimo a garanzia dell'uniformità delle misurazioni: la parità della richiesta sanguigna da parte di aree coinvolte nella stessa misura da un determinato compito. Nulla garantisce in realtà che lo svolgimento di un processo richieda una ripartizione dell'irrorazione sanguigna omogenea per ciascuna delle aree coinvolte<sup>16</sup>.

L'assunzione g., infine, non contempla il caso in cui aumenti e diminuzioni contestuali del CBF all'interno di un'area circoscritta possano neutralizzarsi vicendevolmente, non facendo così emergere alcuna variazione rilevabile nel CBF complessivo, malgrado l'effettivo coinvolgimento dell'area nel processo in esame<sup>17</sup>.

L'insieme di queste osservazioni mostra la complessità dei fattori di cui si deve tener conto in sede sperimentale semplicemente per garantire l'affidabilità degli indici metabolici - indici su cui poggia ogni ulteriore interpretazione dei dati. Protocolli perfezionati e metodi per potenziare il rapporto segnale/rumore permettono di limitare gli effetti di questi fattori. Si tratta nondimeno di debolezze metodologiche non trascurabili nel momento in cui esse si inseriscono nella catena inferenziale complessiva che va dalla singola misurazione di un parametro fisiologico all'interpretazione funzionale dell'attività rilevata.

Questa rassegna conclude l'analisi dei vincoli più propriamente empirici sull'impianto inferenziale degli studi di *brain imaging*. Le questioni teoriche più feconde si pongono tuttavia al livello successivo, quello dei modelli funzionali impiegati per la selezione dei dati fisiologici.

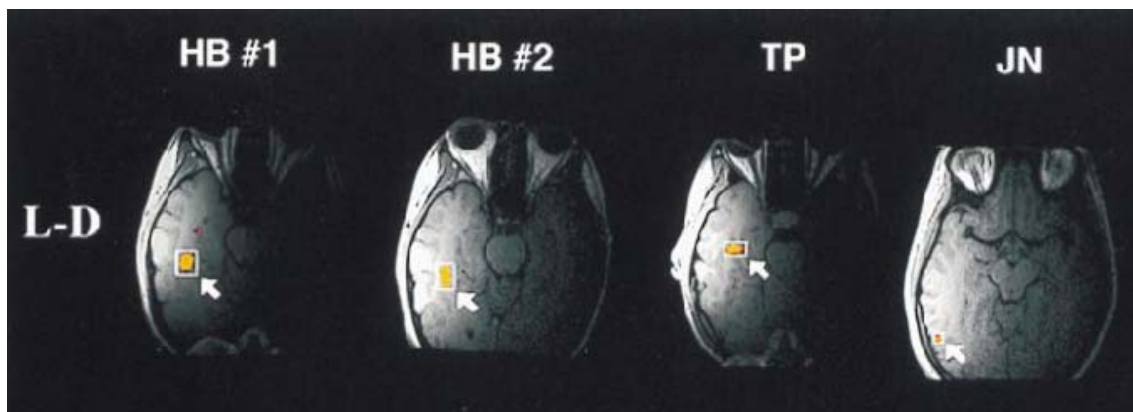


## 5. Il metodo sottrattivo: come scomporre i processi cognitivi ?

*“Se fosse possibile vedere attraverso la scatola cranica e se la zona maggiormente eccitata fosse luminosa, si potrebbe seguire, in un uomo intento a pensare, lo spostamento incessante di questo punto luminoso, in un continuo cambiamento di forma e dimensione, e circondato da una zona d’ombra più o meno fitta che occuperebbe tutto il resto degli emisferi”*

Pavlov (1927)

La citazione da Pavlov, straordinariamente profetica per gli anni in cui fu formulata, tratteggia in modo sommario, ma abbastanza preciso, l'idea soggiacente a tutti gli studi contemporanei di *brain imaging*. Un aspetto di questa citazione merita attenzione: l'idea che l'attività cerebrale legata al pensiero dell'individuo è osservata nei suoi movimenti continui attraverso la scatola cranica, sarebbe "circondata da una zona d'ombra più o meno fitta che occuperebbe tutto il resto degli emisferi". È singolare notare come tale osservazione si applichi perfettamente alle immagini delle mappe di attivazione riportate dalla letteratura di studi di *brain imaging*.



[Differenza nel segnale BOLD (blood-oxygenation level-dependent signal) durante fissazione passiva di lettere e fissazione passiva di cifre in uno studio di risonanza magnetica funzionale condotto da POLK & FARAH (1998)]

Che il pensiero non sia *un punto luminoso in continuo movimento all'interno di una zona d'ombra* è conseguenza triviale del fatto che non esiste un cervello che si attivi solo in presenza di "singoli eventi mentali"<sup>18</sup>. O, inversamente, che la nozione di un "singolo evento mentale" è un'idealizzazione ingenua priva di valore scientifico. La metafora di Pavlov è deviante, ma suggerisce una condizione essenziale per lo studio dell'attività neurale sottesa ai processi cognitivi: la *leggibilità*.

Le immagini "pavloviane" che ricorrono nella letteratura del *brain imaging* sono il frutto di manipolazioni atte ad isolare differenze di attivazione tra le varie condizioni imposte al soggetto, manipolazioni che rappresentano l'unica risorsa disponibile per estrarre dei dati "leggibili" o "interpretabili" dall'insieme della normale attività cerebrale del soggetto.

L'intero impianto metodologico del *brain imaging* si fonda su strategie per la selezione di dati leggibili dalle mappe metaboliche dell'attività cerebrale – strategie che nella loro forma più diffusa prendono il nome di *metodi sottrattivi*. Scopo precipuo della sottrazione, nella sua applicazione al *brain imaging*, è quello di isolare nel modo più preciso possibile l'attività neurale indotta da singole operazioni eseguite durante lo svolgimento di un dato compito: segregare, cioè, l'informazione "rilevante" da tutti quegli effetti secondari dovuti alla normale attività del cervello. Su che cosa debba intendersi per informazione "rilevante" si gioca tutta la validità dell'impianto metodologico del *brain imaging*<sup>19</sup>.

Per capire il ruolo dei metodi di sottrazione occorre fare un passo indietro nella storia delle scienze cognitive. In origine, il metodo sottrattivo costituiva uno degli espedienti teorici più fecondi sviluppati nell'ambito della psicologia sperimentale e dello studio dei tempi di reazione. La sua invenzione risale alla figura del fisiologo olandese F.C. DONDERS e ai suoi studi sulla velocità dei processi mentali<sup>20</sup>. Donders aveva sviluppato una prima rudimentale tecnica sottrattiva che, nelle sue intenzioni, avrebbe permesso di 'isolare le differenti operazioni compiute da un soggetto' nel corso di una determinata attività cognitiva. Consideriamo il seguente esempio. Un soggetto viene sottoposto alla presentazione casuale di due diversi stimoli luminosi e istruito a reagire premendo un pulsante soltanto al presentarsi di uno di essi: l'intuizione di Donders fu quella di

ipotizzare che, sottraendo dal tempo totale impiegato per riconoscere lo stimolo e premere il pulsante il tempo della pura reazione allo stimolo, si sarebbe potuta isolare la durata dell'operazione mentale corrispondente alla "discriminazione" dei due stimoli luminosi. Analogamente, istruendo il soggetto a reagire attraverso due differenti pulsanti ai due stimoli, Donders riteneva di poter isolare per via sottrattiva la durata dell'operazione relativa alla "selezione fra due risposte alternative", e così via.

Il metodo fu ripreso e perfezionato da S. STERNBERG un secolo più tardi. L'idea di Donders, pur nella sua apparente efficacia, dipendeva da un'assunzione estremamente costosa sul piano metodologico e già aspramente criticata dai suoi contemporanei: l'idea che le operazioni svolte dal soggetto si succedessero le une alle altre in modo strettamente seriale. Sternberg propose un perfezionamento del metodo sottrattivo – detto dei "fattori additivi" – e basato su un'assunzione meno forte e ciononostante, come vedremo in seguito, non meno problematica di quella di serialità: il presupposto dell'*a d d i t i v i t à* o della pura inserzione.

Sternberg riuscì a mostrare che determinate manipolazioni delle variabili in gioco in un esperimento di classificazione di stimoli visivi alteravano in modo sistematico i tempi di reazione del soggetto.

Consideriamo l'esempio di un soggetto istruito a reagire al presentarsi su uno schermo di una lettera appartenente ad una sequenza precedentemente memorizzata. Sternberg poté dimostrare che i tempi di reazione relativi a questo compito rimanevano invariati a seguito di determinate manipolazioni degli stimoli e che variavano invece sistematicamente a seguito di altre manipolazioni. Nella fattispecie, manipolazioni dell'aspetto visivo dello stimolo (deterioramento dei contorni, deformazione della lettera..) non influenzano altro che le operazioni di decodifica visiva lasciando inalterati i tempi relativi alle altre operazioni coinvolte, quali per esempio il confronto mentale dello stimolo con la sequenza precedentemente memorizzata. Viceversa, manipolazioni della lunghezza della sequenza da memorizzare producono un rallentamento dei tempi di confronto mnemonico, senza influenzare quelli di decodifica visiva.

La scoperta di Sternberg consisteva principalmente nel fatto che si potessero *dissociare* le diverse operazioni coinvolte in un dato processo sfruttando la loro

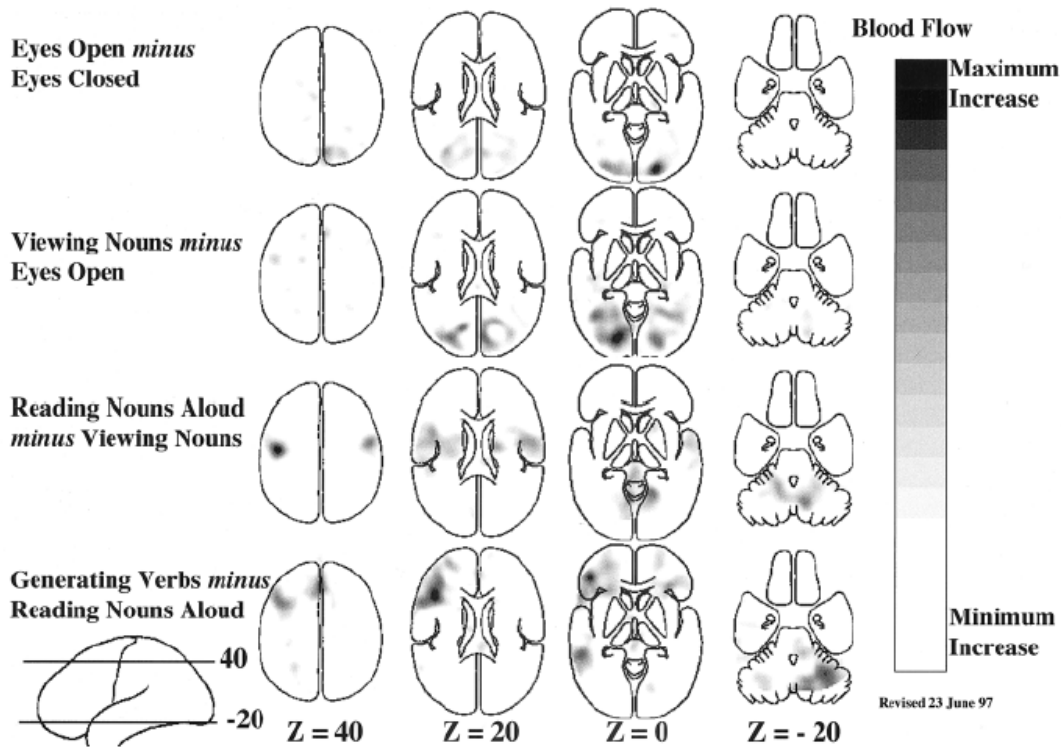
*insensibilità* a manipolazioni selettive di determinate variabili (oggi diremmo delle variabili "ortogonali"). L'ipotesi dei fattori additivi, in altre parole, altro non significava che la possibilità di individuare variabili influenti *selettivamente* su singole operazioni e non su altre.

L'importazione del metodo sottrattivo e dei fattori additivi nell'impianto teorico dei *brain imaging studies* non fu affatto passaggio scontato. Essa è dovuta principalmente alla collaborazione fra il neurologo M. RAICHLE e lo psicologo M. POSNER, erede quest'ultimo della tradizione di studi sulla cronometria mentale e pioniere dell'applicazione della PET alla ricerca delle basi neurali della cognizione<sup>21</sup>.

Le strategie adottate da Posner ricalcano di pari passo i protocolli di Donders e di Sternberg. Il primo principio adottato da Posner (corrispondente alla sottrazione nella sua forma più semplice) consiste nel sottrarre voxel per voxel, nelle mappe di attivazione rilevate dalla PET, i valori relativi allo stato di controllo (*baseline*) da quelli relativi allo stato "attivato". Per esempio, sottraendo dai valori ottenuti chiedendo al soggetto di fissare uno stimolo visivo i valori di controllo registrati nello stesso soggetto mentre tiene gli occhi chiusi, si può verosimilmente ipotizzare di poter isolare l'attività cerebrale correlata alla sola fissazione passiva dello stimolo visivo.

Il principio è semplice: per identificare nelle mappe di attivazione l'effetto di singole operazioni, abbiamo bisogno di un criterio per isolare tale effetto dal resto dell'attività misurata. Si tratta di concepire compiti che differiscano fra loro il meno possibile (o - per ipotesi - di una sola operazione<sup>22</sup>) tali cioè da generare schemi di attivazione "quasi" identici. Confrontando i valori delle due condizioni, è in principio possibile identificare - dalle loro differenze - quali aree di attività siano associate esclusivamente all'operazione che distingue le due condizioni.

Uno studio successivo dello stesso Posner<sup>23</sup> mostra come generalizzare la sottrazione ad una *gerarchia* di condizioni, in cui ogni condizione è utilizzata come controllo per la condizione successiva.



[RAICHLÉ (1998)]

Sottraendo tra loro i pattern ottenuti nelle condizioni "fissazione di un nome" e "occhi aperti" – Raichle ipotizza – dovremmo essere in grado di segregare l'attivazione relativa al puro riconoscimento visivo di stimoli linguistici, sottraendo da "lettura di un nome ad alta voce" "percezione visiva di un nome" l'attivazione relativa alla produzione verbale richiesta dalla lettura e, infine, sottraendo "lettura di un nome ad alta voce" da "produzione spontanea di un verbo" le aree relative alle modulazioni attenzionali richieste dall'elocuzione spontanea.

Le mappe di attivazione ottenute tramite strategie sottrattive rappresentano un caso esemplare di risultati sperimentali *carichi di teoria*.

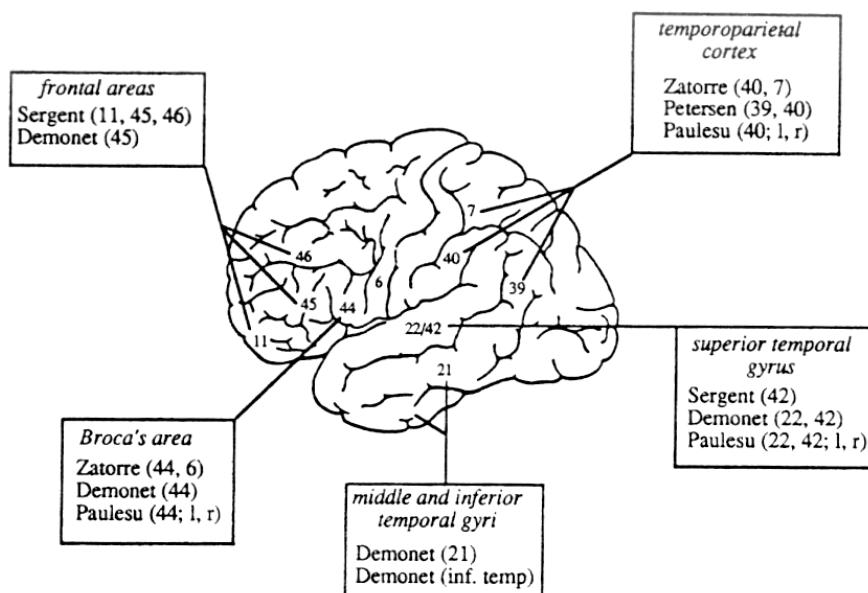
Gli schemi di attivazione ricavati tramite il metodo sottrattivo dipendono *totalmente* dalla scelta del modello funzionale impiegato per analizzare e scomporre il compito effettuato dal soggetto: *diversi* modelli funzionali applicati alle stesse rilevazioni produrranno dati finali (mappe di attivazione sottratte) diversi. Per questa ragione la

scelta delle variabili da manipolare e del modello funzionale soggiacente costituisce il principale nodo teorico della metodologia del *brain imaging*.

Si è ritenuto a lungo che lo studio della fisiologia o dell'anatomia funzionale potesse fornire vincoli indipendenti sull'architettura dei processi cognitivi. Nel caso del *brain imaging* siamo piuttosto di fronte a esperimenti di psicologia cognitiva *applicata* a dati fisiologici. I dati ottenuti tramite strategie sottrattive rappresentano più un test sull'applicabilità dei modelli psicologici come *filtri* per dati di natura fisiologica che non risultati definitivi sulla natura specifica del trattamento e delle specifiche proprietà anatomo-funzionali di determinate popolazioni neurali.

A questa estrema sensibilità dei risultati rispetto al modello adottato sono da ricondurre le palesi divergenze che si riscontrano nella letteratura.

Esemplare a questo proposito è una recente rassegna di POEPEL<sup>24</sup> su una serie di studi di *PET* volti a isolare quali aree corticali siano coinvolte nel trattamento fonologico. Dai risultati complessivi di questi cinque studi – che fanno uso di protocolli sperimentali omogenei e di diverse strategie sottrattive – emergono ben 22 aree non contigue con un'attivazione importante in corrispondenza di operazioni di decodifica fonologica.



[Aree coinvolte nel trattamento fonologico secondo cinque studi, POEPEL (1996)]

Di queste aree, otto mostrano una attivazione significativa in più di uno studio; nessuna di queste 22 aree, per contro, risulta attivata in tutti e cinque gli studi e nessun sottoinsieme di tali aree è condiviso in almeno quattro su cinque studi. Malgrado questa palese divergenza, ognuno di questi studi è riuscito a individuare dati neuropsicologici coerenti con i propri risultati: cinque distinti casi di lesione compatibili con i dati risultanti da ciascuno di questi cinque studi.

La relativa incommensurabilità delle conclusioni di questi studi – che giungono a cinque ipotesi diverse di localizzazione del trattamento fonologico – rappresenta, se non una minaccia, una pesante pregiudiziale sull'affidabilità del metodo. Salvo il caso di differenze macroscopiche fra i soggetti esaminati, le ragioni della divergenza possono soltanto dipendere da caratterizzazioni discordanti di ciò che si intende per 'trattamento fonologico' (ovvero da modelli funzionali che classificano come tale tipi di trattamento non identici) e dagli effetti macroscopici che derivano da tale caratterizzazione sulle strategie sottrattive adottate.

La visualizzazione di determinate mappe di attivazione piuttosto che di altre dipende quindi in modo essenziale dal tipo di modello impiegato per la scomposizione del compito assegnato al soggetto. Lungi dall'essere meri dati fisiologici, ho suggerito che tali mappe di attivazione sono piuttosto da considerare dati 'ibridi' che risultano dall'applicazione a questi ultimi di modelli psicologici in qualità di *filtri*.

Riassumiamo le considerazioni di questo paragrafo.

I metodi sottrattivi sono intesi in origine come strumenti per l'individuazione e la segregazione di sottocomponenti funzionali a partire dai tempi di reazione del soggetto: sotto l'ipotesi dell'esistenza di processi sensibili in modo selettivo alla manipolazione di singole variabili (ipotesi di additività), i metodi sottrattivi sono diventati lo strumento principe della psicologia cognitiva per l'individuazione di sottosistemi funzionali.

Con l'avvento del *brain* imaging, da metodo per la scomposizione funzionale dei fenomeni della cognizione, la sottrazione è divenuta un metodo per lo studio della localizzazione delle basi neurali di questi ultimi.

Questa migrazione metodologica dal campo degli studi dei tempi di reazione al campo delle ricerche sulle basi neurali della cognizione è gravida di conseguenze. L'idea di

sottrazione impiegata nel campo del *brain imaging* ha mantenuto un'identità puramente nominale con la sottrazione impiegata negli studi comportamentali e ha introdotto di conseguenza una nuova serie di assunzioni, legate a un dominio di indagine profondamente diverso da quello originale. Il paragrafo seguente intende valutare le conseguenze sul piano epistemologico di questa tacita migrazione metodologica.



## 6. Vicissitudini della nozione di sottrazione: antiche e nuove assunzioni

L'importazione del metodo sottrattivo nel cuore dell'impianto inferenziale del *brain imaging* – come è stato rilevato da alcuni studiosi<sup>25</sup> – solleva problemi epistemologici sconosciuti alla psicologia sperimentale classica. La differenza sostanziale fra l'oggetto di studio nei due casi – là l'esistenza di tempi di reazione sensibili in modo selettivo a determinate manipolazioni delle variabili, qui l'esistenza di modificazioni locali dell'attività neurale sensibili a tali manipolazioni – richiede un esame dettagliato del significato di tale trasposizione metodologica.

L'assunzione portante dei metodi sottrattivi classici, come ho ricordato nel paragrafo precedente, è l'ipotesi di additività, vale a dire la possibilità di isolare singole variabili che non influiscano sulla totalità dei processi in gioco, ma esclusivamente su specifiche operazioni.

Che cosa significa ora “isolare” queste operazioni a partire dalle mappe di attivazione neurale ottenute tramite *brain imaging*, anziché dai tempi di reazione di un soggetto?

Possiamo affermare che “Isolare” tali operazioni corrisponde a individuare strutture cerebrali che godano di *tre precise proprietà*. Ognuna di queste proprietà rappresenta un vincolo necessario affinché le strategie sottrattive si possano applicare efficacemente ai dati ottenuti dal *brain imaging*.

- La prima di queste proprietà è quella che chiamerò *irrelevanza del contesto*. Tale proprietà richiede che ogni variabile considerata manifesti il suo effetto in modo identico *indipendentemente* dal contesto dell'attività in cui viene modificata. Se la stessa operazione (lo stesso tipo di trattamento) è coinvolta da due processi differenti, si deve cioè assumere che la manipolazione delle variabili che la riguardano sortisca un identico effetto in entrambi i processi. Prendiamo un esempio concreto. Supponiamo di voler effettuare una scansione dell'attività cerebrale di un soggetto nelle seguenti condizioni: la lettura ‘mentale’ di una parola (1) e la lettura della stessa parola ‘ad alta voce’ (2). Ciò che ci aspettiamo di ottenere dalla sottrazione di

tali condizioni è il pattern di attività cerebrale correlata *esclusivamente* all'articolazione motoria della parola. Ci aspettiamo, in altre parole, che il resto dell'attività cerebrale richiesta dalle due operazioni rimanga pressoché inalterato: se cioè processi secondari sono coinvolti dalle due condizioni sperimentali, ci attendiamo che il loro effetto si annulli a vicenda grazie alla sottrazione. Questa aspettativa dipende da un'ipotesi di irrilevanza del contesto nella misura in cui esclude che processi non coinvolti dalla lettura possano avere nella condizione (1) e nella condizione (2) effetti *diversi*, tali cioè da non annullarsi vicendevolmente. Se un'area è specializzata – per esempio – nella discriminazione del timbro e abbiamo ragioni sufficienti per ritenere che operazioni di discriminazione del timbro non siano coinvolte né dalla lettura 'mentale' di una parola né dalla sua lettura ad alta voce, ci aspetteremo che l'attività di quest'area si annulli tramite la sottrazione di (1) da (2), in quanto la sua attività sarà statisticamente *la stessa* in entrambe le condizioni.

Alla proprietà di irrilevanza del contesto corrisponde all'incirca la proprietà che Fodor ha chiamato 'specificità di dominio'<sup>26</sup>. Determinate strutture cognitive sarebbero state selezionate per rispondere in modo efficace ed automatico a tipi *specifici* di problemi. La specificità del loro dominio di applicazione sarebbe appunto una delle condizioni richieste per il loro funzionamento rapido e quasi automatico. Fodor sostiene che tali sistemi ('moduli') ci siano stati trasmessi filogeneticamente e siano pertanto iscritti nell'anatomia funzionale del cervello: strutture cerebrali specializzate per un determinato tipo di trattamento non possono essere quindi coinvolte da altri tipi di trattamento.

Quale sia l'estensione di sistemi a specificità di dominio nel cervello umano è tuttora oggetto di dibattito. Da un lato, abbiamo sostenitori della cosiddetta 'modularità massiva' (*massive modularity*), secondo i quali la maggior parte delle nostre capacità cognitive, comprese quelle più complesse, non è mai costituita da sistemi astratti, flessibili e adatti ai tipi più diversi di contenuto, ma al contrario è rappresentata da un insieme di sistemi ad alta specificità di dominio<sup>27</sup>. Dall'altro lato abbiamo studiosi, Fodor *in primis*<sup>28</sup>, che sostengono che la specificità di dominio appartiene esclusivamente a sistemi periferici, mentre le funzioni 'alte' della cognizione sarebbero caratterizzate da proprietà di generalità e astrazione.

Nel caso del nostro principio di irrilevanza del contesto, non è necessario adottare posizioni estreme come quelle fodoriane per concepire l'esistenza di sistemi che violino tale principio. Basti pensare alla possibilità di sistemi che godano di una proprietà di interdipendenza funzionale, sistemi che cioè, pur avendo una specializzazione funzionale determinata, possano essere reclutati da altri sistemi per svolgere tipi di trattamento che si discostano da quelli a cui sono naturalmente 'dedicati'<sup>29</sup>. Nell'esempio di cui sopra, l'ipotetica area dedicata alla decodifica del timbro, pur senza prestarsi indiscriminatamente a qualsiasi tipo di trattamento, potrebbe essere coinvolta dalla lettura 'ad alta voce' per svolgere determinate routine assenti nella lettura 'mentale'. Il risultato della sottrazione in questo caso evidenzerebbe un'attività rilevante in un'area 'specializzata nella discriminazione del timbro' senza che per questo alcun processo di discriminazione del timbro sia coinvolto.

Ogni ipotesi di specializzazione funzionale di determinate aree dovrà tener conto di questa possibilità di non-additività la quale, se non esplicitata, può condurre a conclusioni devianti sull'interpretazione dei pattern di attivazione neurale rilevati. È verosimile che la maggior parte delle divergenze nei risultati di studi di *brain imaging* volti ad isolare le basi neurali dello stesso tipo di operazioni possa essere ricondotta a situazioni di questo genere, situazioni in cui cioè variabili ritenute *irrilevanti* non lo siano nella stessa misura per i vari processi considerati<sup>30</sup>.

- La seconda proprietà richiesta a garanzia della validità dei metodi sottrattivi applicati al *brain imaging* è quella dell'*additività* propriamente detta. Si tratta a prima vista della stessa proprietà descritta nel caso degli studi di cronometria mentale, ma la sua natura cambia radicalmente nel momento stesso in cui la si applica allo studio dei pattern di attivazione cerebrale anziché di semplici tempi di reazione. Tale proprietà richiede che la manipolazione di determinate variabili, utili alla caratterizzazione di un dato sottosistema, non produca alcun effetto rilevante sull'insieme degli altri sottosistemi: nel caso in cui così non fosse, la manipolazione di una variabile finirebbe per riverberare in tutta l'architettura del sistema<sup>31</sup>.

Tale proprietà è legata al principio descritto da Fodor come 'impenetrabilità' o 'incapsulatezza informazionale' (*informational encapsulation*), principio secondo il

quale l'attività di alcuni sottosistemi sarebbe impenetrabile da parte dell'informazione disponibile alla totalità del sistema. Tale assunzione si giustifica in Fodor attraverso l'argomento che determinati sistemi per poter funzionare in modo rapido ed efficace non devono essere suscettibili di modulazioni da parte dell'informazione disponibile ad altri sottosistemi.

Ora, se nel caso della cronologia mentale esiste un criterio esplicito per valutare l'ipotetica additività di determinate variabili (ad esempio, la linearità dell'aumento dei tempi di reazione rispetto a modifiche lineari di tali variabili), nel caso del *brain imaging* nessun criterio permette di decidere se la manipolazione di una variabile sia effettivamente 'ortogonale' e soprattutto se a livello neurale essa non risulti in una modulazione nell'attività di altri sottosistemi.

Il problema è tanto più rilevante quanto maggiore è la discrepanza tra il modello funzionale ipotizzato per la scomposizione funzionale di un dato processo e la natura effettiva del trattamento neurale che si ritiene ad esso correlato : una minima modifica dal punto di vista del modello funzionale tra la condizione di controllo e la condizione di interesse può cioè tradursi in una modifica macroscopica nella distribuzione dei pattern di attivazione, senza che ciò sia dovuto all'intervento di particolari 'operazioni' o 'sottosistemi specializzati', ma semplicemente al *riverbero* di tale manipolazione in un sistema altamente interconnesso come quello del cervello.

Alcuni studiosi hanno suggerito a questo proposito un rischio di *circolarità* nell'applicazione del metodo sottrattivo agli studi di *brain imaging*<sup>32</sup>. La sottrazione isolerà *né più né meno* variazioni dell'attività cerebrale locale - sia che questa dipenda realmente dall'intervento selettivo di specifici sottosistemi specializzati sia che essa sia il risultato di semplici effetti di riverbero. Tale riverbero può sussistere tanto a livello psicologico - nel caso di non additività fra le variabili in gioco nel protocollo sperimentale - quanto a livello neurofisiologico - nel caso di interazioni fra popolazioni neurali contigue determinate, ad esempio, da fenomeni di *feedback*<sup>33</sup>). Non esiste, in conclusione, alcun modo di interpretare l'attività risultante dai metodi sottrattivi se non in termini di quanto previsto dal modello psicologico utilizzato.

- La terza proprietà richiesta perché i metodi sottrattivi permettano di isolare l'attività di popolazioni neurali responsabili di specifiche forme di trattamento è la *connessione topologica e la compattezza strutturale* di tali popolazioni: solo profili di attivazione concentrati in aree compatte e connesse (*clustered*) della corteccia cerebrale sono infatti in grado di emergere globalmente dalle immagini sottrattive. Attivazioni più diffuse o distribuite in circuiti non contigui rischiano di diventare indiscernibili una volta sottoposti a sottrazione, soprattutto in presenza di un debole rapporto segnale-rumore.

Abbiamo ricordato in precedenza che gli studi microstrutturali sulla connettività dei circuiti neurali hanno permesso di isolare un ampio numero di aree compatte e fittamente interconnesse al loro interno<sup>34</sup>. Che a tali aree debba tuttavia spettare una specializzazione funzionale *in virtù* di tale proprietà di compattezza strutturale è un'ipotesi da sottoporre ad esame empirico.

È interessante nondimeno rilevare come l'idea di 'area', intesa come popolazione neurale compatta ed isolabile sotto il profilo anatomico, sia tendenzialmente privilegiata già a partire dal metodo impiegato per la selezione dei dati negli studi di *brain imaging*. Se parte fondamentale della ricerca del *brain imaging* è volta all'individuazione di *aree specializzate nella realizzazione di particolari tipi di trattamento*, è naturale sospettare che la localizzazione di popolazioni compatte e topologicamente connesse giochi un ruolo esplicativo importante all'interno di tali studi.

Dietro all'ipotesi che gran parte dei sottosistemi dedicati a particolari tipi di trattamento presenti delle basi neurali non soltanto *specializzate* ma anche *anatomicamente isolabili* sembra celarsi il retaggio di una visione aprioristica dei rapporti tra mente e cervello, una visione che ha attraversato l'intera storia della neuropsicologia fin dalle sue origini e che periodicamente riappare negli studi di anatomia funzionale del cervello.

MARSHALL ha battezzato il ritorno di questa concezione col nome di "n u o v a o r g a n o l o g i a" (*new organology*<sup>35</sup>) o "nuova frenologia": la potremmo definire come la fede *a priori* nel fatto che ad un'area definita ed anatomicamente isolabile della corteccia cerebrale debba corrispondere un centro specializzato per lo svolgimento di una determinata forma di trattamento. O ancora come l'idea che, ad un livello di analisi

dato, a ciascuno dei sottosistemi descritti dalla psicologia cognitiva dovrà corrispondere una popolazione neurale dedicata o che, viceversa, una descrizione esaustiva dell'anatomia funzionale del cervello dovrà esplicitare che tipo di trattamento corrisponde a ciascun area corticale.

Tale assunto sembra essere un tacito presupposto di larga parte degli studi contemporanei di *brain imaging* (per lo meno di un numero considerevole di studi che si riconoscono nella prima (A) delle due classi di strategie di indagine descritte nel secondo paragrafo del presente articolo, quella relativa alla *localizzazione corticale* di funzioni postulate a livello psicologico).

Nel loro complesso, queste tre assunzioni impongono vincoli estremamente forti sulla natura dei fenomeni studiati e dei processi neurali che li realizzano. Rispetto alle assunzioni che governano l'uso di metodi sottrattivi nell'ambito della psicologia tradizionale, l'uso degli 'stessi' metodi nel campo degli studi di *brain imaging* comporta ben più che una semplice migrazione metodologica. Si tratta piuttosto di una radicale ridefinizione della nozione di 'isolabilità' di un sottosistema funzionale che non solo è raramente tematizzata nella letteratura epistemologica dedicata, ma che spesso viene trascurata, con gravi conseguenze sul piano esplicativo, da larga parte della letteratura sperimentale.

## 7. Localizzare e scomporre: quale ruolo per le nuove metodologie?

Lichtheim - uno dei grandi padri della moderna neuropsicologia - fu aspramente attaccato al principio del XX secolo per l'ingenuità della sua teoria dei "centri corticali" del linguaggio. Tale teoria riconduceva le varie famiglie di deficit linguistici ricorrenti in pazienti afasici a lesioni di singole aree cerebrali, considerate come centri deputati a specifiche funzioni, o all'interruzione delle connessioni corticali fra tali aree<sup>36</sup>. Si può affermare senza esagerazione che lo sforzo principale della neuropsicologia cognitiva, a partire dagli anni '50, sia stato proprio quello di sbarazzarsi da ipotesi ingenuie di correlazione fra area cerebrale e funzione – sforzo che si manifesta in un affinamento degli strumenti di analisi funzionale dei processi cognitivi e soprattutto una sempre minore fiducia nell'identificazione diretta fra *locus* cerebrale della lesione e deficit funzionale<sup>37</sup>.

Ironia della sorte ha voluto che un energico ritorno a ipotesi 'frenologiche' dell'organizzazione del cervello sia stato suscitato dall'uso delle più recenti tecnologie di visualizzazione dell'attività cerebrale. La possibilità di raffrontare i modelli della psicologia classica con gli schemi di attività cerebrale registrata in un soggetto ha portato a una proliferazione di studi volti all'individuazione di strutture neurali sistematicamente coinvolte in precisi tipi di trattamento dell'informazione.

Dalle osservazioni fatte nel corso di questo articolo è possibile trarre un bilancio provvisorio del ruolo che i risultati di questi studi vanno assumendo nell'impianto esplicativo delle scienze cognitive.

Le critiche emerse dai paragrafi precedenti hanno cercato di mettere in luce i rischi che derivano dal considerare i metodi di *brain imaging* come *lo* strumento definitivo per stabilire che tipo di specializzazione funzionale competa ad una determinata popolazione neurale e, viceversa, a quali popolazioni neurali siano da ricondurre determinati tipi di trattamento.

Contro l'idea che il *brain imaging* possa realizzare l'auspicato connubio fra architetture funzionali dei processi cognitivi e anatomia funzionale del cervello giocano due importanti fattori.

Da un lato, il fatto che le mappe di attivazione ottenute tramite metodi sottrattivi risultano estremamente sensibili alla scelta del modello funzionale adottato per descrivere l'insieme dei processi coinvolti – al punto che l'individuazione delle basi neurali di un dato tipo di trattamento può divergere notevolmente in funzione di minime differenze nelle strategie scompositive. Tale radicale dipendenza dal modello - e il pesante squilibrio che ne consegue fra teoria e osservazione - ha portato taluni ad affermare che il *brain imaging* funzionale, più che una metodologia delle neuroscienze cognitive, debba intendersi come psicologia cognitiva *sotto mentite spoglie*.

D'altro canto, le assunzioni legate alle strategie di rilevamento e selezione dei dati (strategie che ho denominato nel loro insieme 'metodi sottrattivi') impongono alla struttura dei fenomeni studiati dei vincoli estremamente costosi: un'organizzazione anatomo-funzionale del cervello in termini di "aree di elaborazione" compatte, altamente specializzate e funzionalmente indipendenti, "organi" anatomicamente isolabili la cui attività resterebbe sostanzialmente immutata rispetto ai possibili contesti funzionali in cui sono coinvolti e tali da non mostrare modulazioni rilevanti con altre aree nel trattamento di determinati tipi di informazione.

Senza nulla togliere al ruolo euristico del *brain imaging*, queste restrizioni limitano l'aspirazione che vuole fare di tali metodologie – allo stato attuale del loro sviluppo e della riflessione epistemologica sul loro impiego – la chiave per la costruzione di 'modelli neurali della cognizione': il progresso nella conoscenza delle basi neurali dei fenomeni cognitivi dovrà continuare ad avvalersi di dati neuropsicologici, psicologici e soprattutto di descrizioni microfunzionali (come quelle sulla neurofisiologia del sistema visivo in specie diverse dall'uomo) che esplicitino la natura del trattamento effettuato da singole popolazioni neurali<sup>38</sup>.

Possiamo sintetizzare in una formula la conclusione generale della presente analisi: i metodi di *brain imaging* hanno comportato una modifica profonda della nozione di 'isolabilità' nell'ambito dello studio della cognizione.

L'idea di 'isolabilità' di sottosistemi funzionali – un tempo prerogativa della sola psicologia cognitiva – è venuta progressivamente a indebolirsi per accogliere fra i suoi criteri anche proprietà quali la 'localizzabilità corticale', criteri che cioè esulano dal



campo di indagine strettamente comportamentale. La ‘migrazione’ di strategie sottrattive dalla cronometria mentale agli studi di *brain imaging* è da considerare verosimilmente come la causa principale di tale indebolimento. È lecito in ultima analisi chiedersi che ‘valore aggiunto’ rappresenti la localizzabilità di sottosistemi funzionali rispetto alla loro semplice caratterizzazione in termini di trattamento dell’informazione. L’esistenza di pattern di attività neurale altamente selettiva e localizzabile nella corteccia cerebrale sta diventando un criterio riconosciuto in modo sempre più unanime per la ridefinizione dell’architettura funzionale dei fenomeni della cognizione. L’impiego di ipotesi di localizzazione in asserzioni quali “il processo osservato coinvolge ampiamente aree deputate alla decodifica fonologica” o, “nessun’area correlata a processi di categorizzazione visiva sembra essere coinvolta nel compito effettuato dal soggetto” o, più in generale, “l’attivazione dell’area X suggerisce la partecipazione di un trattamento di tipo Y al processo studiato” tradisce – in mancanza di precisi paletti epistemologici – una fiducia crescente in ipotesi ‘frenologiche’ di stretta correlazione fra struttura e funzione.

I metodi di *brain imaging* portano con sé il rischio di conclusioni ingenuie di questo genere - conclusioni che, qualora svincolate da un modello dettagliato del tipo di trattamento svolto da determinate popolazioni neurali, non forniscono alcun contributo stringente per un’ulteriore comprensione dei processi studiati<sup>39</sup>.

Un primo passo verso una scienza più rigorosa dei rapporti tra struttura e funzione richiederà uno sforzo teorico che permetta di tradurre tali asserzioni di ‘coinvolgimento di un’area cerebrale’ in ipotesi empiricamente testabili e in modelli computazionali espliciti sul ruolo che particolari circuiti neurali rivestono nell’architettura generale di un processo.

---

<sup>1</sup> BECHTEL & RICHARDSON (1993).

<sup>2</sup> SHALLICE (1984).

<sup>3</sup> GRABOWSKI & DAMASIO (1996), BECHTEL & MUNDALE (1996).

<sup>4</sup> MASSARO & COWAN (1993).

<sup>5</sup> La storia dei metodi non invasivi di visualizzazione dell'attività cerebrale è in realtà ben più antica di quella delle tecniche di *brain imaging*. Basti pensare all'uso fatto per diversi decenni di tecniche quali l'EEG ed i Potenziali Evocati (PE). Se il *brain imaging* ha fatto parlare di *nascita* di una nuova metodologia, ciò è dovuto principalmente al fatto che tali tecniche, coniugate ad efficaci strategie di selezione dei dati, hanno permesso per la prima volta di individuare direttamente l'attività di specifiche strutture cerebrali, laddove i Potenziali Evocati, misurando l'attività elettrica a livello dello scalpo, erano in grado di fornire informazioni estremamente imprecise sulle strutture corticali responsabili dell'attività registrata. Cfr. GEVINS (1996). Per ulteriori dettagli sulla natura della misura nei metodi di *brain imaging*, si veda il paragrafo 3 del presente articolo.

<sup>6</sup> POSNER & ROTHBART (1994).

<sup>7</sup> POSNER, PETERSEN, FOX & RAICHLE (1988).

<sup>8</sup> KOSSLYN (1999).

<sup>9</sup> I criteri tradizionalmente utilizzati in neurofisiologia per isolare aree cerebrali distinte sono principalmente: 1. la *connettività*, ovvero l'esistenza di popolazioni neurali compatte e fittamente interconnesse al loro interno, distinte da altre popolazioni cui sono collegate tramite 'connessioni lunghe', 2. la *citoarchitettura*, ovvero l'esistenza di caratteristiche omogenee nella morfologia delle cellule, 3. la *selettività*, ovvero la risposta omogenea di un insieme di cellule a specifiche proprietà dello stimolo, 4. l'*organizzazione topografica*, ovvero la corrispondenza tra l'organizzazione spaziale interna di una popolazione neurale e l'organizzazione spaziale dei recettori negli organi sensoriali (ad esempio, nel caso della visione, l'organizzazione spaziale della retina). Presi nel loro insieme, questi quattro criteri hanno permesso di definire singole aree corticali (come V1, V2 o MT nel caso del sistema visivo). La molteplicità di tali criteri e i loro possibili conflitti sono altresì alla base delle principali divergenze nella comunità neuroscientifica sull'esistenza e le frontiere di specifiche aree corticali.

<sup>10</sup> PRICE & FRISTON (1997), TONONI et al. (1998), KOSSLYN (1999). Il seguito di questo articolo sarà dedicato in modo particolare, anche se non esclusivamente, a studi che si rifanno allo schema (A).

<sup>11</sup> Una correlazione precisa fra immagini funzionali del metabolismo cerebrale indotta da particolari compiti e anatomia è possibile solo utilizzando contestualmente PET e immagini provenienti da risonanza magnetica applicata allo stesso soggetto. Nel caso dell'*fMRI*, la possibilità di sovrapporre immagini funzionali e immagini anatomiche del cervello del soggetto attraverso la stessa tecnologia permette di ricostruire con maggior precisione le coordinate spaziali delle aree attivate. Sul problema della risoluzione spaziale e del confronto interindividuale, si confrontino MAZZIOTTA (1996) e WOODS (1996), pp. 313ss.

<sup>12</sup> Si distingue a questo proposito una *state related PET*, da un *event related fMRI*. Mentre l'*fMRI* "rapida" è in grado di misurare autonomamente variazioni che occorrono in tempi dell'ordine del centinaio di millisecondi, la *PET* può fornire informazioni attendibili sull'attivazione "rapida" di aree cerebrali soltanto attraverso protocolli estremamente accurati, strategie sottrattive per l'eliminazione del rumore (sulle quali torneremo oltre) e l'integrazione dei risultati con misurazioni provenienti da tecniche più accurate sotto il profilo temporale,

---

benché spazialmente molto più imprecise, come ad esempio quella dei potenziali evocati (*EP*) – cfr. CHERRY & PHELPS (1996).

<sup>13</sup> Nel caso dell’fMRI ci limiteremo a valutare l’affidabilità di un indice come il segnale BOLD (*blood-oxygenation level dependent signal* - SERGENT (1994)) e non indici derivanti da protocolli più avanzati come il MION (basato sul volume del flusso sanguigno anziché sul tasso di ossigenazione).

<sup>14</sup> Non vanno trascurati tuttavia problemi relativi alle cause specifiche delle variazioni vascolari rispetto all’attività neurale che le genera, come ad esempio il paradosso ai fini della misura che sussiste nel rapporto fra incremento metabolico e incremento del flusso sanguigno. Cfr. WANDELL (1999), p.147.

<sup>15</sup> KOSSLYN (1999), 1291; WANDELL (1999).

<sup>16</sup> SIDTIS et al. (1998), 495.

<sup>17</sup> WANDELL (1999).

<sup>18</sup> GRABOWSKI & DAMASIO (1996).

<sup>19</sup> POSNER & ABDULLAEV (1996); GRABOWSKI & DAMASIO (1996).

<sup>20</sup> Cfr. STERNBERG (1969): all’eredità di Donders e alla cronometria dei processi mentali è dedicato l’intero volume degli *Acta Psychologica* del 1969.

<sup>21</sup> POSNER (1978), POSNER & ROTHBART (1994).

<sup>22</sup> POSNER & ABDULLAEV (1996).

<sup>23</sup> POSNER & RAICHLE (1994).

<sup>24</sup> POEPEL (1996).

<sup>25</sup> SERGENT (1994), FRISTON et al. (1996).

<sup>26</sup> Questo e i rimandi successivi a Fodor, ove non specificato ulteriormente sono da considerarsi rimandi a FODOR (1983).

<sup>27</sup> Si vedano i saggi raccolti in: HIRSCHFELD & GELDMAN (1994) e il *working paper* di SPERBER (2000).

<sup>28</sup> FODOR (2000).

<sup>29</sup> Si veda l’interessante distinzione proposta da SWINDALE (1998) tra ‘modules’ e ‘polymaps’.

<sup>30</sup> Per riprendere i casi studiati da POEPEL (1996), possiamo concepire in astratto un’attività cerebrale corrispondente alle basi neurali del “trattamento fonologico”, ma che cosa viene definito propriamente “trattamento fonologico”? Il termine può designare un insieme complesso e non necessariamente omogeneo di routine il cui ruolo cambia sensibilmente a seconda delle operazioni richieste al soggetto.

<sup>31</sup> SIDTIS et al. (1998).

<sup>32</sup> VAN ORDEN & PAAP (1997).

<sup>33</sup> FRISTON et al. (1997).

<sup>34</sup> Si veda la nota n.9. Nel caso del sistema visivo, il riferimento essenziale sull’organizzazione gerarchica delle aree corticali desunta dallo studio della connettività rimane FELLEMAN & VAN ESSEN (1991).

<sup>35</sup> MARSHALL (1980).

<sup>36</sup> LICHTHEIM (1885) cit. in SHALLICE (1988).

<sup>37</sup> CARAMAZZA (1986), p. 43. SHALLICE (1988) parla a questo proposito di svolta ‘ultracognitiva’ della neuropsicologia.

<sup>38</sup> STUFFLEBEAM & BECHTEL (1997).

<sup>39</sup> WANDELL (1999).

## Riferimenti Bibliografici

- BECHTEL, W. & RICHARDSON, R.C. (1993) – *Discovering Complexity*, Princeton (NJ), Princeton University Press, 1993.
- BECHTEL, W. & MUNDALE, J. (1999) – Multiple Realizability Revisited, *Philosophy of Science*, 66 (1999), 175-207.
- CARAMAZZA, A. (1986) – On Drawing Inferences about the Structure of Normal Cognitive Systems from the Analysis of Patterns of Impaired Performance: The Case for Single-Patient Studies, *Brain and Cognition*, 5 (1986), 41-66.
- CHERRY, S.R. & PHELPS, M.E. (1996) – Imaging Brain Function with Positron Emission Tomography, in: TOGA & MAZZIOTTA (1996).
- FODOR, J.A. (1983) – *The Modularity of Mind*, Cambridge, MA: MIT Press (1983).
- FODOR, J.A. (2001) – *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*, Cambridge, MA: MIT Press, July 2000.
- FELLEMAN, D.J. & VAN ESSEN, D.C. (1991) – Distributed Hierarchical Processing in the Primate Cerebral Cortex, *Cerebral Cortex* 1 (1991) 1-148.
- FRISTON, K.J. et al. (1996) – The Trouble with Cognitive Subtraction, *NeuroImage*, 4:2 (oct 1996), 97-104.
- FRISTON, K.J. et al. (1997) – Psychophysiological and Modulatory Interactions in Neuroimaging, *NeuroImage*, 6:3 (oct 1997), 218-229.
- GEVINS, A. (1996) – Electrophysiological Imaging of Brain Function, in: TOGA & MAZZIOTTA (1996).
- GRABOWSKI, T.J. & DAMASIO, A.R. (1996) – Improving functional imaging techniques: the dream of a single image for a single mental event, *Proceedings of the National Academy of Sciences of the United States of America*, 93 (1996), 14302-14303.
- HIRSCHFELD & S. A. GELMAN (eds) (1994). *Mapping the Mind: Domain specificity in cognition and culture*, New York: Cambridge University Press.
- LICHTHEIM, L. (1885) – On Aphasia, in: *Brain*, 7 (1885), 433-484.
- KOSSLYN, S.M. (1994) – *Image and brain: the resolution of the imagery debate*, Cambridge (MA), MIT Press, 1994.
- KOSSLYN, S.M. (1999) – If Neuroimaging is the Answer, What is the Question?, *Philosophical Transactions, Biological Sciences*, 344 (jul 1999).
- MARSHALL, J.C. (1980) – The New Organology, *Behavioral and Brain Sciences*, 3 (1980), 23-25.
- MARSHALL, J.C. (1984) – Multiple Perspectives on Modularity, *Cognition*, 17 (1984), 209-242.

- MASSARO, D.W. & COWAN, N. (1993) – Information Processing Models: Microscopes of the Mind, in: *Annual Review of Psychology*, 44 (1993), 383-425.
- MAZZIOTTA, J.C. (1996) – Mapping Strategies. Time and Space, in: TOGA & MAZZIOTTA (1996).
- POEPPPEL, D. (1996) – A Critical review of PET studies of Phonological Processing, *Brain and Language*, 55:3 (dec 1996), 317-351.
- POLK, T.A. & FARAH, M.J. (1998) – The neural development and organization of letter recognition: Evidence from functional neuroimaging, computational modeling and behavioral studies, *Proc. Natl. Acad. Sci. USA*, 95 (feb 1998), 847-852.
- POSNER, M.I. (1978) – *Chronometric Explorations of Mind*, Hillsdale (NJ), Erlbaum, 1978.
- POSNER, M.I. & ABDULLAEV, Y.G. (1996) – Mapping Strategies. What to Image ? Anatomy, Circuitry and Plasticity of Human Brain, in: TOGA & MAZZIOTTA (1996).
- POSNER, M.I., PETERSEN, S.E., FOX, P.T. & RAICHLER, M.E. (1988) – Localization of cognitive operations in the human brain, *Science*, 240 (1988), 1627-1631.
- POSNER, M.I. & RAICHLER, M.E. (1994) – *Images of Mind*, New York, Scientific American Library, 1994.
- POSNER, M.I. & ROTHBART, M.K. (1994) – Constructing Neuronal Theories of Mind, in: C. KOCH, J.L.DAVIS (eds.) – *Large-Scale Neuronal Theories of Brain*, Cambridge (MA), MIT Press, 1994.
- PRICE, C.J. & FRISTON, K.J. (1997) – Cognitive Conjunction: A New Approach to Brain Activation Experiments, *NeuroImage*, 5:4 (may 1997), 261-270.
- RAICHLER, M.E. (1998) – Behind the scenes of functional brain imaging: A historical and physiological perspective, *Proc. Natl. Acad. Sci. USA*, 95 (feb 1998), 765-772.
- SERGENT, J. (1994) – Brain-Imaging Studies of Cognitive Functions, *Trends in Neuroscience*, 17.6 (1994), 221-227.
- SHALLICE, T. (1984) – More Functionally Isolable Subsystems but Fewer “Modules”, *Cognition*, 17 (1984), 243-252.
- SHALLICE, T. (1988) – *From Neuropsychology to Mental Structure*, Cambridge (UK), Cambridge University Press, 1988.
- SIDTIS, J.J., STROTHER, S.C., ANDERSON, J.R. & ROTTENBERG, D.A. (1999) Are Brain Functions Really Additive?, *NeuroImage*, 9:5 (may 1999), 490-496.
- SPERBER, D. (2000) *In defense of massive modularity* – working paper, <http://www.dan.sperber.com/modularity.htm>
- STERNBERG, S. (1969) – The discovery of processing stages: extensions of Donders methods, *Acta Psychologica*, 230 (1969), 276-315.
- STUFFLEBEAM, R.S. & BECHTEL, W. (1997) – PET: Exploring the Myth and the Method, *Philosophy of Science*, 64 (1997 – Suppl.), S95-S106.

- SWINDALE, N. (1998) – Modules, polymaps and mosaics, *Current Biology*, 8 (1998), R270-R273.
- TOGA, A.W. & MAZZIOTTA, J.C. (1996) (eds.) – *Brain mapping. The Methods*, San Diego, Academic Press, 1996.
- TONONI, G., MCINTOSH, A.R., RUSSELL, D.P. & EDELMAN, G.M. (1998) – Functional Clustering: Identifying Strongly Interactive Brain Regions in Neuroimaging Data, *NeuroImage*, 7:2 (feb 1998), 133-149.
- VAN ORDEN, G.C. & PAAP, K.R. (1997) – Functional Neuroimages Fail to Discover Pieces of Mind in the Parts of the Brain, *Philosophy of Science*, 64 (1997 – Suppl.), S85-S89.
- WANDELL, B.A.. (1999)– Computational Neuroimaging of Human Visual Cortex, *Annual Reviews in Neuroscience*, 22 (1999), 145-173.
- WOODS, R.P. (1996) – Correlation of Brain Structure and Function, in: TOGA & MAZZIOTTA (1996).