



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

## DOTTORATO DI RICERCA IN

PHILOSOPHY, SCIENCE, COGNITION, AND SEMIOTICS

Ciclo 37

**Settore Concorsuale:** 11/E1 PSICOLOGIA GENERALE, PSICOBIOLOGIA E  
PSICOMETRIA

**Settore Scientifico Disciplinare:** M-PSI/01 PSICOLOGIA GENERALE

ATTENZIONE SPAZIALE Uditiva NELLO SPAZIO  
CIRCOSTANTE, EFFETTI DEGLI EMICAMPI ANTERIORE E  
POSTERIORE E DELLA PRATICA MUSICALE

**Presentata da:** Giuseppe Fuca

**Coordinatore Dottorato**

Claudio Paolucci

**Supervisore**

Elena Gherri

**Co-Supervisore**

Renata Galatolo

Esame finale anno 2025



<b>PREFAZIONE .....</b>	<b>5</b>
<b>I - DAL SUONO ALL'ATTENZIONE.....</b>	<b>8</b>
1.1. LA SCENA Uditiva E LA CODIFICA DEI SUONI .....	8
1.2. L'ATTENZIONE DURANTE L'ANALISI DI UNA SCENA Uditiva .....	17
1.3. I PRIMI MODELLI TEORICI DELL'ATTENZIONE: IL PARADIGMA DELL'ASCOLTO DICOTICO .....	21
1.4. LA TEORIA DEL CARICO PERCETTIVO E LA SUA APPLICAZIONE ALLA MODALITÀ UDITIVA .....	28
<b>II - L'ATTENZIONE SPAZIALE Uditiva.....</b>	<b>37</b>
2.1. IL PARADIGMA DI SUGGERIMENTO SPAZIALE .....	40
2.2. APPLICAZIONI E SVILUPPI NELLA MODALITÀ Uditiva.....	47
2.3. APPLICAZIONI IN STUDI MULTIMODALI .....	58
2.4. STUDI CON INSERIMENTO DI STIMOLI DISTRATTORI.....	63
<b>III - TRAINING MUSICALE E IMPATTO SULLE FUNZIONI COGNITIVE .....</b>	<b>71</b>
3.1. SULL'ATTENZIONE .....	85
3.1.1. L'attenzione durante l'ascolto e l'esecuzione musicale .....	86
3.1.2. Rassegna delle evidenze disponibili.....	101
3.2. SU ALTRE FUNZIONI COGNITIVE .....	118
3.2.1. Funzioni esecutive.....	119
3.2.2. Linguaggio, intelligenza, emozioni e competenze sociali.....	129
<b>IV - STUDIO 1: ATTENZIONE Uditiva NELLO SPAZIO CIRCOSTANTE, DIFFERENZE TRA GLI EFFETTI DEGLI EMICAMPI ANTERIORE E POSTERIORE .....</b>	<b>134</b>
4.1. INTRODUZIONE .....	134
4.2. DOMANDE E IPOTESI DI RICERCA.....	138
4.3. ESPERIMENTO 1: ATTENZIONE TRANSIENTE.....	140
4.3.1. Materiali e metodologia .....	141
4.3.2. Analisi dati .....	149
4.3.3. Risultati .....	151
4.3.4. Discussione .....	156
4.4. ESPERIMENTO 2: ATTENZIONE SOSTENUTA.....	162
4.4.1. Materiali e metodologia .....	164
4.4.2. Analisi dati .....	166
4.4.3. Risultati .....	167
4.4.4. Discussione .....	170
4.5. CONFRONTO TRA I DUE COMPITI E DISCUSSIONE GENERALE.....	172
4.6. CONCLUSIONI .....	177
4.7. LIMITAZIONI E FUTURE DIREZIONI DI RICERCA .....	179
<b>V - STUDIO 2: TRAINING MUSICALE E ATTENZIONE SPAZIALE UDITIVA NELLO SPAZIO CIRCOSTANTE.....</b>	<b>183</b>

5.1. INTRODUZIONE .....	183
5.2. DOMANDE E IPOTESI DI RICERCA.....	190
5.3. MATERIALI E METODOLOGIA .....	192
5.4. ANALISI DATI.....	196
5.5. RISULTATI .....	198
5.6. DISCUSSIONE GENERALE .....	206
5.7. CONCLUSIONI .....	212
5.8. LIMITAZIONI E FUTURE DIREZIONI DI RICERCA .....	213
<b>VI - RIFERIMENTI .....</b>	<b>215</b>

## Prefazione

Questo lavoro, come forse risulterà evidente fin dalla lettura del titolo, è caratterizzato fortemente da numerose e diverse *intersezioni*. La prima e originaria intersezione è da ricercare nel background formativo e professionale di chi scrive. Infatti, la varietà degli studi compiuti e degli interessi che li hanno ispirati e accompagnati fino ad oggi, ha generato e probabilmente continuerà sempre a generare nuove possibili intersezioni. Ambiti forse apparentemente così distanti come l'ingegneria meccanica, la musica, la chitarra classica, l'acustica, la pedagogia e l'insegnamento, mi hanno costantemente stimolato a trovare possibili collegamenti e fattori comuni. Il pensiero di fondo era che, a prescindere dalla professione che avrei poi svolto, tutto mi potesse tornare utile al momento opportuno. In ultimo, l'approdo al corso di dottorato di ricerca con un curriculum in scienze cognitive, che con la redazione di questo lavoro volge al suo termine, esemplifica quantomai efficacemente il carattere di *intersezione* che in qualche modo è stata la costante del mio percorso fino ad oggi. Le stesse scienze cognitive si possono infatti ragionevolmente collocare all'intersezione fra alcune discipline a vocazione umanistico-qualitativa ed altre di carattere più scientifico-quantitativo.

Come forse molti possono ricordare, da reminiscenze di insiemistica più o meno lontane nel tempo, l'intersezione tra due o più insiemi spesso porta ad un insieme più piccolo di quelli che ne hanno dato origine. Non solo, si può infatti dare anche il caso di intersezione nulla, laddove non si trova alcun elemento comune. È quindi un'operazione delicata che comporta quasi sicuramente, da parte di tutte le componenti, un *rinunciare a qualcosa*, perlomeno in un primo momento.

Poniamo l'esempio della mia professione di insegnante di musica nella scuola secondaria: essa, per essere svolta efficacemente, necessita del contributo delle competenze epistemologiche, pedagogiche, didattiche, organizzative, relazionali e docimologiche. Tutte sono necessarie, ma nessuna è sufficiente per raggiungere l'obiettivo. Tutte sono necessarie, ma ciascuna deve fornire quante più *intersezioni*

possibili alle altre. Nessuna può vantare una superiorità preconstituita, tale da imporre condizioni più gravose alle altre, cosicché talune risultino infine snaturate.

Questa ricerca di nessi e collegamenti, che il senso comune non fatica a definire ideale e quantomai auspicabile, eppur tuttavia assai ardua e talvolta, di fatto, quasi impraticabile; proprio essa ha mosso chi scrive a “lasciare la scuola”, che da oramai quindici anni mi vede impegnato come insegnante, per “tornare a scuola” da studente in questo corso di dottorato. La constatazione empirica dei tanti e diversi benefici che la pratica musicale può produrre in chi vi si accosta, unita alla altrettanto empirica constatazione del recente rapido declino delle capacità attentive di buona parte degli studenti delle nostre scuole dell’obbligo, mi ha spinto a rimettermi in gioco per cercare di comprendere, io per primo, le evidenze scientifiche alla base dell’uno e dell’altro fenomeno. La convinzione qualitativa che la pratica musicale potesse avere un impatto positivo sull’attenzione ha incontrato la mia inclinazione quantitativa, ricevuta durante gli studi scientifici, e il lavoro di questi tre anni è in fondo un possibile tentativo di *intersezione*.

L’interdisciplinarietà nella ricerca, quantunque caldeggiata da più parti con lodevoli e condivisibili argomentazioni, risulta ancora oggi di non facile attuazione, per tutta una serie di motivazioni che bene sono state riassunte durante i corsi di formazione dottorale che abbiamo avuto il piacere di frequentare. Per darsi, essa ha bisogno di un *luogo* adatto, costituito non tanto da particolari forme o architetture, quanto da strutture non rigidamente divise per ambito disciplinare. Ma ancor più essa ha bisogno di *persone*, che a quella interdisciplinarietà siano sinceramente aperte e che stimino quelle *intersezioni* come il risultato scientifico più prezioso, perché proprio esse sono la rappresentazione più eminente della complessità del reale.

Al valore scientifico di questo lavoro di ricerca, di cui non compete a me la valutazione, sicuramente va aggiunto il valore del *luogo* che mi ha ospitato – il Dipartimento di Filosofia e Comunicazione - e delle *persone* che mi hanno accompagnato e supervisionato con pazienza, con grandissima competenza, senza snaturare il contributo che potevo offrire, ma cercando, appunto, *intersezioni*.

I primi capitoli del presente lavoro costituiscono il quadro teorico entro il quale si è sviluppato l'intero percorso di ricerca dottorale. Partendo da alcune definizioni di ordine generale, la trattazione procederà verso aspetti sempre più particolari che risulteranno poi quelli più pertinenti rispetto all'oggetto delle ricerche presentate.

L'attenzione, che già nel senso comune suscita non poca curiosità e nondimeno stimola riflessioni e convinzioni talvolta anche contrastanti fra loro, verrà quindi affrontata alla luce della grande mole di ricerche condotte dalla metà del secolo scorso fino al presente. Il punto di messa a fuoco si farà via via sempre più preciso, restringendo progressivamente il campo di osservazione alla modalità uditiva e alla componente spaziale. L'intersezione tra queste due sotto-categorie ha infatti dato vita ad un particolare filone di ricerca, che sarà oggetto privilegiato del presente lavoro.

I capitoli centrali rendono invece conto dell'attività di ricerca sperimentale svolta durante il percorso dottorale, illustrandone i presupposti, i metodi, i risultati e le relative interpretazioni proposte. Al termine di ognuno degli studi presentati, si trovano inoltre esposte alcune considerazioni circa possibili future direzioni di ricerca, a partire da quanto emerso e dal confronto con la letteratura esistente.

## I - Dal suono all'attenzione

### 1.1. La scena uditiva e la codifica dei suoni

L'esperienza quotidiana, vissuta attraverso l'udito, è spesso estremamente ricca di suoni, diversi per numerose caratteristiche di base fra le quali anche la loro collocazione spaziale. L'insieme di tutti gli eventi sonori presenti in un determinato ambiente e in un determinato tempo può essere definito *scena uditiva* (Bregman, 2002 in Levitin, 2002, pp. 213–248), prendendo in prestito questo termine dal lessico visivo, che con esso usa descrivere tutto ciò che si para davanti agli occhi.

Orientarsi e interagire efficacemente all'interno di questo *panorama* sonoro, significa fondamentalmente saper rispondere a due domande su quanto stiamo sentendo: *cos'è* e *dov'è* (Cherry, 1953). In altre parole, saper organizzare il *tutto* sonoro in *parti*, definite da un luogo e da un senso.

Questa capacità di raggruppamento e attribuzione di senso prende il nome di *analisi della scena uditiva*, e senza dubbio la gran parte di ciò che attualmente conosciamo sull'argomento si deve al pluriennale lavoro di Albert S. Bregman che, sia prima che dopo la pubblicazione del celebre volume “Auditory Scene Analysis: The Perceptual Organization of Sound” (1990) ha dedicato la quasi totalità della sua attività di ricerca alla comprensione di questo fenomeno.

Secondo l'autore, l'analisi della scena uditiva è una capacità che l'essere umano, così come altre specie, possiedono fin dalla nascita nei suoi tratti fondamentali (Bregman, 2015), e ha per oggetto le caratteristiche fisiche dei suoni che sono effettivamente percepiti dal nostro sistema uditivo. È importante, infatti, ricordare che la specie umana, così come le altre specie animali, ha alcune soglie acustiche che già riducono di fatto il numero massimo potenziale di suoni percepibili. Per quanto riguarda le frequenze, le normali soglie inferiori e superiori sono di 20 Hz e 20.000 Hz (Yost & Killion, 1997). Più in basso si parla di infrasuoni e più in alto di ultrasuoni, essi possono essere percepiti da altre specie animali, come ad esempio roditori e cetacei, ma non dalla specie umana. Riguardo alla potenza sonora,

misurata in Decibel, essa diminuisce all'aumentare della distanza dalla fonte, seguendo la seguente legge matematica:

$$P_r = P_f - 10 * \text{Log}_{10}(4\pi r^2)$$

La potenza originaria della fonte sonora ( $P_f$ ) e la potenza residua ( $P_r$ ) ad una certa distanza ( $r$ ) sono espresse in Db. Ad esempio, supponiamo che all'interno della scena uditiva sia presente una persona che parla con un tono di voce normale, che i ricercatori attestano mediamente tra i 55 e i 66 Db (Pearsons et al., 1977). Allontanandoci di un solo metro da questa persona, la potenza scende già a 49 Db, a dieci metri di distanza, la potenza scende a 29 Db, che equivale ad un lieve sussurro all'orecchio. A quaranta metri di distanza, la potenza si riduce a 17 Db, poco più che il rumore di un passaggio della mano su un foglio di carta. Si comprende, quindi, come anche la potenza sonora e la distanza dalla fonte siano fattori che influenzano significativamente la qualità e quantità dei suoni oggetto dell'analisi.

L'insieme delle caratteristiche acustiche di ogni suono, che costituisce quindi l'oggetto dell'analisi a carico del nostro sistema uditivo, è descritto dal tipo e dalla forma di *onda* che esso produce. Il suono, infatti, è essenzialmente una vibrazione, prodotta da una fonte, (Stevens & Davis, 1938) che si propaga attraverso un mezzo (tipicamente l'aria ma anche acqua, gas o materiali solidi). Il movimento di ognuna delle particelle di aria, investita da questa vibrazione, può essere descritto graficamente da un'onda, che ha una forma complessa poiché formata da varie componenti. L'analisi e la scomposizione di queste onde complesse prende il nome di *analisi spettro-temporale* (Bregman, 1990; Darwin & Carlyon, 1995; Darwin, 1997), poiché essa si occupa di stabilire le diverse frequenze (suoni acuti e gravi), ampiezze (suoni forti e deboli) e la loro diversa mistura (timbri sonori che aiutano a distinguere fonti diverse ma appartenenti ad una stessa categoria, ad esempio le voci umane). Anche il fattore tempo ha un'importanza cruciale in questa analisi poiché, come avremo modo di vedere più avanti, una delle specificità dei suoni è data proprio dal loro dispiegarsi nel tempo; dal fatto, cioè, che l'informazione in essi contenuta non è interamente e immediatamente disponibile.

Quando una fonte genera un suono, il conseguente movimento di particelle può raggiungere l'orecchio e inizia un processo di acquisizione e conversione di modulazioni di pressione meccanica in risposte neurali (Møller, 1961; 2012). L'orecchio esterno è responsabile della raccolta dei suoni, l'orecchio medio della loro amplificazione e l'orecchio interno della trasduzione dell'energia meccanica in attivazioni neurali. La particolare forma del *padiglione auricolare*, infatti, favorisce la raccolta e il convogliamento dei suoni all'interno del *canale uditivo*, che funge anche da elemento isolante nei confronti del successivo elemento, chiamato *timpano*. Esso è costituito da una membrana, al confine tra orecchio esterno e medio, e risponde alle sollecitazioni in ingresso vibrando coerentemente con esse. L'energia meccanica passa poi all'orecchio medio che, attraverso l'azione reciproca di *martello*, *incudine* e *staffa*, viene amplificata (in caso di suoni deboli) o ridotta (in caso di suoni troppo forti). In questa fase vengono bilanciate le diverse impedenze acustiche tra aria ed acqua, poiché allo stadio successivo il mezzo di propagazione dell'onda cambierà, passando appunto da aria a liquido. La *finestra ovale* della coclea funge poi da interfaccia verso l'orecchio interno, una piccola cavità chiusa contenente liquido. La funzione dell'orecchio interno può essere assimilata a quella della retina per la vista, infatti esso sovrintende alla trasduzione di ogni vibrazione meccanica in stimoli nervosi. I movimenti degli strati della *coclea*, costituita da tre canali paralleli (*timpanico*, *vestibolare* e *medio*) vengono tradotti in segnali neurali nell'*organo di Corti*, che si estende sulla parte superiore della *membrana basilare*, ricca di cellule specializzate chiamate *ciliate*. I segnali inviati da queste cellule (in particolare le cellule ciliate interne) attivano un gran numero di neuroni del nervo acustico. Ogni fibra di questo nervo è connessa con una sola cellula ciliata, inoltre ogni frequenza attiva in modo predominante soltanto alcune cellule, che risultano quindi specializzate per determinate frequenze caratteristiche. L'attivazione nervosa dipende poi anche dall'intensità del suono, variando in modo proporzionale ad essa in modo che suoni più intensi attivano un maggior numero di cellule ciliate e con maggior velocità. Questa organizzazione tono-topica delle strutture che elaborano diverse frequenze e ampiezze d'onda viene mantenuta per tutto il percorso nervoso e fino alla *corteccia uditiva primaria*. Qui

i neuroni vengono organizzati per bande di uguale frequenza, specializzate cioè nell'elaborazione di stimoli con determinate frequenze sonore, a partire dalla frequenza fondamentale che risulta essere quella predominante a livello di intensità. Vi sono poi sottoinsiemi di neuroni che codificano i livelli di intensità e le modulazioni temporali del segnale (se cioè un suono ha una sua continuità o meno in un dato tempo).

Per rispondere, quindi, alla domanda sul *cosa* sia la fonte che ha generato un particolare suono (Alain et al., 2001), è necessario distinguerlo dalla massa sonora complessiva che raggiunge il nostro orecchio. Questa operazione implica la segregazione delle componenti spettrali specifiche di ciascuna fonte, secondo meccanismi comparativi basati su alcune dicotomie basilari come uguaglianza/differenza e ripetizione/variazione. Ad esempio, elementi sonori caratterizzati da un attacco comune o da una comune variazione di intensità tenderanno ad essere percepiti come se fossero originati dalla stessa sorgente (Culling & Summerfield, 1995a; 1995b; Best et al., 2007). Inoltre, suoni che hanno relazioni armoniche forti fra loro (tipicamente intervalli di unisono, ottava, quinta giusta e quarta giusta), così come quelli che mostrano una continuità di frequenza nel tempo, tendono anch'essi ad essere raggruppati insieme (Bregman, 1990; Darwin & Carlyon, 1995; Carlyon, 2004).

Anche la domanda sul *dove* si trova la singola sorgente sonora può trovare risposta grazie all'analisi spettro-temporale. Nella modalità uditiva, questa è una domanda che assume un significato e una valenza particolari, se si pensa anche solo al fatto che circa la metà del campo uditivo complessivo non può giovare dell'aiuto della vista nel lavoro di localizzazione delle fonti. Infatti, come detto in precedenza, il campo visivo ha un'apertura di circa 120° binoculari e 30° monoculari aggiuntivi per ciascun lato sul piano orizzontale (Strasburger & Popper, 2002), invece l'udito ha un'apertura costante a 360°. Il *dove* uditivo necessita quindi di una stima assai complessa, alla quale contribuiscono in misura diversa anche fattori come la variazione di intensità nel tempo e la differenza temporale inter-aurale ITD, definita come la differenza tra l'istante in cui un suono raggiunge l'orecchio destro e

l'orecchio sinistro (Klumpp, Eady, 1956; Moore, 1991). Infatti, il momento esatto in cui un'onda sonora raggiunge le nostre orecchie è diverso se l'origine del suono è situata in posizioni diverse. L'informazione sull'origine spaziale del suono viene quindi ricavata sulla base del confronto temporale tra gli istanti in cui esso raggiunge l'uno e l'altro orecchio. A questo dato si aggiunge anche quello relativo al confronto tra le intensità con le quali il suono raggiunge le due orecchie, in questo caso si parla di ILD ovvero di differenza di livello inter-aurale. Esso è importante ai fini del rilevamento della distanza della fonte, poiché l'orecchio più vicino ad essa verrà investito con maggior intensità, al contrario di quello opposto, che invece riceverà un segnale meno intenso.

A tal proposito, interessante è il contributo di Teraoka e colleghi (2023), che studiano la stima del tempo necessario a entrare in contatto con la fonte sonora (time-to-contact), al variare della posizione relativa tra ascoltatore e stimolo. Il setting sperimentale era costituito da un sistema di cinque paia di altoparlanti posti a distanze progressive dal partecipante. A partire dai 34 cm di distanza minima, con incrementi di 34 cm si arrivava ad un massimo di 170 cm. La seduta del soggetto poteva ruotare, così che risultava possibile testare diverse angolazioni rispetto alla fila di altoparlanti che invece restava fissa. Gli angoli testati erano 0°, 45°, 90°, 135° e 180°. Lo stimolo uditivo, costituito da rumore rosa (un suono le cui frequenze basse hanno potenza maggiore rispetto a quelle alte), ricreava l'effetto di avvicinamento passando da tre altoparlanti consecutivi nello spazio nell'arco di 2 secondi. Ad esempio, il suono poteva essere presentato sulla terza coppia di altoparlanti, a distanza di 102 cm dal soggetto, per poi venire rapidamente presentato sulla seconda coppia (68 cm di distanza) e sulla prima (34 cm di distanza). Allo stesso modo, il suono poteva partire dalla quarta coppia e terminare sulla seconda, o dalla quinta e terminare sulla terza coppia. Il soggetto sentiva quindi l'avvicinarsi del suono e doveva rispondere su una pulsantiera, stimando il momento in cui il suono avrebbe raggiunto il suo corpo e premendo in quell'istante un pulsante. In un secondo esperimento, con il medesimo setting, gli stimoli erano invece presentati su una sola coppia di altoparlanti, e ai partecipanti era richiesto di stimare la distanza della fonte sonora. I risultati del primo esperimento mostravano

che, in linea generale, i tempi riportati erano sottostimati rispetto a quelli reali, confermando la letteratura esistente (Shaw et al., 1991; Rosenblum et al., 1993; Neuhoff, 2001). Tuttavia, nella condizione in cui il soggetto riceveva gli stimoli esattamente da dietro ( $180^\circ$ ), la performance di stima temporale risultava diversa dalle altre, mostrandosi meno dipendente dalla misura effettiva di tempo. Altre analisi mostrano comunque che complessivamente non vi sono differenze di performance tra le diverse direzioni indagate, e l'unica lieve specificità per la direzione  $180^\circ$  si rileva nel caso di tempi effettivi di contatto di 3 secondi, per i quali la stima risulta quasi la più accurata rispetto alle altre direzioni. Gli autori suggeriscono che i suoni provenienti direttamente da dietro potrebbero avere una valenza emozionale negativa (Asutay et al., 2015; Frankowska et al., 2020). Lo spazio posteriore, tuttavia, non comprendeva soltanto la posizione  $180^\circ$  (direttamente dietro) ma anche una posizione laterale posteriore ( $135^\circ$ ), per la quale non è stato riscontrato lo stesso effetto. La conclusione suggerita, nonostante il richiamo a evidenze precedenti, risulta poco supportata dai dati raccolti, poiché non risulta chiaro il motivo per cui lo spazio posteriore non ha un effetto uniforme sulla percezione degli stimoli come emozionalmente negativi e quindi più capaci di richiamare velocemente risorse attentive. Similmente, non vengono proposte conclusioni precise riguardo al secondo esperimento di stima della distanza spaziale del suono. Ad ogni modo, lo studio contribuisce ad una migliore comprensione del processo di stima della posizione dei suoni, poiché essa risulta alla base della stima del tempo di contatto. Infatti, la stima del tempo è strettamente connessa con la stima dello spazio, e la percezione di variazioni spaziali in senso di avvicinamento o allontanamento risulta di fondamentale importanza ai fini del tempo necessario per entrare in contatto con la fonte sonora.

Secondo il pensiero di Bregman, l'analisi della scena uditiva avviene in due stadi successivi e qualitativamente diversi. Il primo stadio, chiamato *primitivo* o *automatico* è situabile a livello sensoriale (processo bottom-up), non risponde a nessun obiettivo o compito predeterminato. Esso avviene quindi involontariamente e costantemente, il suo scopo è quello di scomporre il tutto sonoro in flussi separati, caratterizzati da elementi acustici comuni (frequenze fondamentali, attacco,

intensità, costanza temporale di alcune delle precedenti). Questi flussi costituiscono poi il materiale su cui andrà a lavorare il secondo stadio di elaborazione, volontario, basato sulla conoscenza pregressa, sulla memoria, sulle aspettative create da tutte le informazioni disponibili. In questo stadio avviene quindi un processo top-down, controllato, in cui l'attenzione forma delle unità sonore di senso, chiamate *oggetti* percettivi e dispiega le sue risorse sugli oggetti formati, in vista del raggiungimento di un obiettivo o il completamento di un compito.

Bregman, qui, si rifà anche alla teoria della *Gestalt* (Kofka, 1922; Wertheimer, 1938; Palmer, 1992 per la sua applicazione alla percezione), la quale postula, per l'ambito della percezione, che la totalità di quanto si percepisce in un dato tempo non è costituita dalla semplice sommatoria dalle singole attivazioni sensoriali, ma da qualcosa in più che permette di comprendere possibili forme emergenti e, in ultima istanza, la forma nella sua interezza. Essa è passata al grande pubblico attraverso la frase-slogan "il tutto è più della somma delle singole parti" (Zerbetto, 1998). Un approccio gestaltico al problema della formazione di oggetti percettivi permette di aprirsi ad una visione meno rigida sull'interazione reciproca fra le diverse informazioni fisico-acustiche codificate al livello sensoriale, e ancor più sulla loro integrazione volta alla formazione di unità di senso. Questa attribuzione di senso, appunto, necessita di operazioni ben più complesse di una semplice comparazione, addizione o accostamento di frequenze o intensità o timbri. Lo stadio di alto livello, infatti, secondo Bregman, si avvale non solo dell'informazione sui flussi percettivi fornita dallo stadio primitivo, ma anche di elementi cognitivi quali la memoria, la familiarità e l'aspettativa, che permettono di riorganizzare i flussi secondo possibili criteri di senso, alcuni dei quali sono soggettivi, altri invece sono perlomeno intersoggettivi.

Qualche decennio più tardi, anche Fritz e colleghi (2007) in un suo lavoro di revisione sull'attenzione uditiva, esporrà nuove evidenze anche neuroscientifiche a sostegno di questo impianto teorico (Scheich et al., 1998; Sridharan et al., 2007; Hafter et al., 2007; Xiang et al., 2008). Si approfondisce, qui, la definizione dei due stadi di elaborazione, e si cerca di entrare ancor più profondamente nel merito di

cosa avvenga automaticamente e cosa, invece, richieda l'intervento dell'attenzione. Da queste evidenze sembra emergere una visione più complessa dello stadio primitivo automatico, qui ridefinito come *pre-attentivo*. Infatti, vi sono evidenze che confermano come questo stadio lavori effettivamente in automatico, non prevedendo necessariamente alcun intervento di alto livello. La formazione dei flussi, che anche secondo Bregman avviene attraverso l'applicazione di dicotomie fondamentali come ripetizione/variazione, vive un momento cruciale quando si riscontrano appunto delle variazioni, soprattutto se rapide e concentrate in un tempo relativamente breve. Questo avvenimento sembra implicare e quasi richiamare un intervento dell'attenzione, che attira risorse su quanto sta accadendo. A sua volta, poi, l'attenzione sembra modulare la sensibilità uditiva, orientandola a favore delle caratteristiche o delle zone interessate da questi eventi inaspettati. Vi sarebbe, quindi, una continua e dinamica influenza tra i due stadi, che non esclude il fatto che, in un dato tempo, lo stadio pre-attentivo possa funzionare indipendentemente dall'attenzione; ma, una volta richiamata, l'attenzione è in grado di modulare le abilità percettive, piegandole in qualche modo allo scopo per il quale sta lavorando.

Per quanto detto sopra, gli autori definiscono lo stadio pre-attentivo *semi-automatico*, diversamente da Bregman, dal momento che si trovano prove neurali di questa influenza dell'attenzione sulla sensibilità percettiva di basso livello. È quindi un processo dinamico quello appena descritto, poiché istante per istante vi è un aggiornamento dei flussi, delle ripetizioni e variazioni, degli interventi attentivi e delle successive modulazioni orientate di sensibilità acustica. Inoltre l'attenzione, protratta nel tempo su un determinato oggetto percettivo, contribuisce a creare proprio quelle *aspettative* che sono alla base del processo di formazione di nuovi oggetti percettivi.

Applicando quanto appena detto al caso specifico della posizione delle sorgenti sonore (Rogers & Bregman, 1993) si può pensare che la localizzazione di una fonte emerga ad uno stadio sensoriale-percettivo, grazie al processo di segregazione e comparazione di elementi spettro-temporali, nella costante ricerca di qualche regolarità. Una volta avvenuta la stima della posizione spaziale della fonte,

interviene il controllo da parte dell'attenzione che segue la fonte sonora nel tempo, prendendo la sua posizione e il suo flusso come riferimento. Ci sono, infatti, evidenze (Shinn-Cunningham, 2008; Best et al., 2008) che mostrano come la performance di attenzione spaziale uditiva migliora nel tempo, man mano che il soggetto individua oggetti percettivi, li prende a riferimento e controlla costantemente il loro svolgersi temporale.

In seguito, Darwin (2008), basandosi su alcune evidenze da studi sulla detezione di voci umane in presenza di suoni distraenti da altre direzioni (Woods & Colburn, 1992; Hill & Darwin, 1996), propone una diversa visione del processo di localizzazione spaziale, ipotizzando un terzo stadio a se stante e deputato esclusivamente a questo obiettivo. Egli sostiene che la differenza inter-aurale (ITD) complessiva dei suoni in entrata non sia sufficiente per attribuire una posizione precisa alle fonti sonore. Il sistema uditivo, invece, opererebbe prima una segregazione delle varie frequenze, e successivamente calcolerebbe la differenza ITD separatamente per ogni singola frequenza. Il processo di localizzazione avverrebbe, quindi, attraverso la comparazione di tutti questi ITD e le frequenze caratterizzate da ITD uguali verrebbero attribuite alla medesima posizione spaziale. Darwin porta, a sostegno di questa ipotesi, la constatazione empirica che la localizzazione spaziale dei suoni, una volta acquisita, è un dato estremamente stabile nel tempo. Inoltre, una volta localizzato, se ne percepisce per intero lo spettro delle frequenze, il che potrebbe significare che tutte le sue frequenze sono state correttamente attribuite alla medesima fonte e alla medesima posizione spaziale.

Si arricchisce, così, il dibattito sulla posizione reciproca e sul rapporto tra processi di basso livello (sensoriali e percettivi) e di alto livello (attenzione, memoria, cognizione) che costituisce uno dei temi di ricerca più affascinanti e allo stesso tempo più complessi e sempre bisognosi di nuove evidenze per una più profonda comprensione.

## 1.2. L'attenzione durante l'analisi di una scena uditiva

Come si può evincere da quanto esposto in precedenza, l'abilità di districarsi tra la moltitudine di stimoli uditivi nella quale siamo abitualmente immersi è di fondamentale importanza per portare a termine azioni anche banali nella vita quotidiana. Tale abilità implica, quindi, una efficace analisi spettro-temporale a livello sensoriale e una altrettanto efficace ed efficiente allocazione di attenzione allo scopo di indirizzare la percezione verso le unità di senso sonoro più utili al raggiungimento di un dato obiettivo. In questo paragrafo vedremo più nel dettaglio *quando, come e perché* (in senso causale ma anche finale) l'attenzione interviene durante l'analisi della scena uditiva. Il discorso avrà un particolare accento sul dato spaziale, che rappresenta senza dubbio la lente privilegiata attraverso la quale guardare a questi processi che, altrimenti, richiederebbero trattazioni ben più ampie del presente lavoro.

L'attenzione è un costrutto psicologico molto ampio e dalle molte e diverse sfaccettature. La si può definire, come riporta l'APA Dictionary of Psychology (VandenBos, 2007), uno "*stato durante il quale le risorse cognitive sono focalizzate su alcuni aspetti dell'ambiente circostante, piuttosto che su altri, e il sistema nervoso centrale si trova nella condizione di poter prontamente rispondere a determinati stimoli*". L'attenzione, che gran parte delle evidenze in letteratura collocano idealmente all'intersezione fra percezione e cognizione, è una funzione quanto mai necessaria per portare a termine la quasi totalità delle azioni che la vita quotidiana richiede. L'operazione di *selezione* che essa opera sulla totalità degli stimoli di un dato ambiente, risponde da un lato al problema della limitatezza delle risorse allocabili in un dato momento (Kahneman, 1973), dall'altro ci permette di perseguire con una certa efficacia (Vecera & Rizzo, 2003) lo scopo che in quel momento riteniamo prioritario (Theeuwes, 2014). La selezione degli elementi ritenuti rilevanti in vista del perseguimento di un obiettivo, comporta necessariamente la soppressione, o perlomeno la parziale inibizione, di tutti quegli elementi che invece sono ritenuti distraenti rispetto all'obiettivo preposto (Murphy et al., 2016). Si tratta sempre di una selezione, ma di segno opposto rispetto alla

prima, ed entrambe le tipologie sono cruciali per portare a termine il compito o l'azione che ci siamo prefissati in un dato tempo.

L'attenzione sembra operare su *oggetti* più che su singole caratteristiche fisiche (Shinn-Cunninghama, 2008). L'efficacia e l'efficienza del processo di formazione di questi oggetti percettivi impatta, a sua volta, sull'efficacia e l'efficienza dell'allocazione di risorse attentive sulla scena uditiva, o su parte di essa. Ma quando interviene l'attenzione? Per rispondere a questa domanda possiamo immaginare che i tanti flussi sonori presenti in una scena uditiva siano idealmente in una sorta di *competizione* per attirare la nostra attenzione. Soltanto alcuni di essi riescono ad emergere dal paesaggio sonoro complessivo, per mezzo di determinate peculiarità che ne facilitano lo stagliarsi sullo sfondo.

La prima possibile peculiarità è rappresentata dal chiaro emergere di una specifica caratteristica fisica, che distingue un flusso da tutti gli altri. Quella che, nel lessico scientifico specifico è chiamata *saliienza* (vedasi, ad es., Kayser et al., 2005; Huang & Elhilali, 2017 per la modalità uditiva). Se un flusso sonoro si distingue dagli altri perché è il più acuto (ad es. il cinguettio di un uccellino o il suono sottile ma quasi fastidioso di un'apparecchiatura elettronica), avrà delle ottime possibilità di catturare la nostra attenzione. Similmente, se nella scena un flusso si distingue per la sua sovrastante intensità (il suono di un aereo a bassa quota, un'ambulanza, l'insieme dei clacson di una strada molto affollata), esso probabilmente *ruberà la scena* agli altri suoni. Ma le possibilità di emergere in questo modo naturale e non voluto non si esauriscono alle sole caratteristiche acustiche dei suoni. Infatti, anche la rilevanza personale per l'ascoltatore può costituire un elemento di *saliienza quasi automatica*; un esempio potrebbe essere il sentir pronunciare il proprio nome, o il sentir parlare la propria lingua madre durante una passeggiata in un paese straniero (Moray, 1959; Wood & Cowan, 1995; Conway et al., 2001). Una ulteriore possibilità di attirare la nostra attenzione è costituita dall'insorgenza di una variazione, rapida e significativa, di una certa caratteristica acustica di base. Infatti, in una scena caratterizzata da suoni prevalentemente acuti, l'insorgenza di un suono grave (ad es. la chiusura di un portone a causa del vento o la caduta improvvisa di

un oggetto voluminoso) produrrà probabilmente una allocazione di attenzione. L'elemento temporale, che connota fortemente la modalità uditiva, anche in questo caso risulta determinante. La salienza, quindi, è un attributo di elementi *esterni* al soggetto, capace proprio dall'esterno di suscitare l'attenzione all'*interno* dello stesso.

Oltre alla salienza, può accadere che vi sia un orientamento intenzionale dell'attenzione del soggetto verso un determinato flusso o verso una specifica caratteristica acustica, ritenuti funzionali al completamento di un compito (Kidd et al., 2005; Best et al., 2007; Maddox et al., 2012). Ad esempio, un soggetto potrebbe essere interessato a seguire la melodia di una canzone. Questa volontà attribuisce maggiore importanza ai flussi caratterizzati da bande di frequenza acute come quelle di un violino; non solo, se il soggetto ha una qualche conoscenza o memoria di quella canzone, tenderà a ricercare nei flussi sonori le stesse relazioni fra suoni vicini (dette, nel lessico musicale, *intervalli*) che costituiscono poi lo svolgersi orizzontale della melodia nel tempo.

Un'altra domanda importante riguarda il *come* interviene l'attenzione, cosa accade a livello neurale e a livello sensoriale una volta che essa ha dispiegato le sue risorse su alcuni elementi della scena uditiva.

Poniamo un esempio simile al precedente: mentre una persona sta guidando, sente in lontananza la sirena di un'ambulanza. Si tratta, quindi, di un flusso con una salienza esterna al soggetto, molto ben connotato perché sicuramente presente nella memoria. Non solo, la conoscenza del codice della strada implica che si debba identificare nel più breve tempo possibile la posizione del veicolo di soccorso, e che gli si debba lasciare strada libera. Il primo passo che compie l'attenzione è quello di raggruppare i flussi opportuni (i singoli suoni che compongono il tipico segnale della sirena di ambulanza italiana, che hanno rispettivamente 392 Hz e 660 Hz di frequenza). Ancora non è però sufficiente per formare l'oggetto percettivo, poiché questi due suoni, per essere identificati come *sirena di ambulanza* devono essere riprodotti in un preciso ordine temporale: il ciclo singolo dura 3 secondi ed

è così suddiviso (dal manuale del produttore di sirene per il mercato italiano Uccino Produzioni Elettroniche, n.d.):

- suono a 392 Hz per la durata di 1/3 del ciclo (1000 ms)
- suono a 660 Hz per la durata di 1/18 del ciclo (167 ms)
- suono a 392 Hz per la durata di 1/18 del ciclo (167 ms)
- suono a 660 Hz per la durata di 1/18 del ciclo (167 ms)
- suono a 392 Hz per la durata di 1/3 del ciclo (1000 ms)
- suono a 660 Hz per la durata di 1/18 del ciclo (167 ms)
- suono a 392 Hz per la durata di 1/18 del ciclo (167 ms)
- suono a 660 Hz per la durata DI 1/18 del ciclo (167 ms)

Qui il fattore tempo risulta determinante, insieme al fattore aspettativa e familiarità pregressa. Infatti, forse non basteranno il primo e il secondo suono, ma molto probabilmente già dopo il terzo, e quasi sicuramente dopo il quarto, i sistemi di alto livello rileveranno la struttura del segnale e formeranno così l'oggetto percettivo *sirena di ambulanza*. Una volta che l'attenzione si è *sintonizzata* su questo oggetto, lo segue nel tempo, e un primo effetto comportamentale di questo seguire la fonte nel tempo si manifesta nel rilevamento dinamico della posizione attuale dell'ambulanza. Infatti, come evidenziato da alcuni studi (Buschman & Miller, 2007; Elhilali et al., 2007) l'attenzione agisce potenziando la rappresentazione neurale delle frequenze e dei gruppi di suoni riconducibili al tipico segnale della sirena. La competizione dei flussi per l'attenzione risulta quindi non più ad armi pari, almeno per il tempo necessario al completamento del compito, poiché vengono in qualche modo privilegiati alcuni flussi a scapito di altri.

Lo stadio attentivo volontario sembra, poi, responsabile della definizione dei vari *piani* sonori (Fritz et al., 2007a; 2007b ), cioè contribuisce alla percezione di uno *sfondo*, sopra il quale si stagliano alcune *figure* (mutuando il lessico dalla modalità visiva). Un dettaglio non di poco conto, che rende questa definizione più complessa, è che i rapporti tra figura e sfondo possono cambiare nel tempo, anche in un tempo brevissimo. Non sono quindi staticamente definiti come per un'immagine, ma ciò

che ad un certo istante è parte dello sfondo, può improvvisamente emergere come figura protagonista, e viceversa ciò che era inizialmente emerso come figura può sfumare e finire in piani sonori idealmente *posteriori*.

### **1.3. I primi modelli teorici dell'attenzione: il paradigma dell'ascolto dicotico**

La quasi totalità degli studi pionieristici sull'attenzione, temporalmente situati tra gli anni '50 e '60 del secolo scorso, ebbe come oggetto la modalità uditiva. Gli anni che seguirono, videro un'ascesa di interesse per la modalità visiva, per poi assistere ad un singolare ritorno di interesse per l'udito verso la fine del secolo scorso. Al crescente interesse per lo studio dei processi mentali implicati nell'attenzione, corrispose un rapido e sorprendente sviluppo tecnologico dei dispositivi di registrazione e riproduzione audio. Le applicazioni belliche sempre più sofisticate, così come, nel dopoguerra, la grande spinta alla diffusione sovranazionale della musica e della comunicazione radiotelevisiva, attirarono verso il mondo sonoro molte e ingegnose menti di allora. Videro così la luce i primi registratori a nastro multi-traccia che, ben presto, valicarono i confini delle sale di registrazione e degli studi televisivi e radiofonici, per trovare nuove applicazioni anche nella ricerca psicologica. Il crescente livello di fedeltà, offerto dalle nuove tecnologie, fu un elemento fondamentale per innalzare anche il livello di affidabilità e replicabilità delle misure adottate e raccolte durante gli esperimenti.

Il primo, celebre paradigma sperimentale, concepito per studiare l'attenzione uditiva nello spazio circostante, fu l'*ascolto dicotico* (Cherry, 1953; Broadbent, 1956), nel quale ai partecipanti venivano presentati due stimoli sonori simultaneamente, uno all'orecchio destro e l'altro al sinistro. La presentazione avveniva solitamente attraverso l'uso di cuffie, e ai soggetti era poi richiesto di svolgere uno fra una varietà di possibili compiti su quanto ascoltato, a seconda delle diverse domande di ricerca e dei diversi tipi di attenzione studiati. In un compito di attenzione selettiva, ad esempio, si chiedeva al soggetto di prestare attenzione soltanto ad uno dei due stimoli, ignorando l'altro; mentre in un compito di

attenzione divisa, la richiesta era di fare attenzione a entrambi gli stimoli contemporaneamente.

Gli stimoli sonori erano costituiti quasi sempre da voci umane, che potevano poi differire o meno fra loro per caratteristiche sonore (timbro, intensità, altezza) o per contenuto (diverse parole o sillabe) o per entrambi. Grazie ai dispositivi di registrazione e riproduzione descritti sopra, era infatti divenuto possibile variare o combinare caratteristiche sonore e contenuti. Ai soggetti era talvolta richiesto di riconoscere il lato dal quale veniva presentato lo stimolo target, di annotare la descrizione dello stimolo target al termine della prova oppure, nel caso di stimoli contenenti parole o sillabe, di ripetere quanto ascoltato (Hugdahl, 2011 per una rassegna aggiornata delle varie declinazioni del paradigma)..

Per quanto riguarda l'attenzione selettiva, gli studi di riferimento sono quelli di Broadbent (1952; 1954), Cherry (1953) e Poulton (1953), che mostrarono senza particolari eccezioni che le differenze acustiche di base tra i segnali, così come la loro diversa provenienza spaziale, ne facilitavano la discriminazione. Uno degli esperimenti di Cherry prevedeva la presentazione simultanea di due messaggi vocali sovrapposti ad entrambe le orecchie, e la richiesta ai partecipanti era quella di riportare uno dei due messaggi al termine della prova. In questa situazione, i partecipanti trovavano assai difficoltoso distinguere le parole dei due messaggi, ed erano necessarie molti riascolti prima di poter riportare per intero un messaggio. Spesso, gli errori riguardavano parole che potevano ragionevolmente essere associate al contesto generale del messaggio, e questo è stato interpretato come un intervento dell'attenzione di alto livello. In un altro esperimento, invece, un messaggio veniva presentato all'orecchio destro e un diverso messaggio all'orecchio sinistro. I partecipanti venivano istruiti su quale fosse l'orecchio cui prestare attenzione e il compito era quello di riportare il contenuto del messaggio corrispondente. I risultati mostravano che il messaggio dal lato atteso veniva riportato, mentre non si ricordava pressoché niente del messaggio dal lato non atteso, il che portava l'autore a concludere che l'attenzione avesse efficacemente selezionato il lato desiderato, ignorando ciò che veniva presentato altrove. Tra le

possibili differenze nei messaggi da riportare, quella che risultava più di aiuto era proprio la posizione nello spazio. Non solo, Cherry scoprì che la selezione era ancora più efficace se i partecipanti venivano precedentemente informati su quale delle due posizioni sarebbe stata quella a cui prestare attenzione. Lo stesso non avveniva se, invece, questa stessa informazione veniva data dopo l'ascolto. Negli stessi esperimenti, fu poi chiesto ai partecipanti di ripetere ad alta voce il contenuto dello stimolo presentato sul lato richiesto (compito di *shadowing*). I risultati osservati mostrarono che, dello stimolo concorrente, non veniva ricordata se non qualche caratteristica grossolana, come ad esempio se la voce era maschile o femminile.

L'evidenza di un meccanismo in grado di selezionare alcuni stimoli, e altrettanto in grado di scartarne altri in base a determinate caratteristiche ritenute in partenza non necessarie al completamento del compito richiesto, fu riassunta e interpretata da Broadbent (1958) con la *teoria del filtro*. Essa postula che, dal momento che le caratteristiche fisiche dei suoni risultano influenti sul meccanismo di attenzione selettiva, si verrebbe a costituire un filtro, prima del quale tutto il segnale audio viene velocemente analizzato a livello sensoriale (elaborazione in parallelo), e dopo il quale soltanto gli stimoli o le caratteristiche ritenute determinanti per il compito vengono ulteriormente processate a livello percettivo (elaborazione in serie e più precisa della precedente). L'autore offrì una visione dell'azione umana basata sul concetto di *elaborazione di informazioni*. Egli prese a prestito dal frizzante e dinamico mondo delle comunicazioni alcuni assunti chiave. Il primo è che ogni funzione umana si può esprimere in termini di *flusso di informazioni* che giunge all'organismo, viene elaborato e reimpresso nell'ambiente esterno. Il secondo è che la mente umana può essere assimilata ad un *canale di comunicazione*, unico e limitato nella quantità di informazioni processabili in un dato tempo. Il terzo, discendente dal secondo, è che la mente sembra regolata da principi di *economia di risorse*, che proteggono il sistema dal sovraccarico e filtrano l'ammontare di informazione che complessivamente si affaccia ai nostri sistemi periferici di comunicazione con l'ambiente, selezionando soltanto una quantità limitata e funzionale di informazioni. Questa *selezione precoce*, che quindi può avvenire

anche secondo il criterio della posizione spaziale degli stimoli da processare, implicherebbe l'esistenza di un *luogo* di ritenzione a brevissimo termine delle informazioni in ingresso, che le mantiene in memoria soltanto per il tempo necessario al filtraggio. L'informazione che non supera il filtro selettivo, viene definitivamente persa. Nei suoi esperimenti, egli chiedeva ai partecipanti di prestare attenzione ad un solo orecchio, mentre ad entrambi venivano presentati stimoli uditivi continui. Così come accadeva negli esperimenti di Cherry, anche qui l'informazione presentata all'orecchio non atteso veniva completamente sostanzialmente ignorata e, al momento di riportare quanto ascoltato, i partecipanti non riuscivano a ricordarne né il contenuto né alcune caratteristiche acustiche salienti.

Le ricerche di Broadbent hanno costituito per decenni un punto di riferimento, influenzando il dibattito e la direzione di gran parte delle successive ricerche sull'attenzione.

Ad ogni modo, è opportuno considerare che le ordinarie situazioni di vita quotidiana implicano una complessità talvolta notevolmente superiore a quella di un compito di ascolto dicotico. Ad esempio consideriamo la circostanza in cui una persona sta parlando con alcuni amici a breve distanza e, contemporaneamente, altre persone intorno stanno facendo altrettanto; ciò che nel gergo scientifico viene definito "effetto cocktail party" (Cherry, 1953; Darwin & Carlyon, 1995 per una rassegna). In una tale situazione, gli stimoli sonori sono ben più di due, hanno origine da ben più di due direzioni e ci è richiesto di seguire ben più di un flusso alla volta, ignorando talvolta ben più di un solo flusso distraente. L'aspetto che tuttavia risulta rilevante per la ricerca in materia è che, per la prima volta, si indaga più di una posizione nello spazio, iscrivendo così definitivamente il dato spaziale nell'elenco delle caratteristiche dei suoni che possono essere manipolate e studiate, sia dal punto di vista della percezione, sia da quello dell'attenzione.

Negli anni seguenti, la teoria del filtro fu oggetto di nuovi studi che, se da un lato non misero direttamente in discussione il suo impianto generale – la limitatezza della capacità di processamento e la conseguente necessaria esistenza di un filtro

per selezionare la parte rilevante del tutto costituito dall'informazione in ingresso - , dall'altro posero in discussione il posizionamento temporale del suddetto filtro. Ci si chiedeva, cioè, se veramente la selezione avvenisse precocemente e soltanto sulla base di caratteristiche acustiche dei suoni, o se invece essa si verificasse in un tempo relativamente più lontano dall'insorgenza degli stimoli, potendo forse giovare di qualche operazione già avvenuta su di essi o su qualcuna delle loro caratteristiche. Un primo punto a favore di questa seconda visione fu portato da Moray (1959), che mise a punto una variante del compito di ascolto dicotico, nella quale presentava simultaneamente due flussi di parole, uno all'orecchio destro e un altro all'orecchio sinistro, chiedendo ai partecipanti di riportare le parole pronunciate dal lato atteso. Egli mostrò come i partecipanti, pur con l'indicazione preventiva di quale sarebbe stato il lato al quale fare attenzione, riuscivano comunque a riconoscere se dall'altro lato veniva pronunciato il loro nome. Il resto degli stimoli presentati dal lato non atteso venivano invece quasi sempre dimenticati, ad eccezione del proprio nome. L'evidenza suggerì all'autore che qualche operazione sugli stimoli da ignorare fosse quindi avvenuta comunque. Ulteriore evidenza a sostegno dell'intervento più tardivo del filtro attentivo fu portata dagli esperimenti di Anne Treisman (1960; 1964), che mostrò come il contenuto semantico di quanto ascoltato era rilevante ai fini della performance di ascolto dicotico e di shadowing. Infatti, ai suoi partecipanti furono fatte ascoltare voci umane che leggevano passi di una novella oppure paragrafi di pubblicazioni tecnico-scientifiche. Venivano, inoltre, sempre istruiti su quale fosse l'orecchio al quale fare attenzione e talvolta su entrambi veniva presentata la stessa voce con letture di novelle, talvolta invece la stessa voce con lettura di novelle da un lato e lettura scientifica dall'altro. I risultati mostravano che la condizione più gravosa per i partecipanti era quella di ripetere il passaggio di una novella quando dall'altro orecchio si ascoltava la stessa voce narrante un altro passo della stessa novella. Invece, risultava più facile ripetere quando, all'orecchio non atteso, veniva presentata la lettura scientifica. Anche in questo caso, vi era evidenza che qualche operazione di più alto livello era avvenuta prima dell'arrivo al filtro attentivo, nella fattispecie un riconoscimento di uguaglianza/diversità semantica.

Tuttavia, per una definizione formale di *teoria tardiva dell'attenzione*, si deve far riferimento ai lavori di Deutsch & Deutsch (1963), i quali postularono che, contrariamente a quanto sostenuto da Broadbent, l'informazione uditiva venisse processata per intero e ad alto livello, e venisse così a formarsi una graduatoria di importanza, utile per la successiva selezione. Essi sostengono che, in condizioni normali, lo stimolo che raggiunge il massimo livello di importanza è quello che entrerà in memoria ed eventualmente elicerà una risposta motoria. Questa attribuzione di importanza, evidentemente, suppone operazioni di più alto livello rispetto a quelle meramente sensoriali, in contrasto con la teoria della selezione precoce.

In ogni caso, è da notare che nei precedenti esperimenti, soltanto in una limitatissima percentuale di prove i partecipanti erano in grado di fare affermazioni su quanto udito dal lato non atteso. Ne consegue che le interpretazioni allora proposte avrebbero dovuto essere trattate con cautela, e talvolta gli stessi autori, come nel caso della Treisman, cercarono un punto di caduta accettabile fra i due estremi, sostenendo che il filtro potesse agire non già come un rubinetto con soltanto due posizioni estreme – aperto o chiuso – ma piuttosto come un sistema di attenuazione rispetto agli stimoli che non rispondevano a quanto atteso. Il che risultava compatibile con i risultati allora osservati e, nello stesso tempo, lasciava aperto il dialogo sulla questione della selezione precoce o tardiva.

Il dibattito ha appassionato intere generazioni di studiosi fino al nostro secolo, e sono state portate evidenze anche da studi di neuroscienza cognitiva a sostegno dell'una o dell'altra ipotesi. Ad esempio, Woldorff & Hillyard (1991) somministrarono esperimenti di ascolto dicotico in cui i soggetti prestavano attenzione selettiva a sequenze di toni presentati rapidamente ad un orecchio, ignorando toni di altezza diversa presentati all'orecchio opposto. Le misure ERP durante il compito mostravano una positività della componente P20-50 nei primi 50 millisecondi in zone frontali, in risposta ai toni presentati dal lato atteso. Queste attivazioni in risposta alla sequenza di toni target erano coerenti con le prestazioni di discriminazione a livello comportamentale. I risultati portano gli autori a

sostenere l'ipotesi di selezione precoce, secondo cui l'attenzione selettiva può filtrare e orientare l'elaborazione degli stimoli prima del completamento dell'analisi percettiva.

Di contro Deouell e colleghi (2007) proposero tre esperimenti, in uno dei quali presentarono una serie di stimoli in diverse posizioni dell'emicampo destro, la maggioranza dei quali nella stessa posizione, mentre alcuni in posizioni diverse. In alcuni blocchi di prove, poi, non veniva presentato alcun suono. Il compito richiesto ai partecipanti era però di ignorare tutti i suoni e fare attenzione ad un video muto, presentato su uno schermo. La registrazione dell'attività cerebrale, effettuata con risonanza magnetica, mostrò che i suoni attivavano comunque alcune regioni del giro temporale superiore, anche se il compito richiedeva di ignorarli e concentrarsi sullo stimolo visivo. I risultati suggerirono agli autori che un'analisi percettiva venisse portata a termine, a prescindere dal focus dell'attenzione e dalle richieste del compito da svolgere.

Nel tempo si farà sempre più viva l'ipotesi di una terza via, che contiene e prova a integrare elementi delle due precedenti, ma introduce anche elementi di novità (Johnston & Dark, 1986; Navon, 1989; Driver, 2001; Loui et al., 2005). Essa ipotizza che l'attenzione agisce dinamicamente su più livelli in un ideale continuum tra sensoriale/elementare e cognitivo/associativo, selezionando caratteristiche rilevanti e processandole interamente agli stadi più sensoriali, ma anche combinando le caratteristiche selezionate per formare prima dei flussi e poi degli oggetti percettivi coerenti, ad uno stadio più cognitivo e associativo. Alcune evidenze di attivazioni neurali di componenti relative all'attenzione uditiva, generate ad uno stadio temporale propriamente pre-attentivo, ma con modulazioni tipicamente attentive, sono state interpretate come una conferma di questa ipotesi intermedia. Ad esempio, Woldorff e colleghi (1998) mostrarono che, in un compito di ascolto dicotico, i toni devianti presentati dal lato atteso hanno elicitato sia un campo magnetico non corrispondente, sia una negatività non corrispondente notevoli nei primi 150-200 millisecondi, al contrario di ciò che avveniva per i suoni presentati dal lato non atteso. Questi risultati portano gli autori a ipotizzare che

anche i processi di analisi spettro-temporale di basso livello non siano completamente automatici, ma possano essere influenzati da meccanismi di selezione e soppressione, a seconda della direzione del focus attentivo.

In queste ricerche si fa sempre più chiaro il concetto di *oggetto percettivo*, che per la modalità uditiva significa una unità di senso sonoro, caratterizzata dalla uguaglianza e/o invarianza di alcuni elementi acustici di base. La capacità di formare questi *oggetti* è un processo cruciale (Shinn-Cunningham, 2008) dal quale dipende l'efficacia dell'attenzione selettiva uditiva nei contesti di vita reale. Se questa abilità è incerta o compromessa, anche il filtraggio delle fonti sonore risulterà altrettanto incerto e produrrà quindi un notevole aggravio a carico delle risorse cognitive e attentive.

#### **1.4. La teoria del carico percettivo e la sua applicazione alla modalità uditiva**

Un passo ulteriore, nella comprensione profonda dell'analisi della scena uditiva e del rapporto reciproco fra stadio automatico-sensoriale e stadio volontario-attentivo, è stato fatto cercando un possibile perimetro entro il quale poter ragionevolmente applicare alla modalità uditiva la celebre *teoria del carico percettivo* (Lavie & Tsal, 1994; Lavie, 1995; 2005; 2010), originariamente enunciata e studiata per la vista. Essa postula che il successo o il fallimento della selezione operata dall'attenzione dipende dalla difficoltà di elaborazione del compito da affrontare. In altre parole, si afferma che il livello di carico percettivo, così come quello del carico cognitivo, determina direttamente l'efficacia del processo di selezione e del concomitante processo di inibizione degli elementi distraenti.

Una delle implicazioni più interessanti di questa teoria, che ha ricevuto non poche conferme negli oramai quasi trenta anni dalla sua enunciazione (Murphy et al., 2016), è che il nostro sistema percettivo ha una capacità di elaborazione non infinita, al contrario essa è limitata e non costante nel tempo. Altro aspetto cruciale è che la capacità massima disponibile in un dato tempo non dipende dal nostro controllo volontario, così come sembra non dipendere dalla nostra volontà

l'ammontare di risorse reclutate in un certo tempo. Nel frattempo, tuttavia, tutta l'informazione disponibile viene automaticamente elaborata a basso livello, almeno fin quando detta capacità percettiva non si esaurisce. Il ruolo del controllo volontario, secondo i postulatori della teoria, non sarebbe quindi tanto di stabilire *quanta* attenzione allocare, quanto piuttosto *quali* informazioni privilegiare e analizzare più nel dettaglio (formazione di oggetti percettivi), in funzione del compito prefissato e quali, invece, scartare perché ritenute non rilevanti per il compito. Il carico assorbito dall'informazione rilevante risulta perciò decisivo per stabilire l'ammontare residuo della capacità percettiva totale, che rimarrà a disposizione e verrà impiegato automaticamente per l'elaborazione delle informazioni irrilevanti.

Un esempio estremo potrebbe essere quello di un compito con un carico percettivo basso (riconoscere la voce della mamma senza altre voci distraenti). Esso richiederà un basso carico percettivo, lasciando così molte risorse a disposizione del processamento di altri suoni dell'ambiente, che sicuramente non sono rilevanti per il presente compito, ma sui quali saremo in grado di avere un buon livello di dettaglio, proprio perché il carico residuo era alto. Al contrario, in condizioni di carico percettivo alto (ad esempio riconoscere la voce della mamma in mezzo ad una strada molto affollata di persone) il compito esaurirà tutta, o quasi tutta, la nostra capacità percettiva, e non resteranno risorse residue per poter elaborare altri elementi sonori presenti (i vari suoni provenienti dalle automobili, dai negozi ecc.).

Come già precedentemente accennato, la teoria fu originariamente enunciata a partire da evidenze nella modalità visiva, anche se gli autori vedevano potenziali applicazioni anche alla modalità uditiva (Lavie & Tsal, 1994). Nonostante ciò, gli studi che hanno provato a testare l'applicabilità della teoria all'udito non sono molti e, soprattutto, portano risultati contrastanti. Un interessante lavoro di Murphy e colleghi (2017) analizza le evidenze disponibili, dalle quali emerge un quadro frammentario e variegato.

Innanzitutto, risulta necessario stabilire un perimetro entro il quale vista e udito possano essere ragionevolmente comparati per quanto riguarda l'attenzione e la

percezione. Infatti, l'udito è considerato da più parti come il nostro "sistema di allertamento rapido" (Spence & Driver, 1994; Dalton & Lavie, 2004; Spence & Santangelo, 2009), aperto costantemente a 360° per monitorare tutto l'ambiente sonoro, diversamente da altri sensi che invece possono focalizzarsi su porzioni di spazio più ristrette. Già da questo dato, si potrebbe ipotizzare un diverso trattamento degli stimoli irrilevanti, perlomeno tra vista e udito. Infatti, se si considera la prospettiva evoluzionistica applicata ai sensi, l'elaborazione uditiva di stimoli non rilevanti può aver contribuito positivamente alla sopravvivenza della specie (si pensi a tutti i potenziali pericoli provenienti da zone non coperte dal campo visivo, che sono rilevati solo grazie all'udito). Inoltre, vi sono evidenze (Caparos & Linnel, 2010) che mostrano come, nella modalità visiva, un carico percettivo crescente produce un restringimento spaziale dell'area di interesse, cosa che non sembra ragionevolmente replicabile per l'udito, se si tiene ferma la sua funzione di allertamento rapido a 360°. Chiaramente le differenze fra udito e vista non si esauriscono assolutamente con queste considerazioni, che tuttavia sono state selezionate perché specifiche e funzionali al tema del rapporto fra processi di basso e di alto livello nella elaborazione degli stimoli provenienti dall'ambiente.

Iniziando dagli studi sull'ascolto dicotico – di cui abbiamo parlato in precedenza – gli autori si pongono il problema del se, e fino a che punto, i risultati potessero essere interpretati come una conferma di quanto asserito circa il carico percettivo. Essi mettono sotto la lente di osservazione quanto accade ai suoni presentati dal lato non atteso (suoni distraenti rispetto al target). Normalmente, essi vengono ignorati, ma talvolta se ne riconoscono e ricordano alcune caratteristiche particolari (il già citato studio di Moray, 1959 mostrava come il proprio nome era riconosciuto anche dal lato non atteso). Questo fenomeno è stato replicato anche in tempi più recenti (Wood & Cowan, 1995; Cowan & Wood, 1997; Conway & Kane, 2001; Rivenez et al., 2006; Bronkhorst, 2015 per una rassegna aggiornata). Recentemente, Gamble e Luck (2011) hanno realizzato uno studio sui potenziali evento-correlati (ERP) indagando una situazione di ascolto dicotico con un compito di rilevamento del suono target (se presente fra quelli casualmente scelti e presentati alle due orecchie). Le attivazioni della componente N2ac e la contemporanea positività della

componente N2 controlaterale hanno portato gli autori a concludere che ciò che veniva osservato era una allocazione di risorse per il compito di discriminazione tra suono atteso e non atteso.

Queste evidenze, se da un lato potrebbero portare supporto alla teoria del carico, dal momento che normalmente i suoni distraenti ricevono pochissima elaborazione; dall'altro suggeriscono alcune cautele nel trarre conclusioni. Infatti, i risultati che mostrano talvolta una elaborazione di alto livello sui suoni distraenti, uniti al fatto che il carico percettivo era costante (non manipolato a livello sperimentale), forniscono indicazioni contrastanti circa una possibile e indistinta applicazione dei suddetti principi alla modalità uditiva.

Vi sono stati, in seguito, studi che hanno provato alcune tipologie di manipolazione del carico percettivo nei loro paradigmi sperimentali. Le tipologie di manipolazione sono state essenzialmente tre: variare la quantità di oggetti percettivi nella scena, variare il grado di similitudine fra stimoli target e non-target, variare la quantità di operazioni richieste sugli stimoli.

Per la prima tipologia, è utile riportare i risultati di uno studio (Fairnie et al., 2016) nel quale ai partecipanti venivano presentati contemporaneamente diversi suoni, provenienti da diverse direzioni nello spazio. Il suono target (il verso di un leone o, in alternativa, di un cane) era presentato insieme ad altri suoni non-target costituiti sempre da versi di animali. Il carico è stato manipolato variando il numero di suoni non-target presenti sulla scena e inserendo, nella metà delle prove, il suono di un'automobile nella posizione di volta in volta più lontana dal suono target. Il compito richiedeva sia di determinare quale fosse stato il suono target (leone o cane), sia di riportare la presenza o meno del suono dell'automobile. I risultati, in accordo con la teoria del carico, mostrano come l'accuratezza di percezione del suono dell'automobile si riduceva all'aumentare del numero di suoni non-target. Sembra, cioè, che all'aumentare del carico richiesto dai suoni potenzialmente rilevanti per il compito (tutti i versi degli animali presentati) si riducano le risorse disponibili per elaborare i suoni distraenti.

Una possibile critica all'interpretazione appena proposta, che porterà poi altri ricercatori ad adottare la seconda tipologia di manipolazione (il grado di similitudine tra stimoli target e non-target), sorge dalla considerazione che un aumento degli oggetti nella scena potrebbe ridurre la capacità di discriminazione di ciascun oggetto, sia esso target o non-target. Non si potrebbe, cioè, ascrivere con certezza la riduzione di sensibilità al suono dell'automobile ad un aggravio delle risorse allocate sugli stimoli potenzialmente target, perché tale riduzione potrebbe essere semplicemente il risultato dell'aumento complessivo degli oggetti da analizzare. Per cercare di escludere questa eventualità, alcuni studi (ad es. Murphy et al., 2013), hanno mantenuto invariato il numero di oggetti sonori tra situazione ad alto e basso carico percettivo, andando invece a manipolare il grado di similitudine tra target e non-target. In un compito di tipo *flanker*, venivano presentate in rapida successione sei lettere in posizione centrale, pronunciate da una voce femminile. Fra queste sei lettere c'era quella da identificare, mentre le altre cinque erano da considerare non-target. Il distrattore era costituito dalla stessa lettera target, presentata però lateralmente e pronunciata da una voce maschile. In condizioni di basso carico, le lettere non-target erano distanti a livello fonemico (il target era la lettera P mentre il non-target era la lettera X), mentre in condizioni di alto carico, le lettere non-target erano più vicine a livello fonemico (ad es. la lettera C, H, G, J, K), il tutto considerando che la lingua parlata era l'inglese. I risultati mostravano una chiara riduzione di performance di rilevamento in condizioni di alto carico, la grandezza dell'effetto di congruenza non era differenze tra condizioni ad alto e basso carico. Anche in questo caso, quindi, i risultati erano solo parzialmente coerenti con gli assunti della teoria del carico percettivo.

La terza tipologia di manipolazione, ovvero la variazione delle operazioni richieste sugli oggetti target, consiste nel presentare la stessa quantità e lo stesso tipo di stimoli in entrambe le condizioni (alto e basso carico percettivo) ma di richiedere un numero di operazioni diverso per completare il compito. Un esempio di esperimenti del genere è quello di Alain & Izenberg (2003) che hanno presentato ai loro partecipanti due flussi identici (uno per ciascun orecchio) costituiti da una serie di suoni intonati e non intonati, dando preventivamente istruzione su quale dei due

orecchi fosse quello cui prestare attenzione. La condizione di basso carico richiedeva esclusivamente di identificare suoni più brevi degli altri, mentre quella ad alto carico richiedeva, oltre alla precedente operazione, di riconoscere anche se quei suoni fossero intonati o non intonati. In questo studio veniva presa come misura della presenza di elaborazione di suoni inattesi la cosiddetta MisMatch Negativity, una componente elettrofisiologica, normalmente elicitata in presenza di devianze rispetto ad una serie costante di suoni. I risultati riguardo all'insorgenza dei suoni più brevi mostrano una attivazione minore in condizioni di alto carico, interpretabile come un minor processamento dei suoni devianti e in accordo, quindi, con la teoria del carico. Anche in questo caso, però un possibile elemento confondente è dato dal fatto che i due compiti erano qualitativamente diversi e le risorse richieste per completarli attingono a meccanismi diversi. È quindi possibile che, rispetto alla condizione di basso carico, in quella con doppio compito la performance sulla rilevazione della durata sia stata peggiore perché non era solo quella la priorità nell'analisi. Anche in questo caso, quindi, non si poteva asserire con certezza che i suoni irrilevanti avevano ricevuto una minore elaborazione.

Altri studi hanno adottato un paradigma a doppio compito (Santangelo et al., 2007), uno di rilevamento in posizione centrale e un altro, occasionalmente inserito tra le prove, di discriminazione in posizione laterale con la presentazione di un cue uditivo. Il carico veniva così modulato richiedendo ai partecipanti di svolgere soltanto la discriminazione oppure entrambi i compiti. I risultati mostravano un ridotto effetto del cue, quando i partecipanti dovevano svolgere entrambi i compiti, e questo è stato interpretato dagli autori come un indice dell'esaurimento della capacità percettiva dato dal compito centrale, che non lasciava risorse residue per completare al meglio anche il compito di discriminazione uditiva in posizioni laterali.

Anche l'intervallo di tempo tra gli stimoli (ISI) è stato manipolato per capire se potesse portare effetti compatibili con la teoria del carico (Gomes et al., 2008; Neelon et al., 2011) con risultati non omogenei.

Considerando l'argomento nel suo insieme, l'applicabilità della teoria del carico percettivo alla modalità uditiva è ancora incerta. Probabilmente alcune differenze sostanziali tra vista e udito nella elaborazione degli stimoli potrebbero essere parte della spiegazione di questi risultati inconsistenti. Alcune evidenze (Azzopardi & Cowey, 1993) mostrano infatti come la vista abbia un meccanismo di selezione spaziale molto forte, che permette una rappresentazione molto dettagliata e focalizzata della porzione spaziale rilevante per il compito. L'udito, invece, sembra non altrettanto in grado di focalizzare la percezione su un'area ristretta, al contrario esso opera costantemente una segregazione e integrazione di tutti i flussi sonori in ingresso, fornendo poi all'attenzione il materiale su cui concentrare risorse (Bregman, 1990; Shinn-Cunningham, 2008). Sembra quindi probabile che il sistema uditivo non sia in grado di allocare tutte le risorse percettive sul compito da svolgere, anche se esso è pesante a livello di elaborazione richiesta, ma resti sempre una quota di risorse occupata nel rilevamento di basso livello. Questo aspetto potrebbe avvalorare l'idea dell'udito come il principale sistema di allertamento rapido, proprio per questa sua costante attività di monitoraggio su tutte le direzioni, che lo distingue dagli altri sensi. Come avremo modo di vedere più avanti, anche i risultati del presente lavoro di ricerca potrebbero essere interpretati in tal senso.

La questione del rapporto reciproco tra elaborazioni di basso e di alto livello non cessa di attirare nuovi sforzi di ricerca. Infatti, uno studio molto recente (Symonds et al., 2020) si è posto proprio l'obiettivo di indagare come viene distribuito l'ammontare totale delle risorse percettive durante l'analisi della scena uditiva. In questo studio ERP, il carico è stato manipolato attraverso il numero dei compiti da completare. Uno di essi era una discriminazione uditiva (il più gravoso), l'altro era una rilevazione visiva durante una breve clip. In entrambe le condizioni di carico, sono state elicitate le componenti relative al processamento degli stimoli non attesi, portando gli autori a concludere che la differenza di carico non ha influito sui risultati osservati. Inoltre, l'attivazione sulla componente P3b sembra suggerire un processo di ricerca del target a due stadi successivi, nel quale alcune caratteristiche degli stimoli che aiutano la segregazione, permettono di liberare risorse per

un'analisi di più alto livello. Si rafforza, quindi, l'ipotesi di interazione tra processi di basso e alto livello proprio durante la formazione dei flussi, ma anche l'idea di una distribuzione omogenea dell'attenzione sull'intera scena uditiva. Il carico generato dai flussi a livello basso viene bilanciato dai processi attentivi volontari di alto livello e c'è una continua occupazione e liberazione di risorse dovuta a questa interazione; le risorse libere vengono poi impiegate su tutta la scena uditiva. anche il lavoro di Fiebelkorn e Kastner (2019) approfondisce questa costante mappatura dell'ambiente circostante proponendo che i sistemi sensoriali scansionino nuovamente tutti gli stimoli con una frequenza tipica delle onde teta (dai 3 agli 8 Hz) fornendo una mappa aggiornata e contribuendo a stabilire e aggiornare il *peso specifico* di ciascuno stimolo, in ragione degli obiettivi che i sistemi di alto livello considerano volontariamente prioritari in quel dato istante. In questa attività di scansione *ritmica* dell'ambiente, anche le singole porzioni di spazio possono ricevere alternativamente un maggiore o minore peso, che si traducono poi in una maggiore elaborazione o in uno spostamento di attenzione verso altre posizioni.

A supporto della teoria del carico cognitivo e della inibizione agli stadi tardivi, viene portata l'evidenza che il segnale audio inatteso (dipendente dal tempo) è stato tracciato comunque nella memoria di lavoro, quindi sembra che tutto il segnale venga processato indistintamente (sia esso target e non target). Più i suoni sono diversi (in altezza, intensità, timbro, spazio) più è facile segregarli. Una segregazione poco efficace, come afferma la teoria del carico, richiama molto carico attentivo, e se il carico è impiegato sulla segregazione, non rimane molto a disposizione per l'analisi di alto livello su caratteristiche più fini. Infine, gli autori suggeriscono il pensiero che l'attenzione venga allocata in modo preferenziale sull'intero compito e che la quantità di risorse rimanenti, allocate sui suoni irrilevanti, può essere alterata da fattori legati allo stimolo (come la chiarezza del suo flusso sonoro in relazione agli stimoli concorrenti).

Come l'essere umano gestisce e risolve quotidianamente il problema degli stimoli uditivi concorrenti - siano essi appartenenti allo stesso campo semantico o meno provenienti da diverse direzioni nello spazio circostante - è tutt'oggi una questione

aperta. Alcune recenti evidenze hanno mostrato l'esistenza di specifici loci e relative componenti neurali che sembrano attivarsi quando vi è la necessità di risolvere il *problema cocktail party* (O'Sullivan et al., 2015; Zhang et al., 2019; Liu et al., 2024). Ciononostante, i risultati sono ancora frammentari, e necessitano di un ulteriore sforzo di integrazione e di ampliamento dei campioni sperimentali per poter guadagnare in solidità. Ad ogni modo, essi hanno il pregio indiscutibile di contribuire ad approfondire questo dilemma che, ad ogni nuovo studio, fa affiorare nuovi particolari.

Dal quadro di evidenze appena descritto, emergono alcuni aspetti sulla dimensione spaziale dell'udito e dell'attenzione uditiva, che hanno costituito il perimetro di partenza del presente lavoro di ricerca sperimentale. La ricerca sull'attenzione che, come abbiamo visto, partì prevalentemente dall'udito per muovere i suoi primi passi, raggiunse però la sua maturità negli studi sulla modalità visiva, a livello metodologico, di varietà e complessità dei paradigmi sperimentali, di accuratezza nelle misure e di coerenza e consequenzialità nella discussione dei risultati. Quando poi, in tempi più recenti, l'ambito uditivo riacquistò un po' del suo originario fascino, si è cercato in prima battuta di traslare su di esso il bagaglio scientifico acquisito precedentemente in ambito visivo. L'inconsistenza e frammentarietà delle evidenze disponibili, se da un lato ha fatto sorgere non poche perplessità circa i presupposti fondamentali di quella traslazione, dall'altro ha dato nuova spinta alla ricerca sull'udito, sui suoi rapporti con la vista, sui suoi specifici meccanismi e sulla sua funzione per la specie umana.

## II - L'attenzione spaziale uditiva

Come già evidenziato in precedenza, la collocazione degli stimoli nello spazio circostante, siano essi visivi, tattili o uditivi, è un dato molto importante ai fini dell'analisi dell'ambiente che ci circonda. Dopo aver affrontato la questione del *rilevamento* della posizione dei suoni che ci circondano, a che livello avviene e attraverso quali meccanismi, i paragrafi seguenti intendono indagare i meccanismi alla base della *selezione* di un certo spazio, operata dall'attenzione. Perché e come avviene questa selezione? È questo un meccanismo sopra-modale o modalità-specifico? Come fare a catturarne l'azione in contesti sperimentali?

Dalla definizione di carattere generale di attenzione, proposta in precedenza, discende che uno degli *aspetti* sui quali possono essere focalizzate le risorse cognitive è rappresentato dallo *spazio*. La collocazione spaziale degli elementi di un determinato ambiente costituisce quindi uno dei possibili criteri di selezione che un individuo opera sulla totalità degli stimoli che lo circonda. Viene così a delinearsi una particolare tipologia di attenzione: *l'attenzione spaziale*. Essa è definita come la capacità di selezionare una determinata porzione dello spazio circostante e di elaborarne prioritariamente gli stimoli (Posner, 1980; Carrasco, 2011; Van Ede et al., 2012).

Si potrebbe pensare all'attenzione spaziale come all'obiettivo di una fotocamera, attraverso il quale il fotografo sceglie se e quanto aumentare o ridurre l'area da includere nel suo scatto, in funzione di ciò che egli ritiene importante in quella scena (Eriksen & Yeh, 1985). In tal caso, ciò che ricade fuori dal campo dell'obiettivo risulta escluso dalla fotografia, e il livello di dettaglio sarà tanto maggiore quanto minore è l'area ripresa dall'obiettivo.

Un'altra visione proposta per descrivere la medesima capacità (Pinker & Downing, 1985) mutua il concetto matematico di gradiente e lo applica al caso specifico, postulando che l'efficienza di elaborazione delle informazioni sia massima nel

centro ideale della porzione di spazio selezionata, e decresca all'aumentare della distanza da esso.

In ogni caso, numerose evidenze concordano sul fatto che l'ampiezza spaziale della selezione abbia un impatto non solo sulla efficacia di elaborazione delle informazioni, ma anche sui meccanismi neurali attraverso i quali esse vengono elaborate (Luck, 1995; Mangun, 1995 per la modalità visiva). Secondo questa ipotesi, a livello neurologico l'individuo mette in atto processi diversi: una selezione spaziale ampia, a causa della maggiore quantità di dettagli presenti, implicherebbe un meccanismo orientato più alla globalità; al contrario una selezione spaziale ristretta sarebbe accompagnata da un meccanismo orientato più ai singoli dettagli. Appunto nella modalità visiva si parla di attenzione *diffusa* nel caso di una selezione spaziale ampia, e di attenzione *focale* in presenza di una selezione spaziale ristretta (Wright & Ward, 2008).

La selezione spaziale può avvenire in modi diversi in base alla volontarietà o involontarietà dell'allocazione di attenzione (Prinzmetal et al., 2009; Bowen et al., 2023). Infatti, si può dirigere la nostra attenzione volontariamente verso una porzione di spazio, poiché essa è ritenuta rilevante ai fini del completamento dell'azione prefissata o perché soggetta ad una certa aspettativa precedentemente originata. Si può dare, però, anche il caso in cui la nostra attenzione sia involontariamente catturata da un oggetto situato in una certa porzione di spazio: la si distingue dalla precedente per il fatto che non è orientata *verso* un obiettivo ma *da* uno stimolo esterno o una sua caratteristica saliente. Un esempio di queste due diverse modalità è rappresentato da una situazione di guida di un veicolo: l'attenzione volontaria seleziona una specifica porzione dello spazio anteriore, poiché essa fornisce informazioni determinanti sulla direzione da tenere e, di conseguenza, orienta tutte le azioni necessarie a mantenere la direzione desiderata; al contrario l'attenzione involontaria può essere catturata da un forte suono in arrivo da posizioni laterali o posteriori, che contribuisce a riorientare la selezione spaziale precedente, per ricomprendere l'area dalla quale è stato originato lo stimolo.

L'attenzione spaziale permette di selezionare informazione proveniente da modalità sensoriali diverse (Macaluso et al., 2000; Zuanazzi & Noppeney, 2020). Se i primi studi consideravano singolarmente ciascuna delle modalità e ipotizzavano quindi modelli unimodali di orientamento dell'attenzione separatamente per vista, udito e tatto; a partire dall'ultima decade del secolo scorso è cresciuto l'interesse verso i meccanismi intermodali che si instaurano al momento della selezione spaziale (Driver & Spence, 1998). come vedremo più avanti nello svolgersi del capitolo, sebbene sussistano specificità unimodali nella rappresentazione dello spazio circostante, sono tuttavia evidenti anche meccanismi sopra-modalità che, attraverso il dato spaziale, integrano informazioni provenienti dalle singole modalità.

Le condizioni fisiche di partenza per la selezione spaziale non sono infatti identiche per i diversi sensi. La vista, ad esempio, ha già una limitazione spaziale intrinseca costituita dal campo visivo, che in condizioni normali di visione binoculare misura circa 120° di ampiezza orizzontale sovrapposta per i due occhi, con circa 30° periferici a destra e a sinistra, di competenza del solo occhio interessato (Strasburger & Poppel, 2002). All'interno di questa porzione spaziale, che ha anche limiti sul piano verticale, la vista può operare infinite selezioni diverse, tuttavia, se l'area rilevante per l'azione da compiere ricade al di fuori di questi limiti, è necessario compiere l'azione meccanica di muovere il capo. Sul piano orizzontale, quindi, circa la metà dei 360° intorno al soggetto risultano automaticamente inibiti, senza alcuno sforzo volontario. Inoltre, la vista può operare una ulteriore inibizione meccanica, costituita dalla chiusura di uno o di entrambi gli occhi.

Lo stesso non può dirsi dell'udito, che per sua natura è costantemente e inevitabilmente panoramico in condizioni tipiche (Schafer, 1977; Scharf, 1998). Ogni selezione di una limitata porzione di spazio dovrà quindi necessariamente avvenire ad un livello più alto (Styles, 2005). Sia per questa sua costante apertura a 360°, sia per il suo ruolo in ottica evolucionistica (Fairnie et al., 2016), l'udito non ha la stessa capacità della vista di restringere la selezione spaziale ad ampiezze molto piccole (Murphy et al., 2017).

Viene così a costituirsi il concetto di *attenzione spaziale uditiva*, che può quindi essere definita come la capacità, volontaria o involontaria, di focalizzare la percezione uditiva verso una specifica porzione di spazio.

La natura e le caratteristiche specifiche dell'udito descritte sopra hanno reso lo studio dell'attenzione spaziale uditiva un tema di ricerca a sé stante che, dai primi studi fino alle più recenti evidenze, non cessa di esercitare attrazione su un numero crescente di studiosi.

## **2.1. Il paradigma di suggerimento spaziale**

Per poter dare qualche indicazione sul *quando* avviene una selezione spaziale, risulta ancora necessario partire da studi sulla modalità visiva, che per primi sollevarono la questione sul perché avvenga una selezione spaziale da parte dell'attenzione. Dal momento che tutti i compiti richiedevano di fare attenzione ad un oggetto target, ci si chiedeva sostanzialmente se l'attenzione fosse diretta alla posizione specifica o all'oggetto specifico (Styles, 2005). Se si accetta che l'attenzione sia basata sull'oggetto, e che quindi la sua posizione venga in qualche modo codificata come una sua caratteristica, cosa accade nell'intorno di quella posizione? Quale tipo di elaborazione ricevono gli stimoli che ricadono in prossimità dello stimolo target? A questa domanda hanno provato a rispondere alcuni studiosi (Eriksen & Eriksen, 1974; Tipper, 1985; Tipper & Cranston, 1985; Tipper & Driver, 1988) che mostrarono come gli stimoli non-target presentati in posizioni molto vicine al target ricevevano comunque una quota di attenzione, evidenziata dal fatto che i partecipanti riuscivano a ricordarne alcune caratteristiche, al contrario degli stimoli non-target in posizioni lontane. Inoltre, la selezione risultava più difficile tra stimoli vicini (come già visto negli studi di ascolto dicotico).

Un ulteriore passo avanti nello studio dell'attenzione spaziale fu possibile grazie agli studi di M. Posner (1971; 1978, 1980). Egli introdusse un nuovo paradigma sperimentale per studiare come si sposta l'attenzione spaziale nella modalità visiva, che prese il nome di *suggerimento spaziale* (molto più noto nella sua originale

definizione inglese di *spatial cuing*). Una delle novità fondamentali introdotte da questa metodologia è da ricercare proprio in questo diverso punto di messa a fuoco, per cui la selezione di porzioni dello spazio, operata dall'attenzione, assume una centralità che prima non aveva.

Concepito originariamente dall'autore e dai suoi colleghi per lo studio sulla modalità visiva, l'esperimento classico prevede di posizionare uno schermo davanti ai partecipanti, muniti di una pulsantiera per rispondere agli stimoli che verranno presentati. Variando e controllando alcuni parametri sperimentali, è possibile studiare, con lo stesso paradigma, diversi aspetti dell'attenzione. Infatti, lasciando ai partecipanti libertà di movimento oculare durante lo svolgimento del compito, si andrà a studiare l'*attenzione overt o esplicita*, ovvero accompagnata da un comportamento osservabile (appunto il movimento degli occhi, o *saccade*, verso gli stimoli target); se invece si richiederà ai partecipanti di mantenere fisso lo sguardo in posizione centrale, ciò che verrà elicitato sarà l'*attenzione covert o implicita*, ovvero esclusivamente mentale e non riscontrabile attraverso comportamenti esteriori. Similmente, se il suggerimento spaziale viene presentato in posizione centrale, si parla di *attenzione endogena*, ovvero di un orientamento volontario in conseguenza dell'informazione ricevuta; se invece il suggerimento viene presentato già in una delle due possibili posizioni laterali, si parla di *attenzione esogena*, cioè orientata dall'esterno a causa delle caratteristiche salienti di un certo stimolo.

Nel caso specifico in cui l'interesse sia quello di studiare l'attenzione endogena e implicita, lo schermo è idealmente suddiviso in tre zone: la zona centrale che servirà come punto di partenza e per la presentazione dei suggerimenti (cue); quelle laterali (destra e sinistra) che invece serviranno per la presentazione degli stimoli target (tipicamente racchiusi in box che ne delimitano il perimetro di pertinenza). In ogni singola prova di questo esperimento, il partecipante è inizialmente istruito a fissare un punto centrale dello schermo (sul quale viene presentato un simbolo, tipicamente un “+” per facilitare il mantenimento dello sguardo in quella posizione). Viene inoltre richiesto di non effettuare movimenti oculari per tutta la durata del compito.

In quella stessa porzione centrale di schermo viene poi presentato il suggerimento o *cue* (spesso, ma non sempre, costituito da una freccia) che indica con buona probabilità la direzione dalla quale verrà presentato lo stimolo target, sul quale i partecipanti dovranno completare un compito. Il cue, successivamente, scompare e, dopo un breve intervallo di tempo, viene presentato il target in una delle due aree laterali. Il target può essere una forma geometrica, un oggetto, una lettera dell'alfabeto così come una parola. A questo punto, nel caso di un compito di semplice detezione, i partecipanti dovranno rispondere all'apparizione del target premendo un tasto nel più breve tempo possibile. Invece, nel caso di un compito di discriminazione, la richiesta sarà quella di rispondere con il tasto corrispondente al target effettivamente presentato in quella singola prova.

Se il cue indica e anticipa correttamente la posizione del target, si parla di *prova valida*, se invece indica una posizione diversa rispetto a quella in cui verrà presentato il target, si parla di *prova invalida*. Infine, può darsi il caso in cui il cue indichi entrambe le posizioni possibili, allora si parla di *prova neutrale* o semplicemente *neutra*. Il flusso degli eventi della singola prova, sia essa valida, invalida o neutra, è rappresentato nella figura 1.

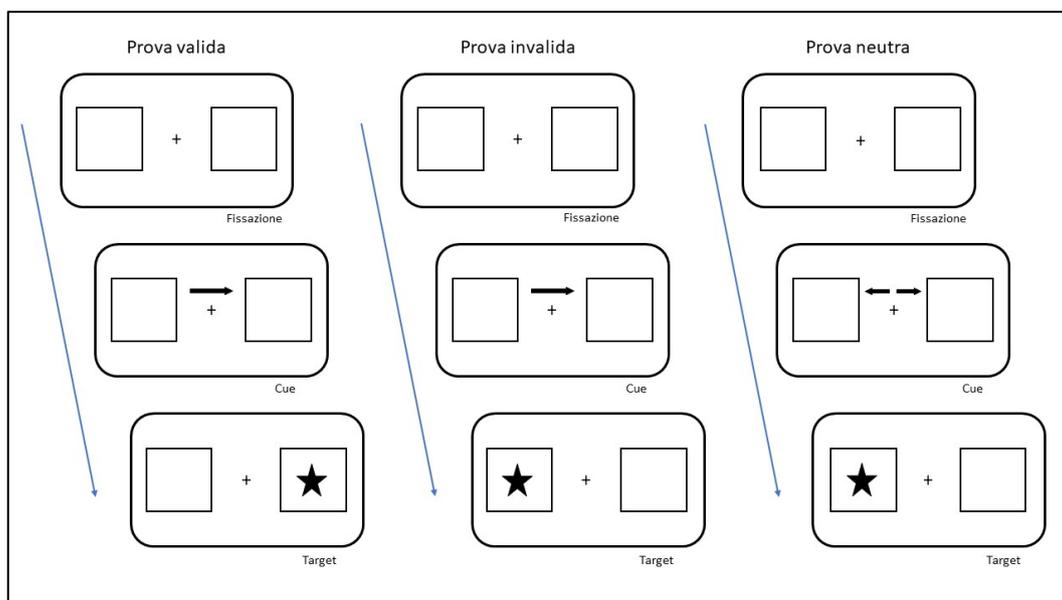


Figura 1: Flusso degli eventi di singole prove in un compito di tipo Posner

Le variabili indipendenti, manipolate dallo sperimentatore, sono quindi la direzione del cue e la posizione del target; mentre le variabili dipendenti, misurate per studiare l'attenzione, sono i tempi di risposta (spesso abbreviati con RT ed espressi solitamente in millisecondi) e l'accuratezza delle risposte (espressi in percentuale di risposte corrette o, alternativamente, in percentuale di errori).

Per quanto riguarda i tempi di risposta, solitamente si osservano risposte mediamente più rapide per le prove valide, valori intermedi per le prove neutre e risposte più lente per le prove invalide. Questo andamento, secondo un'interpretazione oramai consolidata, indica una facilitazione dovuta alla conoscenza preventiva della posizione del target: l'attenzione ha quindi selezionato in anticipo una determinata porzione di spazio, e questa selezione ha costituito un *beneficio* al momento di completare il compito richiesto. Nelle prove invalide, invece, i tempi di risposta più alti indicano un *costo* attentivo, dovuto alla necessità di riorientare l'attenzione verso l'effettiva posizione del target, in quanto la direzione del cue non era corretta. Il partecipante ha dovuto quindi spostare la propria attenzione in una nuova posizione, per poi svolgere il compito, e i tempi di risposta più lenti sono proprio la conseguenza di questo spostamento. Le prove neutre, infine, rappresentano una situazione di riferimento ideale, nella quale il partecipante non opera nessuna selezione spaziale e quindi deve esercitare un controllo su tutta l'area potenzialmente interessata dal compito. Si suppone, quindi, che in questa condizione la performance rilevata non sia ascrivibile ad un intervento dell'attenzione volontaria teso a concentrare le proprie risorse in una porzione precisa di spazio. Normalmente, infatti, sia i tempi di risposta che le accuratèzze percentuali sulle prove neutre sono in posizione intermedia tra prove valide e prove invalide, tale che il beneficio attentivo risulta dalla differenza di performance tra le prove neutre e le prove valide, mentre il costo attentivo risulta dalla differenza tra le prove invalide e le prove neutre.

Un'altra misura molto interessante è rappresentata dalla differenza tra i tempi di risposta delle prove invalide e quelli delle prove valide, che prende il nome di *effetto attentivo*: essa infatti fornisce un'indicazione quantitativa complessiva degli effetti

che la selezione spaziale operata (corretta o non corretta) provoca sul tempo necessario a svolgere un compito.

Queste grandezze non sono necessariamente correlate tra di loro: infatti si può dare il caso di un esiguo beneficio e contemporaneamente di un significativo costo attentivo; così come può avvenire che i benefici superino i costi o che entrambi i valori siano simili e piccoli oppure ancora simili ma grandi. Ognuna di queste situazioni ha poi una o più interpretazioni possibili, tenuto conto della variabilità individuale che può avere un impatto talvolta significativo sull'andamento osservato. Un esempio di andamento che può essere considerato tipico per questo paradigma sperimentale è descritto nel grafico in figura 2, i cui valori sono puramente indicativi.

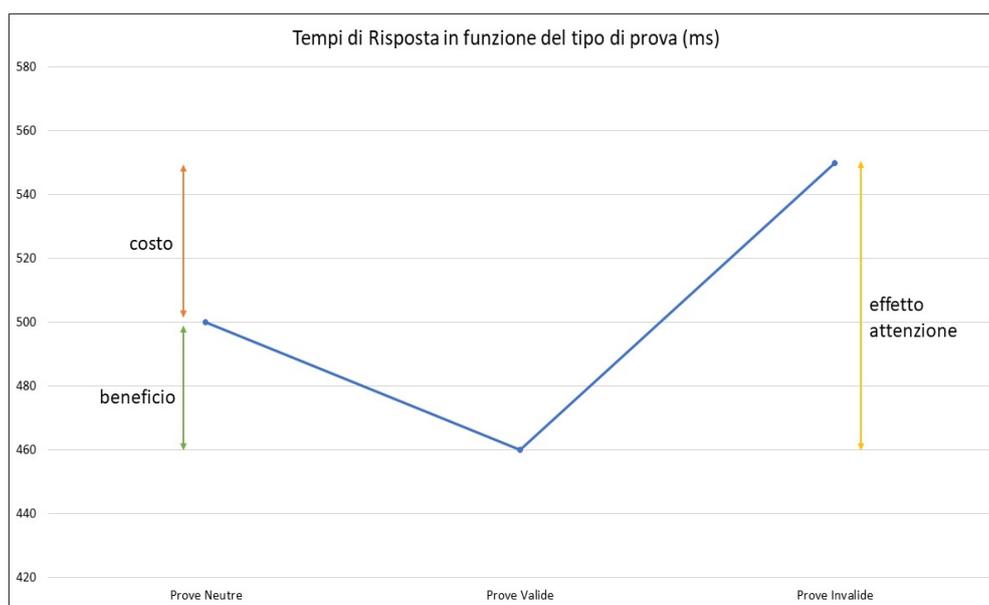


Figura 2: Andamento dei tempi di risposta in funzione del tipo di prova

Vi sono, inoltre, alcuni parametri sperimentali che possono essere manipolati al fine di valutarne l'effetto sull'andamento generale appena descritto. Uno di questi è la percentuale delle prove valide sul totale delle prove somministrate, che originariamente Posner settò all'80%. La variazione del rapporto tra prove valide e invalide, come è stato dimostrato (Arjona et al., 2016; Lou et al., 2022), può avere

un effetto sui risultati dell'esperimento: infatti sembra che all'aumentare della percentuale di prove valide, aumenti anche l'effetto attentivo complessivo e che quindi sia ipotizzabile un costante adattamento neurale - soprattutto per quanto riguarda l'attenzione e l'aggiornamento della memoria di lavoro – teso a fronteggiare l'imprevedibilità degli eventi esterni. Un altro parametro di grande influenza è l'intervallo di tempo che intercorre tra la presentazione del cue e quella del target, convenzionalmente definito *SOA* (Stimulus Onset Asynchrony). Già Posner e Cohen cominciarono a studiarne l'effetto sull'andamento canonico dei tempi di risposta (1984), e anche più recentemente alcuni autori (Benso et al., 1998; Green & Woldorff, 2012) hanno dimostrato che con intervalli minimi si ha un effetto attentivo complessivo molto basso, che poi inizialmente cresce di pari passo con l'aumento del SOA, fino ad attestarsi su valori massimi per intervalli cue-target di durata intermedia. Se si aumenta ulteriormente il SOA, l'effetto attentivo torna a diminuire. Gli autori hanno interpretato i dati ottenuti affermando che, con intervalli troppo brevi tra cue e target, il punto di messa a fuoco dell'attenzione rimanga quello iniziale poiché non vi sarebbe il tempo necessario per spostarsi nella posizione indicata. Vi è poi una rosa di valori intermedi, per i quali il soggetto riesce effettivamente a mettere a frutto l'informazione ricevuta dal cue e ne trae un beneficio (nel caso di prove valide) o un costo (nel caso di prove invalide). Infine si ipotizza che il nuovo calo dell'effetto attentivo con SOA lunghi sia dovuto al superamento di una soglia temporale, oltre la quale l'attenzione si dispiegherebbe nuovamente in modo diffuso su tutta la scena, vanificando il vantaggio derivante dalla selezione spaziale operata precedentemente grazie al cue. Questa dinamica si osserva anche sui tempi di risposta mediamente osservati in questi compiti, che con SOA assente o minimo tendono ad essere mediamente lenti, per poi divenire via via più veloci con valori crescenti di SOA, fino ad un punto di inversione di tendenza, dopo il quale tornano a crescere.

Negli anni che seguirono la pubblicazione dei primi lavori di Posner e colleghi, ci fu grande interesse attorno al nuovo paradigma sperimentale da loro introdotto, e furono fatti molti tentativi di espansione, in diverse direzioni di ricerca; ancora oggi – e il presente lavoro ne è una prova – esso non cessa di essere allo stesso tempo un

punto di riferimento e un punto di partenza di nuove indagini per coloro che sono interessati a studiare l'attenzione spaziale, sia essa visiva, uditiva, tattile o multimodale. Una possibile variante fu ideata da Rizzolatti e colleghi (1987) che si proposero di mettere in questione alcune interpretazioni allora dominanti circa i risultati tipicamente osservati nel compito di suggerimento spaziale. In particolare, essi si chiedevano se i costi attentivi dovuti allo spostamento da una posizione all'altra dello schermo fossero effettivamente dovuti, come si pensava, ad un effetto di inibizione agente sull'emicampo opposto a quello su cui si porta l'attenzione (Hughes & Zimba, 1985). Un'altra domanda di ricerca riguardava lo stesso concetto di *movimento dell'attenzione*: se cioè ci fosse o meno qualcosa che, effettivamente, si muoveva a livello senso-motorio oppure a livello più alto. E anche ammettendo di aver individuato la sede di questo movimento, restava da capire secondo quali leggi matematiche avvenisse, se in modo costante (Remington & Pierce, 1984) o se con un certo gradiente di velocità (Pinker & Downing, 1985).

Per rispondere a tutte queste domande di ricerca, Rizzolatti e colleghi congegnarono una variante del compito Posner, nella quale non c'erano più soltanto due possibili posizioni in cui presentare il target. Il setting prevedeva infatti quattro posizioni target, arrangiate in quattro modi diversi per i quattro esperimenti somministrati. Nel primo esperimento, le quattro posizioni erano disposte in orizzontale nella parte alta dello schermo, nel secondo invece erano disposte sempre orizzontalmente ma nella parte bassa; nel terzo erano disposte in verticale e tutte sulla destra; nel quarto infine tutte sulla sinistra e sempre in verticale. Queste disposizioni avrebbero permesso ai ricercatori di indagare gli effetti dell'attraversamento del meridiano orizzontale e verticale, oltre che raccogliere dati sui costi attentivi in funzione della distanza tra posizione attesa e posizione reale del target. Il 70% delle prove erano valide, ed erano previste anche prove neutre. I risultati mostrarono che, anche con più posizioni, vi era sempre un beneficio sui tempi di risposta e sulle accuratezze, dovuto all'informazione preventiva sulla posizione cui prestare attenzione. Anche il costo dovuto al riorientamento dell'attenzione su posizioni non attese era evidente e significativo. Esso, inoltre, cresceva all'aumentare della distanza tra posizione attesa e posizione reale del target, e si osservava un costo maggiore nel momento

in cui le due posizioni richiedevano di attraversare i meridiani e portarsi negli emicampi opposti.

I dati riportati suggeriscono agli autori di scartare sia l'ipotesi di un movimento dell'attenzione indipendente dalla distanza da coprire, sia l'ipotesi di una inibizione completa dell'emicampo opposto. Inoltre, anche l'assunzione di una velocità costante dell'attenzione non risultava supportata dai dati, che invece facevano pensare più ad un gradiente di velocità. Restava da capire a cosa fosse dovuto l'effetto dell'attraversamento dei meridiani e, più in generale, in cosa consistesse il movimento dell'attenzione. Gli autori, ispirandosi ad alcune ipotesi provenienti da studi precedenti (Rosenbaum, 1980; Rosenbaum & Kornblum, 1982; Rosenbaum et al., 1984), proposero che l'attenzione si sposta quando vengono programmati dei movimenti oculomotori per portare la vista sulla posizione del target. L'idea era che, in questa programmazione, la precedenza venisse data alla direzione, e in seguito si codificasse la distanza da coprire. Da questo assunto discende anche una possibile spiegazione dei costi attentivi dovuti allo spostamento in altra posizione: infatti, questo costo sarebbe dovuto alla cancellazione di un programma oculomotore e la conseguente riprogrammazione di un nuovo movimento verso la posizione reale del target.

Si fa strada, quindi, l'idea di un forte nesso tra l'attenzione e il senso della vista, che riceveva alcune conferme empiriche dall'osservazione che il comportamento tipico di chi vuole portare la sua attenzione su una determinata posizione è quello di muovere il capo, o almeno gli occhi, verso quella precisa posizione. Come vedremo in seguito, questo impianto teorico riceverà alcune conferme anche a livello neurologico, oltre che comportamentale.

## **2.2. Applicazioni e sviluppi nella modalità uditiva**

In questo paragrafo saranno riassunte alcune tra le più importanti evidenze riguardo all'applicazione del paradigma di Posner alla modalità uditiva, in studi unimodali. Le evidenze per la modalità visiva resteranno un costante punto di riferimento e un

termine di paragone che non di rado ispirerà domande di ricerca, disegni sperimentali, discussioni e conclusioni anche sull'attenzione uditiva.

Alcuni tentativi non portarono gli stessi risultati osservati per la modalità visiva, infatti in questi studi non si osservava alcuna facilitazione derivante dall'informazione spaziale ricevuta attraverso un cue uditivo. Ad esempio, Buchtel e Butter (1988) studiarono diverse combinazioni unimodali e multimodali su un esperimento di suggerimento spaziale con due posizioni alternative, a sinistra e a destra della linea mediana del corpo dei partecipanti. Nell'esperimento puramente uditivo, un cue, presentato alternativamente a sinistra, a destra o in entrambe le posizioni, precedeva la presentazione di un target uditivo, e ai partecipanti era richiesto di rilevare il suono target, rilasciando un pulsante. Nell'80% dei casi il cue era valido e anticipava correttamente la posizione del target. I risultati mostrarono che i tempi di risposta non dipendevano dalla validità della prova, ed erano più veloci nelle prove neutre. I risultati portarono gli autori a sostenere che, senza la necessità di orientare lo sguardo (in modo esplicito o implicito) non vi potesse essere uno spostamento attentivo apprezzabile. Questo poteva spiegare perché fossero stati riportati effetti attentivi con cue visivi, al contrario dei cue uditivi. Il risultato replicava quanto già evidenziato in un esperimento dello stesso Posner (1978). Negli stessi anni, Mazzucchi e colleghi (1983) trovarono un effetto attentivo sui tempi di risposta a stimoli uditivi con cue uditivi, sebbene di modesta entità e con presentazione da un solo lato anziché due.

In quegli stessi anni, Rhodes (1987) portò a termine tre esperimenti, nei quali dispose alcuni altoparlanti a copertura di tre diverse aree attorno al soggetto. In un caso, gli altoparlanti furono posizionati a copertura dell'intero spazio circostante il soggetto. Il compito prevedeva di presentare uno stimolo uditivo, costituito da un tono puro, su uno degli altoparlanti disponibili; ai partecipanti era richiesto di nominare il numero dell'altoparlante sul quale veniva presentato lo stimolo sonoro e di mantenere la propria attenzione su quella posizione, dal momento che una ripetizione sullo stesso altoparlante era più probabile rispetto ad un cambio di altoparlante (con percentuali di probabilità dal 29% al 40% a seconda degli

esperimenti). Anche in queste condizioni, che differiscono dall'originale paradigma di Posner, i risultati indicavano una facilitazione dovuta all'informazione spaziale data dal cue uditivo. Invece, nelle prove con spostamento attentivo si assisteva ad un rallentamento dei tempi di risposta che appariva direttamente proporzionale alla distanza idealmente percorsa dall'attenzione fino ad un valore soglia oltre il quale i tempi di risposta si assestavano. Questo particolare andamento, osservato in particolare nell'esperimento con altoparlanti distribuiti tutto attorno ai partecipanti, portò la Rhodes a pensare che l'attenzione si spostasse in modo continuo e lineare fino ad un certo valore di spostamento angolare, per poi cambiare del tutto modalità e passare a salti discreti di spazio. Sebbene questo studio abbia rappresentato un importante passo avanti nella comprensione del fenomeno, restavano però alcuni assunti che non trovavano dimostrazione in quanto osservato. Infatti, non si poteva provare che i suoni, fungenti in questo caso da cue periferici per gli stimoli immediatamente successivi, avessero effettivamente catturato l'attenzione dei soggetti (attenzione di tipo esogeno, poiché la probabilità di un secondo stimolo nella stessa posizione non superava il 40%), né che essi avessero mantenuto il loro focus nella posizione su cui si erano portati, né che lo spostamento attentivo dovuto alla presentazione del target in altra posizione fosse avvenuto prima della localizzazione spaziale o in conseguenza di essa. Il compito, infatti, era di semplice localizzazione, e ai partecipanti non era richiesta un'operazione specifica sulla posizione target.

Tutte queste considerazioni portarono Mondor & Zatorre (1995) a proseguire sulla strada tracciata, cercando di dissipare questi ed altri dubbi riguardo alle evidenze fino a quel momento disponibili. I partecipanti erano seduti davanti ad un sistema di cinque altoparlanti, disposti a distanza di 45° l'uno dall'altro, a copertura dei 180° dell'emicampo anteriore. Un iniziale e breve cue costituito da rumore bianco veniva presentato esattamente nella posizione sulla quale i partecipanti erano invitati poi a completare un compito di discriminazione tra un suono puro (artificialmente ottenuto e costituito da un'onda sinusoidale con una determinata frequenza) e un suono complesso (simile ai suoni reali, formato cioè da una serie di suoni puri concomitanti, ciascuno con un peso relativo diverso). Inoltre, per

controllare e assicurarsi di conoscere l'effettiva posizione del focus attentivo iniziale, i ricercatori inserirono anche una sequenza di fissazione, nella quale veniva presentato un suono con intensità variabile su una precisa posizione, e ai partecipanti era richiesto di riconoscere quando la variazione era discendente. Inoltre, essi variarono anche l'intervallo di tempo tra cue e target (SOA) per studiarne l'effetto sui risultati. Insieme a questo primo esperimento, ne furono condotti altri cinque, inserendo altre manipolazioni fra cui anche l'inserimento di prove invalide per studiare la distribuzione del costo attentivo al variare della distanza tra cue e target. Complessivamente, i risultati portarono gli autori a sostenere l'ipotesi di un gradiente di attenzione, centrato sulla posizione attesa, per il quale la precisione di elaborazione è massima e poi decade molto velocemente per distanze brevi e meno velocemente per distanze lunghe. Inoltre, sembra che un tempo maggiore per focalizzare l'attenzione sulla posizione indicata dal cue porti a performance migliori, anche se le spiegazioni di questo fenomeno potrebbero essere più di una. Infatti, il maggior tempo potrebbe portare maggior concentrazione di risorse oppure potrebbe portare ad un restringimento e affinamento della posizione. Le conclusioni di Mondor e Zatorre, benché molto articolate e aperte a possibili sviluppi, costituirono un primo quadro generale di riferimento per i successivi studi sull'attenzione spaziale uditiva.

Lo studio di Rorden & Driver (2001) propone ai partecipanti due compiti di discriminazione uditiva, uno con cue uditivi periferici non predittivi della posizione del target, l'altro con cue centrali simbolici che anticipavano la posizione target inducendo un orientamento endogeno dell'attenzione. Le posizioni target coprivano l'intero emicampo anteriore, con particolare attenzione al confronto tra gli effetti attentivi nella porzione destra e sinistra, nonché ai possibili effetti del transito attraverso il meridiano centrale. Venivano registrati i tempi di risposta al variare delle condizioni sperimentali. In entrambi gli esperimenti, i partecipanti risultavano in grado di portare la propria attenzione sulla posizione desiderata, nel primo caso si trattava di attenzione esogena mentre nel secondo di attenzione endogena. I tempi di risposta erano più rapidi in posizione attesa, mentre aumentavano in maniera lineare all'aumentare della distanza tra posizione attesa ed effettiva del target.

Mentre si confermava l'efficacia di cue uditivi centrali e periferici ai fini dell'orientamento attentivo, l'effetto del meridiano non risultava univoco tra gli esperimenti e tra le condizioni. I dati, infatti, sembrano indicare con chiarezza soltanto gli effetti dell'attenzione al variare della distanza tra posizione attesa ed effettiva del target, lasciando dubbi sulla possibilità di un effetto aggiuntivo dovuto all'attraversamento del meridiano, cosa che invece era stata evidenziata in precedenza per la modalità visiva (Rizzolatti et al., 1987).

Negli anni più recenti, l'oramai consolidato paradigma di suggerimento spaziale di Posner ha conosciuto alcuni ampliamenti e rivisitazioni, in studi sulla modalità visiva, tattile, uditiva e cross-modali, esclusivamente comportamentali o di neuroimmagine, essendo quest'ultima la protagonista di gran parte delle ricerche recenti sull'attenzione spaziale uditiva. I grandi progressi nella comprensione dei meccanismi neurali che sovrintendono all'allocazione di risorse attentive nello spazio uditivo circostante (Teder-Sälejärvi et al., 1999; Fritz et al., 2007; Gamble & Luck, 2011; Golob & Holmes, 2011; Mesgarani & Chang, 2012; Lee et al., 2013; Kong et al., 2014; Lewald & Getzmann, 2015; Karhson et al., 2015; Deng, 2019; Rotaru, 2024; Calcus, 2024 per una recensione aggiornata) hanno lasciato aperte alcune domande di fondo, oltre che dischiudere nuovi orizzonti e conseguentemente nuove questioni che stimolano il dibattito attuale.

Una questione cruciale riguarda il rapporto tra processi bottom-up (di livello sensoriale, guidati dalle caratteristiche salienti degli stimoli) e quelli top-down (di livello cognitivo, guidati dalla volontà e dagli obiettivi prefissati). Benché sia oramai superato il pensiero che essi operino su due piani paralleli e non comunicanti (Driver, 2001; Awh et al., 2012), non si spegne la curiosità dei ricercatori circa i meccanismi che regolano questa oramai accertata interazione, che segue il momento in cui l'attenzione seleziona un oggetto percettivo o una posizione nello spazio (Petersen & Posner, 2012; Jääskeläinen & Ahveninen, 2014; Kaya & Elhilali, 2017) e provoca una alterazione momentanea delle abilità percettive relative alle caratteristiche o alla posizione di interesse (vedi paragrafi precedenti sull'analisi della scena uditiva e il ruolo dell'attenzione).

Inoltre, ancora più specifico e non meno affascinante è l'apparente conflitto che si genera quando l'attenzione deve concentrare risorse su una determinata posizione rilevante per il compito e, contemporaneamente, deve garantire un corretto monitoraggio degli stimoli in ingresso da ogni direzione. Questo *dilemma stabilità/flessibilità* (Liljenström, 2003) ispirerà i più recenti lavori sull'attenzione spaziale uditiva, che costituiscono attualmente lo stato dell'arte e il riferimento di partenza per il presente lavoro di ricerca.

Un interessante contributo, che amplia gli orizzonti della comprensione del paradosso appena menzionato, arriva da uno studio di Golob e colleghi (2016), ripreso e ampliato da Scheuerman e colleghi (2018). Per la prima volta, infatti, si assiste al tentativo di elaborare un modello teorico di calcolo che possa descrivere in modo affidabile il modo in cui l'attenzione uditiva si distribuisce nello spazio circostante. Il modello si ispira all'approccio basato sui vincoli, utilizzato nelle applicazioni dell'intelligenza artificiale. Il modello, basato su evidenze comportamentali, assume come vincolo principale la limitatezza delle risorse che l'attenzione volontaria può allocare in un determinato tempo, e immagina che entrambi gli stadi di elaborazione (top-down e bottom-up) costruiscano costantemente una loro propria *mappa* descrittiva. La *mappa di obiettivi* stabilisce le priorità e alloca risorse in ragione degli obiettivi e delle richieste dei compiti da svolgere. La *mappa di salienza*, nel frattempo, si occupa di monitorare in parallelo tutti gli stimoli circostanti e, quando necessario, ha la capacità di orientare verso uno stimolo spostando momentaneamente il focus attentivo di alto livello, che era finalizzato al completamento di un compito preciso. La mappa di salienza è costituita da tutti quegli elementi che, come già precedentemente evidenziato (vedasi, in proposito, anche Coensel & Botteldooren, 2010), possono catturare la nostra attenzione: frequenza, intensità, timbro, posizione, familiarità e significato del contenuto. Tutti i possibili contributi di queste caratteristiche vengono qui integrati nella suddetta mappa, e si sposta l'accento sul modo in cui i due modelli e le due mappe portano ai risultati comportamentali osservati negli esperimenti.

Il setting sperimentale dello studio prevedeva l'utilizzo di cuffie, attraverso le quali venivano presentati due possibili stimoli uditivi costituiti da rumore bianco, modulato in ampiezza a due diverse frequenze (25 e 75 Hz). Le posizioni virtuali erano cinque, scalate di 45° di azimut a copertura dell'intero emicampo anteriore. Ai partecipanti era richiesto di discriminare quale dei due stimoli fosse stato presentato in ogni prova. All'inizio di ogni blocco, venivano istruiti sulla posizione alla quale prestare attenzione per tutta la durata del blocco. Erano previste quindi prove valide (84% di probabilità) e prove invalide (16% di probabilità, distribuita sulle quattro posizioni non-target). I dati dei tempi di risposta vengono poi analizzati introducendo una terza mappa, chiamata *mappa di priorità*, risultante dalla somma pesata dei contributi delle due mappe precedenti, e rappresentativa dell'andamento dei risultati comportamentali in ogni posizione spaziale prevista dall'esperimento. I risultati indicano, come già in studi precedenti, che le prove valide elicitano risposte più veloci, e che vi è un costo di spostamento dell'attenzione nel caso di prove invalide. L'andamento di questi costi segue una legge non lineare, con un gradiente di forma quadratica, che vede una iniziale crescita dei tempi di risposta per posizioni vicine a quella attesa, per poi assestarsi ed in alcuni casi invertire la tendenza in posizioni estreme, tornando cioè a diminuire. Gli autori suggeriscono che, se si dovesse considerare soltanto il contributo dei processi top-down volontari, l'andamento dei risultati avrebbe dovuto essere un gradiente lineare, con decremento progressivo della performance all'aumentare della distanza dalla posizione attesa. Come spiegare, invece, l'inversione di tendenza per le posizioni estreme? Ecco che potrebbe sussistere un effetto di monitoraggio costante bottom-up, il cui effetto di salienza sui tempi di reazione si fa sentire maggiormente nelle posizioni estreme, nelle quali invece le risorse attentive di alto livello allocate sono sempre più scarse. La mappa di priorità risulterebbe quindi dalla somma di due diversi gradienti che rappresentano i due contributi appena descritti. Un aspetto interessante è che, in posizione centrale, il peso della componente di alto livello sembra essere minore rispetto a quello di basso livello, al contrario delle posizioni laterali, per le quali il contributo top-down è invece il più pesante a livello di calcolo. Non viene data in proposito una

interpretazione precisa, tuttavia alcuni dei risultati degli esperimenti presentati nel presente lavoro di ricerca potrebbero essere compatibili con l'andamento appena descritto, e potrebbero indicare un ancoraggio speciale in posizione centrale anteriore, forse dovuto alla postura del soggetto e all'informazione visiva disponibile, supponendo che sia la linea mediana del corpo, sia quella dello sguardo, siano centrate proprio in quella posizione. Ad ogni modo, gli autori concludono che il modello di calcolo è compatibile con i risultati raccolti, e che quindi può costituire un affidabile strumento per descrivere l'interazione tra processi top-down e bottom-up nell'allocatione di attenzione nello spazio circostante.

Lo stesso autore, insieme ad altri colleghi (Golob et al., 2017), continua nell'indagine sui gradienti di attenzione spaziale nel dominio uditivo, alla ricerca di una replica dei risultati ottenuti pochi anni prima. Infatti, negli studi che avevano indagato precedentemente l'andamento dei costi dovuti allo spostamento dell'attenzione uditiva tra posizione attesa e posizione reale (Rhodes, 1987; Spence & Driver, 1994; Mondor & Zatorre, 1995; Rorden & Driver, 2001) il pattern era risultato diverso, approssimabile ad un gradiente lineare, con aumento maggiore dei tempi di risposta nelle posizioni vicine a quella attesa, ma senza alcuna stabilizzazione o inversione di tendenza all'aumentare della distanza.

In un recente studio (Golob & Mock, 2020), gli autori utilizzano gli stessi stimoli e il medesimo setting utilizzati negli studi precedenti: i partecipanti dovevano indossare delle cuffie e il compito era quello di discriminare tra due possibili suoni target, presentati in una delle cinque possibili posizioni virtuali ricreate in cuffia. Ad ogni blocco, i partecipanti erano istruiti su quale posizione fosse quella attesa. Da notare che soltanto tre delle cinque posizioni sono state utilizzate come posizione attesa: la posizione centrale anteriore e le posizioni laterali ( $+90^\circ$  e  $-90^\circ$  di azimut, esattamente davanti all'orecchio destro o sinistro). I suoni erano preceduti da un cue uditivo non spaziale, e nei diversi esperimenti è stato manipolato anche il SOA (2,4 secondi o 1,2 secondi). Anche il numero delle posizioni è stato manipolato tra gli esperimenti, infatti nel terzo esperimento soltanto tre delle cinque posizioni erano ingaggiate, scegliendo alternativamente

due combinazioni di posizioni per studiare gli effetti dello spostamento dell'attenzione in zone vicine e in zone lontane dalla posizione attesa. La probabilità della prova valida era all'84%, mentre le prove invalide erano il 16% restante, diviso equamente tra le posizioni non attese.

Complessivamente, i risultati mostrano che l'andamento dei tempi di risposta, al variare della distanza tra posizione attesa e reale, segue un profilo di gradiente quadratico piuttosto che lineare. Infatti, ad eccezione delle prove in cui l'attenzione era sostenuta in posizione laterale destra, tutti gli esperimenti evidenziano che i tempi di risposta sono minori nelle prove valide, mentre aumentano rapidamente nelle prove invalide con spostamenti piccoli ( $\pm 45^\circ$  di azimuth) per poi tendere a stabilizzarsi e, in alcuni casi (soprattutto quando la posizione attesa era quella centrale) a invertirsi per spostamenti verso le posizioni laterali estreme. Gli autori, alla luce di questi risultati e di quelli ottenuti in precedenza, rafforzano l'idea che i processi top-down distribuiscono l'attenzione prevalentemente sulla posizione attesa, con risorse decrescenti all'aumentare della distanza da essa. Contemporaneamente, i processi bottom-up potrebbero essere distribuiti in modo prioritario nelle posizioni più distanti dal focus attentivo (come suggerito anche da Bahcall & Kowler, 1999 per la modalità visiva), oppure agire con sensibilità uniforme su tutta la scena uditiva. La prima ipotesi sembra quella che meglio potrebbe descrivere i risultati in quanto, per spostamenti verso posizioni estreme, i tempi di risposta sono quasi confrontabili con quelli delle prove valide. In questa pubblicazione, perciò, si protende per un effetto di confine, che facilita le posizioni più esterne della porzione di spazio indagato, anche se si lascia aperta la possibilità che, indagando tutti i  $360^\circ$  attorno al soggetto, possa prevalere invece una distribuzione uniforme dei processi di monitoraggio. Anche questo aspetto, come vedremo più avanti, avrà la sua importanza nell'interpretazione dei risultati degli esperimenti proposti nel nostro lavoro di ricerca. Infine, al contrario degli studi simili, che non ravvisavano alcuna asimmetria negli effetti attentivi tra destra e sinistra, qui l'andamento a partire dalla parte destra risulta diverso, più piatto e caratterizzato da un effetto attentivo generalmente minore. Questo vantaggio destro potrebbe essere supportato, secondo gli autori, da alcune evidenze in paradigmi di

ascolto dicotico (Kimura, 1967; Hugdahl, 2003; Yurgil & Golob, 2010) che riportavano una migliore performance di rilevamento e memoria uditiva per gli stimoli presentati a destra. Ad ogni modo, non è completamente chiaro il motivo di questa differenza sui tempi di risposta, anche perché le evidenze portate a supporto di questi specifici risultati riguardano un contesto sperimentale assai diverso dal presente, sia per complessità che per tipologia di stimoli adoperati.

In un ulteriore studio (Golob & Mock, 2019) gli autori adottano il medesimo compito e lo stesso setting sperimentale per indagare se vi fossero differenze nella distribuzione dell'attenzione spaziale uditiva tra due gruppi di età diverse (il primo gruppo con una media di 25 anni, il secondo con una media di 75). Il compito consisteva sempre in una discriminazione tra due target uditivi costituiti da rumore bianco a due diverse frequenze, presentati solitamente in una posizione attesa e conosciuta dai soggetti prima dell'inizio del singolo blocco di prove (si intendeva quindi reclutare l'attenzione sostenuta). Occasionalmente, il target veniva presentato in posizioni diverse, richiedendo così ai partecipanti di spostare l'attenzione sulla posizione effettiva di presentazione dello stimolo. La probabilità delle prove valide era, complessivamente, dell'84%. L'esperimento prevedeva inoltre anche la manipolazione delle possibili sequenze di prove valide e invalide, poiché l'interesse era quello di comprendere se vi fossero differenze tra giovani e anziani nel recupero del focus sulla posizione attesa dopo una prova invalida. I risultati mostravano che l'andamento dei tempi di risposta in funzione della distanza angolare tra posizione attesa ed effettiva era ancora una volta approssimato al meglio con un gradiente quadratico, senza differenze tra giovani ed anziani. Invece, si riscontrava una differenza tra i gruppi al variare della lunghezza delle sequenze di prove valide dopo una prova invalida. In particolare, gli anziani mostravano più difficoltà e lentezza in sequenze molto brevi, caratterizzate da una sola prova valida dopo una invalida. Gli autori, oltre che portare nuove conferme a sostegno della forma quadratica del gradiente di attenzione spaziale uditiva, suggeriscono che l'età possa avere un impatto sulla velocità con la quale si recupera la posizione attesa dopo uno spostamento attentivo.

Ancora lo stesso gruppo di lavoro (Golob et al., 2021) ha continuato ad indagare l'attenzione spaziale uditiva, mantenendo invariato il setting e la tipologia di compito ma, al contempo, spostando l'accento sullo studio dello sforzo, espresso come decremento misurato di vigilanza nel mantenere l'attenzione sostenuta su una posizione, in rapporto ai costi di spostamento verso altre posizioni. I risultati di questo studio confermano che l'andamento dei costi attentivi di spostamento tra posizione attesa e posizione reale ha una forma di gradiente quadratico, replicando sostanzialmente quanto riscontrato in precedenza. Tuttavia, dai dati non emerge una chiara inversione di tendenza per le posizioni estreme, come invece era accaduto nel precedente studio del 2020; infatti, la curva dei tempi di risposta, dopo una ripida crescita in posizioni vicine a quella attesa, tende ad appiattirsi e soltanto negli esperimenti n. 3 e 4 si assiste ad un nuovo decremento dei tempi di risposta in posizione estrema, e soltanto verso destra.

Nel complesso, gli studi di E. J. Golob e colleghi rappresentano, a tutt'oggi, un importante punto di riferimento per la ricerca sull'attenzione spaziale uditiva, essi infatti hanno il merito di provare a fornire una descrizione più completa possibile del fenomeno indagato, sia dal punto di vista della quantità di posizioni indagate e dello spazio coperto (5 posizioni e 180° di spazio anteriore, anche se solo 3 potevano essere oggetto di attenzione sostenuta da parte dei partecipanti), sia dal punto di vista dei passi avanti nella comprensione del ruolo reciproco dei processi bottom-up e top-down. Da questi studi, infatti, emerge un possibile quadro teorico che potrebbe costituire un punto di partenza per le ricerche future in materia.

Ultimamente, uno studio di Popov e colleghi (2023) utilizza due compiti di discriminazione uditiva, per portare nuove evidenze a supporto dei forti nessi tra vista e allocazione di risorse nello spazio circostante, anche in presenza di un compito uditivo. I partecipanti hanno svolto due compiti: un classico compito di ascolto dicotico, con cue uditivo (un La puro a 440 Hz) e due target costituiti da sillabe in lingua tedesca, seguito da un compito di discriminazione che prevedeva l'utilizzo di 8 altoparlanti, scalati di 30° di azimut a copertura dell'intero spazio circostante. Anche in questo caso, veniva presentato un cue (sempre un La puro) su

uno degli otto altoparlanti, e ai partecipanti era richiesto di mantenere l'attenzione in quella posizione fino alla presentazione del target e al conseguente compito di discriminazione. Non erano previste né prove invalide né neutre, ma l'angolazione degli altoparlanti veniva variata per tre volte, in modo tale che in totale fossero coperte 24 posizioni spaziali, scalate di 15° l'una dall'altra. Venivano poi registrati i tempi di risposta, oltre che acquisiti i dati elettroencefalografici. Anche l'attività oculomotoria veniva registrata, per valutarne l'impatto su un compito di questo tipo. I risultati comportamentali sui tempi di risposta non mostravano differenze significative tra le varie posizioni, invece si confermava la lateralizzazione delle oscillazioni di banda alfa a testimonianza dell'orientamento spaziale dei partecipanti. Inoltre, i dati sull'attività oculomotoria confermano che la direzione dello sguardo (effettiva o anche solo programmata ma non eseguita) è fortemente connessa con l'orientamento spaziale anche uditivo. In altre parole, anche un cue uditivo spazialmente orientato elicitazioni di movimenti oculomotori verso la direzione corrispondente. Questi movimenti sembrano quindi non tanto una conseguenza dell'avvenuto orientamento, ma una parte sostanziale dello stesso processo di orientamento spaziale, che quindi appare almeno in parte sopra-modale. Vi è qui anche una estensione delle evidenze sull'attività oscillatoria di banda alfa all'emicampo posteriore al soggetto, che sembra comportarsi alla stregua di quello anteriore per quanto riguarda questo particolare fenomeno neurale.

### **2.3. Applicazioni in studi multimodali**

Non pochi studi hanno indagato l'attenzione spaziale uditiva adottando stimoli appartenenti a diverse modalità, ispirati dalla constatazione che le situazioni di vita reale sono eminentemente multimodali. La capacità di descrizione del reale, così come la generalizzazione dei risultati ottenuti, dipendono infatti anche dal tentativo di integrare più modalità in contesti sperimentali (Styles, 2005).

Nel panorama degli studi sull'integrazione tra vista e udito, un posto di rilievo è senza dubbio occupato dagli studi di Spence & Driver (1994), che indagarono il modo in cui attenzione visiva e uditiva collaborano attraverso la dimensione

spaziale. Essi misero a punto un esperimento nel quale sia stimoli visivi che uditivi (nel caso specifico si trattava di parole) venivano presentati alternativamente a destra o a sinistra rispetto ai soggetti partecipanti. Risultava così possibile indagare se l'orientamento endogeno dell'attenzione potesse essere controllato in modo indipendente per le due modalità o se, invece, vi fosse un controllo sopra-modale. Il setting prevedeva che, in entrambe le posizioni, fossero sistemati un monitor e un altoparlante, mentre al centro vi era un altoparlante che poteva riprodurre rumore bianco. Il compito era quello di fissare lo sguardo su uno dei due monitor e contemporaneamente fare attenzione e poi riportare le parole che si sentivano da un certo altoparlante. Quattro erano le condizioni sperimentali: infatti il monitor poteva mostrare una persona che, alternativamente, pronunciava esattamente le parole che si stavano sentendo su uno degli altoparlanti oppure muoveva le labbra casualmente; inoltre, l'altoparlante cui prestare attenzione poteva essere dalla stessa parte del monitor atteso o dalla parte opposta. I risultati mostravano la massima facilitazione quando l'altoparlante e il monitor attesi erano dalla stessa parte, e la persona stava pronunciando proprio le parole udite, mentre la performance peggiorava quando l'informazione uditiva proveniva dalla direzione opposta. Ad ogni modo, quest'ultima situazione portava comunque risultati migliori rispetto al caso in cui il monitor mostrava movimenti labiali incomprensibili. Gli autori conclusero che fosse ipotizzabile una certa sinergia tra attenzione visiva e uditiva, e che quindi l'attenzione endogena non operasse in modo indipendente per le due modalità. Proprio la selezione spaziale sembrava un *collante* di livello superiore tale che, una volta portata l'attenzione in una certa porzione di spazio, tutta l'informazione ivi contenuta fosse elaborata prioritariamente, sia quella visiva che quella uditiva.

Gli stessi autori, in studi successivi (Spence & Driver, 1996; 1997), ebbero modo di rafforzare questa ipotesi con nuove evidenze che mostravano come l'orientamento attentivo fosse più forte e portasse migliori performance quando l'informazione rilevante, uditiva e visiva, proveniva dalla stessa direzione. Risultava comunque possibile dividere l'attenzione tra le due modalità e tra le due posizioni, ma i risultati erano peggiori. L'insieme delle evidenze disponibili portò

gli studiosi a concludere che, anche se non si può parlare di un unico meccanismo attentivo sopra-modale, si riscontrano degli evidenti nessi tra attenzione uditiva e visiva, attraverso la selezione spaziale.

Anche Sach e colleghi (2000) realizzarono un esperimento di localizzazione spaziale con cue visivi e target uditivi, utilizzando però le cuffie per la presentazione dei suoni. Per simulare le diverse posizioni spaziali, agirono sull'intervallo di tempo inter-aurale (ITD) che permette di riprodurre la sensazione di suoni provenienti da posizioni laterali attraverso il calcolo della differenza tra il tempo impiegato dal suono a raggiungere l'uno e l'altro orecchio. Ai partecipanti era richiesto di riconoscere se il suono era stato presentato in posizione attesa (suggerita da un cue visivo presentato su un monitor) o in posizione diversa. Si trattava quindi di una scelta secca uguale/diverso. Anche in questo caso, i risultati sulle prove valide furono migliori di quelli sulle invalide, come testimoniato dalle risposte più rapide e accurate. Questo studio contribuì ad espandere il campo di applicazione del paradigma di Posner ai suoni presentati in cuffia, dimostrando che la manipolazione dell'ITD – al fine di ottenere una simulazione di posizioni spaziali - dava luogo agli stessi effetti osservati utilizzando altoparlanti dislocati nello spazio. Tuttavia, se da un lato la spazializzazione tridimensionale dei suoni in cuffia - negli anni ulteriormente evoluta grazie ai progressi del digitale - offre indubbi vantaggi a livello pratico nella predisposizione dei setting laboratoriali, dall'altro mostra alcuni elementi di criticità in prospettiva ecologica. Infatti, come recentemente affermato (Heine et al., 2020) il suono riprodotto attraverso le cuffie dà la sensazione di essere originato nella testa dell'ascoltatore, e viene a mancare tutta la parte di suono dovuta alla sua interazione con l'ambiente in cui è prodotto e nel quale è inserito anche l'ascoltatore. Questa interazione, che in qualche modo "arricchisce" il suono prodotto dalla sorgente, sembra portare ad una migliore discriminazione precoce tra suoni rilevanti e irrilevanti.

Kean & Crawford (2008) adottarono anch'essi una versione adattata del paradigma di Posner, presentando ai partecipanti cue uditivi lateralizzati, che anticipavano la posizione di target visivi costituiti da led colorati, disposti alla destra e alla sinistra

di uno schermo sul quale era presentata una luce di fissazione. Durante il compito di rilevamento, venivano misurate le velocità dei movimenti saccadici dell'occhio destro verso la posizione target. La probabilità di prova valida è stata manipolata, dando luogo a tre condizioni possibili, con valori percentuali di 20%, 50% e 80%. I risultati mostrano generalmente una risposta saccadica più veloce per le prove valide, in tutte e tre le condizioni sperimentali. Poiché i risultati di questo primo esperimento non potevano portare con certezza a pensare che l'attenzione visiva si fosse spostata verso la posizione del cue uditivo (venivano infatti misurati soltanto i movimenti saccadici), in un secondo esperimento, gli autori inserirono un compito di discriminazione da compiere sulla posizione target (valutare l'elevazione positiva o negativa del led luminoso rispetto alla posizione orizzontale). In questo caso, i risultati sull'accuratezza delle risposte mostravano un effetto principale della validità del cue, portando quindi gli autori a sostenere che un meccanismo attentivo inter-modale rendesse possibile l'orientamento dell'attenzione visiva a partire da un cue visivo.

A completamento del presente paragrafo, risulta utile dare nota di alcuni interessanti studi di neuroscienza, basati su esperimenti simili a quelli visti fino ad ora, mirati ad investigare quali attività neurali possano essere ritenute responsabili, in tutto o in parte, dei processi legati al dispiegamento delle risorse attentive nello spazio circostante. A questo proposito, è da menzionare il lavoro di Banerjee e colleghi (2011) che si pone l'obiettivo di verificare se il meccanismo di allocazione dell'attenzione nello spazio abbia una natura sopra-modale (come sostenuto da Farah et al., 1989; Macaluso et al., 2003; Shomstein & Yantis, 2006; Störmer et al., 2009; Kerlin et al., 2010) oppure si possano distinguere meccanismi specifici per ogni singola modalità, oppure ancora vi siano interazioni tra meccanismi modalità-generalisti e modalità-specifici (Spence & Driver, 1996; Eimer et al., 2002; Santangelo et al., 2010). Partendo dalle precedenti evidenze sul ruolo delle onde di banda alfa parieto-occipitali (onde di frequenza compresa tra gli 8 e i 15 Hz) nella soppressione degli stimoli visivi posizionati in porzioni di spazio da ignorare, l'autrice si pone l'obiettivo di verificare se questo tipo di fenomeno si registra anche per un compito di suggerimento spaziale puramente uditivo, mettendo comunque

in parallelo un secondo esperimento puramente visivo per confrontarne i risultati. Il compito uditivo prevedeva due possibili posizioni, una alla destra e l'altra alla sinistra di un monitor centrale, distanti 14° di azimuth dalla linea di mezzo. Un cue uditivo precedeva la presentazione di un target su uno dei due altoparlanti oppure su entrambi. I partecipanti, diversamente dal tipico compito di suggerimento spaziale, erano invitati a rispondere soltanto alle prove valide, pertanto le invalide fungevano esclusivamente da distrattori. Il cue, in questo esperimento, era presentato su entrambi gli altoparlanti, e il dato spaziale derivava dalla codifica di due sillabe, ciascuna delle quali indicava una delle due posizioni. L'esperimento puramente visivo utilizzava lo stesso setting, con la differenza che gli stimoli erano presentati sullo schermo centrale. I risultati mostravano chiaramente attività di onde alfa, lateralizzate a seconda della porzione di spazio da ignorare, sia nel compito uditivo che in quello visivo. Tuttavia, si notavano leggere ma chiare differenze di attività neurale tra i diversi compiti, in particolare dopo i 700 ms dalla presentazione del cue, momento in cui si assisteva ad una divergenza nella topografia, seppur limitata e ricompresa nell'area della corteccia parietale destra. Questo andamento particolare veniva interpretato dagli autori come una sostanziale conferma dell'esistenza di un meccanismo sopra-modale, condiviso tra vista e udito, responsabile dell'allocazione spaziale di attenzione, interagente però con un secondo sistema modalità-specifico, che entra in azione in stadi leggermente più tardivi dell'elaborazione. Ancora una volta, l'ipotesi dell'interazione tra meccanismi risultava la più accreditata.

Recentemente, uno studio di A. Wilsch e colleghi (2020) propone ai partecipanti due esperimenti di discriminazione con suggerimento spaziale, uno con cue uditivo e target visivo e l'altro con cue visivo e target uditivo, per valutare il ruolo delle oscillazioni di banda alfa (8-13 Hz) e delta (1-3 Hz) nei processi che anticipano la presentazione di un target. In particolare, gli autori erano interessati a comprendere se le modalità visiva e uditiva condividessero gli stessi processi o avessero delle specificità nell'anticipazione spaziale e/o temporale del target. Nei due esperimenti erano previste prove valide e neutre ma non invalide, poiché lo scopo dello studio non era direttamente quello di studiare lo spostamento attentivo, bensì di registrare,

attraverso il magneto-encefalogramma, l'attività neurale nell'arco di tempo tra la presentazione del cue e del target. I cue davano informazione sulla direzione cui prestare attenzione e, nel caso dell'esperimento sull'anticipazione temporale, davano anche un'indicazione ritmica che prediceva il momento in cui sarebbe stato presentato il target. I target uditivi scelti per la discriminazione erano costituiti da due serie di tre toni, una ascendente e una discendente, e ai partecipanti era richiesto di riconoscere quale delle due fosse stata presentata nella singola prova. Oltre alle misure elettrofisiologiche, venivano anche raccolti i tempi di risposta per la successiva analisi. I risultati mostravano il tipico beneficio attentivo delle prove valide rispetto alle prove neutre, evidenziato dai tempi di risposta più rapidi. Inoltre, la registrazione delle oscillazioni neurali alfa e delta porta gli autori a suggerire che l'anticipazione di tipo spaziale abbia forti nessi con la modalità visiva, mentre le aspettative temporali si manifestano molto più chiaramente nella modalità uditiva. Infatti, la lateralizzazione delle onde alfa, riconosciuta precedentemente come un indicatore di allocazione di attenzione spaziale, è evidente nel compito con target visivi, e ha una correlazione con i tempi di risposta. Nel compito uditivo, invece, si osservava una migliore sincronizzazione temporale con le onde di banda delta, osservata nel momento in cui i partecipanti si preparavano a ricevere il target. Questa sincronizzazione aveva una significativa correlazione con la velocità di risposta. Gli autori, pur sottolineando alcune limitazioni e la necessità di futuri approfondimenti, concludono che il cervello sembra utilizzare diverse bande di oscillazione a seconda delle richieste del compito da svolgere e che le modalità uditiva e visiva mostrano differenze nel modo in cui si reclutano risorse in anticipazione dell'esecuzione di un compito, a seconda delle dimensioni (spaziale e/o temporale) implicate.

#### **2.4. Studi con inserimento di stimoli distrattori**

La presenza di stimoli concorrenti, come abbiamo visto in precedenza, porta ad una maggiore necessità di risorse, sia per l'elaborazione di basso livello che per l'attenzione. Quest'ultima dovrà sovrintendere sia alla selezione dell'informazione

rilevante, sia all'inibizione di quella irrilevante ai fini del compito previsto. Best e colleghi (2007) hanno indagato l'interazione tra le modalità uditive e visive, in un esperimento nel quale i cue visivi davano la direzione cui fare attenzione, e i partecipanti dovevano svolgere in quella posizione un compito di discriminazione uditiva con la contemporanea presenza di distrattori. I risultati mostrano ancora una volta come l'informazione preventiva sulla posizione cui fare attenzione, e il conseguente suo orientamento, facilita il completamento di un compito, anche se questo non appartiene alla stessa modalità con la quale si è ricevuta l'informazione spaziale.

Un altro studio (Arbogast & Kidd, 2000) portò ulteriori variazioni sul tema del medesimo paradigma sperimentale, studiando l'attenzione spaziale uditiva in presenza di distrattori con informazione simile a quella del target. Questo tipo di mascheramento viene chiamato *informazionale* dal momento che il fattore di possibile confusione riguarda, appunto il contenuto dello stimolo. L'obiettivo dello studio era infatti quello di simulare in laboratorio una situazione reale, nella quale può accadere non di rado che i segnali uditivi in competizione per la nostra attenzione appartengano ad una stessa area di significato (voci umane oppure suoni dell'ambiente stradale, oppure ancora suoni di un ambiente specifico come potrebbe essere una fabbrica).

L'altro tipo di mascheramento tipicamente adottato negli esperimenti, in contrapposizione ma anche contemporaneamente al precedente, è quello *energetico*, per il quale il fattore confondente tra stimolo target e distrattori riguarda l'intensità o la sovrapposibilità di frequenze simili.

Tornando allo studio in questione, gli autori approntarono un sistema di sette altoparlanti a copertura dei 180° dell'emicampo anteriore rispetto ai soggetti, ai quali era chiesto di identificare uno stimolo target, distinguendolo dai distrattori, presentato contemporaneamente. I partecipanti erano istruiti sulla posizione cui fare attenzione per la durata del singolo blocco di prove, posizione che poi veniva cambiata nel blocco successivo. Erano previste prove valide e invalide, dipendentemente dal fatto che il target venisse presentato o meno nella posizione

attesa. Al contempo, venivano presentati anche i distrattori sugli altri altoparlanti. Target e distrattori erano costituiti entrambi da una serie di toni puri, presentati in rapidissima successione. La differenza fra target e distrattore stava nel fatto che la serie corretta presentava durate regolari mentre la serie distraente presentava durate e altezze variate casualmente. L'obiettivo, quindi era anche in questo caso di confrontare la performance sulle prove valide e sulle invalide, attraverso la misurazione dei tempi di risposta e delle accuratezze percentuali. I risultati confermarono l'andamento atteso, in quanto la prestazione sulle prove valide era statisticamente superiore a quella sulle prove invalide, sia sui tempi di risposta che sulle accuratezze. Quantitativamente parlando, l'effetto attentivo fu superiore a quello riscontrato da Mondor e Zatorre (1995), e gli autori suggerirono che la presenza di distrattori e la conseguente maggior difficoltà del compito potrebbero aver contribuito a spiegare tale differenza. Inoltre, gli autori suggeriscono che l'attenzione spaziale uditiva, una volta concentrata su una posizione, funzioni come un filtro audio, che amplifica il segnale dalla posizione attesa (forse proprio in termini di rapporto segnale-rumore) e attenua quello nelle posizioni non attese. La performance sulle prove invalide, che vede risposte più lente e un andamento compatibile con l'ipotesi di gradiente, potrebbe rivelare anche un altro meccanismo messo in atto sulle prove invalide. Infatti, dal momento in cui i partecipanti accertavano che il target non era in posizione attesa, un meccanismo di ricerca seriale avrebbe comportato l'analisi di ogni singola posizione, per confrontare ogni stimolo con quello atteso. Questo modo di procedere avrebbe dovuto portare a tempi di risposta ancora più lunghi di quelli osservati. Per spiegare questo fenomeno, gli autori proposero che l'attenzione uditiva, una volta disancorata dalla posizione attesa, si distribuirebbe in modo diffuso su tutta la scena e tratterebbe i segnali in parallelo anziché in serie. La presente interpretazione, sempre secondo gli autori, potrebbe spiegare sia i tempi di risposta non eccessivamente diversi tra prove valide e invalide, ma anche le differenze di accuratezza, significative ma non così accentuate come ci si aspettava. Era la prima volta che, sulle accuratezze, si osservavano differenze più marcate tra prove valide e invalide, e il fatto portò

Arbogast e colleghi a concludere che l'azione filtrante operata dall'attenzione spaziale uditiva fosse maggiormente evidente in presenza di compiti più gravosi.

Qualche anno più tardi, il medesimo gruppo di lavoro (Kidd et al., 2005) proseguì nell'esplorazione dei meccanismi che permettono di selezionare uno stimolo target in presenza di distrattori concomitanti. Stavolta, però, si trattava di frasi simili, pronunciate da tre diverse persone, da tre possibili posizioni nell'emicampo anteriore. Tre furono le condizioni sperimentali indagate: la prima prevedeva di non dare ai partecipanti alcun suggerimento preventivo, né sulla la frase target, né sulla posizione cui prestare attenzione; la seconda anticipava invece soltanto l'informazione sulla frase target, mentre la terza anticipava la posizione attesa. La presente manipolazione diede luogo a tre diversi esperimenti, nei quali furono raccolti e analizzati i tempi di risposta e le accuratezze, con prove valide e invalide. I risultati mostrarono che, nella prima e più incerta condizione, le prestazioni furono le peggiori e paragonabili a quelle che si possono riscontrare in un compito di focalizzazione istantanea dell'attenzione su una zona o su uno stimolo obiettivo. Nella seconda situazione (nella quale il cue anticipava quale delle tre frasi fosse l'obiettivo) i risultati sulle prove valide furono migliori rispetto alle invalide, anche se la posizione non era conosciuta a priori. Infine, quando i partecipanti erano istruiti proprio sulla posizione, anziché sulla frase, la performance complessiva fu la migliore dei tre esperimenti, sia sulle valide che sulle invalide. Inoltre, la differenza di accuratezza fu la più accentuata rispetto agli altri due esperimenti. In linea con le precedenti conclusioni, Kidd e colleghi suggerirono che, per catturare al meglio gli effetti dell'attenzione spaziale uditiva, i contesti sperimentali dovessero essere più probanti di quelli visti fino a quel momento e che la gravosità dei compiti da completare mettesse ancor più in risalto l'azione di tali meccanismi. Si fa sempre più strada l'idea, già sostenuta da tempo (Eriksen & S. James, 1986; Buchtel & Butter, 1988; Scharf et al., 1987; Rhodes, 1987; Benso et al., 1998; Green & Woldorff, 2012) che i compiti di semplice detezione o localizzazione con stimoli brevi non siano la condizione ottimale per un'effettiva allocazione di risorse attentive.

Un altro significativo contributo arriva dagli studi di Teraoka e colleghi (2020) che hanno studiato la capacità di distinguere parole pronunciate nella madrelingua, in presenza contemporanea di parole distraenti. Inoltre, gli autori erano interessati a comprendere come l'intervallo temporale tra cue e target potesse influire sull'accuratezza con la quale si riportavano le parole desiderate. Il setting prevedeva un sistema di dodici altoparlanti, sistemati tutto attorno al soggetto, a distanza di 30° di azimut l'uno dall'altro, anche se soltanto cinque di essi venivano utilizzati per la presentazione degli stimoli target. Il compito richiedeva ai partecipanti di portare l'attenzione su un certo altoparlante e di riportare la parola pronunciata in quella posizione. Una voce femminile pronunciava la parola target, mentre voci maschili pronunciavano le parole distraenti sugli altri altoparlanti, e viceversa. La posizione target veniva suggerita da un cue uditivo costituito da rumore bianco della durata di 50 ms, mentre l'intervallo temporale cue-target veniva variato tra i diversi blocchi di prove, da un minimo di 100 ad un massimo di 1000 ms. I risultati riportati, riguardanti soltanto le accuratezze percentuali al variare dell'intervallo cue-target, mostrano come la capacità di riconoscere le parole in presenza di altri stimoli distraenti sia minima per intervalli molto brevi, e successivamente cresca e si attesti su valori massimi per intervalli di 500 ms, per poi decrescere lievemente per intervalli di 1000 ms. L'andamento riportato, in linea con quanto precedentemente riscontrato da Mondor e Zatorre (1995) porta gli autori a suggerire che sia necessario un certo tempo per portare completamente le risorse attentive su una determinata posizione spaziale. Questo tempo potrebbe influire sulla capacità di riconoscere le parole presentate in una data posizione. La sorpresa riguarda invece il lieve decremento dell'accuratezza per intervalli lunghi, in contrasto con i risultati precedenti, possibilmente dovuto – secondo gli autori – ad una perdita di focus dell'attenzione, in conseguenza dell'eccessivo tempo intercorso tra cue e target. Infine, non vengono qui riscontrate differenze tra emicampo anteriore e posteriore riguardo alle misure di accuratezza. Una limitazione dei presenti risultati riguarda la mancanza di dati sui tempi di risposta, in assenza dei quali non si può scartare l'ipotesi che accuratezze maggiori siano dovute, in tutto o in parte, ad un maggior tempo di risposta. Inoltre, anche il confronto tra emicampi diversi risulta parziale,

dal momento che resta la possibilità che tempi di risposta diversi per emicampi diversi abbiano influito sui risultati di accuratezza, per alcuni o per tutti gli intervalli inter-stimolo adoperati.

Deng e colleghi (2020), interessati a studiare le attivazioni neurali e le onde alfa parieto-occipitali durante un compito uditivo, adottano un setting sperimentale con cinque posizioni target, che permetteva loro di verificare l'ipotesi che l'intensità di tali onde cresca in corrispondenza di posizioni vicine a quella target, per poi diminuire gradualmente al variare dello spostamento attentivo da compiere. I partecipanti, in un compito di attenzione spaziale, dovevano identificare una sequenza target di tre sillabe pronunciate da una specifica direzione, ignorando una sequenza distraente di tre sillabe simili, presentata da un'altra direzione. All'inizio di ogni prova, sullo schermo appariva un punto di fissazione. Due secondi dopo, un cue uditivo veniva riprodotto da una delle cinque posizioni possibili, per anticipare la posizione della successiva sequenza target. Gli attacchi della sequenza target e della sequenza distraente erano scalati di 200 ms, consentendo di apprezzare separatamente le rispettive risposte neurali. La sequenza target poteva precedere quella distraente, così come poteva darsi la condizione opposta. Oltre ai dati EEG, venivano misurate le accuratezze percentuali delle risposte, separatamente per ciascuna delle posizioni e per numero di sillabe correttamente riportate. I risultati comportamentali indicano che, quanto la sequenza target precedeva quella distraente, l'accuratezza sulla serie target era massima per la prima sillaba, e diminuiva per la seconda e terza sillaba. Quando invece la sequenza distraente precedeva quella target, l'andamento si invertiva, e l'accuratezza era massima sulla terza sillaba. Non vengono registrate differenze di performance tra le varie posizioni. I dati EEG, che mostravano una lateralizzazione graduale dell'intensità dell'attivazione di banda alfa, suggeriscono agli autori l'esistenza di una correlazione tra potenza delle onde alfa e distanza spaziale da coprire. Inoltre, l'asimmetria osservata tra le variazioni delle oscillazioni alfa tra emisfero destro e sinistro, simile a quanto osservato in studi sulla modalità visiva, suggerisce e conferma l'esistenza di una componente di elaborazione spaziale sopra-modale.

La situazione “cocktail party”, che vide la luce con i primi studi di ascolto dicotico e che nel tempo ha visto un ampliamento e delle posizioni spaziali e degli stimoli presentati, assume nel tempo sempre maggior peso negli studi sull’attenzione spaziale uditiva, in quanto le si riconosce un livello di difficoltà e di rappresentazione della realtà ottimale e preferibile rispetto ad altri compiti. Tuttavia, non mancano autori (Gamble & Luck, 2011) che ne pongano in questione l’alto potenziale di rappresentazione di situazioni reali. Infatti, la conversazione con più persone dovrebbe avvenire idealmente secondo turni di parola tali che non sarebbero previste, se non eccezionalmente, voci concomitanti tra le quali doversi orientare. La presentazione contemporanea di più stimoli, quindi, andrebbe a sovrastimare la difficoltà di una tipica conversazione reale e non risultando così ecologica come spesso si ritiene. Inoltre, altri autori (Darwin, 2008) sostengono che il parlato non debba essere trattato al pari degli altri stimoli uditivi, poiché potrebbero sussistere reti e connessioni neurali apposite, che ne facilitano la comprensione anche in contesti rumorosi. Anche la memoria sembra essere implicata nel riconoscimento delle parole e più in generale nella modulazione delle risorse attentive (Berti et al., 2003; Dalton et al., 2009; Golob et al., 2017) quindi i risultati degli studi in questione andrebbero interpretati anche alla luce di questi contributi senso-cognitivi specifici. Recentemente uno studio di Lewald & Getzmann (2015) raccoglie alcune evidenze neuroscientifiche a sostegno di questa elaborazione privilegiata del parlato, che convince gli autori a non utilizzare stimoli uditivi contenenti parole o frasi, proprio per evitare questo possibile elemento di confusione sui risultati ottenuti. Anch’essi, come vedremo nello specifico introducendo il nostro primo studio, indagano gli effetti dell’attenzione in una situazione di tipo cocktail party, ma gli stimoli scelti sono alcuni versi di animali considerati familiari per una popolazione non selezionata di soggetti.

Come si può evincere dalla presente rassegna di evidenze, la ricerca, attraverso la rivisitazione di alcuni paradigmi di riferimento come quello di ascolto dicotico e quello di suggerimento spaziale, ha allargato notevolmente gli orizzonti della comprensione dei processi – comportamentali e neurologici – che stanno alla base dell’allocazione di risorse attentive nello spazio circostante.

L'interazione tra modalità diverse, che vede il senso della vista esercitare una certa predominanza sulla selezione spaziale anche uditiva, sembra ricevere nuove e intriganti conferme, a testimonianza delle intuizioni di coloro che per primi avanzarono questa ipotesi, pur senza l'ausilio degli strumenti e delle tecnologie che la ricerca contemporanea può vantare. Inoltre, si riscontra un promettente ritorno di interesse per la modalità uditiva e per tutte le sue implicazioni e specificità, tra le quali senza dubbio si può annoverare la capacità di percepire stimoli a tutto tondo.

Lo studio dell'attenzione spaziale uditiva, come già si poteva intuire anche solo dalle tre parole che la compongono e descrivono, mostra con forza in questa rassegna la necessità di far convergere competenze ed evidenze da aree diverse del sapere scientifico. È infatti dall'intersezione di studi sull'attenzione, sulla localizzazione spaziale e sulla percezione uditiva che questa *materia di confine* nasce e si sviluppa con una sua specificità; ogni nuovo contributo in ciascuna di queste tre aree d'origine, se letto in prospettiva intersecante, spinge in avanti l'intera materia.

### III - Training musicale e impatto sulle funzioni cognitive

La seconda parte del presente lavoro intende offrire uno sguardo particolareggiato sul corpo di evidenze disponibili circa le possibili relazioni tra la pratica musicale e alcune funzioni cognitive fondamentali e trasversali, tra le quali l'attenzione occupa indubbiamente un posto di rilievo.

Risulta necessario, tuttavia, formulare alcune premesse, che possano fungere da chiavi di lettura per orientarsi nella vastità ed estrema varietà degli studi che hanno avuto come principale interesse quello di valutare gli effetti dell'acquisizione di competenze musicali sulla più generale acquisizione di abilità cognitive. Infatti, non di rado la frammentarietà dei contenuti di ricerca presentati – ancorché appartenenti ad una medesima area di interesse – può indurre un certo disorientamento, dovuto talvolta alla mancata percezione di un filo conduttore, un criterio guida, una chiave di interpretazione comune. In assenza di questi strumenti, anche un insieme di numerose evidenze potrebbe portare alla sensazione di mancanza di *evidenza*, poiché i mille rivoli della ricerca, se non opportunamente intersecati, potrebbero apparire soltanto nella loro singolarità.

Come è avvenuto nel precedente capitolo, anche qui l'intento è quello di mettere gradualmente a fuoco la materia, dai suoi aspetti fondamentali fino alle direzioni più specifiche e particolari, affinché possa esser percepito effettivamente un *corpo* di evidenze che, pur con le sue diverse e necessarie articolazioni, lascia trasparire i suoi tratti unitari e distintivi.

La prima e principale premessa riguarda il concetto stesso di *training* musicale. Il termine inglese, adottato anche in questo lavoro soprattutto per uniformità con quanto riportato comunemente in letteratura, non ha infatti un significato univoco nel contesto della ricerca (Swaminathan & Shellenberg, 2018). Talvolta ci si riferisce ad una forma di allenamento tecnico-pratico che porti quindi a sviluppare alcune abilità specifiche sullo strumento musicale utilizzato (compresa la voce). Talvolta invece ci si riferisce ad un allenamento delle abilità di ascolto musicale, o

ancora all'apprendimento di nozioni teoriche riguardo la musica. Infine, si dà anche il caso in cui il termine *training* si riferisca ad attività più articolate, che coinvolgano due o più delle precedenti aree di competenza. Anche il livello di strutturazione e formalità delle diverse attività che sono andate sotto il nome di *music training* è quanto mai eterogeneo (Tervaniemi et al., 2018). Vi sono esempi di apprendimento non strutturato e informale (tipicamente in contesti familiari o di gruppo, caratterizzati da processi prevalentemente spontanei e imitativi), come anche esempi di apprendimento non strutturato in contesti formali (quando, ad esempio, si agisce in contesto scolastico ma senza una chiara linea guida pedagogica e didattica), ma anche apprendimenti strutturati e non formali (ispirati a precise intenzioni e metodologie ma svolti in contesti familiari o extra-scolastici), infine apprendimenti strutturati e formali (tipicamente la formazione scolastica e accademica).

Già da questa carrellata, che raccoglie in breve ed anticipa la molteplicità di declinazioni del training musicale rinvenuta in letteratura, emerge un forte elemento di frammentarietà, che inevitabilmente si ripercuote sui risultati ottenuti e sulla loro solidità e replicabilità. Si può ipotizzare che essa sia dovuta, almeno in parte, a difficoltà di natura pratica e logistica della ricerca, che non di rado deve conciliare esigenze contrastanti tra loro come l'ottimizzazione della metodologia, il tempo e i fondi a disposizione, la rapidità di reclutamento di un campione adeguato. Tuttavia, a monte dei pur possibili limiti pratici e logistici, non si può scartare l'ipotesi che ci sia una sorta di *zona grigia* metodologica, dovuta probabilmente alla scarsità di intersezioni tra aree diverse ma egualmente implicate del sapere scientifico. Le discipline psicologiche e neurologiche da un lato, quelle musicali dall'altro, con l'aggiunta di una sorta di *convitato di pietra* rappresentato dalla pedagogia e dalla didattica applicate all'apprendimento musicale. Potrebbe essere sufficiente dare uno sguardo critico a quanto esposto finora a proposito delle diverse declinazioni che ha assunto il training musicale, per comprendere chi tra questi tre attori principali ha avuto la peggio nella maggioranza dei casi. Che si trattasse di un disegno di ricerca correlazionale o che prevedesse un intervento attivo, che si portasse avanti una ricerca trasversale o longitudinale; uno dei nodi cruciali resta

sempre lo stesso: la qualità pedagogica e didattica dell'apprendimento cui hanno preso parte i *musicisti* di cui parlano le molte ricerche che costellano il panorama scientifico in materia. L'estrema varietà delle evidenze disponibili andrebbe quindi letta anche alla luce della variabilità di coerenza interna dell'apprendimento di cui si voglia valutare l'effetto su altre competenze che non siano quelle strettamente disciplinari. Inoltre, tale variabilità costituisce di fatto un ostacolo significativo alla possibilità di replica degli studi e dei risultati precedentemente ottenuti, sia in chiave confermativa che falsificatoria, poiché viene a mancare un decisivo elemento di controllo sperimentale. Quanto detto finora, lungi dall'intento di svalutare i molti risultati ottenuti, vorrebbe rappresentare un possibile strumento per leggere, tra le righe e le pieghe dei tanti contributi, tutta la potenzialità dell'apprendimento musicale per lo sviluppo delle abilità cognitive trasversali, finora solo parzialmente emerso. Sicuramente, se i futuri sforzi di ricerca saranno mirati anzitutto al raggiungimento di un punto di intersezione più solido e condiviso tra tutte le aree del sapere coinvolte, i benefici dell'apprendimento musicale – spesso riscontrati con chiarezza a livello empirico – potranno ricevere maggiori e più qualificate conferme a livello scientifico (Mortari, 2010; Mortari & Tarozzi, 2010; Mortari & Ghirotto, 2019).

Un'altra doverosa premessa riguarda un aspetto non meno cruciale, che si interseca con quanto detto sopra riguardo alle diverse tipologie di training musicale riscontrate in letteratura. Infatti, assai spesso si incontra, leggendo le tante pubblicazioni in materia, il termine *musicisti*. Solitamente, lo si utilizza per definire il gruppo sperimentale di coloro che, all'interno o al di fuori dello studio in questione, hanno preso parte ad un training musicale per un certo tempo, presente o passato. Ma cosa caratterizza un *musicista*? Cosa lo distingue dai *non-musicisti*? Questa domanda di fondo è stata anch'essa oggetto di ricerca specifica (Gembris, 1997; Hallam & Prince, 2003), poiché evidentemente una risposta univoca e solida non è semplice da formulare. Un interessante lavoro comparativo di Zhang e colleghi (2020) risulta di grande aiuto nella mappatura dei molti e diversi criteri adottati nella ricerca dell'ultimo decennio per definire il gruppo dei musicisti. Attraverso l'elaborazione di un modello a tre componenti per la definizione di

musicista (Zhang et al., 2018) essi analizzano e categorizzano in dettaglio tutti i criteri utilizzati dai ricercatori per reclutare il proprio gruppo sperimentale. Le tre componenti, alle quali gli autori fanno afferire la totalità di questi criteri, sono: abilità musicali, identità percepita e predisposizione.

La prima caratteristica che, comunemente, aiuta a connotare un musicista è la capacità di suonare uno strumento musicale. Suonando, poi, si possono manifestare anche alcune abilità più specifiche come, ad esempio, il senso del ritmo, la capacità di comprendere e interpretare la musica, esprimere e comunicare pensieri e sentimenti attraverso le note di un brano musicale, saper suonare insieme ad altri (Hallam, 2010). Queste abilità sono solitamente il risultato di un processo di insegnamento-apprendimento (Sloboda, 2005), per questo uno dei principali criteri di selezione dei musicisti è l'ammontare del training musicale ricevuto (Rickard & Chin, 2017). Questo primo criterio trova sostegno in diverse evidenze neuroscientifiche, che mostrano significativi cambiamenti cerebrali, sia anatomici che funzionali in coloro che hanno ricevuto molti anni di istruzione (Gaser & Schlaug, 2003; Herholz & Zatorre, 2012; Du & Zatorre, 2017; Habibi et al., 2017). Tuttavia, questo criterio potrebbe essere necessario ma non sufficiente (O'Neill, 2002), dal momento che non sono pochi coloro che hanno acquisito abilità musicali anche notevoli, senza aver ricevuto alcuna istruzione musicale formale (Gagnè & McPherson, 2016). A questo scopo, si trova in Lamont (2002) una interessante distinzione tra *musicisti che suonano* (la cui abilità non proviene da una qualche forma di istruzione formale) e *musicisti formati* (la cui abilità è il risultato del training formale ricevuto).

Un secondo gruppo di criteri riguarda l'identità percepita (da sé stessi o da altri) di musicista. Infatti, *l'essere un musicista* è un costrutto dipendente anche da fattori sociali e culturali (Hallam, 2016), come la professione che si svolge o il titolo di studio conseguito, per cui si usa definire, ad esempio, un *medico* come colui che ha ottenuto una laurea in medicina e svolge un'attività professionale nel settore sanitario. Non sempre accade lo stesso per i musicisti, per i quali può capitare che una persona si definisca tale soltanto per la professione che svolge, anche se non ha

ottenuto nessun titolo accademico preciso; ma capita anche di persone che non si definiscano musicisti pur avendo conseguito titoli accademici, per il fatto che essi non svolgono attività professionali legate alla musica. Risulta perciò maggiormente rilevante l'auto-percezione di sé come di un musicista (Mills, 2010; Perrenoud & Bataille, 2017) poiché, come già sottolineato da O'Neill (2002) l'identità di musicista non si presta ad essere il semplice risultato di una sommatoria di attributi, ma coinvolge anche la percezione sociale e personale.

La terza componente, chiamata dagli autori *predisposizione*, offre già un possibile equivoco, poiché sembra alludere a *disposizioni preesistenti* verso la musica, che comunque sono state anch'esse oggetto di dibattito scientifico (Tan et al., 2014; Schellenberg, 2015 e Gingras et al., 2015 per le considerazioni più recenti). In realtà ci si vorrebbe riferire a tutti quegli aspetti, esterni e corollari al training musicale, che potrebbero avere un impatto più o meno significativo sullo sviluppo di abilità musicali (Bright, 2006; Rickard & Chin, 2017; Lima et al., 2020). Inizialmente, infatti, la valutazione delle abilità musicali avveniva tramite test attitudinali abbastanza orientati al tracciamento di competenze tipicamente acquisite durante il training (McPherson & Hallam, 2009; McPherson, 2015). Recentemente, alcuni ricercatori (Ollen, 2006; Müllensiefen et al., 2014; Zhang & Schubert, 2019) hanno cercato di formulare batterie di test che andassero ad intercettare, oltre alla storia formativa dei soggetti o le effettive abilità attualmente possedute, anche altri fattori come l'assiduità e la varietà degli ascolti musicali, le emozioni legate all'ascolto e alla produzione musicale (MacGregor & Müllensiefen, 2019), l'esposizione alla musica in età precoce (Cogo-Moreira & Lamont, 2018), gli elementi del contesto familiare che potrebbero aver contribuito ad avvicinare un soggetto alla musica. Da queste ricerche sono scaturiti l'Ollen Music Sophistication Index e il Goldsmith's Music Sophistication Index, quest'ultimo adottato anche nel presente lavoro di ricerca come uno dei criteri per la definizione del gruppo sperimentale dei musicisti. Ad ogni modo, non c'è ancora un consenso solido sui metodi quantitativi di valutazione delle abilità e competenze musicali (Law & Zentner, 2012).

L'analisi comparativa di Zhang e colleghi, attraverso questo modello a tre componenti, ha portato a rintracciare quei criteri di definizione sui quali vi è maggior convergenza tra i ricercatori. I risultati mostrano l'emergenza di un criterio sul quale si è accumulato un significativo consenso, che stabilisce mediamente in sei anni la soglia minima di training musicale necessario per essere annoverati tra i *musicisti*. Segue un secondo criterio basato sul numero di ore settimanali spese nella pratica musicale, con un valore soglia di un'ora a settimana, e il livello di istruzione frequentata, che nel 49% dei casi era il terziario (da notare qui che, a seconda degli stadi nei quali era basata la ricerca, il termine terziario può riferirsi a livelli anche diversi, come il college o l'università). L'identikit del *musicista*, risulta quindi essere quello di una persona con almeno sei anni di training musicale alle spalle, che abbia praticato almeno un'ora a settimana sul proprio strumento (o voce) e che abbia frequentato istituti di formazione terziaria. Da notare che non è stato comunemente utilizzato il criterio del possesso di un qualche titolo accademico, poiché implicitamente si riconosce e si accetta che un sufficiente livello di competenza musicale possa essere acquisito anche attraverso altri percorsi e altri contributi.

Tra le righe di questa preziosa analisi, traspare ancora una volta un'ambiguità ingombrante, che riguarda cosa si intenda per *training musicale*. Un interessantissimo paragrafo della pubblicazione di Zhang e colleghi elenca tutti i modi diversi in cui i ricercatori hanno definito la formazione musicale indicata come criterio di assegnazione al gruppo dei musicisti: viene usato il termine *music training* nel 37% dei casi, *formal music training* nel 33% dei casi, *music lesson* nel 15% dei casi, poi *formal music lessons*, *musical experience*, *degree in music* e *Suzuki method* in percentuali marginali. È significativo che, solo nel caso del metodo Suzuki (ovvero nel 3% dei casi), si faccia riferimento ad una metodologia ben precisa, codificata e quindi anche replicabile indipendentemente dal contesto culturale e sociale specifico. A questo 3% potremmo ragionevolmente sommare il 2% dei casi in cui ci si riferisce ad un titolo accademico (*undergraduate degree*), e concludere che solo nel 5% dei casi si potrebbero rintracciare i contenuti e le

metodologie pedagogiche e didattiche adottate nella formazione musicale cui hanno preso parte i partecipanti agli studi in questione.

Andando ancora più in profondità, affiora tuttavia un'altra questione decisiva: in cosa consiste la *competenza musicale*? Da cosa è composta? E, strettamente collegata alla precedente, è un'altra domanda di importanza notevole: quali competenze possano dirsi *specifiche* dei musicisti e quali, invece - pur avendo a che fare con il suono, l'udito, la percezione e le abilità cognitive e motorie - sono più *generali*, cioè riscontrabili e utili non solo durante la pratica musicale ma in contesti assai più ampi di vita reale? Dalle possibili risposte a queste domande scaturirebbe sia una più completa e coerente definizione di *musicista*, sia un affinamento dei criteri per selezionare - tra tutti i *training* oggetto di ricerche - quelli che sono maggiormente ancorati agli stati dell'arte in ambito pedagogico e didattico musicale e, in definitiva una più profonda comprensione dell'impatto che la formazione musicale può avere sulle funzioni cognitive trasversali. Anche se le due questioni mostrano un certo grado di interconnessione, è opportuno scinderne i significati e, soprattutto, gli effetti sulla materia di interesse del presente lavoro di ricerca.

Quanto alla prima domanda, è necessario rifarsi al profondo ripensamento pedagogico riguardo all'educazione musicale, avviato negli ultimi quattro decenni (Delfrati, 2008), che ha contribuito a ridefinire in modo più articolato il concetto di *competenza* musicale. A tal proposito, Deriu (2000), raccogliendo l'eredità di autorevoli predecessori, postula che essa si componga di quattro parti essenziali: la *percezione*, la *comprensione*, l'*esecuzione* e l'*invenzione*. Ognuna di queste quattro componenti si interseca con le altre, sia nei momenti ricettivi (durante l'ascolto di un brano), sia nei momenti produttivi (durante l'esecuzione sul proprio strumento o durante l'attività compositiva), così che nessuna di esse può ritenersi necessaria e sufficiente. Un altro aspetto che già Serafine e colleghi (1986) aveva intuito, è che si può riscontrare un certo grado di competenza musicale anche in soggetti che non hanno ricevuto o partecipato ad una formazione specifica e *intenzionale*. Proprio come accade ai bambini durante l'apprendimento del linguaggio, che comincia ad essere praticato ben prima dell'acquisizione dei più elementari rudimenti

grammaticali; così accade attraverso l'esposizione alla musica, che fin dal grembo materno pazientemente forma un sistema di ricorrenze e di strutture che il bambino, appena ne è in grado, mette in opera producendo ritmi con le mani o con oggetti, e imitando alcuni spunti melodici (Frances, 1984; Sloboda, 1988; Addessi et al., 1995; Penhune, 2011). Questo processo, che Sloboda (1988) definisce *acculturazione*, è dato dall'interazione dell'individuo con l'ambiente sonoro che lo circonda, che a seconda delle diverse culture musicali forma diverse sensibilità specifiche, tali che Frances (1984) parla del sistema musicale tonale come la "lingua materna degli occidentali", intendendo che essa, al di sopra anche delle stesse lingue nazionali, rappresenta un *linguaggio comune* che può essere riscontrato anche nelle competenze informali dei bambini che si avvicinano a intonare semplici melodie.

Ciò che si può aggiungere, attraverso un apprendimento intenzionale, risulta quindi la *consapevolezza della competenza*, che rende possibile l'astrazione, la replica o la manipolazione di strutture, elementi distintivi, particolari estrapolati da un certo materiale sonoro. Questa auto-coscienza porta progressivamente anche alla conoscenza dei processi cognitivi che hanno reso possibile determinate azioni tipicamente musicali come l'esecuzione di un brano, la sua interpretazione, la percezione e comprensione del materiale sonoro utilizzato in un preciso genere o da un preciso autore, la composizione di brani originali. In altre parole (Serafine, 1986; Sloboda, 1988), comprendere, riprodurre ed interpretare un brano musicale significa quindi ri-costruire il processo mentale che lo ha generato, e avere la capacità di intervenire attivamente su di esso, operando anche variazioni in alcuni passaggi.

È quindi ipotizzabile che sia proprio questa auto-consapevolezza dei processi mentali ad aprire la possibilità di trasferire competenze verso altri domini più generali, e che un apprendimento musicale completo, ancorato fortemente alle evidenze psicopedagogiche e didattiche, possa contribuire a rafforzare questo trasferimento. Evidenziarlo nei suoi effetti, è invece tutt'altra materia, di competenza della ricerca neuropsicologica, in un costante dialogo tra le diverse

scienze, finalizzato ad una sempre migliore e più profonda comprensione di ciò che è dominio-specifico della competenza musicale e ciò che, invece è dominio-generale (Carey et al., 2015), e che l'apprendimento musicale può contribuire a rafforzare o recuperare.

Ammessa, quindi, la suddetta definizione di competenza musicale come un sistema interconnesso formato da quattro aree fondamentali (che rappresenta ancora oggi lo stato dell'arte in ambito pedagogico e che informa, di conseguenza, i processi di insegnamento-apprendimento), possiamo tentare di rispondere alla seconda domanda fondamentale, su cosa sia, o non sia, specifico della competenza musicale.

Se per l'esecuzione e l'invenzione (attività creativa e compositiva) non si rintracciano in letteratura argomenti che ne mettano in dubbio l'appartenenza specifica alle competenze musicali, diverso e molto più articolato è il dibattito riguardo quella che Deriu chiama *percezione*. Come alcuni ricercatori hanno evidenziato (Schellenberg, 2015), essa non si identifica con la capacità uditiva, tale che i musicisti sarebbero migliori dei non musicisti quanto a sensibilità uditiva di base. Piuttosto, in una singolare consequenzialità di pensiero con quanto affermato da Shinn-Cunningham (2008) riguardo agli *oggetti percettivi*, la percezione musicale si riferisce alla capacità di cogliere relazioni di alto livello tra gli oggetti precedentemente formati dalla collaborazione tra i sistemi bottom-up e top-down. Questa capacità si svolge nel tempo (Best, 2008): è infatti attraverso lo scorrere del tempo che si percepiscono quelle che in gergo musicale si chiamano *cellule, frasi, strutture, forme*. Le cellule sono unità melodiche o ritmiche basilari, delle quali si può percepire la coerenza interna e la ripetizione o variazione. Esse possono poi costituire unità via via più complesse, caratterizzate da una serie stabile di cellule e frasi, che possono dare infine vita a strutture e infine forme definite e riconoscibili. La percezione musicale è spesso confusa o sovrapposta ad alcune abilità come la discriminazione delle diversità di altezza, timbro, durata dei suoni, che invece andrebbero più correttamente considerate abilità di dominio più generale, come è stato abbondantemente discusso in precedenza riguardo all'analisi della scena uditiva e ai processi percettivi e attentivi. La competenza percettiva musicale non

ha come oggetto le caratteristiche fisiche del suono (Overy & Molnar-Szakacs, 2009, Pfordresher, 2012), ma le possibili sequenze, gli andamenti, l'intenzione di seguirli, la loro previsione, la loro rappresentazione, la loro astrazione e riproduzione. Ad esempio, per quanto riguarda la capacità di discriminare le diverse altezze (ovvero le diverse frequenze fondamentali di ogni suono), vi sono evidenze (Abel et al., 2016) che mostrano come la padronanza o meno di questa competenza abbia un impatto su domini che vanno ben oltre quello strettamente musicale, con ricadute sulla percezione del linguaggio (Sares et al., 2018) e, più in generale, sull'accuratezza dell'analisi di qualsivoglia scena uditiva reale (Slater et al., 2015). Le stesse evidenze sembrano indicare che il training musicale possa influire positivamente su queste competenze, indicandole tuttavia con chiarezza come dominio generali. Anche Miendlarzewska & Trost (2014), quando si riferiscono alle abilità ritmiche, mostrano come esse si possano rintracciare in contesti ben più ampi di quello specifico musicale, come ad esempio nella comunicazione sociale, che per avvenire ordinatamente necessita, secondo gli autori, della percezione del suo proprio *ritmo*, che coinvolge l'udito ma anche la voce e il corpo in generale. Recentemente, uno studio di Niarchou e colleghi (2022), interessato principalmente a rintracciare, nel genoma, le componenti che potrebbero essere associate con l'abilità di sincronizzarsi e muoversi su un certo ritmo (definita dagli stessi autori come una componente fondamentale della musicalità); hanno trovato che una certa parte dell'architettura genetica correlata con questa abilità sembra, in realtà, condivisa con altre abilità molto più generali come la funzione del respiro, le funzioni motorie la temporizzazione e l'elaborazione della velocità di esecuzione di azioni come la camminata. Anche in questo caso, quella che è comunemente accettata come una abilità specifica musicale mostra in realtà forti nessi con abilità ben più generali.

Rispetto alle altre tre competenze musicali fondamentali (comprensione, esecuzione, invenzione) la percezione è stata indubbiamente la destinataria privilegiata della quasi totalità degli studi neuro-psicologici di settore, e si è venuta formando l'idea che l'essenza dell'essere *musicali* risieda prioritariamente in capacità uditive. A testimonianza di quanto appena detto, si noti che i primi test che

intendevano misurare la cosiddetta *attitudine musicale* (Seashore et al., 1956; Gordon, 1965) erano interamente costituiti da prove uditive di discriminazione di spunti melodici e ritmici, con la sola aggiunta di prove di memoria che richiedevano di riportare alcuni elementi appena ascoltati. Questa tendenza ha avuto probabilmente una spinta anche dagli studi, passati e recenti, sul cosiddetto *orecchio assoluto* (Deutsch, 2013) ovvero quella competenza che permette di stabilire con precisione l'altezza dei suoni, senza l'aiuto di un suono di riferimento (in tal caso si parla di *orecchio relativo*). Come sottolineato da Gingras e colleghi (2015) la rilevanza di questa abilità nella pratica musicale, anche a livelli professionali, è assai discutibile, e ciò sarebbe confermato anche dal fatto che soltanto una piccola percentuale di musicisti la possiede (Gregersen et al., 1999). Già Karma (2007) evidenzia alcune importanti criticità di un simile approccio alla valutazione dell'attitudine musicale, reclamando l'adozione di test che misurino non tanto le abilità uditive di base (discriminazione dei parametri del suono), quanto l'abilità di cogliere strutture sonore, avvicinandosi significativamente a quanto gli studiosi in campo musicale indicano come proprio e specifico della competenza percettiva musicale. Inoltre, Gingras e colleghi, nel fare una rassegna di ciò che biologicamente potrebbe definire l'attitudine musicale, arrivano a concludere che, in futuro, le misurazioni di questa capacità dovrebbero tenere conto anche di aspetti legati alla produzione (esecuzione) di melodie e ritmi, intercettando alcuni dei tasselli mancanti della competenza musicale, che può essere rintracciata sia in coloro che hanno svolto a vario titolo un training musicale, sia in coloro che potrebbero mostrarla, magari in misura più contenuta, anche in assenza di training (come ad esempio in tenera età).

Come si intuisce, non vi è unanime consenso su cosa sia esattamente l'attitudine musicale, e nemmeno sul fatto che essa esista realmente (Schellenberg, 2015). Non è ancora chiaro neanche il rapporto reciproco tra elementi genetici e ambientali nell'emergenza di questa attitudine. Ciò che, invece, risulta evidente dalla revisione della letteratura è che, se l'attitudine riguarda quasi esclusivamente abilità uditive, l'essenza della competenza musicale risulterebbe quindi schiacciata sulla dimensione percettiva, considerando le altre espressioni cognitive, performative e

creative come una conseguenza di ciò che sembrerebbe precederle almeno a livello temporale ed evolutivo.

Questa visione contrasta fortemente con quanto suggerito da diverse altre fonti, attinenti ai diversi settori scientifici implicati, che mostrano un coinvolgimento multidimensionale assai più complesso, durante l'ascolto musicale (Maes et al., 2014; Fortuna & Nijs, 2019), durante la comprensione musicale e l'attività compositiva, e a maggior ragione durante la produzione (Overy, 2009). Già nelle prime manifestazioni di competenza musicale di soggetti che non hanno ricevuto ancora alcun tipo di training (Delalande et al., 1984; 2016), si possono identificare tutte le componenti della competenza musicale, presenti e operanti nei gesti e nelle intenzioni dei bambini, che dimostrano una certa capacità di estrarre significato da ciò che ascoltano, di operare astrazione e imitazioni soprattutto quando l'informazione uditiva è accompagnata da quella visiva e corporea corrispondente. Essi, inoltre, al contrario di quanto si credeva in precedenza, sono in grado non solo di produrre suoni imitando semplici cellule melodiche o ritmiche, ma anche di inventare sequenze sonore, avvertendo regolarità e variazioni in quanto essi stessi producono e operando quindi consapevoli selezioni sul materiale sonoro prodotto. Tutta questa attività, come si può intuire, implica un coinvolgimento sensoriale, cognitivo, motorio ed affettivo-emotivo che va ben oltre la modalità uditiva. Ed è proprio in questa costante integrazione che potrebbe identificarsi, in definitiva, una possibile radice comune della competenza musicale (Penhune, 2011) il che avrebbe riscontri anche in evidenze neuroscientifiche (Zatorre et al., 2007; Miendlarzewska & Trost, 2014; Wang & Agius, 2018; Van Vugt et al., 2021) che mostrano differenze tra musicisti e non musicisti in aree di riconosciuta valenza sensomotiva.

Da quanto appena emerso, quindi, discende che l'*attitudine musicale*, ammesso che esista, si dovrebbe manifestare attraverso atti complessi, ben più articolati della semplice discriminazione tra toni o ritmi diversi; tali atti complessi implicherebbero probabilmente strumenti di valutazione altrettanto complessi e, forse, non soltanto quantitativi (Karma, 2007).

A tal proposito, molto interessante è la ricerca di J. Tafuri, riportata inizialmente in uno studio (Tafuri & Villa, 2002) e poi raccontata in uno dei suoi libri, dal titolo quasi provocatorio “nascere musicali” (2007). L’autrice, partendo dalle evidenze esistenti circa il funzionamento dell’orecchio del bambino fin dalla 24° settimana di gravidanza (Porzionato, 1980; Deliège & Sloboda, 1995), conduce una ricerca longitudinale per valutare quanto una ricca esperienza sonora e musicale pre e post-natale possa influire sulla produzione sonora del neonato nei primi mesi di vita. 119 madri gestanti hanno preso parte alla sperimentazione, iniziata durante la gravidanza e protratta per ben sei anni, che le ha viste partecipare ad un piano strutturato di attività musicali con cadenza regolare fatto di vocalizzazioni, canti, ascolti, girotondi e danze, uso di strumenti a percussione. I genitori dovevano poi continuare le attività a casa, possibilmente tutti i giorni. Dopo la nascita del bambino/a, i genitori raccoglievano alcune registrazioni audio/video delle loro produzioni sonore spontanee, oltre a compilare appositi diari di osservazione. Il gruppo sperimentale è stato affiancato da un gruppo di controllo costituito da madri gestanti che non avrebbero partecipato alle attività musicali, ma soltanto alla raccolta dati audiovisivi e alla compilazione dei diari. I risultati delle analisi dati, condotte da due ricercatori indipendenti, hanno evidenziato una sostanziale differenza quantitativa e qualitativa tra le produzioni sonore dei bambini dei diversi gruppi. In particolare, del gruppo di controllo, mediamente soltanto tre bambini su dieci producevano sporadicamente vocalizzazioni entro i primi 8 mesi di vita, anche se era presente una risposta motoria alle attività proposte, visibile attraverso sorrisi e movimenti di arti superiori e inferiori. Nei bambini del gruppo sperimentale, invece, si riscontrava una maggiore quantità e varietà di produzioni sonore, simili a esplorazioni con intenzione comunicativa o anche soltanto di auto-intrattenimento. I risultati portarono gli autori a concludere che i bambini, sollecitati e accompagnati da esperienze musicali nella vita pre e neonatale, esplorano la voce e producono delle vocalizzazioni che possiamo considerare "proto-canti", che si manifestano più precocemente, con più abbondanza e con maggiore ricchezza musicale rispetto a quelle dei coetanei non altrettanto sollecitati. Troviamo qui un contributo importante, che aiuta a non confondere tra gli effetti degli aspetti

puramente genetici e di quelli ambientali, poiché questi ultimi iniziano a interagire con i primi ben prima della nascita e risulta perciò molto problematica una effettiva separazione dei contributi singoli. Ciò che, spesso, va sotto il nome di *attitudine*, dovrebbe forse essere ripensato non in termini di predisposizione, ma più coerentemente come il risultato di una interazione precoce con un ambiente musicalmente persistente, ricco e vario.

Infine, si registra un crescente consenso verso l'inquadramento della musica come *linguaggio incarnato* (Perlovsky, 2015; Clarke et al., 2015; D'Ausilio et al., 2015; Clark & Tamplin, 2016; Leman, 2016; Dell'Anna et al. 2021) o come *sapere incarnato* (Matyja & Schiavio, 2010; Korsakova-Kreyn, 2018; Leman et al., 2018), riconoscendo in essa efficacemente esemplificato e realizzato quel rapporto bidirezionale tra corpo e mente, intravisto già ai primi del novecento negli studi e nell'opera didattica musicale di J. Dalcroze (2008), e poi ripreso nelle ricerche di F. Delalande (1993) e studiato diffusamente negli ultimi decenni (si riportano, come esempi più attinenti alla materia specifica, i lavori di Lakoff & Gallese, 2005; Johnson, 2007; Tucker, 2010). Anche in Cross (2014) si rinviene il tentativo di trascendere i confini acustici e percettivi, descrivendo la musica non più come una "manifestazione di andamenti del suono", bensì come un "processo comunicativo interattivo". In particolare, Koelsch et al. (2019) raccolgono quanto precedentemente mostrato da Clark riguardo all'ipotesi di un *circuito di causalità sensomotoria*, applicandolo alla percezione e alla produzione musicale. Essi sostengono che, durante questi momenti (che spesso possono essere anche contemporanei) dell'esperienza musicale, sia costantemente in atto un processo circolare per il quale vengono generate previsioni sensomotorie circa le caratteristiche ritmiche o melodiche di quanto si sta ascoltando e/o producendo. Queste previsioni stimolano i sistemi cognitivi a programmare azioni motorie (a cui può essere dato seguito o meno) che tendono a sincronizzare il corpo con quanto si sta ascoltando (movimenti della testa, delle mani, delle dita, dei piedi ecc...), il tutto in un continuo processo reciproco di aggiornamento e inferenza, nel quale anche le azioni prodotte entrano a far parte dell'analisi sensoriale successiva, e le emozioni

provocate da quanto ascoltato e prodotto fungono anch'esse da mediatori dell'intero processo (Cox, 2016).

Questa visione complessa, interagente, multidimensionale dell'esperienza musicale (Loui & Guetta, 2018), potrebbe contribuire alla ridefinizione di ciò che è effettivamente e coerentemente *specifico* della competenza musicale e ciò che, invece, essa recluta, coinvolge, stimola di tutte le funzioni sensoriali, cognitive, motorie e affettive, ma che attiene a domini più generali. È in quelle interazioni virtuose viste in precedenza, e nell'affiorare della auto-consapevolezza di questi processi, che andrebbe ricercato il possibile *transfer* di competenza dall'area musicale all'area più trasversale.

### **3.1. Sull'attenzione**

Il presente paragrafo si apre con una disamina su ciò che, durante l'ascolto e ancor più completamente durante la pratica esecutiva musicale, può indicare la necessità e l'effettivo reclutamento di risorse attentive. Ci chiediamo quindi quante e quali tipologie di risorse attentive sono necessarie per lo svolgimento dei molti e diversi compiti connessi con il fare e l'ascoltare musica. Se empiricamente non vi è dubbio che, per praticare questa disciplina a qualunque livello, sia necessario prestare attenzione a molti e diversi stimoli, non soltanto sonori; risulta però opportuno dimostrare, per quanto l'evidenza disponibile ce lo consente, che ciò che si constata empiricamente ha una certa corrispondenza con quanto effettivamente accade a livello sensoriale, motorio, cognitivo, neurale e comportamentale.

La prima parte della presente sezione si occuperà, quindi, di evidenziare il ruolo dell'attenzione nelle sue componenti selettiva, sostenuta e divisa, durante i gesti, i pensieri, le azioni e le interazioni tipici della pratica musicale. Questa parte intenderebbe fornire quindi un valido fondamento sperimentale, propedeutico alla lettura e alla comprensione delle più significative evidenze a sostegno di un impatto della pratica musicale sull'attenzione in una o più delle sue forme e manifestazioni. Saranno inoltre messi a confronto i diversi paradigmi, le scelte sperimentali adottate dai ricercatori per studiare l'effetto dei diversi training musicali, i compiti utilizzati

e infine i risultati e la loro discussione, in uno sforzo costante di far intravedere un quadro globale, aldilà dei singoli contributi e del loro particolare microcosmo di competenza.

### **3.1.1. L'attenzione durante l'ascolto e l'esecuzione musicale**

Un possibile denominatore comune delle ricerche selezionate per la presente rassegna specifica potrebbe essere costituito dal fatto che la grande maggioranza di esse considera l'ascolto e/o la produzione musicale (individuale o d'insieme) come oggetto di studio a sé stante. A differenza degli studi esaminati nei prossimi paragrafi, che considerano l'esperienza musicale come un dato acquisito poiché sono essenzialmente mirati a confrontare musicisti e non musicisti su tutta una serie di funzioni sensoriali, cognitive, affettive, emotive e relazionali; qui si rintraccia il tentativo di addentrarsi nei gesti, i pensieri, i comportamenti e le azioni caratteristiche della pratica musicale. Ognuno di essi, infatti, se elevato ad oggetto di studio, può rivelare molto dei meccanismi che lo hanno generato.

Durante l'ascolto di brani musicali di qualsivoglia genere o stile, sono le competenze di percezione e comprensione ad essere prioritariamente reclutate. Esse rimandano intuitivamente ai sistemi sensoriali, cognitivi e alla loro interazione, anche se, come abbiamo visto in precedenza a proposito del *circuito di causalità sensomotoria*, in realtà le aree coinvolte sono più ampie. Risulta qui necessario richiamare, se pur brevemente, quanto detto nel precedente capitolo a proposito dell'analisi della scena uditiva, della segregazione del *tutto uditivo* in flussi distinti e caratterizzati da elementi acustici di base comuni come le frequenze, i timbri, le intensità, le posizioni spaziali (Bregman, 1994; Shinn-Cunningham, 2008). Questi flussi, segregati ad un livello sensoriale, vengono elaborati al livello attentivo, che li seleziona in base a determinati obiettivi e forma *oggetti percettivi*, che poi segue nel tempo (Best et al., 2008), alterando localmente la sensibilità acustica di base per ancorarla meglio a quanto si sta seguendo.

Questo processo, che è di carattere generale in quanto messo in atto costantemente da tutti gli individui e con qualsiasi stimolo uditivo, acquista una particolare valenza

quando si ascoltano brani musicali. È qui che possiamo rinvenire gli effetti di quella competenza specifica, che ad esempio potrebbe raggruppare suoni diversi ma temporalmente vicini, dando così origine alla percezione del *tempo musicale* e del *ritmo*, ovvero della organizzazione temporale degli eventi sonori in pulsazioni regolari (tempo) e suddivisioni (ritmo). Questo tipo di processo è stato riconosciuto come eminentemente attentivo (De Freitas et al., 2014; Grahn, 2012), poiché si tratta di selezionare e combinare, nel tempo, diversi oggetti percettivi in modo da formare unità di più larga scala, come appunto le cellule ritmiche e la frasi musicali. Si tratta, quindi, di una forma assai complessa di analisi della scena uditiva, che richiede all'ascoltatore di operare continuamente separazioni e fusioni, tramite le quali possono emergere aspetti orizzontali (linee melodiche e/o ritmiche, crescendo e diminuendo, variazioni agogiche) e verticali (piani musicali, armonie). Alternativamente, o contemporaneamente, l'attenzione selettiva e divisa sono reclutate a questo scopo (Loui & Guetta, 2018) anche se sono riscontrabili effetti del training pregresso (Loui & Wessel, 2007) e del conseguente apprendimento di schemi, estratti ed astratti dai precedenti ascolti, che vanno a formare un sistema di aspettative temporali con un significativo interessamento della memoria uditiva (Fritz et al., 2007; Sridharan et al., 2007). L'evidenza di queste aspettative emerge anche dagli studi sul paradigma di suggerimento, ampiamente discusso nel precedente capitolo, attraverso i quali è apparso che un suggerimento (cue) uditivo in una certa frequenza (un tono puro o un suono reale con la sua chiara ed evidente frequenza fondamentale) può facilitare il riconoscimento di suoni con la medesima frequenza (Haftner & Saberi, 2001). Questo effetto è stato riscontrato anche quando il target non condivideva necessariamente aspetti acustici con il cue (Haftner et al., 1993), portando all'idea che la sensibilità uditiva potesse essere orientata e aumentata non solo da ciò che effettivamente viene ascoltato, ma anche da ciò che si attende. Quando il cue presenta le stesse caratteristiche acustiche del target, si può pensare ad un effetto di orientamento esogeno, mentre si parla di attenzione endogena quando il cue anticipa soltanto qualche informazione sul target, e riduce il grado di incertezza riguardo ad esso. Ad un livello di elaborazione più alto dell'elaborazione attentiva, quello cioè che ha come oggetti le cellule

melodiche/ritmiche, le frasi, le strutture, le armonie e infine le forme, si è potuto riscontrare un simile effetto di aspettativa, dovuta in questo caso alla conoscenza pregressa della musica tonale. Essa ha portato a risposte più rapide in compiti nei quali ai partecipanti era richiesto di riconoscere armonie conosciute, alternate alla presentazione di accordi inaspettati per i quali le risposte erano, invece, più lente. (Bharucha & Stoeckig, 1986; Marmel et al., 2008). Lo stesso effetto di aspettative tonali è stato riscontrato, sorprendentemente, anche in compiti di elaborazione visiva (Escoffier & Tillmann, 2008) e anche in soggetti che non avevano alle spalle alcun tipo di training musicale formale (Bigand, 2003), e ciò potrebbe indicare che l'esposizione alla musica tonale (che rappresenta la quasi totalità degli ascolti della popolazione occidentale, fin dai primi anni di vita) produce implicitamente una conoscenza e, conseguentemente, delle aspettative su quanto si ascolta. Anche i risultati dello studio di Ventura e colleghi (2018), interessati a mettere in relazione le misure di musicalità con misure di abilità cognitive più generali, somministrano diverse batterie di test anche percettivi, e trovano conferma del fatto che la sensibilità tonale si rintraccia anche in coloro che non hanno svolto training musicale significativo. La familiarità con le armonie, le cadenze e gli accordi maggiori e minori, riconducibili alla tonalità, può essere dovuta anche soltanto alla continua e prolungata esposizione. È importante notare che, per interpretare bene questi risultati, non dovremmo riferirci alla musica tonale in senso stretto, cioè al sistema di regole armoniche costituito gradualmente a partire dal XVII secolo, adottato dalla quasi totalità delle musiche colte occidentali fino alla metà del XIX secolo circa, insegnato e praticato in tutti i conservatori del mondo attraverso corsi specifici. Di questo sistema complesso, si possono trovare alcune tracce nella musica commerciale, tuttavia essa attinge anche a sistemi diversi, come ad esempio quello modale, che è precedente al sistema tonale e che non può essere identificato come specifico della cultura musicale occidentale. Vi è poi da considerare che la musica colta rappresenta mediamente solo una piccola parte degli ascolti cui siamo esposti quotidianamente. Per tutti questi motivi, è necessario interpretare con cautela i precedenti risultati, infatti è probabile che, adottando stimoli sonori più specifici della musica tonale in senso stretto, si possano riscontrare maggiori

differenze tra musicisti e non musicisti, dovute a familiarità di ascolto che sarebbero, almeno in parte, differenti.

Tuttavia, lo studio di Loui & Wessel (2007) mostra che, in un compito di attenzione selettiva verso la progressione delle note più acute di alcuni accordi, i musicisti sono più influenzati dalla presentazione di accordi inaspettati rispetto ai non musicisti, e ciò è testimoniato da un diverso andamento dei benefici e dei costi attentivi, questi ultimi maggiori nei musicisti. In questo caso, gli autori suggeriscono che il training musicale abbia avuto un effetto sui risultati osservati. Si trovano, qui come in altri studi di neuroscienza (Justus & List, 2005; List et al., 2007) conferme del fatto che la percezione musicale sia una forma complessa di analisi di una scena uditiva.

Da quanto detto, emerge con forza che la dimensione temporale dell'attenzione gioca un ruolo di assoluto primo piano. La musica, infatti, si svolge nel tempo, e attraverso di esso divengono progressivamente evidenti tutti i suoi contenuti di senso, che l'ascoltatore può percepire attraverso un'operazione di agglomerazione di stimoli secondo criteri e gerarchie (Longuet-Higgins & Lee, 1982; Povel & Essens, 1985; Imberty, 1993; Fitch, 2013). Altri lavori di ricerca (Savage et al., 2015; Ravignani et al., 2016; Kots et al., 2018) suggeriscono che la tendenza ad organizzare i suoni, individuando primariamente una periodicità fondamentale e poi una possibile suddivisione metrica, potrebbe essere ancorata a basi biologiche e quindi costituire una costante per la specie umana, che il training e/o la sola esposizione alla musica potrebbe quindi rafforzare.

Questa operazione di scomposizione e ricomposizione, definita da più parti con il termine inglese *chunking* (Talamini et al., 2017) è richiesta al suo massimo grado durante l'ascolto musicale, durante il quale si devono trovare strategie efficienti per gestire tutto il materiale sonoro in entrata, sia nella sua dimensione orizzontale che verticale.

La componente ritmica della musica, poi, esercita un'azione potente sull'attenzione, tanto che Pato e colleghi (2002) hanno riportato che i partecipanti

al loro studio riconoscevano meglio le altezze di alcuni suoni quando questi venivano presentati a intervalli di tempo regolari. Vi è, qui, una singolare assonanza con la Teoria dell'Attenzione Dinamica (Jones, 1976; Jones & Boltz, 1989; Jones, 2018), la quale postula che l'attenzione abbia delle fluttuazioni regolari nel tempo, secondo precise frequenze che danno origine a diversi livelli di sensibilità agli stimoli presentati in corrispondenza di tali pulsazioni o lontano da esse. La teoria ha ricevuto recentemente nuova linfa grazie agli studi sulle oscillazioni neurali di banda alfa, beta, gamma e delta, di cui abbiamo dato notizia nel precedente capitolo. Richiamiamo qui alcune evidenze particolarmente orientate alla percezione della pulsazione musicale, tra le quali il lavoro di Fujioka e colleghi (2015) che individua le oscillazioni di banda beta (intorno ai 20 Hz) come le responsabili della percezione della pulsazione sia durante l'ascolto che durante l'immaginazione musicale.

Significativamente, l'area motoria è la più interessata durante i processi di sincronizzazione ritmica (Nozaradan, 2013) e ha una forte influenza sull'attività riscontrata nella corteccia uditiva, indicando che i sistemi motori potrebbero guidare l'attenzione uditiva durante questo tipo di compito (Morillon & Baillet, 2017). Lo stesso tipo di fenomeno è stato evidenziato anche per la percezione del linguaggio, che presenta svolgimento e modulazioni temporali simili alla musica (Ding et al., 2017). Questi risultati hanno ispirato anche la composizione di musiche appositamente pensate per accentuare la corrispondenza con le oscillazioni cerebrali di banda alfa, beta, gamma e delta, con l'intenzione di utilizzarle in contesti di potenziamento o riabilitazione cognitiva (James et al., 2017). Sul fenomeno della sincronizzazione, che può esplicarsi sia cominciando a seguire una data pulsazione (da più parti valutato attraverso il *frequency following response*) sia seguendo un certo ritmo (compito di difficoltà anche significativamente maggiore rispetto al precedente), Miendlarzewska & Trost (2014) suggerisce che in questo tipo di attività siano implicate contemporaneamente abilità di anticipazione (a carico dell'attenzione e della memoria di lavoro), coordinazione motoria e previsione (basata sull'apprendimento di strutture e forme della musica).

Andando poi ad indagare quando e come l'attenzione sia coinvolta durante l'esecuzione musicale, dobbiamo innanzitutto premettere che tutte le considerazioni fatte in precedenza per l'ascolto possono ragionevolmente essere applicate anche al momento performativo. Infatti, mentre si suona uno strumento musicale o durante il canto, le abilità percettive nella modalità uditiva sono costantemente reclutate; tuttavia, oltre all'informazione uditiva, devono essere processati anche stimoli provenienti dalle modalità tattile e visiva (Luck & Sloboda, 2007; Koreimann et al., 2014; Lee & Lin, 2015; Bishop & Goebel, 2017). Inoltre, si va a chiudere e realizzare in pienezza il già citato *circuito sensomotorio*, in quanto ciò che si produce sul proprio strumento è il risultato di quel costante processo ricorsivo, che lega il gesto, il suono e il senso in un continuo aggiornamento e adattamento reciproco finalizzato ad ottenere il migliore risultato sonoro complessivo.

Nello specifico, la modalità visiva recluta l'attenzione dei musicisti attraverso diversi stimoli, elaborati con diversi livelli di rilevanza a seconda della situazione specifica in cui si trovano a suonare, e gli autori appena menzionati identificano una serie di possibili foci dell'attenzione visiva durante l'esecuzione. Uno di essi, anche se non necessariamente il primo per importanza, è il proprio strumento, inteso qui non tanto come elemento isolato ma, più propriamente, come parte di un assieme più ampio che comprende anche il corpo del musicista in una unità che è tangibile, sostanziale e produttrice di senso musicale. Infatti, si trae vantaggio dal controllo visivo sulla posizione complessiva dello strumento in relazione al corpo, ma anche su zone specifiche deputate alla produzione del suono (tipicamente le zone in cui agiscono le mani, ma non solo). È noto che una determinata posizione dell'assieme strumento-corpo ha un effetto significativo sulla qualità del suono prodotto, ed è quindi necessario un continuo controllo da parte dell'attenzione visiva, al fine di fornire un costante feedback per l'aggiornamento del circuito sensomotorio. Timmers & Li (2016) riportano anche come la pratica di lungo periodo su strumenti diversi possa influire sulla rappresentazione delle altezze su un piano orizzontale, a testimonianza di quanto il rapporto prolungato con il proprio strumento possa avere effetti anche sull'immaginazione mentale.

In particolare, i musicisti dirigono la propria attenzione visiva verso le proprie mani (anche soltanto una delle due) in maniera tipicamente non continuativa, ma abbastanza puntuale, probabilmente in concomitanza con passaggi di particolare difficoltà tecnica che richiedono maggiore precisione e quindi un controllo esecutivo maggiore, che si può raggiungere quando l'attenzione visiva e uditiva sono selettivamente orientate al medesimo oggetto (Spence & Driver, 1994; 1996).

Un altro focus attentivo visivo è rappresentato dalla partitura scritta, se presente. Nonostante la pratica di eseguire i brani a memoria sia piuttosto diffusa, tra i musicisti professionisti così come in determinati ambiti specifici come il jazz e la musica popolare, la lettura della partitura resta un indiscusso punto di riferimento visivo, attraverso il quale si possono ricavare molte delle informazioni fondamentali per una corretta esecuzione. Un interessante lavoro di Rayner & Pollatsek (1997) si poneva l'obiettivo di misurare l'abilità di lettura a prima vista di un brano musicale attraverso il tracciamento dei movimenti oculari sulla partitura. La coerenza tra i movimenti registrati e l'accuratezza dell'esecuzione (con particolare attenzione al tempo musicale) veniva qui considerata una misura di competenza musicale specifica, che si esplica nella traduzione di informazioni visive in gesti motori, a partire dall'ipotesi che la conoscenza di alto livello degli elementi musicali potesse fungere da mediatrice tra i due estremi. Ciò che effettivamente si misurava era la distanza (calcolata qui non in termini di spazio ma di frazioni di tempo musicale) tra il focus attentivo visivo e il punto corrispondente alle note effettivamente suonate in un dato istante, con più misurazioni ad istanti successivi. Il compito, similmente a quanto accade durante la lettura ad alta voce, richiede di stabilire un compromesso tra l'aumento dell'anticipo visuale e la capacità di ritenzione della memoria di lavoro. Un aumento eccessivo della distanza anticipatoria causerà un sovraccarico della memoria di lavoro, che può comportare la perdita di informazioni retrostanti, con la conseguente interruzione del flusso di elaborazione. Al contrario, un anticipo troppo piccolo comporta un minor carico della memoria di lavoro ma espone ad una maggior probabilità di interruzioni, soprattutto al sopraggiungere di passaggi più complessi. I risultati mostrano che, nella maggioranza delle misure, l'anticipo registrato è del valore musicale di un quarto,

l'occhio quindi è avanti di circa un quarto rispetto a quanto le mani stanno eseguendo. La distanza aumenta per passaggi facili e diminuisce per passaggi difficili, esattamente come nella lettura ad alta voce. La pulsazione, invece, viene processata con maggior anticipo, da un minimo di due battiti ad un massimo di cinque. Risulta quindi che l'attenzione uditiva può dispiegarsi contemporaneamente su più di un livello musicale (melodia e tempo), mentre l'attenzione visiva contribuisce in modo sostanziale all'esecuzione musicale, garantendone la fluidità e la correttezza in caso di lettura di un brano a prima vista.

Ancora un altro studio (Furneaux & Land, 1999) approfondisce il fenomeno, confrontando pianisti amatori ed esperti in un compito di lettura a prima vista, misurando il cosiddetto *eye-hand span*, ovvero, come nel caso precedente, la distanza tra il punto della partitura fissato dallo sguardo e quello eseguito in un certo istante di tempo. Rispetto allo studio di Rayner e colleghi, qui si misura sia la quantità di note di distanza, sia il tempo in valore assoluto, per valutare se vi siano differenze anche sotto questo aspetto. Inoltre, la tracciatura dei movimenti oculari era continua, e non discreta. I brani musicali selezionati erano di difficoltà proporzionale all'esperienza musicale dei due gruppi, per eliminare un elemento di disparità a monte, ma veniva chiesto anche di eseguirli a diverse velocità. I risultati sulla distanza misurata in tempo assoluto, come era lecito attendersi, indicano che, all'aumentare della velocità, la distanza occhio-mano si riduce per entrambi i gruppi, con la differenza che i pianisti esperti riuscivano ad ottenere maggiore accuratezza a parità di velocità, mostrando quindi maggior capacità di ritenzione di informazioni nel medesimo tempo. Il tempo assoluto di anticipo medio, per entrambi i gruppi, è di circa un secondo. Sulla quantità di note di distanza, invece, la differenza tra amatori ed esperti è più evidente, dato che mediamente gli amatori riuscivano ad anticipare di circa due note, al contrario delle 4 dei pianisti esperti. Un dato molto interessante, ai fini dello studio dell'attenzione, è che gli occhi dei partecipanti andavano a zig zag fra i due pentagrammi di cui è composta una partitura per pianoforte, quindi l'elaborazione verticale complessiva dovrebbe avvenire presumibilmente a livello più alto, a conferma dei risultati di altri studi simili, che mostravano come i pianisti leggano separatamente i due pentagrammi, e

l'elaborazione successiva e completa avviene in stadi successivi. Lo stesso fenomeno, secondo gli autori, porta a pensare che, in un medesimo tempo di latenza o *buffer*, i professionisti riescano a elaborare più informazioni, non soltanto una maggior quantità di note singole, ma una più fitta rete di rapporti tra le note, in modo anche predittivo, il che potrebbe essere una conseguenza del training di lungo corso.

Durante l'esecuzione musicale, oltre a focalizzarsi sul proprio corpo in relazione con lo strumento e alla partitura, sono stati identificati altri possibili oggetti dell'attenzione che, con valenze anche diverse, si contendono le risorse complessive. Kenny (2011), studiando il fenomeno dell'ansia da prestazione durante la performance musicale attraverso l'analisi di questionari somministrati a musicisti professionisti, identifica tre principali centri di attenzione; sé stessi, il pubblico e la musica. Quest'ultima risultava essere il più importante e il più ricorrente focus attentivo, e il fenomeno dell'ansia veniva interpretato come uno spostamento dell'attenzione dalla musica stessa verso il pubblico o verso sé stessi. L'autore suggeriva infine che, per mitigare o eliminare stati d'ansia durante l'esecuzione, i musicisti avrebbero dovuto accorgersi dello spostamento attentivo e invertirlo, per tornare a focalizzarsi soltanto sugli aspetti legati alla musica. Un altro passo in avanti è stato compiuto da Buma e colleghi (2015) che, a partire dai precedenti risultati, hanno condotto un'analisi mista qualitativa e quantitativa sulla stessa tematica, chiedendo ai musicisti professionisti di fare sia un brainstorming che un questionario su quali fossero i pensieri ricorrenti durante l'esecuzione musicale, soprattutto quando si avverte una certa pressione. L'analisi cluster dei concetti emersi ha evidenziato che le quattro aree su cui si focalizza più spesso l'attenzione erano: aspetti fisici, pensieri che danno sicurezza, pensieri disturbanti o di paura e, infine, aspetti legati strettamente alla musica suonata. Anche in questo caso, gli aspetti musicali avevano il peso maggiore (48% del totale), e i ricercatori sono andati poi più in profondità, cercando di capire cosa – del tutto musicale – fosse più determinante per l'attenzione durante la performance. All'interno della categoria "musica", gli stessi musicisti identificavano molti aspetti diversi, con un alto grado di variabilità, anche se la gran parte di essi ricade entro il concetto di

*flusso* (Kirchner, 2011). In altre parole, lo stato ottimale viene raggiunto quando si avverte un alto grado di assorbimento dato dal compito, tale che tutte le risorse sono impiegate per il medesimo scopo. In queste condizioni, prevale la gioia di condividere la musica con il pubblico e la percezione di essere all'interno di un evento relazionale positivo. Invece, la fonte di controllo maggiore durante gli stati d'ansia sembra essere la percezione del proprio respiro, la postura e l'ancoraggio a terra, e i musicisti vi portano l'attenzione quando avvertono la necessità di riorientarsi positivamente dopo essere, per qualche ragione, usciti dal flusso di cui si parlava poc'anzi.

Quando poi si tratta di esecuzione d'insieme, sia essa un semplice duetto o la più grande orchestra sinfonica tardo-romantica, i possibili centri di attenzione si moltiplicano, rendendo necessario il contemporaneo esercizio di attenzione selettiva, sostenuta e divisa. Bishop & Goebel (2018) si pongono l'obiettivo di studiare come l'informazione visiva data dai gesti precedenti l'inizio di un brano viene tradotta in informazione operativa riguardo al tempo di esecuzione e al modo in cui sincronizzarsi al primissimo attacco. Il campione era composto da duetti di violini, di pianoforti e misti. Si tratta quindi di considerare questi gesti iniziali alla stregua di cue visivi, che devono veicolare informazioni sia spaziali che temporali a colui che li riceve. Il movimento della testa, degli occhi e delle mani viene qui tracciato attraverso accelerometri e sensori Kinect. L'ipotesi era che la traduzione delle informazioni fosse più chiara nel caso di duetti con lo stesso strumento, a causa della grande conoscenza condivisa. I risultati mostrano come i duetti di violini siano i più sincronizzati, ma in generale appare che i gesti iniziali producono un sistema di codifica abbastanza chiaro e condiviso tra i diversi strumenti. Il battere, che corrisponde alla prima pulsazione, viene tipicamente suggerito attraverso una repentina accelerazione della testa verso l'alto sull'asse verticale (simile al gesto di annuire), mentre la velocità (tempo) del brano viene suggerito dall'ampiezza e dalla periodicità dei gesti iniziali. Gli autori suggeriscono che sia in atto una codifica sia spaziale che temporale dell'informazione visiva, che quindi è in grado di suggerire una codifica specificatamente musicale, che si traduce poi in gesti motori sul proprio strumento. È importante notare che, per comprensibili ragioni di carattere

sperimentale, sono stati qui indagati soltanto i gesti iniziali e quindi precedenti l’inizio del brano. L’osservazione visiva delle performance di insieme potrebbe testimoniare come, anche durante tutto lo svolgimento del brano e in particolar modo verso la sua conclusione, si possano ravvisare gli stessi meccanismi di codifica necessari ad ottenere un risultato musicale complessivo soddisfacente. Infatti, sono molti i momenti in cui l’attenzione del singolo deve dividersi – se non addirittura spostarsi completamente – per andare a ricavare informazioni dai gesti e dai suoni prodotti dai colleghi esecutori.

Si noti, ad esempio, il seguente passaggio di un celebre brano natalizio, arrangiato per coro e pianoforte (Benetti, n.d.).

**JINGLE BELLS**  
T T B B  
(Pianoforte ad libitum)

testo in italiano e arrangiamento:  
Tito Fiorenzo Benetti

Gioioso e scorrevole  
(\*1)

(\*2)

Nel-la Not - te di Na - tal da lon - ta - no sen - ti -  
Dash - ing through the snow, in a one horse o - pen

Drin  
Drin drin drin drin drin drin drin drin drin drin  
Drin Drin drin drin drin drin drin drin drin drin  
Drin Drin drin drin drin drin drin drin drin drin

Pianoforte  
(ad libitum)

Figura 3: Jingle Bells, arr. T.F. Benetti, primo estratto

Il secondo cantante basso (quarto pentagramma) è l’unico dei quattro cantori a dover eseguire una nota sul primo battere. Il secondo tenore e il primo basso (secondo e terzo pentagramma), che entrano in rapida successione con le loro prime note, non potrebbero ricavare il tempo del brano dalla sola nota iniziale prodotta

dal collega che ha iniziato. Infatti, una sola nota non può dare l'informazione sulla scansione temporale da seguire. È necessario avere almeno un altro riferimento e, in questo caso, esso è dato senza dubbio dai gesti precedenti l'inizio del brano. Come vedremo nelle prossime pagine, questa informazione preventiva può arrivare non solo dai colleghi ma, se presente, anche dalla persona che dirige l'ensemble. Andando poco più avanti nell'analisi della partitura, si vede che il primo tenore (primo pentagramma) esegue le sue prime note in fondo alla battuta numero due, cantando la parola "nella". Comunemente si potrebbe dire che egli "entra" a battuta due. In realtà, invece, ha dovuto entrare nel brano ben prima, dovendo ricavare l'informazione sul tempo (velocità) dagli altri cantori. Questo tempo dovrà essere scandito internamente, fino all'esecuzione della prima nota (Freschi & Neulichedl, 2012). Si noti, a tal proposito, che la durata delle note (o pause) musicali non è espressa in termini temporali assoluti (in secondi o suoi multipli e sottomultipli) bensì in termini relativi (frazioni di un intero) e questa relatività si esplica in modo marcatamente *relazionale*.

Infatti, l'interpretazione scelta dal gruppo potrebbe prevedere di non tenere la stessa velocità per tutta la durata del brano, ma la comunicazione non può certo avvenire verbalmente. Si presentano quindi momenti precisi del brano, nel quale l'attenzione visiva e uditiva deve spostarsi decisamente da sé verso l'insieme sonoro, oppure verso la persona che, attraverso gesti codificabili, darà l'informazione sul cambio di velocità. Così può accadere anche per le variazioni dinamiche (intensità del suono), anch'esse presenti in determinati punti del brano, nei quali è necessario far passare l'informazione a tutti gli esecutori attraverso input uditivi e/o visivi dati dall'interazione tra i componenti del gruppo. La fine del brano rappresenta poi un altro momento cruciale, in preparazione del quale spesso si assiste ad un progressivo intensificarsi di relazioni visive tra gli esecutori, propedeutiche al raggiungimento della migliore sincronizzazione possibile. A tal proposito, un altro estratto del brano precedente risulta particolarmente esplicativo:

The image shows a musical score for 'Jingle Bells' by T. F. Benetti, second extract. It features four vocal parts (Two Tenors and Two Basses) and a Piano accompaniment. The score is in G major (one sharp) and 4/4 time. The vocal parts have lyrics: 'Buon O - pen sleight! Buon O - pen sleight! Drin. Sciahh! Drin. Sciahh!'. The piano part has 'clom clom clom clom' lyrics. The score ends with a glissando in all parts.

Figura 4: Jingle Bells, arr. T. F. Benetti, secondo estratto

Dando uno sguardo all'ultima battuta, si nota che tutti gli esecutori (pianista accompagnatore compreso) eseguono contemporaneamente le proprie note sul battere della battuta, e tutte hanno la medesima durata (due quarti). I quattro cantori, successivamente, hanno tutti una pausa da un quarto, che prelude al glissando finale d'insieme. Esso non avrebbe molte probabilità di essere eseguito in sincronia perfetta, senza essere preceduto da una fitta rete di sguardi e gesti reciproci tra i cantori, che permettono ad ognuno di scandire internamente quella pausa allo stesso modo. Spesso, in questi casi, si assiste ad una sorta di "respiro comune" che permette proprio questa preparazione fisica all'esecuzione delle note finali. Tuttavia, la sincronia sulla scansione interna della pausa potrebbe non bastare per lo scopo prefissato, infatti essa da sola non può restituire l'informazione sul tempo (che ha bisogno di almeno due eventi consecutivi per essere codificata). Per questo motivo, anche durante l'esecuzione delle note precedenti potrebbe essere già in atto una comunicazione visiva tra i cantori. L'ipotesi è rafforzata dal fatto che le note che essi eseguono contemporaneamente hanno una durata doppia della pulsazione (due quarti, mentre la pulsazione base di questo brano è di un quarto), il che rende

necessario anche in questo caso scandire e *contare* internamente la corretta durata e assicurarsi che tutti abbiano contato i due quarti allo stesso modo. Sia l'esecuzione delle pause, sia la resa della corretta durata di ogni nota (specialmente quelle più lunghe) rimandano alla rappresentazione uditiva interiore (Brochard et al., 2004; Talamini et al., 2022) che ha per oggetto non eventi sonori fisici che provengono dall'esterno, bensì eventi sonori mentali immaginari, che tuttavia hanno ricadute tangibili nel reale, in quanto informano, anticipano e preparano le azioni motorie successive. I legami tra area uditiva e motoria, infatti, sembrano forti anche a livello di immaginazione mentale, e contribuiscono a *continuare* il flusso del movimento anche in presenza di pause o di note con durata maggiore della pulsazione. Questo fenomeno è stato riscontrato con chiarezza anche in contesti di musica d'insieme (Pecenka & Keller, 2009; Keller & Appel, 2010).

Come anticipato, un altro centro di attenzione per i musicisti che suonano in gruppo è costituito dal direttore/direttrice, in particolar modo dai gesti che compie con le mani e con la testa. L'informazione veicolata attraverso i gesti direttoriali comprende molti aspetti musicali come l'espressività, il carattere, le interazioni tra sezioni della stessa orchestra (Dahl et al., 2010). Luck & Toiviainen (2006) riportano pongono l'accento anche sulla dimensione temporale dell'informazione, indagando come essa viene tradotta e con quale grado di fedeltà. Essi erano interessati a studiare quale fosse la fonte di sincronizzazione primaria, tra quelle disponibili durante la performance orchestrale (gli altri musicisti, il direttore, lo spartito e il senso interno del ritmo) e, manipolando tre delle quattro variabili menzionate, trovarono che si ricava il tempo principalmente dagli altri musicisti, mentre il ruolo del direttore sembra quello di vigilare sul tempo, offrendo a tutti i musicisti un feedback visivo dato dai suoi gesti. Il suo contributo si fa però determinante in alcuni momenti chiave come l'inizio del brano, la fine e ogniqualvolta sia prevista una variazione di qualche parametro fondamentale come la velocità, l'intensità o quando il suono sembra "spostarsi" da una sezione all'altra. Anche Luck & Nte (2008) ha studiato proprio la capacità dei musicisti orchestrali di sincronizzarsi temporalmente con i gesti del direttore, codificando informazioni visive come la posizione delle mani nello spazio (ivi compresa la bacchetta, se

presente), la direzione e la velocità degli spostamenti effettuati. In questo studio erano previsti tre gruppi sperimentali: direttori, musicisti e non musicisti. Il compito prevedeva di osservare una varietà di gesti direttoriali, riprodotti via software su uno schermo, e ai partecipanti era richiesto di sincronizzarsi sul tempo proposto dai gesti e produrre dei battiti come prova dell'avvenuta sincronizzazione. I parametri spaziali e dinamici dei gesti direttoriali venivano manipolati per valutarne l'effetto sulla velocità e accuratezza di sincronizzazione, mentre anche i gesti dei partecipanti erano monitorati da un sistema di motion-capture in tre dimensioni. I risultati mostravano che la velocità e accuratezza di sincronizzazione era predetta efficacemente soltanto dall'esperienza orchestrale pregressa. Una delle limitazioni, ammesse dagli stessi autori, è che l'esperimento manca di alcuni elementi ecologici potenzialmente determinanti, infatti i partecipanti non suonavano il loro strumento e la sincronizzazione era affidata alle sole mani. Inoltre, mancava un centro di attenzione fondamentale costituito dallo spartito, anch'esso attinente alla modalità visiva e oggetto di attenzione divisa durante la performance (Rodrigues et al., 2007).

In uno studio più recente (Rodrigues et al., 2013), un gruppo di musicisti con significativa esperienza orchestrale è stato confrontato con un gruppo di controllo costituito da non musicisti su una serie di compiti attentivi che miravano a far emergere le tre diverse componenti dell'attenzione (selettiva, sostenuta e divisa) sempre nella modalità visiva. L'obiettivo dello studio era quello di valutare i benefici della pratica musicale separatamente per ciascuna delle componenti, infatti sono stati somministrati tre test. Lo studio verrà presentato dettagliatamente più avanti, qui però è da notare l'intento di associare le competenze acquisite durante la pratica orchestrale con l'attenzione divisa nella modalità visiva. Nel compito ad essa dedicato, in ciascuna prova veniva presentata a schermo la lettera "x" da sola, doppia o tripla, e contemporaneamente poteva essere presente la lettera "a" in una zona superiore dello schermo. Ai partecipanti era richiesto, prova dopo prova, di identificare il numero di "x" presentate (premendo il tasto con il numero corrispondente) e, contemporaneamente, di contare interiormente quante lettere "a" fossero state presentate complessivamente nel blocco di prove. Alla fine del blocco,

veniva chiesto ai partecipanti di scrivere questo numero su carta. I risultati mostrano come il gruppo dei musicisti era significativamente migliore dei non musicisti in due misure su tre, segnatamente nel tempo di risposta al numero di “x” e nel computo delle lettere “a”. Questi risultati sono stati interpretati dagli autori come un effetto del costante training attentivo visuale cui sono sottoposti i musicisti, in special modo coloro che suonano insieme. Infatti, come già è emerso in precedenza, l’attenzione visiva è quasi costantemente divisa tra il proprio spartito, l’assieme corpo-strumento, gli altri musicisti e il direttore, al fine di ricavare preziose informazioni da tradurre in azioni motorie finalizzate alla qualità musicale individuale e, soprattutto, collettiva.

### **3.1.2. Rassegna delle evidenze disponibili**

Dopo aver posto le basi che permettono di considerare con ragionevole confidenza la pratica musicale – nelle sue competenze fondamentali di ascolto, comprensione, esecuzione e invenzione – come un’attività strettamente connessa con l’attenzione nelle sue diverse componenti e modalità, vorremmo adesso passare in rassegna tutti gli studi che avevano come obiettivo primario o corollario quello di indagare gli effetti del training musicale sull’attenzione. Quello proposto, quindi, è un passo ulteriore: dopo aver esaminato quando, quanto e come l’attenzione venga reclutata nel momento stesso in cui si pratica una competenza musicale; adesso ci si chiede se il protrarsi nel tempo di questa pratica possa portare a differenze apprezzabili nelle abilità attentive, se confrontate con quelle possedute da un campione casuale di persone che non hanno alle spalle alcun tipo di training musicale significativo.

Lo studio di Nager e colleghi (2003) metteva a confronto un gruppo di direttori d’orchestra, uno di pianisti e infine uno di non musicisti (ogni gruppo era formato da sette partecipanti), per valutare se vi fossero differenze tra gruppi nell’elaborazione spaziale dei suoni. L’esperimento vedeva i partecipanti seduti di fronte ad un sistema di sei altoparlanti, tre posizionati in posizione anteriore frontale e tre in posizione laterale destra. All’interno di ciascun sottogruppo di altoparlanti, la distanza reciproca era di 6° di azimut. Gli altoparlanti in posizione perfettamente centrale (0° di azimut) e perfettamente laterale (90° di azimut destro) erano gli unici

ad essere oggetto di attenzione da parte dei partecipanti. Infatti, essi venivano istruiti, per ciascun blocco di prove, a fare attenzione ad uno dei due per tutta la durata del blocco. Veniva poi presentato un breve rumore rosa (dai 500 ai 5000 Hz) su uno dei tre altoparlanti della zona limitrofa alla posizione attesa, oppure rumore di frequenza maggiore ma probabilità minore (16% del totale degli stimoli), considerato deviante rispetto al target. Ai partecipanti era richiesto di premere un pulsante soltanto nel caso in cui il suono deviante fosse stato presentato in posizione attesa, in tutti gli altri casi era richiesto di non premere alcun pulsante. Oltre all'acquisizione di misure comportamentali classiche nello studio dell'attenzione, come i tempi di risposta e l'accuratezza percentuale, gli autori hanno raccolto anche misure elettrofisiologiche durante tutta la durata dell'esperimento, approssimativamente due ore e mezza. L'analisi dei dati comportamentali non rileva differenze statisticamente significative tra i tre gruppi per quanto riguarda i tempi di risposta (da notare che sono stati analizzati soltanto i dati delle risposte sulle posizioni target, non quelli delle risposte date erroneamente per stimoli presentati su altri altoparlanti). Invece, per quanto riguarda le accuratezze percentuali, sono stati analizzati i dati di tutte le posizioni, compresi gli altoparlanti non target, che davano origine ai cosiddetti *falsi allarmi*, cioè a risposte corrette quanto a discriminazione ma errate quanto alla posizione di presentazione del suono. In questo caso, i risultati mostrano una differenza significativa tra i gruppi, soprattutto tra i due gruppi di musicisti e quello dei non musicisti. Infatti, le percentuali dei direttori d'orchestra erano le più alte, non tanto nelle due posizioni attese, quanto nelle posizioni devianti. I pianisti esibivano risultati intermedi tra i tre gruppi, mentre i non musicisti mostravano le maggiori differenze sui falsi allarmi. Questi risultati portano gli autori a suggerire che la speciale pratica della direzione orchestrale, avvicinata soltanto dalla pratica del pianoforte quanto a estensione delle possibilità di selezione spaziale, può aver contribuito a rafforzare detta selettività, soprattutto in zone periferiche, per le quali si assiste ad un generale degradamento della performance, che però è meno evidente per direttori e pianisti. Per quanto riguarda le misure elettrofisiologiche, la differenza più evidente tra i tre gruppi riguardava la cosiddetta MisMatch Negativity (MMN), che indica l'azione

inibitoria per la zone lontane da quella del target. Mentre nel gruppo dei non musicisti questo effetto risultava debole, nei due gruppi con esperienza era chiaramente apprezzabile. Gli autori interpretano queste differenze come un possibile effetto della grande mole di training, che può aver contribuito ad una diversa elaborazione pre-attentiva, soprattutto nei primissimi momenti che seguono la presentazione di uno stimolo. Sembra, infatti, che se questa elaborazione non superi un certo livello-soglia minimo, non vi sia una cosciente identificazione dello stimolo, e questo livello-soglia sembra essere superato soltanto dai musicisti. Ad ogni modo, è da notare che sia l'esiguità del campione reclutato, sia la scelta di non includere la posizione a sinistra nel setting sperimentale, dovrebbero essere tenuti in considerazione per valutare la solidità dei risultati riportati.

Un altro gruppo di lavoro (Patston et al., 2007), basandosi su alcune evidenze che riportavano migliori abilità visuo-spaziali per i musicisti, probabilmente dovute alla pratica della lettura della partitura (Brochard et al., 2004), conduce due studi per valutare se i musicisti fossero meno soggetti al fenomeno della lateralizzazione dell'attenzione nella modalità visiva. Nel loro primo studio (Patston et al., 2006) essi mettono a confronto musicisti e non musicisti su un compito bisezione manuale di linee (BIT). Mentre i non musicisti dividevano le linee tendendo a sinistra con entrambe le mani, i musicisti erano generalmente più accurati con una leggera tendenza verso destra, mostrando un bilanciamento dell'attenzione maggiore in questo particolare compito visuo-spaziale. Sulla scorta di questi risultati, gli autori conducono un secondo studio (Patston et al., 2007) prendendo a campione un gruppo di musicisti con almeno 8 anni di training e capacità di leggere la partitura, per confrontarlo con un gruppo di non musicisti su un compito simile a quello usato da Brochard e colleghi. Ai partecipanti (tutti destrimani) era richiesto di osservare una linea verticale che appariva per breve tempo sullo schermo, alla quale seguiva la presentazione di un punto alla destra o alla sinistra della linea che, nel frattempo, era scomparsa. Il compito era quello di identificare il lato nel quale era stato presentato il punto, rispetto alla linea che quindi idealmente divideva lo spazio tra destra e sinistra. Due tasti fungevano da pulsanti di risposta per l'una o l'altra eventualità. Erano poi previste due condizioni di difficoltà, a seconda che il punto

venisse presentato più vicino o più lontano alla linea di bisezione. Venivano poi registrati i tempi di risposta e le accuratezze percentuali. I risultati mostravano che nell'emicampo sinistro le accuratezze di entrambi i gruppi erano migliori, mentre sulla destra i musicisti esibivano un minore degrado della precisione. Inoltre, i tempi di risposta del gruppo dei musicisti erano in generale più veloci. L'analisi dei risultati ha portato gli autori a concludere che i musicisti confermavano il maggior bilanciamento attentivo visuo-spaziale, compensando la naturale tendenza migliore a sinistra tipica dei destrimani (Jewell & McCourt, 2000; Nicholls et al., 2004; Nicholls et al., 2005). L'interpretazione sembra trovare conforto anche da studi neuroscientifici che mostrano come il lungo training musicale (soprattutto se iniziato in età precoce) possa produrre modificazioni anatomo-funzionali anche in aree coinvolte nell'attenzione visuo-spaziale come la corteccia parietale superiore (Gaser & Schlaug, 2003; Fan & Posner, 2004). Anche i dati sulle velocità di pressione delle due dita coinvolte nell'esperimento (una della mano destra e una della sinistra) portano gli autori a sostenere che i musicisti abbiano un maggior bilanciamento inter-emisferico. Una possibile causa di questo maggior bilanciamento è da ricercare, secondo gli autori, nella natura fortemente bimanuale della pratica musicale. Infatti, se si eccettuano alcuni strumenti molto particolari come il Kazoo o il theremin, tutti gli altri strumenti musicali prevedono l'uso delle due mani, con compiti solitamente diversi ma contemporanei ed egualmente necessari alla produzione del suono. Gli autori qui si aspettavano che pianisti esibissero questo effetto più degli altri strumentisti, in ragione della particolare posizione spaziale delle mani sullo strumento. Essi, però, non trovano riscontro statistico a questa ipotesi, probabilmente perché ciò che conta ai fini del bilanciamento non è tanto il posizionamento spaziale delle mani rispetto ad un ideale riferimento centrale (peraltro similmente riscontrabile in altri strumenti come l'organo, il clavicembalo, la chitarra, le percussioni, gli xilofoni e i metallofoni) quanto piuttosto alla contemporaneità delle diverse azioni e del sottostante controllo percettivo e motorio richiesto. È quindi probabile che una tale attività, iniziata in età evolutiva e protratta nel tempo, possa portare alle modifiche cerebrali mostrate

in precedenza, oltre che alla osservazione comportamentale di un minore effetto di lateralizzazione nell'attenzione visuo-spaziale.

Lo studio di Strait e colleghi (2010) si pone sulla scia del crescente consenso attorno all'ipotesi che l'attenzione possa modellare e guidare la sensibilità verso caratteristiche acustiche rilevanti per il compito da svolgere, di cui abbiamo reso conto nel precedente capitolo. Parallelamente, molte evidenze suggeriscono una migliore codifica sensoriale da parte dei musicisti. Gli autori, quindi, si chiedono se questi riconosciuti vantaggi in aree di basso livello possano essere legati ad abilità cognitive ben sviluppate e, per poter rispondere a questa domanda, mettono a confronto musicisti e non musicisti su una apposita batteria di test percettivi e cognitivi. Vengono manipolati sia il grado di controllo cognitivo richiesto dal compito, sia il grado di coinvolgimento della modalità visiva e uditiva. I risultati indicano che le maggiori differenze tra i due gruppi si registrano nei compiti uditivi correlati fortemente con abilità cognitive (compito di attenzione uditiva e compito di *backward masking*). Nel primo compito, i partecipanti tenevano fisso lo sguardo su un monitor che mostrava un personaggio dei cartoni animati, e venivano poi presentati casualmente due tipi di stimoli uditivi, uno dei quali richiedeva una rapida risposta attraverso un pulsante, l'altro andava ignorato. Nel secondo compito, basato sugli stessi stimoli visivi, ai partecipanti veniva presentato uno tra tre possibili rumori bianchi distraenti, talvolta preceduto o accompagnato dal suono target (un tono puro). Venivano poi manipolati i livelli di intensità dei suoni target e dei distrattori, per valutare il loro impatto sulla performance di rilevamento. Ai partecipanti, in questo caso, era richiesto di premere un pulsante ogniqualvolta avessero ravvisato la presenza del target. Venivano raccolte misure sui tempi di risposta e sulle soglie in Decibel. La migliore performance dei musicisti su questi due compiti, unita alla insussistenza di differenze tra gruppi sul QI non verbale e sulla memoria di lavoro uditiva, portano gli autori a concludere che il training musicale possa avere avuto un impatto sulle abilità cognitive, e che i vantaggi osservati precedentemente a livello percettivo possano ragionevolmente essere ascritti al miglior controllo cognitivo di alto livello. L'effetto è evidente soltanto nella modalità uditiva e non sembra qui propagarsi alla modalità visiva. Ulteriori

considerazioni e implicazioni di questo studio per quanto riguarda la contingenza tra l'affinamento della percezione uditiva dei suoni musicali e dei fonemi linguistici – che portano gli autori a sostenere che il training musicale possa avere un impatto positivo sulla mitigazione dei deficit linguistici – verranno discussi nel paragrafo dedicato.

Dello studio di Rodríguez e colleghi (2013), già discusso nel paragrafo precedente in merito all'associazione tra pratica musicale e attenzione divisa, occorre menzionare anche la restante parte dei risultati e delle interpretazioni proposte. Infatti, oltre al test sull'attenzione visuale divisa, gli autori somministrano anche altri due compiti. In quello per l'attenzione selettiva, un compito tipico di ricerca visuale, venivano presentati a schermo l'immagine target e, poco sotto, una serie di cinque immagini simili, che potevano includere l'immagine target. Se presente, ai partecipanti era richiesto di premere un tasto il più velocemente possibile. Nel compito di attenzione sostenuta, invece, venivano presentati alternativamente tre numeri target insieme a due distrattori (operatori matematici come + o =) e i partecipanti dovevano rispondere alla presentazione dei numeri premendo il tasto corrispondente. Per evitare che una diversa coordinazione manuale tra gruppi potesse influire sui risultati, gli autori hanno somministrato anche un test di risposta semplice. Le variabili misurate nei test erano i tempi di risposta e le accuratezze percentuali. Oltre alle già citate differenze, che vedevano i musicisti migliori in due misure su tre del test sull'attenzione divisa, si registrano qui differenze significative anche nell'accuratezza del test di attenzione selettiva e nei tempi di risposta del test di attenzione sostenuta. Invece, non si registravano differenze significative per quanto riguarda il test di risposta semplice. Inoltre, da un'analisi più approfondita emerge la significativa correlazione tra l'età di inizio del training e la performance dei musicisti. La pratica musicale, riconosciuta anche qui come altamente multidimensionale e impegnativa potrebbe, secondo l'interpretazione data, aver influito positivamente sulle abilità attentive, soprattutto per quanto riguarda l'efficienza dell'utilizzo delle risorse disponibili istante per istante. La correlazione con l'età precoce di inizio del training va a confermare la letteratura esistente, che già da qualche anno indicava come la neuro-plasticità indotta dalla pratica musicale

potesse portare a modifiche anche comportamentali in aree non solo percettive ma anche più cognitive come l'attenzione. Resta, anche agli autori stessi, una possibile limitazione relativa al fatto che il gruppo di musicisti era composto esclusivamente da membri stabili di orchestre sinfoniche. Si tratta, quindi, di criteri di selezione alquanto stringenti, che implicavano non soltanto l'esclusione di solisti, ma anche la poca varietà del tipo di formazione musicale ricevuta che, in questo caso, è ragionevole ritenere che sia stata per tutti classica e in ambito accademico. Questi fattori possono aver contribuito a spostare l'ago della bilancia verso l'affidabilità sperimentale, a scapito della possibile estendibilità del campo di applicazione dei risultati ottenuti.

Nello stesso anno, uno studio di Kaganovich e colleghi (2013) si pone l'obiettivo di stabilire se il training musicale possa migliorare la codifica della voce umana e se abbia anche delle ricadute positive sull'abilità di ignorare cambiamenti uditivi irrilevanti ai fini del compito da svolgere. Per raggiungere lo scopo, gli autori hanno utilizzato una particolare versione del paradigma di distrazione uditiva (Schroger e Wolff, 1998; 2000) nel quale i partecipanti sono chiamati a discriminare i suoni (vocali e di strumenti musicali) in base alla durata, ignorando allo stesso tempo cambiamenti di timbro. In questo compito, venivano inseriti anche dei suoni digitalmente modificati per ottenere una rotazione dello spettro delle frequenze. Sono state misurate alcune componenti ERP comunemente associate alla elaborazione uditiva precoce (N1) e altre associate alla distrazione e successivo riorientamento sulla categoria rilevante per il compito (P3a, P3b e RON), oltre che le classiche misure comportamentali (tempi di risposta e accuratezza percentuale). Il gruppo dei musicisti veniva quindi confrontato a quello dei non musicisti su tutte queste misure, raccolte durante il test somministrato. I risultati comportamentali mostrano come entrambi i gruppi sono più veloci ed accurati quando operano su suoni vocali, rispetto a quelli strumentali. Il gruppo dei musicisti tuttavia risultava generalmente più accurato in entrambe le condizioni. In presenza di suoni devianti, invece, i musicisti non esibivano differenze di velocità e accuratezza tra suoni vocali e strumentali, al contrario dei non musicisti. La componente N1 risultava sempre più ampia nei musicisti, così come la RON, mentre non si ravvisavano

differenze nelle componenti P3a e P3b. I risultati portano gli autori a suggerire che il training musicale possa migliorare non solo la codifica di suoni musicali, ma anche di quelli vocali, lasciando quindi intravedere una generalizzazione della competenza acquisita con stimoli dominio specifici. Una caratteristica acustica che potrebbe essere all'origine di questa competenza è la codifica della complessità spettrale, già riconosciuta come una possibile specificità dei musicisti (Shahin et al., 2005; Caclin et al., 2005; 2008; Strait et al., 2012). Per quanto riguarda, invece, la questione del riorientamento dell'attenzione dopo stimoli devianti, gli autori suggeriscono che i musicisti potrebbero avere una migliore capacità di resistere a stimoli devianti, se questi appartengono a categorie più familiari. Infatti, non si notano differenze ERP significative tra gruppi quando gli stimoli devianti sono costituiti da vocali. La differenza tra i gruppi sulla componente RON viene invece interpretata come una migliore capacità dei musicisti di disimpegnare l'attenzione da un suono distraente e tornare sul focus, sulla scia di precedenti risultati che riportavano risultati e conclusioni simili (Strait & Kraus, 2011).

Interessante risulta anche lo studio di Ouimet e colleghi (2012), che indaga se l'effetto globale-locale, già riportato nelle ricerche sull'attenzione visiva, possa replicarsi anche nella modalità uditiva e se il training musicale possa eventualmente portare a differenze nell'andamento tipicamente osservato per una popolazione non musicalmente esperta. L'effetto in questione si trova quando sono presenti due livelli di elaborazione attentiva diversi, uno mirato a caratteristiche e oggetti percettivi specifici, l'altro alla percezione di caratteristiche globali e comuni a tutta la scena percettiva. Per la vista, le caratteristiche globali vengono elaborate prima di quelle locali, ma al tempo dello studio non era chiaro se questo effetto fosse ipotizzabile e verificabile anche per l'udito. In aggiunta a questa domanda di ricerca, gli autori si chiedono anche se i musicisti possano mostrare differenze di elaborazione rispetto ai non musicisti, in particolare in un diverso bilanciamento tra elaborazioni locali e globali. A questo scopo, propongono due compiti, uno di attenzione selettiva e l'altro di attenzione divisa. La scelta degli stimoli uditivi rappresenta lo sforzo di uscire dall'alveo della musica tonale occidentale, in quanto le note utilizzate per i compiti appartenevano ad una scala di quattordici note,

artificialmente costruita per evitare gli intervalli tipici della musica tonale. Tuttavia, manca un passo importante in direzione ecologica, poiché queste note non sono state riprodotte da strumenti musicali reali, bensì elaborate e riprodotte digitalmente. Tali suoni, infatti, non sono presenti in natura e non possono quindi essere considerati parte del materiale di apprendimento – attivo o passivo che sia – dell’uno e dell’altro gruppo sperimentale. Il compito di attenzione selettiva richiedeva ai partecipanti di fare attenzione ad una serie di tre gruppi con tre note ciascuno, identificando se l’andamento era ascendente o discendente. L’andamento locale era rappresentato dai singoli terzetti di note, mentre quello globale si percepiva considerando la nota iniziale di ogni terzetto. Due tasti erano utilizzati per le rispettive risposte. Nel compito di attenzione divisa, invece, ai partecipanti era anticipato l’andamento target (ascendente o discendente) del quale essi dovevano poi ravvisare la presenza o l’assenza a livello locale o globale. Da notare che gli andamenti locale e globale potevano essere concordi o discordi tra loro, generando così condizioni compatibili, neutre o incompatibili. I tempi di risposta e le accuratezze percentuali venivano quindi raccolti per ogni condizione. I risultati mostravano che, in generale, l’accuratezza a livello globale era maggiore rispetto al livello locale. I tempi di risposta, invece, erano generalmente più rapidi per l’analisi globale rispetto a quella locale. L’accuratezza dei musicisti era generalmente maggiore rispetto ai non musicisti, con una grossa differenza per l’analisi locale, per la quale i non musicisti avevano un calo drastico di precisione. Le differenze sui tempi di risposta, invece, erano significative soltanto per l’analisi locale, con i musicisti significativamente più veloci dei non musicisti. I dati portano gli autori a concludere che l’effetto di precedenza e facilitazione globale nella elaborazione attentiva è presente anche nella modalità uditiva, e che i musicisti – in ragione della loro specifica competenza uditiva – si mostrano meno soggetti a questo effetto, avendo una capacità di elaborazione locale molto più efficiente ed efficace. Una importante direzione di ricerca futura potrebbe essere, in questo caso, l’adozione di stimoli uditivi reali e non costituiti da toni musicali, bensì da suoni ambientali, i quali assai spesso danno origine a schemi e andamenti riconoscibili, sia a livello locale che globale. Questo permetterebbe di affermare con maggiore affidabilità

che il training musicale comporti un diverso bilanciamento tra elaborazione attentiva globale e locale, poiché si tratterebbe effettivamente di un trasferimento di competenza verso stimoli non musicalmente specifici. Peraltro, questa ipotesi sarebbe rafforzata anche dal fatto che la pratica musicale richiede costantemente, in ogni brano, più livelli di elaborazione. Infatti, anche sul proprio strumento è necessario rendere contemporaneamente l'orizzontalità e la verticalità della musica, che spesso contiene gruppi di note con un proprio andamento interno, da inserire però in un andamento globale di livello superiore.

Roden e colleghi (2014), interessati a valutare gli effetti di un programma di apprendimento musicale in contesto scolastico su una serie di abilità cognitive tra le quali anche l'attenzione, assegnano due diversi gruppi di bambini dai sette agli otto anni a due distinti interventi educativi: uno basato sulla musica e l'altro sulle scienze naturali. Una serie di test socioeconomici, motivazionali e cognitivi, nonché le abilità musicali, sono stati somministrati in ingresso per fornire una efficace base dati di riferimento. I due gruppi sono stati poi seguiti per diciotto mesi e testati tre volte in questo arco temporale. Oltre ad un test sulla velocità di elaborazione delle informazioni (Zahlen-Verbindungs-Test, Vernon, 1993), il compito prescelto per la valutazione dell'attenzione selettiva visuale era il "d2" (Brickenkamp, 2002). Si tratta di un test grafico di cancellazione. Durante il test, ai partecipanti viene chiesto di identificare i target all'interno di sequenze di lettere "d" e "p", segnando (cancellando) il target con una matita. Ogni lettera ha da uno a quattro trattini verticali nelle sue immediate vicinanze, o sopra o sotto il carattere. Solo le "d" accompagnate da esattamente due trattini (due sopra, due sotto o una sopra e una sotto la lettera) devono essere cancellate. Ad essi si aggiungeva una parte del test di "audiation" di E. Gordon (1979; 1989) e un test di abilità cognitiva. I risultati mostrano differenze significative tra gruppi per quanto riguarda l'attenzione selettiva visuale in favore del gruppo di scienze naturali, mentre il gruppo musicale ottiene migliori performance nel test di velocità di elaborazione delle informazioni. I risultati portano una conferma e un rafforzamento delle precedenti evidenze a sostegno di una migliore e più efficiente elaborazione delle informazioni da parte dei musicisti, poiché in questo caso il gruppo di controllo non era passivo ma

seguiva a sua volta un programma specifico. Di contro, la mancanza di un effetto superiore del programma educativo musicale sulle misure di attenzione visiva porta al pensiero che i benefici di tale apprendimento possano essere limitati alla modalità uditiva, come già altre volte è stato evidenziato. In ogni caso, il disegno di ricerca longitudinale e l'equivalenza delle misure in ingresso tra i due gruppi, sono due argomenti a supporto dell'efficacia dell'intervento educativo (sia musica che scienze naturali), dato anche che le misure finali sono mediamente superiori a quelle riportate da studi sullo sviluppo per quella specifica fascia di età. Questo potrebbe avere positive implicazioni sull'opportunità di adozione di questi programmi in ambito scolastico su vasta scala.

Proseguendo, e al contempo addentrandoci nell'ultimo decennio di ricerche, troviamo un importante contributo di Martens e colleghi (2015), che si chiedono se i benefici dell'apprendimento musicale sull'attenzione e sulle funzioni esecutive, riportati fino ad allora in letteratura, siano limitati alla modalità uditiva o si estendano – in qualche forma – anche alla modalità visiva. In particolare, essi studiano la temporalità dell'attenzione (vedi cap. 1 e, in particolare gli studi di Best et al., 2008 e Shinn-Cunningham, 2008) supponendo che i musicisti abbiano particolare esperienza di questa abilità, acquisita durante il training. Uno dei fenomeni che meglio descrivono le abilità attentive nel loro svolgersi temporale è l'*attentional Blink* (Raymond et al., 1992; Martens & Wyble, 2010), ovvero la difficoltà a riportare il secondo di due target, quando essi vengono presentati in rapida successione. Gli autori, quindi, reclutano due gruppi di volontari (musicisti e non musicisti) e somministrano due versioni del test attentional blink (abbreviato successivamente con l'acronimo "AB"), una per la modalità visiva e l'altra per la modalità uditiva. Nel test uditivo, in ogni prova veniva presentata una serie di ventidue suoni in rapidissima successione, costituiti da numeri o da lettere pronunciati da voci registrate. Le lettere (in questo caso due) costituivano i target, mentre i numeri erano i distrattori. Il compito dei partecipanti era quello di identificare se e quante lettere erano presenti nel flusso ascoltato. Infatti, poteva essere presentata solo una lettera, due lettere o nessuna lettera. Per rispondere, i partecipanti dovevano premere i tasti corrispondenti alle lettere nel minor tempo

possibile. Il compito visivo, similmente, prevedeva la presentazione a schermo di una serie di numeri (distrattori) e lettere (target) con brevissimi intervalli temporali, e ai partecipanti era richiesto di rispondere nel medesimo modo dell'esperimento precedente. I risultati nel test uditivo mostrano che i musicisti soffrono meno il fenomeno dell'AB, che risulta sia attenuato che ritardato in funzione dell'intervallo tra distrattore e target. Nel test visivo, invece, si trova l'effetto opposto: i musicisti soffrono maggiormente dei non musicisti. Inoltre, non musicisti con valori di AB alti nella modalità uditiva, tendono a valori alti anche per la vista, cosa che invece non accade per i musicisti. Tutto questo porta gli autori a sostenere che i benefici del training musicale sull'attenzione temporale sembrano limitati alla modalità visiva, e che la similarità tra gli effetti AB tra modalità visiva e uditiva potrebbe essere soltanto frutto di differenze individuali o di elaborazioni attentive distinte per ciascuna delle due modalità, piuttosto che la conseguenza di un funzionamento centrale e sopra-modale. In ogni caso, è da notare che l'adozione di stimoli reali e non specificamente musicali contribuisce positivamente all'ipotesi che le abilità acquisite durante il training possano generalizzarsi, perlomeno nella modalità uditiva. Un altro pregio dello studio, anche se si tratta solo di un suggerimento di ricerca futuro, riguarda la domanda su quali aspetti specifici del training musicale possano influire sull'attenuazione dell'AB, in altre parole si nota l'intenzione di non considerare la formazione musicale come un intero indivisibile, ma come l'insieme di tanti possibili elementi di competenza, ciascuno dei quali potrebbe contribuire in maniera più e meno significativa allo sviluppo di abilità più generali.

Carey e colleghi (2015) compiono un'altra indagine ad ampio raggio, per valutare quanto le competenze acquisite attraverso il training musicale possano generalizzarsi o, piuttosto, restare ancorate ad aree di abilità contigue a quelle musicali. In particolare, gli autori sono qui interessati a comprendere se vi siano differenze tra musicisti e non musicisti riguardo l'attenzione – considerata anche qui una funzione di alto livello costantemente e intensamente reclutata durante il training- e anche riguardo al già discusso impatto dell'attenzione sulle abilità di percezione di caratteristiche acustiche di base. Se cioè, i musicisti possano mostrare diversi meccanismi di influenza reciproca tra processi di alto e basso livello, come

da altre parti suggerito (Tervaniemi et al., 2005; 2009; Marie et al., 2011; Seppänen et al., 2012). Un campione di tre gruppi, di cui un gruppo di violinisti, uno di pianisti e il gruppo di controllo formato da non musicisti, è stato sottoposto ad una batteria di test. Troviamo un compito SIMON di seriazione audio-visuale, un test percettivo di intonazione, un test di analisi della scena uditiva ambientale, accompagnati da un test audiometrico tonale e un test di soglie visive che fungevano da riferimento. Nel compito SIMON, ai partecipanti veniva mostrato a schermo un bottone colorato associato ad un preciso suono, e veniva chiesto di premere il tasto del colore corrispondente. Successivamente, veniva aggiunto un secondo bottone colorato unito ad un altro tono, e i partecipanti dovevano premere i corrispondenti tasti in sequenza. La serie aumentava di numero fino ad un massimo di sei elementi, oppure fino all'errore del partecipante, che doveva quindi riprodurre con i tasti serie sempre più lunghe. Nel test di percezione, veniva in vece chiesto ai partecipanti di valutare la corretta o errata intonazione di un accordo di modo maggiore, eseguito da strumenti ad intonazione fissa (come il pianoforte) o ad intonazione variabile (come ad esempio il violino). Nel test di analisi della scena uditiva, ai partecipanti venivano fatti ascoltare preventivamente uno, due o tre suoni target, che andavano poi riconosciuti durante l'ascolto dell'intera scena uditiva nella quale erano inseriti. Infine, il compito di attenzione uditiva sostenuta prevedeva la presentazione casuale di stimoli target e non-target, conosciuti preventivamente dai partecipanti, i quali dovevano rispondere premendo tasti diversi con mani diverse, a seconda che il suono ascoltato fosse un target o non target. I risultati, specialmente quelli relativi al compito di attenzione sostenuta, non mostrano differenze significative tra musicisti e non musicisti riguardo alle classiche misure di velocità (tempi di risposta) e accuratezza percentuale. L'unica differenza statisticamente significativa riguarda la variabilità dei tempi di risposta (espressa attraverso la deviazione standard), meno marcata nei musicisti rispetto ai non musicisti. Anche il compito di analisi della scena uditiva non rileva differenze tra gruppi, al contrario di alcune evidenze precedenti, già analizzate nel presente lavoro. Tuttavia, sono evidenti differenze nei test percettivi, ma solo quando gli stimoli utilizzati sono attinti dal repertorio sonoro degli strumenti padroneggiati, ma questo risultato era abbastanza

prevedibile. Ad ogni modo, esso porta gli autori a dire che il training specifico su uno strumento ha un impatto sulle abilità percettive di base, orientandone in qualche modo la sensibilità verso i timbri tipici di quel medesimo strumento. A parte questa differenza, violinisti e pianisti mostrano qui performance confrontabili, probabilmente motivate dalla non trascurabile sovrapposizione di apprendimento tra i due gruppi. Infatti, in ambito accademico lo studio del pianoforte è obbligatorio per qualunque studente di violino almeno per due anni. Non vi è quindi soltanto sovrapposizione a livello di training soltanto uditivo, ma anche a livello performativo. Questo fatto, almeno in parte, potrebbe spiegare l'andamento dei risultati dei due gruppi, poiché la distinzione operata in ingresso è almeno in parte artificiosa. Inoltre, almeno per quanto riguarda il compito sull'analisi della scena uditiva, gli stessi autori ipotizzano che la relativa semplicità di esecuzione, se comparata con altri compiti simili, potrebbe aver avuto un peso nella non constatazione di differenze tra gruppi. Ad ogni modo, questi risultati sono tendenzialmente in controtendenza rispetto alla maggioranza della letteratura disponibile, non tanto nella parte riguardante un eventuale trasferimento lontano di competenze (argomento da più parti dibattuto), quanto per i risultati sulle abilità percettive di livello basso, già riconosciute come molto attinenti alla pratica musicale e come ambiti del trasferimento vicino di competenza. Su quest'ultimo aspetto, la relativa semplicità dei compiti utilizzati potrebbe aver influito anche più di quanto ipotizzato dagli autori in discussione generale.

Anche lo studio di Clayton e colleghi (2016), che sarà analizzato ulteriormente nel prossimo paragrafo, mirava a stabilire anche una possibile correlazione fra il training musicale e l'attenzione selettiva visuale, valutata attraverso il compito di *multiple object tracking* (MOT). In questo compito è richiesto ai partecipanti di individuare da uno a cinque puntini su uno schermo, che venivano resi noti all'inizio di ciascuna prova evidenziandoli con il colore verde, rispetto ad altri puntini non target colorati di grigio. Successivamente, tutti i puntini diventavano grigi e si spostavano casualmente sullo schermo. Infine, riapparivano i puntini target verdi e i partecipanti dovevano cliccare con il mouse sopra ognuno di essi nel più breve tempo possibile. I risultati di questo specifico compito non mostrano

differenze tra gruppi, né in condizioni facili (un solo puntino da individuare) né in condizioni gravose (cinque puntini da individuare). Anche in questo caso, quindi, non si ravvisa un trasferimento lontano di competenza verso la modalità visiva, che sicuramente non può essere definita specificatamente musicale, anche se ampiamente coinvolta durante l'ascolto e l'esecuzione. Gli autori suppongono che la variabilità nei disegni sperimentali e nel tipo di compiti e stimoli utilizzati possa in parte aver influito sull'inconsistenza dei risultati riguardo possibili associazioni fra training musicale e attenzione visiva. Inoltre, la letteratura esistente spesso indica la modalità uditiva (non coinvolta in questo compito) come la mediatrice di un possibile trasferimento lontano di competenza.

Vi è poi un interessante studio (Celma-Miralles & Toro, 2019) che si pone l'obiettivo di indagare come il cervello interagisce con suoni separati nello spazio e integra le informazioni per cercare l'eventuale presenza di una pulsazione e di una struttura metrica. Già in precedenza altri ricercatori (Fitch, 2013) avevano mostrato come durante un ascolto musicale, si tenda a estrapolare prioritariamente una pulsazione o *beat*, grazie alla scansione ripetuta della scena uditiva, che rileva regolarità temporali e comincia a mettere in relazione eventi sonori legati da tali regolarità, non solo in modo predittivo ma anche retroattivo. In seguito, si tende a stabilire gerarchie tra eventi sonori regolari, distinti per intensità, ed emerge così il metro, ovvero la scansione temporale di eventi *forti* e *deboli*. Questa capacità di organizzare i suoni secondo gerarchie e strutture metriche è riconosciuta da più parti come una competenza universale della specie umana, con forti basi biologiche (Savage et al., 2015; Ravignani et al., 2016; 2018). Celma-Miralles e colleghi sono quindi interessati a studiare questo fenomeno sia da un punto di vista comportamentale che elettrofisiologico, attraverso la somministrazione di due esperimenti di percezione metrica, di cui uno soltanto uditivo e l'altro cross-modale (uditivo e visivo). Sono inoltre interessati a capire quanto il training musicale possa modulare questa capacità di percezione. Un gruppo di musicisti e un gruppo di controllo vengono sottoposti ad un primo esperimento, nel quale viene presentato un metro ternario in cuffia, in quattro diverse condizioni spaziali: la prima condizione di controllo prevedeva la presentazione di tutti e tre i suoni nella

posizione di 0° di azimut (davanti al partecipante), mentre le altre tre condizioni prevedevano la presentazione del suono in battute da un lato e dei due suoni in levare simmetricamente dall'altro lato (+/- 30°, +/- 60° e +/- 90° di azimut). Ai partecipanti era semplicemente richiesto di fare attenzione ai suoni e, alla fine dell'esperimento, di riportare se avessero o meno sentito l'alternanza destra-sinistra dei suoni e se avessero o meno rilevato una regola per questa alternanza. Nel frattempo, venivano acquisiti i dati EEG. I risultati indicano chiaramente una sintonizzazione ritmica cerebrale, con valori di picco corrispondenti alla frequenza della pulsazione e alternati tra destra e sinistra seguendo la distribuzione spaziale dei suoni. Inoltre, la maggiore ampiezza dei picchi per i musicisti suggerisce agli autori un effetto del training musicale che facilita la percezione sia della pulsazione che del metro, senza peraltro ravvisare differenze che inducano a pensare ad un effetto di lateralizzazione. Dato che questo primo esperimento non poteva distinguere efficacemente i processi percettivi da quelli attentivi, gli autori propongono un secondo esperimento nel quale, oltre alla presentazione degli stimoli uditivi precedenti, veniva presentato anche uno stimolo visivo distraente, costituito da episodi tratti dalla serie "La pantera rosa", proiettati a schermo in modalità muta. Alcune ricerche precedenti, infatti, mettevano in risalto differenze tra musicisti e non musicisti ad un livello pre-attentivo (Vuust et al., 2012) mentre altre mostravano risposte neurali a stimoli spazialmente distribuiti anche quando l'attenzione era fissa su stimoli visivi (Munte et al., 2001; 2003). L'intento, perciò, era quello di verificare se le risposte neurali fossero riconducibili soltanto a processi di basso livello oppure fosse presente anche un intervento dell'attenzione ad un livello più alto. I partecipanti erano quindi invitati a fare attenzione soltanto al video, e alla fine dell'esperimento venivano poste le stesse domande del primo esperimento, accompagnate da un questionario sugli episodi de "La pantera rosa", per assicurarsi che l'attenzione al video fosse stata efficace. I risultati di questo secondo esperimento mostrano attivazioni di picco coerenti con la frequenza del metro e della pulsazione musicale soltanto nel gruppo dei musicisti, anche se non si ravvisano differenze significative tra gruppi sulle quattro condizioni sperimentali. Inoltre, l'analisi comparata degli effetti del primo e secondo esperimento evidenzia

il ruolo dell'attenzione nell'incrementare l'ampiezza delle attivazioni neurali, soprattutto in caso di distribuzioni spaziali intermedie ( $\pm 30^\circ$  o  $\pm 60^\circ$ ) per le quali sembra che la presenza del video e dell'attenzione fissata alla modalità visiva abbia determinato una parziale compromissione dell'abilità di percezione metrica. Gli autori concludono quindi che la percezione del metro, ancor più di quella della pulsazione, è un processo che richiede non solo abilità percettive di basso livello, ma anche abilità superiori di raggruppamento ascrivibili all'intervento dell'attenzione. Maggiore è l'esperienza pregressa in ambito di percezione ritmica e metrica, minore è il carico attentivo richiesto, in quanto l'acquisizione di un determinato schema metrico può entrare a far parte del patrimonio di conoscenza ed essere richiamato e applicato al momento opportuno durante un compito uditivo. Non si può escludere, tuttavia, un possibile contributo della memoria uditiva, che potrebbe aver facilitato in partenza i musicisti nel riconoscimento dello schema metrico, e questo avrebbe potuto essere verificato mettendo a confronto i trend di miglioramento dei due gruppi durante l'esecuzione dei compiti. se infatti vi fossero trend simili, un eventuale ruolo della memoria pregressa sarebbe stato da escludere.

Molto recentemente, uno studio di Krzyżak (2021) riprendeva il dibattito sulla lateralizzazione dell'attenzione e della percezione, confrontando musicisti e non musicisti per valutare eventuali differenze tra gruppi attraverso una variante del compito di ascolto dicotico adottata in precedenti ricerche (Kurkowski, 2007). Il compito, che prevedeva l'utilizzo di cuffie per la presentazione degli stimoli sonori, era suddiviso in tre condizioni distinte: nella prima non si richiedeva ai partecipanti di fare attenzione ad un lato specifico, ma di ripetere tutte le sei parole monosillabiche riprodotte; nella seconda e nella terza condizione, invece, si richiedeva ai partecipanti di fare attenzione ad uno specifico lato (sinistra o destra) e di ripetere soltanto le parole riprodotte dal lato atteso. I risultati del test nella condizione di assenza di focus attentivo, mostrano una migliore performance a destra per i non musicisti, mentre i musicisti ottengono una performance del tutto simile tra destra e sinistra. Invece, nelle condizioni di attenzione focalizzata e destra o a sinistra, le performance non differiscono tra i gruppi, ma differiscono leggermente per quanto riguarda la performance generale, con maggior accuratezza

nel caso di attenzione rivolta verso destra. L'andamento dei risultati sembra coerente con le evidenze precedenti che indicano una dominanza dell'orecchio destro (emisfero sinistro) per l'elaborazione degli stimoli uditivi relativi al linguaggio (Mukari et al., 2006; Moncrieff, 2009; Romero-Diaz et al., 2011). Per quanto riguarda le differenze tra gruppi, i musicisti mostrano una performance più omogenea tra destra e sinistra nella condizione di assenza di focus attentivo, dovuta probabilmente al training, il quale a sua volta potrebbe aver alterato la tipica risposta neurale agli stimoli uditivi, rendendola meno soggetta alla dominanza emisferica sinistra (Onishi et al., 2002; Elmer, 2013). In condizioni di attenzione focalizzata, invece, il training musicale non sembra aver influito sull'andamento dei risultati.

### **3.2. Su altre funzioni cognitive**

Oltre agli studi che hanno provato a stabilire nessi e correlazioni tra la pratica musicale (variamente intesa) e l'attenzione, troviamo in letteratura numerose evidenze di benefici su una vasta gamma di funzioni cognitive (Swaminathan & Shellenberg, 2018). Il compito del presente paragrafo sarà quindi quello di offrire una rassegna aggiornata delle ricerche sull'argomento, che da quasi cinquant'anni ormai si sforzano di fare sempre nuova e maggior luce su ciò che il training musicale, specialmente se iniziato in tenera età e protratto nel tempo, può apportare a livello comportamentale e neurologico alla persona che lo intraprende. I risultati saranno divisi per macro-aree, poiché molte e diverse sono le direzioni e le abilità (generali e specifiche) per le quali si è tentato di valutare possibili associazioni o nessi causali. Non è semplice, infatti, rintracciare uniformità nelle domande di ricerca, nei metodi, negli strumenti e, di conseguenza, negli esiti e nelle interpretazioni. Questa estrema variabilità, se da un lato risulta problematica e certamente reclama un contatto più stretto tra le diverse professionalità coinvolte in tali ricerche, dall'altro può rappresentare e cogliere in qualche modo la multiformità della competenza musicale. Essa, infatti, ben si discosta da una qualunque forma di *allenamento* specifico di una qualsivoglia abilità (senza nulla togliere alla validità di pratiche del genere), ma abbraccia e coinvolge diverse aree sensoriali, motorie,

cognitive, emotivo-affettive e sociali. Risulta quindi abbastanza comprensibile la necessità di sforzi sperimentali mirati e selettivi, che possano garantire un adeguato controllo e al contempo far emergere un particolare effetto tra i tanti possibili. Nondimeno questa tendenza al *sezionamento* del tutto musicale in parti osservabili e misurabili lascia inevitabilmente sul campo una parte degli effetti che si vorrebbero valutare (Mortari, 2010). Alla futura ricerca spetterà l'arduo compito di trovare paradigmi sempre più adatti a considerare la multiformità della competenza musicale, al fine di evidenziarne più accuratamente gli effetti su altri domini e competenze.

### **3.2.1. Funzioni esecutive**

Le funzioni esecutive sono un insieme interagente di processi cognitivi che permettono di pianificare, attuare e monitorare azioni e comportamenti orientati al perseguimento di un obiettivo (Miyake et al., 2000; Diamond, 2013). Si tratta di processi top-down consapevoli e volontari, che sovrintendono all'organizzazione di pensieri, emozioni e azioni. Essi intervengono in tutti quei casi in cui l'elaborazione percettiva di tipo automatico risulta insufficiente o inefficace e si rende necessario un richiamo di risorse attentive per completare adeguatamente un compito. Le funzioni esecutive fondamentali che la letteratura individua sono: inibizione, flessibilità cognitiva e memoria di lavoro. Questi tre processi sono di cruciale importanza per un corretto sviluppo e un comportamento efficace in tutte le situazioni che la vita quotidiana presenta (Brown & Landgraf, 2010; Diamond, 2016). Tramite esse la persona è in grado di selezionare set di informazioni rilevanti e scartarne altre, mantenere in memoria queste informazioni per il tempo necessario e sufficiente allo svolgimento del compito prefissato, pianificare l'esecuzione di azioni motorie e di pensieri, bloccare risposte comportamentali indesiderate, aggiornare rapidamente il proprio comportamento in ragione delle mutate condizioni esterne.

Da ormai qualche decennio si assiste ad un significativo interesse verso la valutazione dell'impatto della musica su abilità e processi cognitivi più o meno riconducibili a competenze dominio specifiche. Le alte richieste percettive,

cognitive, motorie (Altenmüller et al., 2006; Zatorre et al., 2007; Wan & Schlaug, 2010) ed emotivo-affettive (Salimpoor & Zatorre, 2013; Vuilleumier & Trost, 2015) della pratica musicale hanno indotto non pochi ricercatori a pensare che essa avesse tutte quelle caratteristiche che sono raccomandate (Diamond & Ling, 2016) per lo sviluppo o il recupero delle funzioni esecutive (Hannon & Trainor, 2007; Seinfeld et al., 2013).

L'ascolto, e ancor più l'esecuzione musicale, infatti, implicano e richiedono assai spesso di esercitare flessibilità, al fine di spostare il proprio focus tra stimoli della stessa modalità o anche tra diverse modalità, posizioni spaziali e temporali (Moore & Chen, 2010; Loehr & Palmer, 2011; Okada, 2016). Per questo, alcuni studiosi hanno evidenziato benefici della performance musicale sulla flessibilità cognitiva generale, confrontando esperti e non esperti su compiti ad essa correlati (Bugos et al., 2007; Hanna-Pladdy & MacKay, 2011; Moradzadeh, 2014; Zuk et al., 2014).

Anche il controllo inibitorio è riconosciuto come un fattore determinante durante l'esecuzione musicale individuale e, ancor più, durante la pratica d'insieme. Infatti, il musicista deve costantemente aggiustare molteplici parametri della sua esecuzione per renderla più aderente sia al risultato estetico individualmente atteso (attraverso il circuito di causalità sensomotoria), sia al risultato complessivo della performance di gruppo (Palmer et al., 2013; Jentsch et al., 2014). Anche durante attività tipicamente ritmiche vi è una forte richiesta di abilità inibitorie per permettere alle due mani di svolgere compiti diversi (Vuust et al., 2011; Bugos, 2019), discorso che può essere ragionevolmente allargato ed esteso a tutti gli strumenti che richiedono l'us di entrambe le mani con compiti a volte assai diversi a livello temporale, spaziale e motorio. Queste considerazioni portano a pensare che l'intenso training inibitorio possa in qualche misura generalizzarsi (Moreno & Farzan, 2015), e infatti non mancano evidenze a sostegno di questa ipotesi: Bialystok & DePape (2009) riportano una maggiore velocità di risposta dei musicisti rispetto al gruppo di controllo in due compiti con risoluzione di conflitti, segnatamente un test di *Stroop* uditivo e un compito *Simon* visivo, mentre altri risultati mostrano migliori abilità in compiti *Stop-Signal* (Strait et al., 2010). Travis

e colleghi (2011) riportano invece un effetto della durata del training musicale su un compito di *Stroop* con interferenza tra colori e parole, mostrando migliori risultati per musicisti esperti, se confrontati con amatori. Moreno e colleghi (2011) si spingono a studiare gli effetti longitudinali di un training musicale in età evolutiva, confrontandolo con un training artistico: la somministrazione a intervalli regolari di un test *go/no-go* evidenziava maggiori miglioramenti nel gruppo del training musicale, e si notavano corrispondenze tra le misure comportamentali e quelle elettrofisiologiche associate. Janus e colleghi (2016) trovano che un intervento educativo musicale simile al precedente porta a miglioramenti nel controllo esecutivo in bambini dai quattro ai sei anni, ma ottiene lo stesso effetto in un gruppo parallelo di bambini che frequentavano un corso di francese, non potendo così isolare alcun effetto specifico del training musicale. Ancora, in studi sull'età avanzata, si riportano benefici della pratica pianistica sulle condizioni di interferenza dei test *Stroop*, rispetto ad altre attività ricreative proposte a gruppi di controllo (Seinfeld et al., 2013), mentre i musicisti adulti sono risultati migliori della controparte non esperta in compiti di controllo cognitivo (Amer et al., 2013).

Non meno coinvolta e richiamata, durante la pratica musicale di ascolto e di esecuzione, è la memoria di lavoro: ad esempio mentre i musicisti leggono il proprio spartito, gli occhi cercano quanto più possibile di precedere la mani (Furneaux et al., 1999; Drake & Palmer, 2000) e questa attività di lettura è associata a misure più generali di controllo della memoria (Kopiez & Lee, 2008; Meinze & Hambrick, 2010). Alcuni studi evidenziano una migliore abilità di memoria verbale nei musicisti, confrontati con un gruppo di controllo non esperto (Brandler & Rammsayer, 2003; Franklin et al., 2008). George & Coch (2011) riportano che il training musicale di lungo periodo è associato con migliori performance nella memoria di lavoro visuale e uditiva, con misure sia comportamentali che elettrofisiologiche. Roden e colleghi (2014) attuarono un programma educativo musicale per un anno e mezzo ad un gruppo di bambini di scuola primaria, mettendolo in parallelo con un gruppo di controllo che svolgeva invece un programma di scienze naturali: al termine dell'intervento, il gruppo dei musicisti risultava superiore in compiti di *Span*, utilizzati per misurare le capacità di

immagazzinamento delle informazioni e riconosciuto come strumento di misura della memoria di lavoro. Anche in compiti della famiglia degli *n-back* – che richiedono di giudicare se lo stimolo appena presentato è uguale o diverso dal precedente – gruppi di musicisti sono risultati superiori ai gruppi di controllo in più occasioni (Pallesen et al., 2010; Oechslin et al., 2013). Bergman Nutley e colleghi (2014) adottano un disegno di ricerca longitudinale e seguono musicisti e non musicisti per quattro anni con test ogni due anni (il campione era composto da bambini, adolescenti e adulti fino a 25 anni), misurando le abilità di ragionamento, la velocità di elaborazione e la memoria di lavoro attraverso test comportamentali e misure FMRI. I musicisti mostrano risultati significativamente migliori nella memoria di lavoro verbale e visuo-spaziale che, inoltre, cresce proporzionalmente con il tempo settimanale dedicato alla pratica musicale. I risultati, persistenti anche con il controllo di fattori come l'educazione dei genitori e la quantità di attività scolastiche, indicano che la pratica musicale può sostenere lo sviluppo della memoria di lavoro, soprattutto nell'infanzia e nell'adolescenza.

Ad ogni modo, oltre alle evidenze precedenti, sono stati riportati anche risultati deboli e meno evidenti (Schellenberg, 2011), che hanno favorito un ripensamento e affinamento metodologico, soprattutto per quanto riguarda la necessità di una valutazione separata delle tre componenti fondamentali delle funzioni esecutive, la necessità di controllare alcuni fattori preesistenti potenzialmente confondenti come lo status socioeconomico (Schellenberg, 2015) e, infine, l'urgenza di trovare criteri univoci e affidabili per definire e reclutare i gruppi di musicisti (vedasi, a proposito, l'introduzione al cap. 2).

Un importante passo in questa direzione è stato compiuto da Zuk e colleghi (2014), che reclutano gruppi di musicisti e non musicisti sia in età adulta che scolare, per valutare eventuali associazioni fra il training musicale ricevuto e lo sviluppo delle funzioni esecutive, misurate nelle singole componenti attraverso la batteria DKEFS (Delis et al., 2004). Nel caso dei bambini, la batteria è stata ridotta per adattarla all'età. Durante l'esecuzione dei test, sono state raccolte anche misure FMRI per studiarne la correlazione con le sottocomponenti delle funzioni esecutive. I risultati

mostrano che i musicisti adulti erano migliori della loro controparte nei test di flessibilità cognitiva e memoria di lavoro, mentre nei bambini la differenza significativa era evidente soltanto per la flessibilità cognitiva, accompagnata tuttavia da differenze nella velocità di elaborazione e nella ricchezza e varietà delle attivazioni neurali durante l'esecuzione dei compiti. Le aree cerebrali attivate che distinguevano i bambini con training musicale dai compagni senza esperienza erano, in particolare, la regione fronto-parietale e la SMA (area motoria supplementare), che spesso viene associata all'attenzione e al controllo cognitivo.

Slevc e colleghi (2016) mettono a confronto musicisti e non musicisti su una batteria di ben sei compiti, due per ognuna delle componenti delle funzioni esecutive, nelle modalità uditiva e visiva. In entrata venivano qui controllati la dominanza manuale, lo status socio-economico e le abilità di bilinguismo. La musicalità, invece, era misurata con due strumenti, un questionario di autovalutazione (Ollen, 2006) e il test percettivo MET (Wallentin et al., 2010). I risultati mostrano che l'abilità musicale è correlata fortemente con la memoria di lavoro – sia uditiva che visiva – anche dopo aver controllato i fattori potenzialmente interferenti; invece il legame con l'inibizione risulta qui presente ma non altrettanto evidente, mentre non si riscontra alcuna correlazione con la flessibilità cognitiva. Vi è, qui, una prova del trasferimento di competenza verso un dominio generale, che appare però processo-specifico, legato cioè quasi esclusivamente alla memoria di lavoro.

Nel già citato studio di Clayton e colleghi (2016), il gruppo dei musicisti e il gruppo di controllo hanno svolto, oltre ad un test di ascolto spaziale e il test di tracciamento di oggetti multipli (MOT), anche una batteria di test per la valutazione delle funzioni esecutive. La memoria di lavoro uditiva è stata valutata con il test *backward digit span* appartenente alla batteria WAIS-IV (Hartman, 2009); l'inibizione con un test di interferenza colori/parole, dalla batteria DKEFS e la flessibilità cognitiva con un test di fluidità di disegno, anch'esso appartenente alla batteria DKEFS. I risultati indicano una differenza significativa tra gruppi soltanto nel test della memoria di lavoro, nel quale i musicisti hanno ottenuto performance

migliori. Inoltre, gli autori riportano che le prestazioni individuali sul test MOT, unite all'appartenenza al gruppo dei musicisti, hanno predetto efficacemente i risultati nel test di ascolto spaziale. In generale, si suggerisce che il possibile trasferimento di competenza dall'area musicale ad aree più generali possa essere non diretto, bensì mediato dalle funzioni esecutive, in particolare la memoria di lavoro uditiva (Moreno & Bidelman, 2014).

Questo particolare nesso tra esperienza musicale e memoria di lavoro è stato studiato anche attraverso una metanalisi di molte ricerche in materia (Talamini et al., 2017) che ha rivelato come l'effetto della pratica musicale su questa componente specifica delle funzioni esecutive sia mediamente robusto ( $g = 0.56$ ). Una ulteriore analisi ha rivelato che l'effetto è massimo in presenza di stimoli uditivi tonali, moderato con stimoli verbali e scarso con stimoli visuo-spaziali.

Più recentemente, uno studio longitudinale con bambini prescolari (Bugos & Demarie, 2017) mette a confronto un intervento educativo musicale con un intervento basato sulle costruzioni Lego, misurando l'inibizione attraverso il test MFFT (un test di accoppiamento di figure familiari) e il test di stroop *day/night* (nel quale è necessario inibire una risposta verbale predominante). I risultati complessivi al termine degli interventi mostrano che il training musicale ha portato a risultati migliori nel primo test (indicati dal minor numero di errori del gruppo training musicale, rispetto al gruppo Lego), mentre nel secondo test entrambi i gruppi mostrano miglioramenti, senza differenze inter-gruppo. Le autrici concludono che un intervento musicale in ambito prescolare possa portare benefici nel controllo inibitorio, soprattutto quando esso coinvolge l'area motoria.

Un altro gruppo di ricercatori (Bowmer et al., 2018) adottano il medesimo disegno sperimentale longitudinale con bambini prescolari utenti di scuole dell'infanzia, implementandolo ulteriormente al fine di rafforzare i risultati eventualmente ottenuti. Dopo una misurazione in ingresso delle funzioni esecutive, tre gruppi di bambini prendono parte alla sperimentazione che, in una prima fase vede soltanto il primo gruppo coinvolto in un intervento di educazione musicale, mentre gli altri due partecipavano a sessioni di gioco libero. Al termine di questa prima fase, la

batteria di test veniva ripetuta. Nella seconda fase, invece, il primo gruppo continuava il training musicale e anche il secondo gruppo iniziava, mentre il terzo era assegnato ad un training artistico. Un'ultima misurazione post trattamento chiudeva l'esperimento. La batteria comprendeva sei test, scelti per la misurazione separata delle tre componenti delle funzioni esecutive: il test *Peg Tapping* (Luria, 1966), il *Baby Stroop* (Hughes & Ensor, 2005), il *Dimensional Change Card Sort* (Zelazo, 2006), il *Trucks* (Hughes & Ensor, 2005), il *Tower of London* (Anderson et al., 1996) e infine il test *Spin the Pots* (Hughes & Ensor, 2005). I risultati dopo la prima fase mostrano che il gruppo training musicale è migliore nelle misure di pianificazione e inibizione, mentre nella successiva misurazione si ravvisano miglioramenti in tutti e tre i gruppi, senza però differenze significative tra gruppi. Gli autori interpretano i risultati come una prova dell'efficacia dell'intervento educativo musicale sull'abilità di inibizione, anche se il vantaggio iniziale sembra essere eroso e non più apprezzabile al momento della seconda misurazione. Ciò potrebbe essere dovuto sia alla brevità del training (otto settimane), sia al fatto che il beneficio sull'inibizione potrebbe non essere linearmente proporzionale al tempo speso nel training. Anche Carey e colleghi (2015), infatti, sostengono che in età evolutiva vi sia un vantaggio iniziale significativo dovuto al training, che però non cresce di pari passo con il tempo, e altre attività proposte ai gruppi di controllo potrebbero garantire un effetto simile, perlomeno nelle prime settimane di intervento.

Lo studio di Jaschke e colleghi (2018), sempre con un disegno sperimentale longitudinale, recluta quattro gruppi di bambini di circa sei anni, e ne assegna rispettivamente due ad interventi musicali diversi, uno ad un intervento artistico e l'ultimo gruppo ad una semplice funzione di controllo passivo. I ricercatori seguono il suddetto campione per due anni e mezzo, testando ad intervalli regolari i bambini prima, durante e dopo gli interventi didattici seguiti. I test per le funzioni esecutive comprendevano il *Tower of London*, il *Klingberg Short Term and Working Memory Task* (Klingberg, 1998) e il *go/no-go*. I due gruppi del training musicale mostrano risultati migliori sia sull'inibizione che sulla pianificazione, mentre il gruppo artistico era migliore nelle abilità di memoria visuo-spaziale. Una ulteriore,

interessante analisi, indica che le sottocomponenti delle funzioni esecutive migliorate in queste misurazioni sono risultate mediatrici dei risultati ottenuti dai bambini in test standardizzati per l'apprendimento scolastico. Gli autori suggeriscono quindi che l'educazione musicale precoce in ambito scolastico possa contribuire, attraverso la mediazione dello sviluppo di funzioni esecutive, a migliori risultati scolastici.

Nello stesso anno arrivano altri due contributi importanti che continuano nella direzione appena tracciata: il primo arriva da un gruppo di ricerca cinese (Shen et al., 2019), che propone un intervento musicale di dodici settimane ad un gruppo di bambini in età prescolare, basato su attività di canto, ballo e teoria. Un gruppo di controllo, nel frattempo, segue le attività curricolari tipiche della scuola dell'infanzia. La valutazione delle funzioni esecutive nelle sue tre componenti viene effettuata attraverso una batteria di test adattati all'età: il test *Day/Night Stroop*, il *Dimensional Change Card Sort*, il *Dot Matrix* e infine il *Backward Digit Span*. La prima valutazione viene fatta in ingresso, per poi essere ripetuta immediatamente al termine dell'intervento didattico e a distanza di ulteriori dodici mesi, per misurare la persistenza dell'effetto del trattamento. Da notare che gli autori compiono un interessante e pregevole lavoro di programmazione didattica e pedagogica, al fine di evidenziare e potenziare il nesso tra attività didattiche e funzioni esecutive elicitate. I risultati immediatamente dopo l'intervento, a parità sostanziale di misure in ingresso, mostrano che il gruppo musicale ottiene performance superiori del gruppo di controllo su tutti i test somministrati, evidenziando una traiettoria di miglioramento significativamente più marcata. Dopo dodici mesi, la differenza tra gruppi resta significativa in tutti i test, anche se non si rileva lo stesso gradiente di crescita osservato in precedenza per il gruppo musicale. Infatti, i risultati dei due gruppi crescono abbastanza similmente, e questo può essere interpretato come un effetto della semplice maturazione evolutiva dei bambini. Tuttavia, la differenza persistente tra gruppi porta gli autori a concludere che l'effetto del training musicale sullo sviluppo delle funzioni esecutive è presente e persistente nel tempo, almeno fino a dodici mesi dopo l'intervento didattico. Nonostante questi risultati siano in qualche modo limitati dalla mancanza di un gruppo di controllo attivo, essi

dimostrano che le attività musicali, se adeguatamente congegnate e progettate, sono un valido strumento didattico intra-scolastico per promuovere lo sviluppo delle funzioni esecutive che, a loro volta, sono state più volte riconosciute potenti mediatrici di preparazione ai successivi gradi di istruzione (Zuk et al., 2014; Ribner et al., 2017).

Frischen e colleghi (2019) studiano quali specifiche attività musicali possano promuovere lo sviluppo di singole componenti delle funzioni esecutive. Tre gruppi di bambini prescolari sono assegnati a tre diverse attività: un percorso ritmico, un percorso melodico e un percorso sportivo con funzione di controllo attivo. Intelligenza, condizioni socioeconomiche e tratti di personalità sono invece utilizzati come variabili di controllo. Le funzioni esecutive vengono anche qui valutate separatamente per ciascun processo, attraverso batterie simili a quelle adottate nello studio precedente, con l'eccezione della memoria di lavoro, misurata soltanto nella modalità visuo-spaziale. La batteria viene somministrata in ingresso e dopo la fine dell'intervento didattico. I risultati indicano che il gruppo *ritmico* ha ottenuto performance migliori per quanto riguarda l'inibizione, mostrando un miglioramento superiore sia al gruppo *sportivo* che al gruppo *melodico*. In generale, il gruppo con intervento ritmico è risultato migliore del gruppo di controllo attivo su tutti i compiti. L'andamento sui test di memoria di lavoro e flessibilità cognitiva risulta simile, con miglioramenti più marcati dei gruppi musicali rispetto a quello di controllo attivo, ma tali differenze non risultano statisticamente significative. L'assegnazione casuale ai gruppi e il disegno sperimentale longitudinale porta gli autori a concludere che il training ritmico ha prodotto un miglioramento dei processi di inibizione nei bambini prescolari, e che questo miglioramento, se confrontato con gli altri due gruppi sperimentali, è ascrivibile con ragionevole confidenza all'intervento didattico proposto. Da un punto di vista pedagogico, tuttavia, non si può non ravvisare una certa artificiosità nella separazione dell'attività musicale. La divisione tra attività ritmiche e attività melodiche, infatti, non è così netta come questo contributo sembra suggerire. Basti pensare al canto, che è costituito sì da note musicali, ma ogni nota ha una sua durata e tutte quante le note di un canto seguono un tempo (fatto di accenti forti e deboli) e un ritmo

(costante o variabile a seconda del repertorio o della difficoltà del brano). Ogni cantante, quindi, ha bisogno di esercitare e padroneggiare le sue abilità ritmiche durante l'esecuzione dei brani, esse sono infatti inscindibili dalle abilità melodiche di percezione o intonazione di intervalli. A meno che non si tratti di mero esercizio tecnico – esso infatti si presterebbe meglio ad una distinzione tra ritmo e melodia – la pratica musicale, anche e soprattutto in contesti didattici, ha nella multi-modalità e multiformità di stimoli e processi richiamati la sua caratteristica più unica, distintiva e apportatrice di effetti specifici (Talamini et al., 2017; Wang et al., 2018).

Ultimamente, Chen e colleghi (2022) hanno studiato l'associazione fra la pratica musicale e le singole componenti delle funzioni esecutive in uno studio trasversale su bambini dai sette ai tredici anni di età. In particolare, gli autori sono qui interessati a stabilire se vi sia un periodo ottimale nel quale iniziare a praticare la musica per ottenere i maggiori benefici sullo sviluppo di questi processi cognitivi generali e basilari. Il campione è stato diviso in due gruppi: il gruppo dei musicisti e quello di controllo, e sono stati somministrati test per la valutazione di inibizione, memoria di lavoro e flessibilità cognitiva, precisamente il *go/no-go*, lo *stroop*, il test *Continuous Performance* e quello di *task switching*. Il livello musicale è stato testato attraverso una sottoscala specifica del Goldsmith Music Sophistication Index (Mullensiefen et al., 2014). I risultati di questo primo esperimento mostrano che il gruppo dei musicisti è significativamente superiore al gruppo di controllo nelle misure di inibizione e memoria di lavoro, ma non nelle misure di flessibilità cognitiva, come già altre volte riportato in letteratura. Un secondo esperimento viene condotto soltanto sul campione dei musicisti, che viene diviso ulteriormente in due sotto-gruppi, a seconda che avessero o meno iniziato a praticare la musica prima dei sette anni di età. Infatti, vi sono evidenze che indicano questo come il periodo più sensibile per lo sviluppo delle funzioni esecutive (Penhune, 2011; Steele et al., 2013; Chieffo et al., 2016; Van Vugt et al., 2021). I risultati del gruppo di musicisti con inizio precoce sono stati migliori nelle misure di inibizione e memoria di lavoro e un ulteriore approfondimento e suddivisione dei processi delle funzioni esecutive porta gli autori a suggerire che la precocità di inizio del training ha influito soprattutto sull'inibizione delle risposte e il controllo proattivo, le cui

misure sono risultate significativamente migliori rispetto al gruppo di bambini che aveva iniziato a praticare dopo i sette anni. Questo ulteriore approfondimento può essere considerato coerente con quanto affermato da studiosi delle funzioni esecutive (Friedman & Miyake, 2004), per i quali sussistono differenze interne tra processi cognitivi riconducibili alla stessa macro-area (ad es. diversi sotto-processi all'interno della macro-area dell'inibizione). Considerando anche la naturale traiettoria di maturazione dei singoli processi, è probabile che le attività musicali iniziate in questo periodo sensibile offrano un singolare supporto allo sviluppo di alcuni di questi sotto-processi. Inoltre, i confronti non significativi sulla flessibilità cognitiva, che apparentemente contrastano con evidenze di effetti positivi in età adulta, potrebbero essere dovuti non tanto all'inefficacia della pratica musicale, quanto ad una diversa età di maturazione di questa specifica componente, che si attesta oltre la preadolescenza (Saarikivi et al., 2016).

Al termine di questa rassegna specifica, si segnalano anche studi che indagano l'effetto della pratica musicale non tanto sullo sviluppo, quanto sulla riabilitazione delle funzioni esecutive e attentive, e non pochi sono i risultati positivi riscontrati (Koshimori & Thaut, 2019). Pur con molti nodi ancora da sciogliere e con molta conoscenza ancora da acquisire, queste ulteriori evidenze completano il quadro degli effetti potenzialmente positivi della pratica musicale su alcuni processi dominio generali come le funzioni esecutive e l'attenzione.

### **3.2.2. Linguaggio, intelligenza, emozioni e competenze sociali**

Una delle più robuste evidenze dei benefici del fare musica su abilità più generali riguarda l'area del linguaggio, che comprende la percezione fonemica, le abilità di lettura e comprensione, così come la padronanza della madrelingua e di eventuali altre lingue. A partire dalla dimostrazione di competenze più dominio-specifiche, come la discriminazione di piccole variazioni di altezza nelle melodie (vedasi, ad es., Magne et al., 2006; Moreno et al., 2009), si è cercato di indagare se questa capacità si estendesse verso stimoli verbali. Kraus e colleghi (2014) mostrano una migliore discriminazione di suoni tipici del parlato in bambini che avevano fatto un percorso musicale di due anni. Questi risultati ispirano altre ricerche che mostrano

incoraggianti effetti della pratica musicale nel trattamento della dislessia (Thomson et al., 2013; Flaugnacco et al., 2015) che spesso è accompagnata proprio da una difficoltà di discriminazione delle frequenze. Risultati simili sono stati ottenuti per la discriminazione dell'altezza e intonazione nel parlato, confrontando gruppi di musicisti e non musicisti (vedasi, ad es., Thompson et al., 2004; Wong et al., 2007; Good et al., 2017).

Sempre in età infantile, vi sono associazioni tra abilità musicali, consapevolezza fonologica e percezione della struttura tonale delle parole (Anvari et al., 2002), che suggeriscono un legame forte tra percezione musicale e fonologica. Degè e colleghi (2011) assegnano casualmente tre gruppi di bambini prescolari ad attività musicali, sportive e specificamente fonologiche, per verificare se la pratica musicale potesse essere in grado di sviluppare la percezione del parlato. Dopo venti settimane, il gruppo musicale ottiene gli stessi risultati del gruppo che ha svolto training fonologico specifico, ed entrambi sono migliori del gruppo sportivo.

François e colleghi (2013) mostrano come un training musicale di due anni possa migliorare le abilità di segmentazione del parlato, confrontando il gruppo sperimentale con un gruppo di controllo con attività di disegno. Inoltre, dopo dodici mesi di training musicale, un gruppo sperimentale risultava migliore del gruppo di controllo su abilità di elaborazione pre-attentiva della durata e attacco delle sillabe (Chobert et al., 2014). Ancora, dopo due anni di musica d'insieme, il gruppo dei musicisti mostrava migliori abilità di percezione del parlato in presenza di rumore (Slater et al., 2015), risultato replicato anche in differenti condizioni e con diversi paradigmi, con l'indicazione di una più fedele rappresentazione degli stimoli contenenti il parlato (vedasi, ad es., Parbery-Clark et al., 2009; Strait & Kraus, 2011; Tierney, et al., 2013; S; Ruggles et al., 2014).

Si consolida l'ipotesi formulata qualche anno prima da Patel (2003), che postulava l'esistenza di una sovrapposizione consistente dei processi neurali coinvolti nel linguaggio e nella musica, ipotesi che ha ricevuto poi ulteriore supporto (Bidelman et al., 2013; White et al., 2013). Moreno e colleghi (2015) portarono altre prove a sostegno di questa ipotesi, confrontando longitudinalmente un intervento musicale

con uno linguistico, in bambini prescolari, per quattro settimane. I test che misuravano la capacità di sopprimere stimoli irrilevanti mostrava che entrambi i gruppi riuscivano a concentrarsi meglio sui rispettivi suoni dominio-specifici, e l'effetto persisteva anche dopo dodici mesi dall'intervento didattico. Questo network neurale condiviso sembra comprendere in modo speciale la corteccia prefrontale (Slevc et al., 2015). Gruppi di musicisti, confrontati con gruppi di controllo, sono risultati migliori anche nella rilevazione dell'andamento degli accenti linguistici (Kolinski et al., 2009) o nella percezione di stimoli acusticamente degradati contenenti alcune vocali (Bidelman et al., 2010).

Oltre a queste abilità a livello percettivo, si riportano anche migliori abilità cognitive legate al linguaggio, come una migliore memoria verbale a breve termine, a lungo termine e di lavoro (vedasi, ad es., Franklin et al., 2008; Hansen et al., 2013). Il training musicale è stato anche associato a migliori abilità lessicali nella madrelingua (vedasi, ad es., Forgeard et al., 2008) e nella seconda lingua (Posedel et al., 2012; Talamini et al., 2018).

Moreno e colleghi (2011), nel già citato studio longitudinale su bambini prescolari, confrontano un gruppo con training musicale ed un gruppo con training artistico anche su misure di abilità lessicali, e dopo l'intervento il gruppo musicale ottiene performance migliori rispetto a quelle iniziali, al contrario del gruppo artistico, risultato che permette di attribuire con ragionevole confidenza il miglioramento al training ricevuto. Lo stesso team di ricerca riporta che il gruppo musicale risultava migliore anche nell'abilità di collegare simboli a parole.

Vi sono state anche evidenze più deboli o addirittura nulle (Ruggles et al. 2014; Boebinger et al., 2015; Swaminathan & Schellenberg, 2017) che tuttavia non permettono di mettere in discussione un cospicuo corpus di ricerche che indicano, pur nella varietà e diversità dei paradigmi e delle attività musicali prese in considerazione, un consistente effetto positivo della pratica musicale sullo sviluppo e il potenziamento delle abilità linguistiche, soprattutto a livello percettivo.

Non pochi sono stati i tentativi di indagare l'impatto del training musicale su costrutti ancor più generali come, ad esempio, l'intelligenza. Alcuni studi correlazionali, sia con bambini che con adulti, mostrano che gruppi di musicisti ottengono migliori punteggi delle loro controparti nei test del quoziente intellettivo (Gibson et al., 2009; Schellenberg, 2011), e la durata del training è spesso associata positivamente con la crescita delle misure di intelligenza (vedasi, ad es., Corrigan, 2013; Degé et al., 2014; Swaminathan et al., 2017; 2018). A sua volta, le misure di intelligenza sono spesso indicate come predittori dei futuri risultati accademici e professionali (vedasi, ad es., Deary et al., 2007), e alcuni studi hanno messo in relazione i termini estremi di questa concatenazione, mostrando che il training musicale può essere mediatore di migliori risultati scolastici (vedasi, ad es., Vaughn & Winner, 2000; Gouzouasis et al., 2007), anche controllando gli indicatori socio-economici in ingresso, anch'essi spesso associati allo stesso tipo di risultato a lungo termine (Chung, 2015). Vi sono anche alcune evidenze di effetti del training musicale sull'intelligenza, provenienti da studi longitudinali (Schellenberg, 2004; Kaviani et al., 2014, Jaschke et al., 2018) che, nonostante alcune limitazioni ascrivibili ora alla mancanza di un controllo attivo, ora al mancato controllo di alcune variabili potenzialmente influenti, arricchiscono e consolidano il nesso ipotizzato in partenza.

Ad ogni modo, i risultati non sono costanti e sembrano più evidenti soltanto per alcune misure di intelligenza, piuttosto che per altre. Certamente non si può escludere a priori l'impatto che alcuni fattori preesistenti (genetici o di personalità) potrebbero avere sulla probabilità di iniziare un percorso di formazione musicale (Corrigan & Schellenberg, 2015; Mosing et al., 2016; Swaminathan & Schellenberg, 2018). È altresì evidente che all'interno del concetto di *preesistenza* possano essere ragionevolmente collocati fattori sia genetici che ambientali, e che raramente queste due componenti sono state oggetto di studio o anche soltanto di controllo sperimentale. Come già evidenziato nei paragrafi iniziali di questo secondo capitolo, il cosiddetto *ambiente* comincia ad interagire con la persona addirittura prima della nascita, attraverso stimoli di varia natura. Almeno durante la gravidanza e nei primissimi mesi di vita del neonato, la modalità uditiva ha un ruolo

centrale e determinante per le prime fasi dello sviluppo di molte funzioni cruciali (Tomatis, 2015). Per questo motivo, ogni ipotesi riguardo all'esistenza di fattori genetici che predisporrebbero maggiormente alcuni individui alla musica rispetto ad altri dovrebbe considerare attentamente i fattori ambientali – particolarmente la ricchezza delle interazioni uditive intra-familiari – che hanno caratterizzato le prime fasi di vita dei soggetti presi in esame (Tafuri, 2007). Forse, soltanto uno studio che controlli queste variabili in ingresso potrebbe aspirare a far emergere, se effettivamente vi fossero, fattori esclusivamente genetici che predispongano o meno un soggetto verso attività legate alla musica.

A completamento del quadro che, al meglio della nostra attuale conoscenza, rappresenta lo stato dell'arte circa l'impatto del training musicale sulle abilità cognitive, è importante menzionare le molte evidenze riguardo all'osservazione di cambiamenti anatomo-funzionali in conseguenza di attività musicali, soprattutto in alcuni periodi più sensibili dello sviluppo ma anche in adolescenza e in età avanzata (vedasi, ad es., Kraus & Chandrasekaran, 2010; Moreno, 2015; Benz et al., 2016 e Mado Proverbio, 2019 per una rassegna). Le molte e diverse aree cerebrali coinvolte in questi effetti di neuro-plasticità indotta dal training portano diversi autori a suggerire che l'impatto di questa particolare forma di training potrebbe estendersi ben oltre le abilità e competenze dominio-specifiche.

Il corpo di evidenze discusso in questo capitolo, unito alle molte indicazioni di effetti su aree emotive, affettive e sociali (Laukka, 2007; Overy & Molnar-Szakacs, 2009; Kirschner & Tomasello, 2010; Chanda & Levitin, 2013; Brattico, 2015; Sheppard & Broughton, 2020; Martinez-Castilla et al., 2021; Váradi, 2022), porta a concludere che la pratica musicale - specialmente se iniziata precocemente, portata avanti per almeno sei anni e arricchita da significative esperienze esecutive – può produrre cambiamenti cerebrali e benefici comportamentali in una serie di abilità trasversali, i quali risultano più o meno evidenti a seconda della contiguità con ciò che viene considerato dominio-specifico della musica.

## **IV - Studio 1: attenzione uditiva nello spazio circostante, differenze tra gli effetti degli emicampi anteriore e posteriore**

### **4.1. Introduzione**

Come si evince dalla lettura del capitolo primo di questo lavoro, l'attenzione spaziale è un'abilità cognitiva fondamentale per affrontare al meglio una grande quantità di situazioni quotidiane che ci richiedono di concentrarci su una posizione specifica e di elaborare principalmente e prioritariamente le informazioni da essa provenienti. Altrettanto indispensabile è la capacità di spostare rapidamente il nostro attuale focus attentivo, nel momento in cui dobbiamo elaborare nuove informazioni salienti che provengono da una diversa posizione nello spazio. Un esempio efficace dell'applicazione di questa duplice abilità è una tipica situazione di guida - di un'auto, così come di una bicicletta o di un monopattino - poiché ci viene costantemente richiesto di dirigere la nostra attenzione verso lo spazio davanti e contemporaneamente di monitorare gli stimoli provenienti da altre direzioni. Infatti, dovremmo essere pronti a spostare rapidamente la nostra attenzione quando, ad esempio, arriviamo ad un incrocio ed è necessario controllare le auto che arrivano da diverse direzioni, oppure quando un mezzo vorrebbe superarci da dietro e per questo ci invia segnali visivi o acustici. Anche mentre camminiamo, benché la nostra attenzione sia tipicamente rivolta verso lo spazio anteriore, vi è un costante monitoraggio di ciò che ci circonda, soprattutto per merito dell'attenzione uditiva che per sua natura ha una sensibilità panoramica. Ogni rumore può essere importante, sia per attirare semplicemente la nostra curiosità, sia per avvisarci di un potenziale pericolo proveniente da zone al di fuori del campo visivo attuale.

Possiamo quindi comprendere come, a questa capacità di riorientare l'attenzione su qualcosa che sta accadendo al di fuori del focus corrente e che richiede una reazione immediata, sia attribuita una grande importanza ai fini della sopravvivenza (Erulkar, 1972; Scharf, 1998; Bradley, 2009). Anche quando leggiamo o stiamo insieme agli altri, ci si aspetta che manteniamo l'attenzione in modo efficiente su

una posizione specifica, pur essendo pronti a spostare la nostra attenzione ogni volta che sia necessario.

Sebbene i primi studi sull'attenzione spaziale avessero riguardato quasi esclusivamente la modalità uditiva, i decenni successivi videro una netta maggioranza di ricerche in ambito visivo. Tuttavia, alcune differenze significative tra le due modalità suggeriscono cautela nel confrontare sia i disegni sperimentali che i risultati riportati in letteratura. In primo luogo, è da considerare che la vista ha la possibilità di essere diretta meccanicamente verso una specifica porzione di spazio (corrispondente al campo visivo o ad aree più ristrette all'interno di esso) e questa selezione può essere facilmente modificata muovendo la nostra testa o i nostri occhi. Inoltre, la percezione può essere parzialmente o totalmente interrotta chiudendo uno o entrambi gli occhi. Lo stesso non può dirsi dell'udito, la cui sensibilità è costantemente e inevitabilmente panoramica, il che significa che ogni selezione spaziale dovrebbe avvenire ad uno stadio di elaborazione superiore a quello sensoriale (Styles, 2005). Inoltre, è noto che l'acuità di selezione spaziale nell'udito è minore rispetto alla vista, mentre l'elaborazione uditiva e la rappresentazione degli stimoli provenienti da tutte le possibili posizioni è più uniforme rispetto alla vista (Middlebrooks & Green, 1991; Blauert, 1997; Grill-Spector & Malach, 2004; Healey & Sawant, 2012).

È stato ipotizzato che, nella complessità dei meccanismi di controllo dell'attenzione, il sistema uditivo sia responsabile del monitoraggio dell'ambiente circostante, al fine di rilevare potenziali minacce. L'ascolto spaziale può estendere la nostra sensibilità a posizioni che ricadono al di fuori del campo visivo, sia in termini di angolo che di distanza. Se, infatti, è possibile *vedere ciò che non si sente* - ad esempio osservare persone che dialogano senza tuttavia sentirne le voci, a causa dell'eccessiva distanza e della conseguente dissipazione del segnale audio (vedasi cap. 1) - è altrettanto possibile *sentire ciò che non si vede*, e questo accade assai più spesso, in quanto tutto ciò che accade al di fuori del campo visivo è esperibile principalmente attraverso l'udito, e non meno frequenti sono i casi in cui la vista è impedita al rilevamento di uno stimolo - poiché nascosto o coperto da altri elementi

della scena visiva – e l'unico modo per accertarne l'esistenza e la posizione è affidarsi alla componente uditiva dell'informazione proveniente da quello stesso stimolo. In tal senso, il sistema uditivo potrebbe essere visto come un dispositivo di allerta e un rilevatore di salienza (Schafer, 1977; Heffner & Heffner, 1992; Scharf, 1998; Tajadura-Jiménez et al., 2010; Asutay, 2015). Se si considerano gli studi su altre specie, la stessa funzione difensiva dell'udito sembra essere ampiamente e trasversalmente conservata (Bradley, 2009).

Secondo questa prospettiva, in caso di stimoli provenienti da dietro, la funzione dell'udito è quella di richiamare rapidamente le risorse cognitive e guidare la percezione visiva, nota per essere più accurata nella rappresentazione spaziale degli oggetti (Witten & Knudsen, 2005; Kolarik et al., 2016; Aggius-Vella et al., 2018; 2022). In effetti, l'attenzione delle persone è generalmente focalizzata su ciò che sta accadendo nello spazio visibile, ma quando sentono qualcosa di saliente giungere da dietro, di solito muovono il loro corpo o almeno la loro testa verso di esso (vedasi, ad es., Hofman et al., 1998; Previc et al., 2000; Iwaya et al., 2003). Questi effetti di allertamento degli stimoli uditivi sono generalmente maggiori rispetto quanto osservato per la vista, come riportano alcuni studi sull'attenzione in presenza di suggerimento o *cue* (Posner et al., 1976; Posner, 1978).

Nel panorama delle ricerche sull'attenzione uditiva spaziale, possiamo trovare paradigmi comportamentali e neurofisiologici, volti rispettivamente a indagare le implicazioni pratiche e neurologiche dei meccanismi di allocazione e spostamento di attenzione. Per le principali e più recenti evidenze neuroscientifiche, si vedano, ad es., Lewald (2015), Michalka et al. (2016), Tune et al. (2018), Deng et al. (2019), Wostmann et al. (2019), Popov (2023).

Negli studi comportamentali, oltre ai lavori pionieristici sull'ascolto dicotico (vedasi, ad es., Cherry, 1953; Broadbent, 1958), uno dei paradigmi di riferimento è senza dubbio il suggerimento spaziale (Posner, 1980), originariamente concepito per la modalità visiva. Per la spiegazione dettagliata di questo compito, dei suoi presupposti teorici e delle scelte sperimentali adottate, si rimanda alla lettura del paragrafo 2.4 del presente lavoro. Al solo scopo di richiamo di informazioni

rilevanti per il presente studio, è da ricordare che i risultati di questi esperimenti mostrano tipicamente una facilitazione per le prove valide, con risposte più rapide e accurate, una performance intermedia per le prove neutre e, infine, risposte più lente e meno accurate per le prove invalide. Sebbene nel tempo siano stati utilizzati diversi compiti per studiare gli effetti dell'attenzione spaziale, come il rilevamento e il riconoscimento, alcuni autori indicano che un compito di discriminazione garantisce una maggiore efficacia a motivo della superiore richiesta di risorse attentive (Eriksen & St. James, 1986).

Nei decenni successivi, abbiamo visto un'ampia varietà di disegni sperimentali, sia in studi unimodali (per la modalità visiva, vedasi Klein et al., 1992) che intermodali, a volte aggiungendo più posizioni target rispetto alle due originarie del compito Posner, oppure inserendo distrattori (ad es. Lewald & Getzmann, 2015; Celma-Miralles & Toro, 2019). Anche l'estensione o la semplice variazione della porzione di spazio coperto sono state oggetto di indagine e di manipolazione (ad es. Pinker & Downing, 1985; Spence & Driver, 1994; Mondor & Zatorre, 1995; Cave & Bichot, 1999; Golob et al., 2021).

Secondo quanto normalmente riportato da questi studi, lo spostamento spaziale dell'attenzione elicit performance superiori nelle posizioni attese, con decrementi gradualmente all'aumentare della distanza da esse (Cave, 2015). Questo effetto è indicato come "gradiente di attenzione". La grande maggioranza degli studi che indagano i gradienti di attenzione spaziale hanno coinvolto la modalità visiva (Posner & Cohen, 1984; Pinker & Downing, 1985; Cave & Bichot, 1999) cosicché la comprensione di questo fenomeno nella modalità uditiva risulta ancora scarsa e meno attrattiva. Infatti, il corpo significativo di ricerche che esaminano l'attenzione selettiva uditiva in due posizioni spaziali (in genere una a sinistra e una a destra del meridiano centrato sulla testa dei partecipanti, con diverse disposizioni angolari) non può dare alcuna informazione circa i gradienti di attenzione spaziale, dal momento che sono necessarie almeno tre posizioni per quantificarne l'effetto.

Negli ultimi decenni, gli studi riguardanti i gradienti di attenzione spaziale uditiva hanno riportato andamenti di forma lineare, con tempi di risposta che aumentano

all'aumentare della distanza tra posizione attesa e reale (Spence e Driver, 1994; Mondor & Zatorre, 1995; Rorden & Driver, 2001; Driver, 2001). Recentemente, invece, alcuni ricercatori hanno riportato un gradiente di forma quadratica, con forti diminuzioni di velocità a distanze minori dalla posizione attesa e una graduale inversione di tendenza per le posizioni più lontane, caratterizzate da risposte addirittura più veloci rispetto alle posizioni intermedie (Golob et al., 2017; Golob & Mock, 2020). A volte, questa distribuzione quadratica degli effetti dell'attenzione è stata osservata anche per la vista, ma lo spazio coperto nel setting sperimentale era generalmente inferiore rispetto agli studi sulla modalità uditiva (Müller et al., 2005). Quindi, una questione aperta è se, e come, la porzione di spazio esaminata influisca sugli andamenti dei tempi di risposta registrati.

Ad oggi, inoltre, la ricerca sull'attenzione spaziale uditiva (sia in studi unimodali che cross-modali) si è concentrata principalmente su porzioni più o meno circoscritte dell'emicampo anteriore, sebbene una delle caratteristiche che distinguono l'udito dagli altri sensi sia proprio la sua sensibilità panoramica. Come mostrato in precedenza, questa percezione a tutto tondo ci aiuta a monitorare costantemente anche le porzioni di spazio che non possiamo controllare attraverso la vista. Non molti sono stati gli studi che presentavano ai partecipanti stimoli uditivi provenienti da dietro, ed oltretutto erano principalmente interessati alla localizzazione uditiva piuttosto che agli effetti dell'attenzione spaziale.

#### **4.2. Domande e ipotesi di ricerca**

Tutto ciò premesso, una delle nostre domande di ricerca è se i risultati riportati in precedenza, riguardo gli effetti dell'attenzione spaziale uditiva in varie porzioni dell'emicampo frontale, possano essere replicati o meno considerando l'intero spazio circostante i partecipanti, ovvero coprendo l'intero arco dei 360°. Inoltre, ci stiamo chiedendo se gli effetti attentivi osservati nell'emicampo anteriore possano o meno essere paragonabili a quanto accade in quello posteriore. Secondo la teoria premotoria dell'attenzione, i movimenti corporei pianificati ma non eseguiti permetterebbero di spostare l'attenzione in modo implicito. Nel nostro caso, si

tratterebbe verosimilmente di movimenti della testa volti a raggiungere posizioni che non ricadono nel campo visivo. Tuttavia, vi è un limite oltre il quale non è possibile girare la testa, e questo presupporrebbe una *cono d'ombra* nell'emicampo posteriore, verso il quale l'attenzione non può essere portata. Risulta quindi interessante verificare se questa zona sia o meno svantaggiata rispetto alle altre e se, in generale, l'emicampo posteriore possa portare ad andamenti diversi rispetto a quanto accade per l'emicampo anteriore.

Una questione altrettanto importante riguarda i gradienti di attenzione spaziale uditiva. In particolare, abbiamo voluto testare fino a che punto i recenti risultati, che indicano un gradiente di forma quadratica nell'emicampo anteriore, potrebbero essere generalizzati all'intero spazio circostante.

Le evidenze che indicano come il senso della vista eserciti una certa predominanza nella allocazione di attenzione, con meccanismi di selezione spaziale più forti ed accurati rispetto all'udito (Azzopardi & Cowey, 1993), ci portano ad ipotizzare che le posizioni nell'emicampo anteriore godano di un ancoraggio più forte e spazialmente accurato. Ciò potrebbe comportare un disimpegno più difficoltoso dell'attenzione dalle posizioni attese, ogniqualvolta sia necessario spostarsi verso altre direzioni (Spence & Driver, 1996).

Viceversa, l'ancoraggio alle posizioni attese posteriori potrebbe essere meno forte e preciso senza il contributo dell'informazione visiva e, se questo fosse vero, dovremmo osservare spostamenti attentivi più veloci verso le altre direzioni.

Ci potremmo anche aspettare che i suoni provenienti da dietro elicitino risposte più rapide e forse meno accurate, se si uniscono le considerazioni di cui sopra con altre evidenze che attribuiscono all'udito la funzione di monitoraggio e allertamento rapido.

Per quanto riguarda i gradienti di attenzione spaziale, le evidenze recenti, circa la possibile coesistenza di processi top-down di selezione e processi bottom-up di monitoraggio e allertamento rapido, lasciano intendere che la distribuzione degli effetti attentivi potrebbe conservare un andamento di forma quadratica anche in un

compito che considera l'intero spazio circostante e non soltanto l'emicampo anteriore. Infatti, la funzione di monitoraggio costante del sistema uditivo potrebbe portare ad un livellamento dei tempi di risposta per spostamenti verso posizioni lontane, supponendo che l'effetto della selezione spaziale sia andato gradualmente svanendo all'aumentare della distanza da coprire.

### **4.3. Esperimento 1: attenzione transiente**

Alla luce delle domande di ricerca appena enunciate, abbiamo progettato una variante ad-hoc del compito di discriminazione di Posner, riconosciuto come uno degli strumenti più utili per indagare gli effetti dell'attenzione. Questo paradigma adotta i tempi di reazione e le accuratezze percentuali come variabili dipendenti per la misura di tali effetti. Alcune scelte sperimentali sono state guidate dall'interesse a elicitare un'attenzione endogena di tipo *covert*, quindi i partecipanti erano invitati a non muovere la testa e a fissare lo sguardo su un punto centrale dello schermo. Inoltre, i *cue* o suggerimenti visivi venivano presentati al centro dello schermo e il tipo di *cue* richiedeva uno spostamento dell'attenzione mentale anziché fisico.

Il nostro interesse era quello di considerare l'intero spazio attorno ai partecipanti e il maggior numero possibile di posizioni target e di spostamenti attentivi tra di esse. La disposizione delle sorgenti sonore mirava a confrontare efficacemente gli effetti nell'emicampo anteriore e posteriore, uno dei nostri principali interessi di ricerca.

Un'importante questione metodologica, in questo campo di ricerca, riguarda il modo in cui gli stimoli uditivi vengono presentati. Si è ipotizzata una differenza sostanziale nella percezione del suono erogato in cuffia o attraverso altoparlanti (Maddox et al., 2014; Deng et al., 2019). Le cuffie sono sicuramente più facili da maneggiare e la spazialità del suono può essere ben riprodotta, sia sul piano orizzontale che sul piano verticale. Tuttavia, una tale situazione fa sì che i suoni vengano percepiti come originati dall'interno della testa (Scharf, 1998), producendo un ambiente significativamente meno ecologico. D'altra parte, la riproduzione dei suoni sugli altoparlanti offre una sensazione e una percezione della distanza più naturali, poiché i suoni interagiscono non solo con le orecchie del soggetto, ma con

tutto l'ambiente circostante, comprese le pareti, gli oggetti e i volumi d'aria (Heine et al., 2020). Tuttavia, la gestione di più altoparlanti può rappresentare una non trascurabile difficoltà tecnica, a causa della complessità del controllo hardware e software. Infatti, oltre a dover gestire la sistemazione spaziale rispetto al soggetto (ricercando quindi altezze, distanze e intensità sonore ottimali all'orecchio), la complessità della catena audio (composta da altoparlanti, cavi, interfaccia audio, PC e software dedicati) richiede un attento controllo della latenza che dovrebbe essere idealmente azzerata, per non incorrere in alterazioni indesiderate dei tempi impostati, manipolati o acquisiti durante gli esperimenti. Per questi motivi, al fine di ottenere un ragionevole equilibrio tra le condizioni sperimentali e l'ambiente di vita reale, le posizioni spaziali dei suoni non sono state simulate tramite software. Infatti, una componente significativa dei suoni reali deriva dalla loro interazione con l'ambiente in cui sono prodotti, ed è necessario tener conto della sostanziale differenza di presenza e salienza tra l'erogazione del suono dagli auricolari e dagli altoparlanti. Per lo stesso motivo, sono stati scelti suoni reali anziché toni puri o modulazioni di rumore.

I criteri guida per la scelta del cue sono stati principalmente due: in primo luogo, volevamo garantire la migliore efficacia possibile, consentendo ai partecipanti di spostare rapidamente la loro attenzione e allocare la maggior parte delle loro risorse sul compito di discriminazione uditiva. In secondo luogo, volevamo evitare le interazioni che si verificano quando si utilizzano cue e target uditivi. Ci sono alcune prove che dimostrano come l'orientamento dell'attenzione possa essere influenzato da somiglianze o differenze tra cue e target, riguardo diversi parametri sonori come frequenza, timbro, durata e intensità (Mondor & Terrio, 1998; Mondor & Lacey, 2001; Fritz et al, 2007). Inoltre, l'uso di un cue visivo ci ha permesso di confrontare i gradienti di attenzione uditiva con quelli riportati in altri studi cross-modali.

#### **4.3.1. Materiali e metodologia**

Sono stati reclutati un totale di 30 adulti ( $n=30$ , età media = 28,5; SD = 6,9; 19 femmine; 24 destrorsi). Il protocollo sperimentale è stato approvato in tutte le sue parti dal Comitato di Bioetica dell'Università di Bologna. I partecipanti hanno letto

e sottoscritto i moduli di consenso informato e l'informativa sul trattamento dei dati personali a fini di ricerca, prima di iniziare l'esperimento.

Una prima analisi della potenza statistica è stata eseguita utilizzando il software G-Power (Faul et al., 2007; 2009), assumendo come riferimento per il calcolo il valore di  $\eta^2$  riportato da Golob & Mock (2020) nella sua analisi del gradiente di attenzione spaziale uditiva, ovvero 0.15. Considerando il numero delle condizioni sperimentali e delle prove per ciascuna condizione, il software ha stimato una dimensione totale del campione di 16 partecipanti per ottenere una potenza di 0.95 con un valore di significatività atteso di 0.05 e una dimensione dell'effetto stimata di 0.42. Tuttavia, studi recenti che adottano un paradigma cue-target simile a quello di Posner per studiare i meccanismi di attenzione (Hayward & Ristic, 2013) hanno reclutato un totale di 22 partecipanti (è da considerare, però, che lo studio prevedeva due gruppi sperimentali con analisi di fattori interni ai gruppi e tra gruppi diversi, il che influisce sull'analisi di potenza e sulla grandezza del campione ottimale). Inoltre, le assunzioni del teorema del limite centrale e le sue implicazioni nell'analisi statistica (Gyu et al., 2017) suggeriscono che un campione di 30 soggetti per ciascun gruppo sperimentale dovrebbe garantire che la distribuzione statistica sia approssimabile a quella normale. Alla luce di quanto sopra, abbiamo optato per il criterio più restrittivo, reclutando 30 partecipanti.

Un primo controllo in ingresso sulla capacità uditiva è stato effettuato richiedendo ai partecipanti di compilare un auto-report su eventuali patologie pregresse e/o attuali e nessuno dei partecipanti ha riportato problemi degni di nota. Pertanto, nessun soggetto è stato scartato a causa di problemi uditivi.

Volevamo inoltre controllare il background musicale dei partecipanti, supponendo che una significativa storia di training musicale avrebbe potuto influenzare le prestazioni nel compito previsto. Per questo motivo, sono stati somministrati i questionari dell'inventario Goldsmiths Music Sophistication Index o GOLD-MSI (Müllensiefen et al., 2014), uno strumento di auto-valutazione dell'esperienza e delle abitudini musicali, nella versione italiana proposta da Correia et al. (2022). I partecipanti che hanno riportato più di due anni di training musicale formale sono

stati esclusi da questo studio. Infine, un auto-report sul livello di istruzione completava la base dati in ingresso.

Nel giorno fissato per l'esperimento, i partecipanti si accomodavano su una sedia posta esattamente al centro di una stanza con luce bassa. Durante l'esperimento ci siamo assicurati che nessun rumore, proveniente dall'esterno della stanza o da apparecchiature in dotazione al laboratorio, potesse interferire con il compito proposto. Non abbiamo insonorizzato la stanza, affinché fosse preservata la naturalità della percezione del suono, data dall'interazione dello stesso con l'ambiente reale che lo circonda. Il monitor, posto di fronte al partecipante, è stato utilizzato per presentare i suggerimenti o cue visivi. Sei altoparlanti cablati (Anker Soundcore 2) sono stati collocati con una disposizione circolare attorno al partecipante (a 120 cm di distanza dalla testa) e sono stati utilizzati per presentare gli stimoli uditivi. Gli altoparlanti erano distanti tra loro  $60^\circ$  di azimut, coprendo così l'intero arco dei  $360^\circ$  (vedi figura 5). La loro altezza è stata regolata per adattarsi al livello medio dell'orecchio (Van't Hullenaar et al., 2017). Durante l'esperimento tutti gli altoparlanti sono stati ricoperti da un sottile tessuto nero, per nasconderli alla vista senza tuttavia compromettere la qualità del segnale audio. Infatti, come evidenziato nell'introduzione al presente studio, era ipotizzabile un possibile vantaggio delle posizioni anteriori, dovuto forse al concorso dell'informazione visiva disponibile.

Gli stimoli uditivi sono stati presentati da un'interfaccia audio professionale multicanale (Focusrite Saffire Pro 40), collegata ad un PC desktop della serie ThinkPad di Lenovo, dotato del software Presentation (versione 24.0 di Neuro Behavioral Systems Inc. Berkeley, CA). Un dispositivo di risposta, collegato al PC per la raccolta dei dati, è stato posizionato su un tavolo di fronte al partecipante. I due tasti di risposta sono stati allineati sia con la linea mediana del corpo dei partecipanti che con il centro del monitor. Questo era uno degli espedienti sperimentali volti ad evitare possibili effetti di compatibilità tra la mano che agisce sui tasti e il lato di presentazione dello stimolo, e fa parte di una più articolata strategia di contro-bilanciamento degli effetti della mano dominante. Una

telecamera a circuito chiuso, controllata a distanza dallo sperimentatore, è stata posizionata di fronte ai partecipanti, per controllare ed evitare qualsiasi movimento della testa o degli occhi.

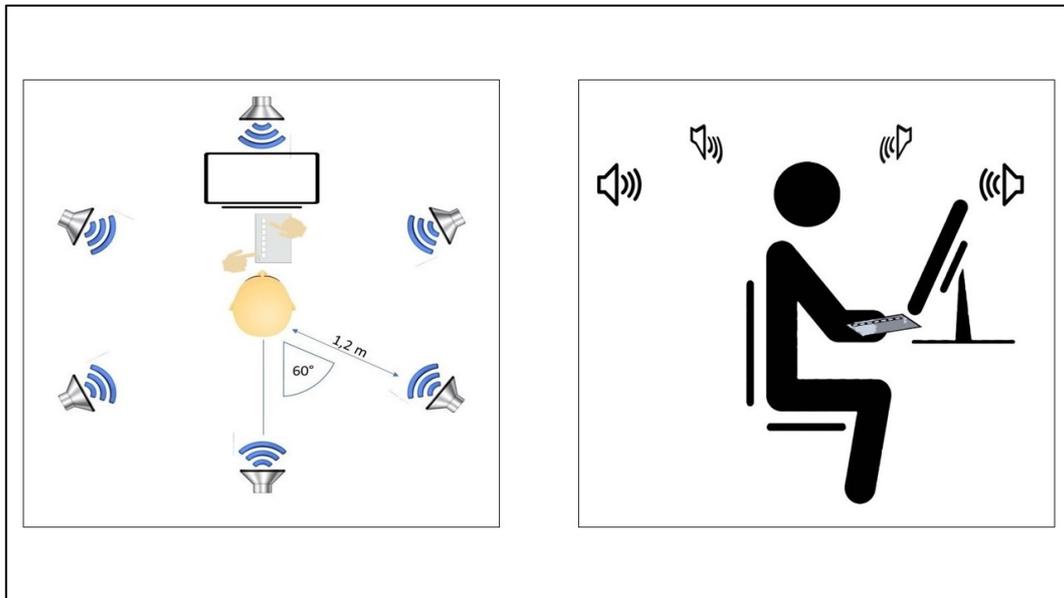


Figura 5: schema del setup sperimentale, visto in pianta (immagine a sinistra) e di profilo (immagine a destra)

Due suoni reali sono stati utilizzati come stimoli target in questo compito di discriminazione uditiva. Il suono originale, una registrazione di uno breve squillo di un telefono d'epoca (concesso a titolo gratuito da BBC Sound Effects Archive, <https://sound-effects.bbcrewind.co.uk>), è stato manipolato con il software Audacity®, un programma di produzione audio gratuito (versione 3.5; <https://www.audacityteam.org>), al fine di ottenere una seconda versione dello stesso suono, con una modulazione in frequenza di -10 punti percentuali. Questi squilli sono prodotti da un campanello posto all'interno del telefono ed attuato in modo elettromeccanico. Ai partecipanti è stato quindi detto che il suono più acuto era il "campanello più piccolo", mentre quello più grave era il "campanello più grande". La pressione sonora all'altezza delle orecchie è stata misurata con un fonometro digitale e tutti i volumi sono stati regolati per ottenere un picco di 75 dB durante l'erogazione del suono (Humes et al., 1996). Poiché il suono originale durava oltre 1000 ms, anche la sua durata è stata manipolata, accorciandone

l'attacco e il decadimento ma preservando l'identità e la riconoscibilità dei parametri sonori chiave, cioè l'altezza e il timbro. Questo intervento ha portato la durata finale a 500 ms, confrontabile con la letteratura precedente che adotta stimoli sonori del mondo reale in compiti uditivi (Kaganovich et al., 2013). L'analisi dei file Wave originale, la normalizzazione, la manipolazione della frequenza e l'accorciamento della durata, sono stati tutti eseguiti con il software Audacity sopra citato.

Gli stimoli visivi erano costituiti da frecce bianche su sfondo nero, con sei diversi orientamenti. Anche se presentate sul monitor del PC, e quindi su un piano verticale, ogni freccia indicava una specifica posizione dell'altoparlante nello spazio circostante. Durante la presentazione, il baricentro della freccia corrispondeva con il centro del monitor e la durata è stata fissata a 500 ms (Benso et al., 1998; Green et al., 2012). Durante gli studi pilota sono state testate diverse forme di freccia o di puntatore tridimensionale, tuttavia le frecce bidimensionali sono risultate le più facili e immediate da comprendere per i partecipanti.

Ogni prova aveva una durata massima di 2300 ms e iniziava con la presentazione del cue visivo (freccia) sullo schermo (durata del cue = 500 ms). L'intervallo cue-target era di 400 ms, dando così luogo ad un intervallo tra presentazioni di 900 ms (convenzionalmente indicato con l'abbreviazione *SOA*, ovvero *Stimulus Onset Asynchrony*). Uno dei due possibili suoni target veniva poi presentato in modo casuale da uno dei sei altoparlanti (la durata del suono, come detto in precedenza, era di 500 ms), aprendo contemporaneamente la finestra di risposta, della durata massima di 1400 ms. I partecipanti rispondevano a ciascun suono premendo uno dei due tasti sulla tastiera apposita, terminando così la singola prova. L'intervallo tra le prove è stato variato in modo casuale da 600 a 1000 ms, al fine di evitare possibili effetti ritmici nelle risposte.

L'esperimento completo era costituito da quattro blocchi di 192 prove ciascuno. Ogni stimolo uditivo aveva la stessa probabilità di essere presentato da ciascuno dei sei altoparlanti (2 stimoli uditivi x 6 posizioni altoparlanti = 12 possibili target). Oltre alle sei frecce, abbiamo aggiunto un cue neutro, rappresentato da un'immagine con tutte le frecce disposte attorno al centro dello schermo (vedi figura 6).

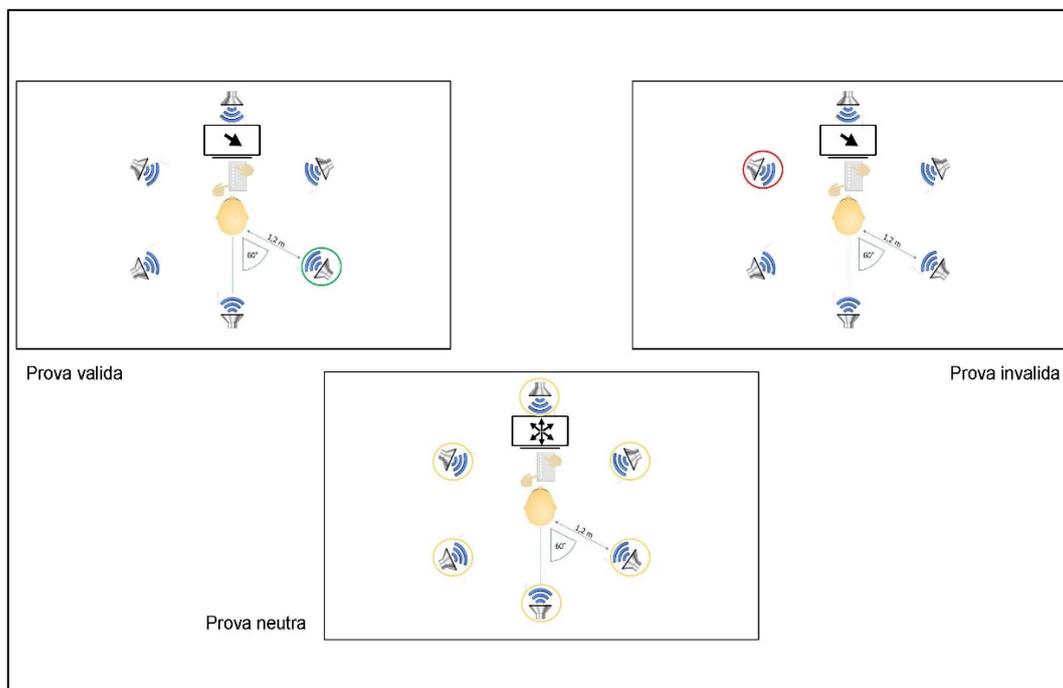


Figura 6: esempi di prova valida, invalida e neutra

Ciò ha prodotto un totale di 84 potenziali coppie cue-target (7 cue x 12 target) completamente randomizzate e presentate in ciascun blocco. Per quanto riguarda la probabilità di prove valide e invalide, in 120 prove il cue-freccia indicava esattamente la posizione del target, in 60 prove invece il cue non indicava esattamente la posizione del target, che quindi veniva presentato in modo casuale in una delle restanti cinque posizioni non indicate (10 prove per posizione). Nelle restanti 12 prove, il cue neutro era seguito da uno dei due target, casualmente presentato in una delle sei posizioni possibili. Ogni partecipante ha quindi completato un totale di 768 prove.

I partecipanti sono stati istruiti a spostare la loro attenzione uditiva in modo implicito - o *covert* - sulla posizione indicata dal cue, senza muovere la testa, il corpo o gli occhi, mantenendo cioè lo sguardo sul punto di fissazione centrale. Sono stati inoltre invitati a non chiudere mai gli occhi durante l'esecuzione del compito. Il loro compito era quello di discriminare i due stimoli uditivi, premendo il tasto in alto per il suono del campanello telefonico più grande, e il tasto in basso per il suono

del campanello più piccolo. Tra un blocco e l'altro, i partecipanti avevano qualche minuto di riposo prima di poter iniziare il blocco successivo.

Al fine di eliminare i possibili effetti di compatibilità spaziale tra lo stimolo uditivo e la risposta, nonché potenziali differenze derivanti dalla diversa mano dominante, ai partecipanti è stato richiesto di invertire la posizione della mano sul dispositivo di risposta prima dell'inizio di ogni blocco di prove successivo al primo. Mentre la mappatura della corrispondenza suono-tasto è rimasta costante per tutto l'esperimento, la disposizione delle mani sui tasti variava, così che in una configurazione la mano sinistra agiva sul tasto in alto e la destra sul tasto in basso, nell'altra configurazione accadeva l'inverso. Pertanto, i partecipanti hanno eseguito un numero uguale di prove con entrambe le configurazioni, ottenendo così un contro-bilanciamento interno.

Prima dell'inizio dell'esperimento, i partecipanti hanno svolto un breve training per familiarizzare con il setting e con il compito. In primo luogo, abbiamo verificato che ogni partecipante fosse effettivamente in grado di percepire correttamente la posizione degli stimoli uditivi presentati da ciascun altoparlante. In questa fase, i suoni sono stati presentati in sequenza circolare su ciascuno dei sei altoparlanti ed è stato chiesto di riportare verbalmente la posizione del suono percepito. La presentazione è iniziata con l'altoparlante centrale anteriore, quindi è passata a tutte le altre posizioni, in senso orario. Tutti i partecipanti sono stati in grado di indicare correttamente le sei posizioni. Successivamente, sono stati invitati a posizionare i loro indici sui tasti di risposta e, dopo aver introdotto la mappatura della corrispondenza tasto-suono, si sono esercitati a rispondere ai due suoni. La fase finale del training prevedeva che i partecipanti completassero un piccolo blocco di 44 prove, con funzione di esercitazione pratica. Ai partecipanti che non raggiungevano almeno il 70% di accuratezza totale in questo primo blocco di pratica, è stato chiesto di svolgerne un altro per consentirgli un'altra opportunità di raggiungere l'obiettivo minimo di accuratezza designato per considerare valido l'esperimento.

A conclusione della procedura sperimentale, abbiamo fornito a tutti i partecipanti un link per la compilazione da remoto dei questionari. Abbiamo utilizzato un codice alfanumerico univoco per ogni partecipante, che ci ha permesso di ricollegare in modo efficiente i dati dell'esperimento e dei questionari, garantendo così l'anonimità durante l'intero processo. I dati dei questionari in formato digitale sono stati raccolti attraverso la piattaforma software Qualtrics (Società Qualtrics, Provo, Utah, USA; versione 2023; <https://www.qualtrics.com>).

Nell'immagine sottostante (figura 7) sono riassunti e schematizzati sia il flusso di lavoro dell'intera procedura sperimentale, sia i dati principali dell'esperimento, inclusa la sequenza degli eventi della singola prova.

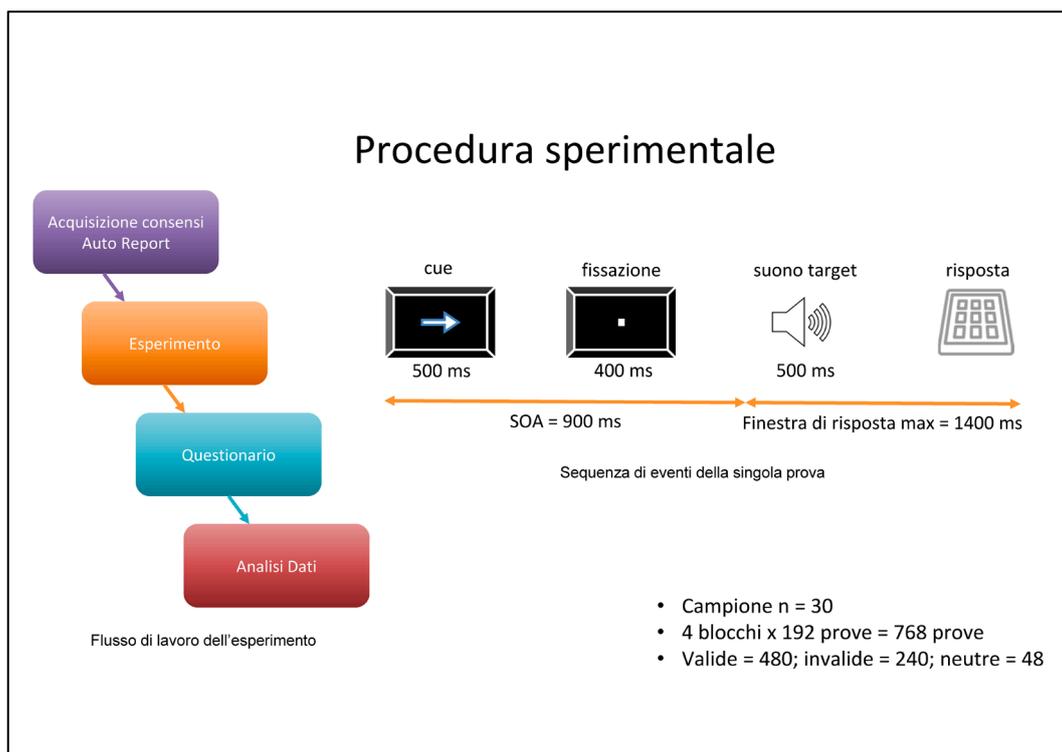


Figura 7: riassunto schematico della procedura sperimentale

### 4.3.2. Analisi dati

Per questo esperimento, sono state raccolte due tipologie di dati, che costituiscono le due variabili dipendenti: il tempo di risposta (TR) e l'accuratezza percentuale delle risposte (ovvero le variabili più comunemente indagate in questo tipo di paradigma), mentre le variabili indipendenti erano, rispettivamente, la direzione del cue visivo (7 cue diversi: 6 possibili direzioni spaziali ed un cue neutro) e la posizione del suono target (6 possibili posizioni, corrispondenti ai 6 altoparlanti predisposti). I dati sono stati esaminati utilizzando l'analisi della varianza per misure ripetute (ANOVA). Abbiamo valutato se l'ipotesi di sfericità fosse soddisfatta utilizzando il test di sfericità di Mauchy. Il grado di libertà veniva quindi aggiornato con l'epsilon corretto Greenhouse-Geisser in caso di violazione dell'assunto di sfericità. Le risposte errate e mancate sono state escluse dall'analisi. Le medie e le deviazioni standard dei tempi di risposta sono stati calcolati separatamente per ciascun partecipante e sono stati esclusi dall'analisi le prove con tempi superiori o inferiori a 3 deviazioni standard dalla media. Dal totale teorico di 23040 prove, risultante dalla somma delle prove di ognuno dei 30 soggetti, sono quindi state escluse 2499 prove, secondo i criteri appena esplicitati. I tempi di risposta e le accuratezze percentuali sono state oggetto di analisi separate.

Una prima ANOVA per misure ripetute è stata eseguita innanzitutto per verificare l'efficacia del compito, secondo una letteratura ben consolidata che indica risposte più rapide e accurate per prove valide, valori intermedi per prove neutre e valori peggiori per prove non valide (Posner & Boies, 1971; Posner, 1984). Inoltre, volevamo confrontare gli effetti dello spostamento dell'attenzione all'interno dello stesso emicampo e tra diversi emicampi, considerando la posizione spaziale reciproca di cue e target. La Figura 8 illustra la nostra definizione degli emicampi rispetto alla posizione del partecipante. Al solo scopo di una maggior chiarezza esplicativa, ad ogni posizione viene qui associato un numero da uno a sei, con il numero uno corrispondente alla posizione anteriore e centrale, senza per questo implicare o rimandare ad alcun tipo di ordinamento o codifica spaziale.

L'analisi prende in esame quindi due fattori intra-soggetto: la validità della prova (quattro livelli: valida, invalida con spostamenti attentivi nello stesso emicampo, invalida con spostamenti attentivi tra emicampi diversi, neutra) e l'emicampo del target (due livelli: anteriore e posteriore), per un totale di otto condizioni sperimentali. La tabella in figura 9 mostra alcuni esempi per ciascuna delle condizioni, seguendo la numerazione appena introdotta, con l'aggiunta della lettera "N" per il cue neutro.

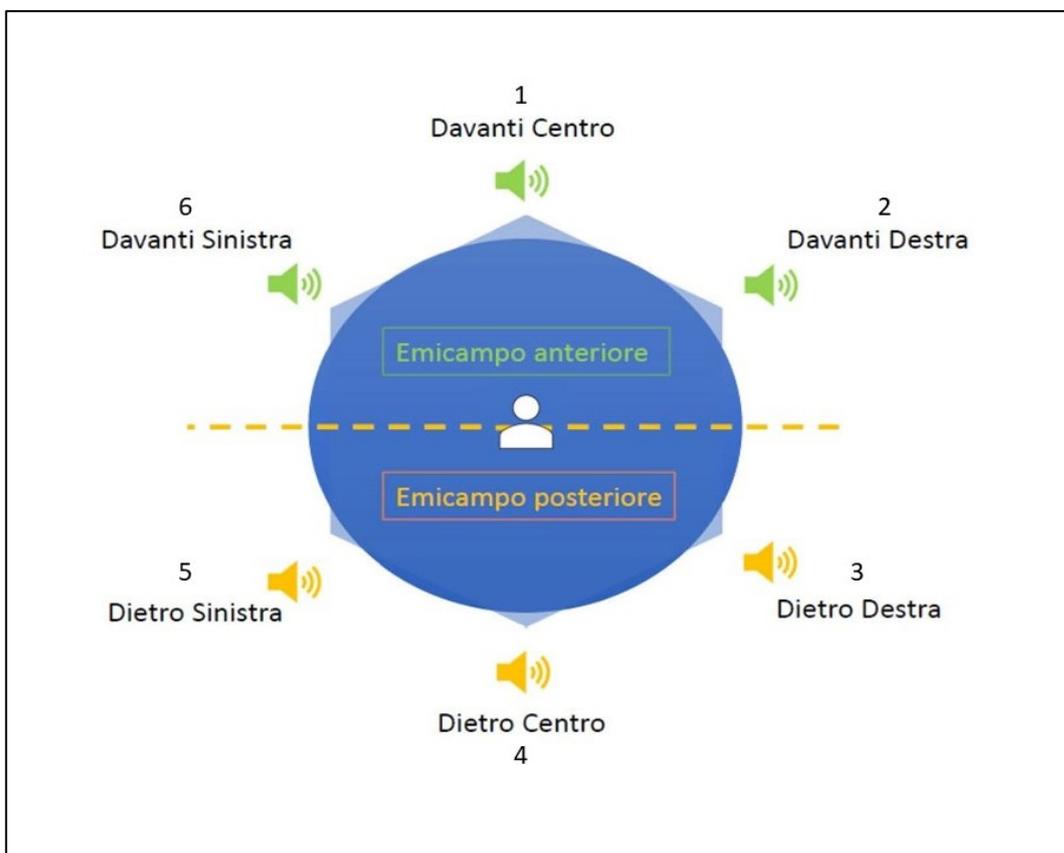


Figura 8: definizione emicampi anteriore e posteriore, in base alla posizione del partecipante

Condizione	Validità prova	Emicampo Target	Esempio
1	Valida	Anteriore	Cue 2-Target 2
2	Valida	Posteriore	Cue 5-Target 5
3	Invalida stesso emicampo	Anteriore	Cue 6-Target 1
4	Invalida stesso emicampo	Posteriore	Cue 4-Target 3
5	Invalida emicampo diverso	Anteriore	Cue 3-Target 6
6	Invalida emicampo diverso	Posteriore	Cue 2-Target 4
7	Neutra	Anteriore	Cue N-Target 6
8	Neutra	Posteriore	Cue N-Target 5

Figura 9: esempi di combinazione cue-target per ognuna delle otto condizioni sperimentali

Nella seconda analisi della varianza, abbiamo indagato le differenze nei costi di spostamento dell'attenzione in funzione della distanza tra posizione attesa (anticipata dalla direzione della freccia) e posizione effettiva di presentazione del target. A questo scopo, tutte le coppie cue-target sono state definite in termini di distanza angolare, ottenendo così un'analisi a fattore singolo denominato *distanza cue-target* (7 livelli:  $-180^\circ$ ,  $-120^\circ$ ,  $-60^\circ$ ,  $0^\circ$ ,  $+60^\circ$ ,  $+120^\circ$ ,  $+180^\circ$ ). Abbiamo assunto valori positivi per spostamenti attentivi in senso orario dalla posizione cue alla posizione target, e valori negativi per spostamenti in senso antiorario. Quindi, ad esempio, se il cue indicava la posizione 3 e il target veniva presentato in posizione 5, lo spostamento angolare è di  $+120^\circ$  di azimut; viceversa poteva darsi che il cue indicasse la posizione 3, ma il target venisse presentato in posizione 2, dando luogo ad uno spostamento di  $-60^\circ$  di azimut.

#### 4.3.3. Risultati

La prima analisi generale, che assumeva come variabile dipendente il tempo di risposta per tutte le condizioni sperimentali, ha mostrato un effetto principale significativo del fattore *validità prova* ( $F(1.97, 57.12) = 18.875$ ,  $p < 0.001$ ,  $\eta^2 = 0.394$ ). Nelle prove valide, il TR medio è stato di 655,8 ms SD 49,5 ms, nelle prove invalide con spostamento nello stesso emicampo è stato di 680,3 ms SD 53,4 ms,

nelle prove invalide con spostamento tra diversi emicampi è stato di 677,7 ms SD 54,5 ms e nelle prove neutre di 692,6 ms SD 68,6 ms. I confronti accoppiati, che adottano la correzione Bonferroni per confronti multipli, hanno evidenziato differenze significative tra i TR di prove valide e invalide con spostamento nello stesso emicampo ( $p < 0.001$ ), tra prove valide e prove invalide con spostamento tra diversi emicampi ( $p < 0.001$ ), tra prove valide e prove neutre ( $p < 0.001$ ) e tra prove invalide con spostamento tra diversi emicampi e prove neutre ( $p = 0.048$ ).

Come previsto, nelle prove valide i TR sono stati i più veloci, mentre nelle prove invalide si mostra un effetto attentivo generalizzato, che rappresenta i costi degli spostamenti tra le posizioni attese e quelle effettive. Questi costi sono resi evidenti dal rallentamento dei tempi di risposta. Sorprendentemente, le prove neutre hanno elicitato le risposte più lente, sebbene in linea teorica ci si aspettasse che fossero ad un livello intermedio fra prove valide e invalide. La difficoltà dei partecipanti in questa condizione è stata evidentemente maggiore rispetto alle altre, e questa evidenza ha suggerito alcune possibili interpretazioni, che saranno discusse di seguito. I risultati sono rappresentati nella figura 10.

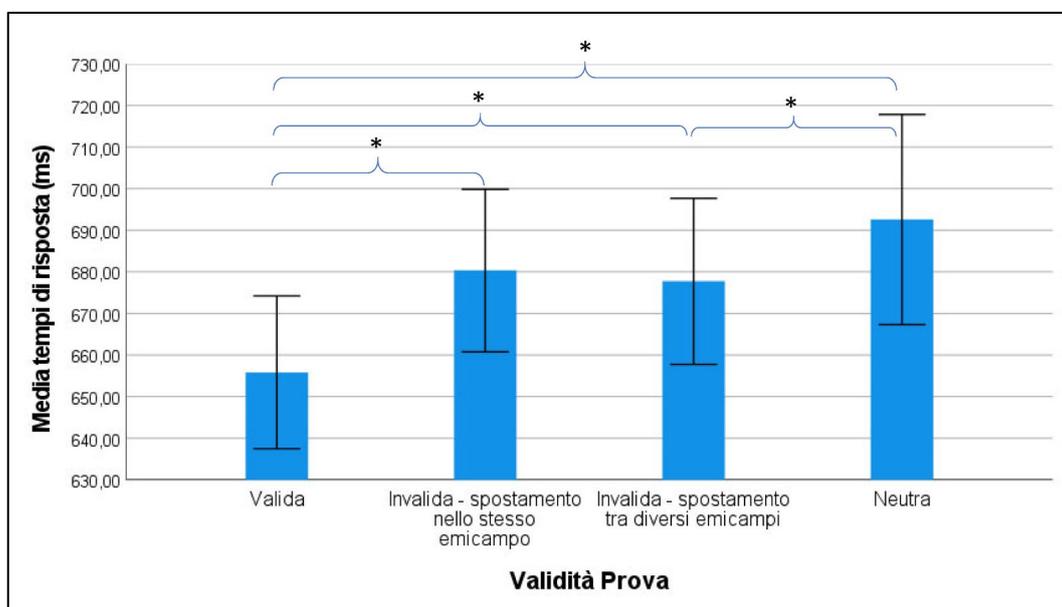


Figura 10: tempi di risposta in funzione del fattore Validità Prova

Il fattore *emicampo target* ha portato ad un ulteriore effetto principale significativo ( $F(1, 29) = 4.497, p = 0.043, \eta^2 = 0.134$ ) indicando risposte leggermente ma significativamente più rapide per i target presentati nell'emicampo posteriore. La media dei TR per le prove con presentazione del target nell'emicampo anteriore è stata di 678,7 ms SD 41,6 ms, mentre per i target presentati nell'emicampo posteriore è stata di 674,5 ms SD 38,7 ms. Riportiamo i risultati nella figura 11.

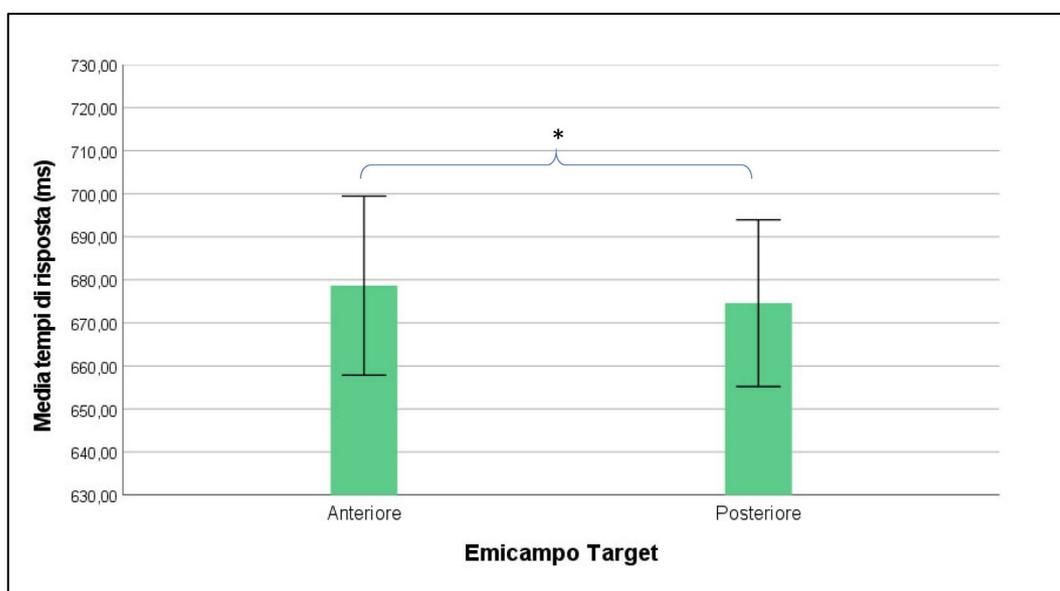


Figura 11: tempi di risposta in funzione dell'emicampo target

Inoltre, c'è stato un significativo effetto di interazione tra i due fattori ( $F(3, 87) = 11.823, p < 0.001, \eta^2 = 0.290$ ). Sono stati effettuati poi i relativi confronti, adottando la correzione di Bonferroni. Nelle prove valide la differenza tra i tempi di risposta con target anteriori e quelli con target posteriori non risulta significativa (657,2 ms SD 53,2 ms contro 654,5 ms SD 45,8 ms,  $p = 0.268$ ). Invece, nelle prove invalide con spostamento all'interno dello stesso emicampo, la differenza tra i tempi di risposta con target anteriori e posteriori è risultata significativa (687,1 ms SD 56,6 ms contro 673,6 ms SD 50,3 ms,  $p = 0.003$ ). Lo stesso è accaduto per le

prove invalide con spostamento tra diversi emicampi (TR con target anteriori 670,6 ms SD 51,9 ms, TR con target posteriori 684,8 ms SD 57,1 ms,  $p = 0.001$ ) e per le prove neutre (TR con target anteriori 699,8 ms SD 72,3 ms, TR con target posteriori 685,4 ms SD 64,7 ms,  $p = 0.004$ ).

Come si evince anche dalla figura 12, ad eccezione delle prove valide per le quali non si ravvisano differenze significative tra diversi emicampi, spostare l'attenzione all'interno dello stesso emicampo è risultato più facile nella zona posteriore, mentre spostarsi tra emicampi diversi era più difficile quando le frecce o cue indicavano posizioni anteriori e i target invece venivano presentati posteriormente. Inoltre, nelle prove neutre i TR con presentazione del target nell'emicampo posteriore erano significativamente più veloci di quelli con target presentati anteriormente.

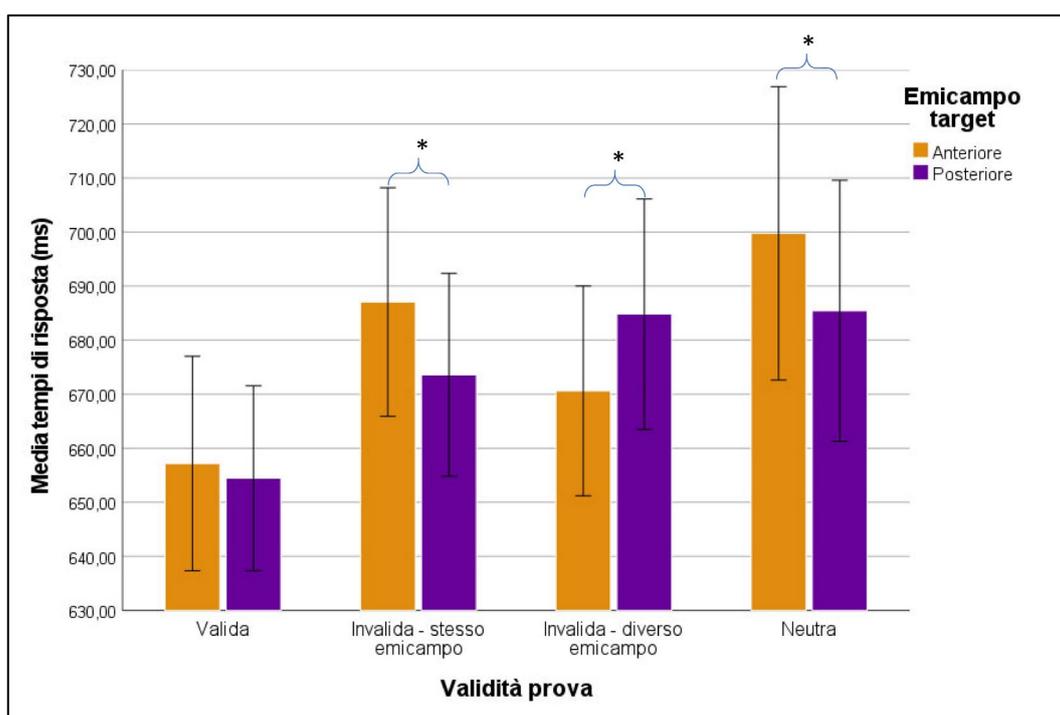


Figura 12: tempi di risposta in funzione di validità prova ed emicampo target

Per quanto riguarda, invece, l'accuratezza percentuale delle risposte, l'analisi specifica non ha portato ad alcuna differenza significativa tra le condizioni sperimentali. L'accuratezza generale dei partecipanti è stata in media di 89,2% SD 6,9%. Né il fattore *validità prova* ( $p = 0.350$ ) né l'*emicampo target* ( $p = 0.316$ )

hanno portato a variazioni statisticamente apprezzabili delle percentuali di correttezza delle risposte. I dati sono riportati in forma di tabella nella figura 13.

<b>Condizione (Validità prova - Emicampo target)</b>	<b>Accuratezza media in %</b>	<b>SD</b>
Prova valida - Target davanti	88,67	6,17
Prova valida - Target dietro	89,83	6,22
Prova invalida stesso emicampo - Davanti	88,47	7,40
Prova invalida stesso emicampo - Dietro	89,37	5,48
Prova invalida emicampo diverso - Da dietro a davanti	88,19	6,83
Prova invalida emicampo diverso - Da davanti a dietro	89,17	7,17
Prova neutra - Davanti	90,14	8,44
Prova neutra - Dietro	90,00	7,63

Figura 13: accuratezze medie % in funzione di Validità prova ed Emicampo target

La seconda analisi, introdotta per studiare i gradienti di attenzione spaziale uditiva, ha mostrato un effetto significativo del fattore di *distanza cue-target* ( $F(6, 174) = 9.581, p < 0.001, \eta^2 = 0.248$ ). Come ci aspettavamo, le risposte sono state più rapide nelle prove valide (distanza angolare cue-target =  $0^\circ$ ), poi si osserva un significativo rallentamento dei tempi di risposta, più pronunciato per le distanze più brevi ( $\pm 60^\circ$ ), un successivo graduale appiattimento della curva e un'inversione di tendenza per le distanze intermedie ed estreme ( $\pm 120^\circ$  e  $\pm 180^\circ$ ) come mostrato nella figura 14. I test di contrasto, infatti, hanno confermato che questo andamento quadratico era statisticamente significativo ( $F(1, 29) = 15.267, p < 0.001, \eta^2 = 0.345$ ) e quindi rappresentava la migliore approssimazione dei dati raccolti.

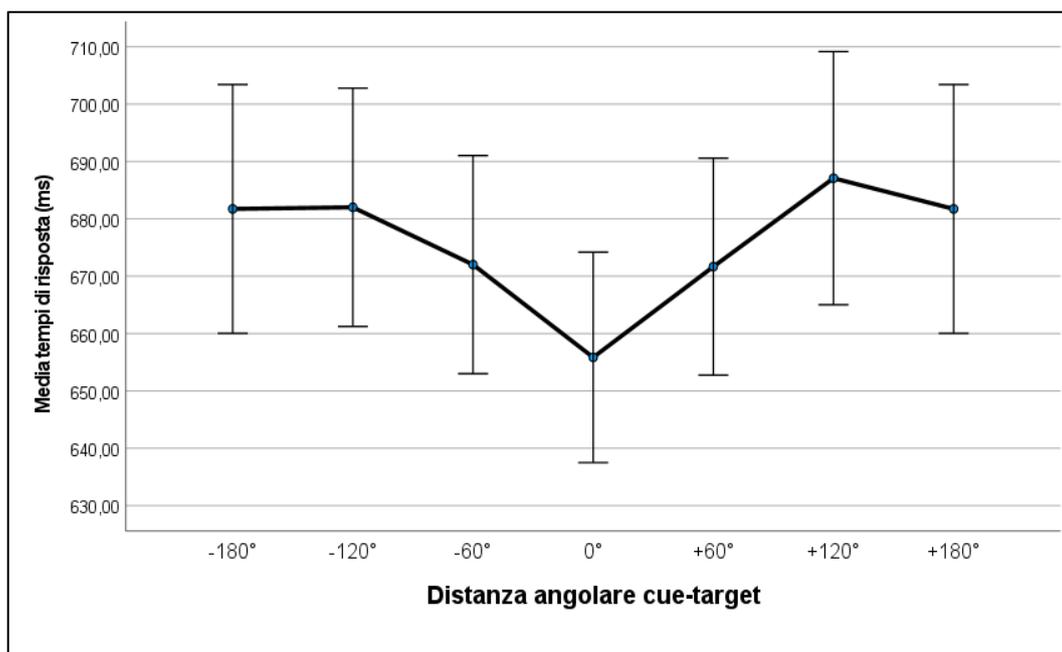


Figura 14: tempi di risposta in funzione della distanza angolare cue-target

#### 4.3.4. Discussione

I risultati osservati nella nostra prima analisi hanno confermato un'evidenza ben consolidata che dimostra come la correttezza dell'informazione preventiva sulla posizione del target porti a tempi di risposta più veloci e, di converso, la necessità di spostare l'attenzione dalla posizione attesa a quella effettiva nella quale viene presentato il target porti a tempi di risposta più lenti. Abbastanza sorprendentemente, per quanto riguarda i TR i risultati indicano che le prove neutre sono state le più impegnative e gravose per l'attenzione, mentre l'andamento tipico osservato in compiti simili le vede in una posizione intermedia tra prove valide e invalide (vedasi, ad es., Posner, 1980). Questo risultato inaspettato può indicare un carico di attenzione supplementare, dovuto alla particolarità della prova neutra nel compito da noi progettato. Infatti, quando i partecipanti ricevevano un cue neutro, veniva chiesto loro di distribuire la loro attenzione sull'intera scena uditiva, perché il successivo suono target avrebbe potuto essere presentato casualmente su uno dei sei altoparlanti. Ci pare, quindi, che questo possa essere un caso limite di attenzione selettiva spaziale, poiché lo spazio selezionato è, in realtà, l'intero spazio

circostante, e l'ampiezza di questa selezione potrebbe aver determinato un aggravio di difficoltà nel richiamo veloce di risorse attentive, una volta identificata la posizione effettiva del target. Un'altra ipotesi che potrebbe contribuire a spiegare questo tipo di risultato sulle prove neutre è un diverso comportamento dei partecipanti, che potrebbero non aver effettuato alcuno spostamento attentivo o selezione spaziale fino al momento della presentazione del target. In tal caso, lo spostamento richiesto avrebbe dovuto essere eseguito durante l'attacco del suono, e questo potrebbe aver allungato i tempi di risposta. Se si accetta questa ipotesi di spostamento attentivo contemporaneo ai primi istanti della presentazione del target, si dovrebbe accettare conseguentemente che la parte di informazione sonora contenuta in quei primi istanti sia stata persa o comunque non percepita per intero. Questo avrebbe potuto ragionevolmente portare ad una maggiore difficoltà di discriminazione tra i due suoni target e quindi ad un decremento di accuratezza, cosa che invece non è accaduta. Perciò, pur non potendo escludere del tutto tale seconda ipotesi, ci sembra che la prima sia maggiormente supportata dai dati raccolti.

Una delle nostre domande di ricerca era se vi fossero differenze negli effetti dell'attenzione spaziale uditiva tra l'emicampo anteriore e quello posteriore. Oltre al dato sull'effetto principale del fattore *emicampo target*, è l'interazione osservata tra i fattori *validità della prova* ed *emicampo target* ad essere maggiormente informativa ai fini del confronto di nostro interesse. Essa è trainata da due differenze significative, entrambe riguardanti le prove invalide. Infatti, con cue invalidi e spostamento attentivo all'interno verso target presentati nello stesso emicampo, le risposte sono state più rapide nell'emicampo posteriore, ovvero a parità di distanza media idealmente percorsa tra la posizione attesa e quella effettiva del target, l'emicampo posteriore ha suscitato risposte più rapide. Similmente, con cue invalidi e spostamenti tra diversi emicampi, le risposte sono state più lente quando i partecipanti sono stati indirizzati verso le posizioni anteriori e poi hanno dovuto spostare la loro attenzione verso posizioni target posteriori. Viceversa, quando si passava dalle posizioni posteriori a quelle anteriori, le risposte sono state più rapide.

Sebbene, al meglio della nostra conoscenza, quelli appena presentati siano i primi dati riguardanti esplicitamente un tale confronto, alcune evidenze attualmente disponibili (si veda Spence et al., 2020 per una rassegna) potrebbero portare a pensare che, nell'emicampo posteriore, la velocità e l'accuratezza delle risposte (o almeno una delle due misure) sarebbero state peggiori. Esistono, infatti, prove che indicano come la vista tenda a guidare le altre modalità a livello attentivo (Posner et al., 1976). Se questo fosse vero, si dovrebbe conseguentemente ipotizzare una performance peggiore nell'emicampo posteriore, dal momento che in quella porzione di spazio non è disponibile alcun riferimento visivo. Inoltre, altre evidenze neurofisiologiche indicano che sono diversi i circuiti cerebrali reclutati per rilevare e rispondere ai suoni presentati in diverse porzioni dello spazio (Previc, 2000; Graziano et al., 1999; 2004; Ho & Spence, 2009; Ocelli et al., 2011).

Alla luce di quanto appena esposto, i dati relativi alle prove invalide potrebbero suggerire la sussistenza di un contributo visivo alla rappresentazione dello spazio uditivo (Rizzolatti et al., 1987). In tal caso, le posizioni anteriori godrebbero di un ancoraggio più forte ed efficiente (Azzopardi & Cowey, 1993), dal quale risulta più difficile disimpegnare l'attenzione (Spence & Driver, 1996). Questo spiegherebbe la maggiore lentezza delle risposte a partire da cue anteriori, sia per spostamenti all'interno dell'emicampo anteriore, sia per spostamenti verso l'emicampo posteriore.

In assenza di differenze tra emicampo anteriore e posteriore, si poteva assumere a priori che spostamenti attentivi caratterizzati da distanze simili avrebbero portato a tempi di risposta simili, cosa che invece non è avvenuta. Anzi, i tempi di risposta per prove invalide con cue anteriori sono generalmente più lenti, a prescindere dalla distanza media da coprire, e questo è sorprendente dal momento che la distanza, per spostamenti all'interno del medesimo emicampo, era inferiore rispetto a quella per spostamenti verso l'emicampo opposto. Ciò accade anche per cue invalidi posteriori: distanze mediamente diverse hanno portato a tempi di risposta simili, suggerendo così l'esistenza di un processo qualitativamente diverso, responsabile dei risultati osservati.

Tuttavia, la consapevolezza della possibile influenza di riferimenti visivi in un compito del genere ci aveva portato ad adottare alcune scelte sperimentali mirate a minimizzare un potenziale vantaggio dell'emicampo anteriore dovuto alla visione diretta delle posizioni degli altoparlanti: infatti la stanza dell'esperimento era buia e tutti gli altoparlanti erano coperti da tessuto nero, cosicché i partecipanti non risultavano in grado di rilevare visivamente gli altoparlanti dell'emicampo anteriore. La maggiore difficoltà del disimpegno da posizioni anteriori può quindi avvalorare l'ipotesi che l'ancoraggio a quelle posizioni sia rafforzato non tanto da ciò che è effettivamente visibile, ma più profondamente da una diversa rappresentazione spaziale, guidata dalla modalità visiva e in qualche modo indipendente dalle informazioni disponibili in un determinato momento.

Ad ogni modo, era comunque presente un residuo di informazione visiva, costituito dal punto di fissazione e dai cue presentati a schermo. Non si può quindi escludere che questa informazione residua abbia avuto un impatto sulla rappresentazione spaziale dell'emicampo anteriore e sul rafforzamento degli ancoraggi alle corrispondenti posizioni. In tal caso, avremmo una conferma che detta rappresentazione sia, almeno in parte, sopra-modale. Per escludere l'ipotesi appena discussa, sarebbe necessario riproporre il medesimo compito, eliminando tanto la fissazione centrale quanto i cue visivi.

Si potrebbe inoltre descrivere lo stesso fenomeno affermando che l'assenza della visione diretta della fonte degli stimoli uditivi non ha portato ad un rallentamento dei tempi di risposta o un calo dell'accuratezza di esecuzione del compito di discriminazione. Infatti, i partecipanti hanno mostrato risposte più rapide dopo essere stati indirizzati nell'emicampo posteriore attraverso i corrispondenti cue, e questo è avvenuto sia in caso di spostamento all'interno dell'emicampo posteriore che in caso di spostamento verso l'emicampo anteriore. Questa diversa lente di osservazione, attraverso la quale possiamo guardare agli stessi risultati, rimanda alla prospettiva evolucionistica che considera l'udito come il nostro sistema di allertamento rapido (Schafer, 1977; Scharf, 1998 ed altri riportati nell'introduzione al presente esperimento). Con alcune ovvie limitazioni, derivanti dal fatto che in

questo compito non sono state riprodotte o emulate situazioni rischiose (i suoni, infatti, non avevano una riconosciuta valenza emotiva negativa), possiamo considerare i cue verso posizioni posteriori, e ancora di più i target presentati posteriormente, come "situazioni di allerta". Quando i partecipanti ricevevano un cue verso l'emicampo posteriore, dovevano dirigere la loro attenzione verso una posizione non controllabile con la vista, in meno di mezzo secondo, e rispondere il più rapidamente possibile. Per di più, soprattutto nelle prove neutre, si potrebbe ammettere che il suono target proveniente dall'emicampo posteriore abbia suscitato un allertamento superiore. Infatti i risultati in quella condizione indicano una differenza significativa tra i target anteriori e posteriori, con risposte significativamente più rapide per questi ultimi. Questa differenza ha contribuito in maniera maggioritaria all'effetto principale dell'emicampo target, riportato nell'analisi principale.

Infine, l'analisi del gradiente di attenzione spaziale uditiva suggerisce alcune considerazioni sui processi top-down e bottom-up nel dominio uditivo. La facilitazione per le prove valide, resa evidente dai tempi di risposta più rapidi, può essere interpretata come l'espressione di un meccanismo di selezione top-down, che sfrutta efficacemente la correttezza dell'informazione fornita dal cue. Questo processo può anche spiegare il decremento di velocità più rapido osservato per gli spostamenti attentivi verso le posizioni target più vicine, poiché la ragionevole conseguenza di un'efficace selezione spaziale è un forte effetto di inibizione intorno alla posizione attesa.

Alcuni autori (ad es. Rizzolatti et al., 1987) indagando il modo in cui l'attenzione si sposta in una determinata porzione di spazio in conseguenza di una necessità di ri-orientamento, hanno proposto due diverse ipotesi per la spiegazione del fenomeno. La prima era che gli spostamenti dell'attenzione fossero caratterizzati da una velocità costante, indipendentemente dalla quantità di spazio da coprire. Secondo questa ipotesi di forma lineare, ci aspetteremmo un decremento costante dei tempi di risposta all'aumentare delle distanze, con effetti attentivi direttamente proporzionali alla distanza angolare tra la posizione attesa e quella effettiva del

target. La seconda ipotesi era che la velocità degli spostamenti dell'attenzione dipendesse dalla quantità di spazio da coprire, implicando così che una diversa porzione di spazio potesse portare a diverse velocità di spostamento. I nostri dati, in linea con le più recenti evidenze (Golob et al. 2017; Golob & Mock, 2020), suggeriscono che l'ipotesi del gradiente di attenzione è quella che meglio rappresenta i risultati ottenuti.

Tuttavia, l'azione del solo processo volontario non può spiegare completamente i risultati osservati, in particolare l'appiattimento e l'inversione di tendenza dei TR per distanze angolari cue-target intermedie ed estreme. La letteratura precedente (vedasi in particolare Spence & Driver, 1994; Dalton & Lavie, 2004; Golob et al., 2017) ha sostenuto che un processo di monitoraggio contemporaneo e costante, dal basso verso l'alto, consente al sistema uditivo di distribuire una quota di risorse su tutti i 360° intorno al soggetto e di riorientare rapidamente l'attenzione verso stimoli che ricadono al di fuori della zona di inibizione sopra menzionata (vedasi Liljenström, 2003 per la definizione di questo *dilemma stabilità/flessibilità*). Se ammettiamo un'azione complementare multilivello di questi due processi, potremmo capire perché il gradiente osservato non ha una relazione lineare con la distanza angolare cue-target. Un forte effetto di inibizione sulle posizioni adiacenti a quella attesa, unito alla concomitante attività di monitoraggio dell'intera scena uditiva potrebbero aver portato al particolare andamento di risultati osservato. Inoltre, non abbiamo riscontrato tempi di risposta simili a quelli delle prove valide per nessuna posizione diversa da quella attesa, come invece riportato in un recente studio (Golob & Mock, 2020) nel quale le distanze cue-target estreme portavano a tempi di risposta paragonabili a quelli delle prove valide. Questo fenomeno potrebbe essere dovuto al fatto che gli autori avevano posizionato gli altoparlanti estremi in posizioni +/- 90° di azimut, ovvero esattamente davanti alle orecchie destra e sinistra. Gli stimoli uditivi presentati direttamente davanti alle orecchie potrebbero infatti risultare più facilmente percettibili rispetto a quelli presentati in posizioni intermedie: in questa posizione infatti sono massime sia la differenza di tempo inter-aurale (ITD) sia la differenza di intensità inter-aurale (IID), responsabili della rilevazione della posizione di stimoli uditivi nello spazio

circostante (Middlebrooks & Green, 1991; Blauert, 1997; Darwin, 2008; Grothe et al., 2010; Schnupp et al., 2011; Trainor, 2015). Stante questa possibile facilitazione per le posizioni davanti alle orecchie, le abbiamo volutamente evitate nella sistemazione circolare degli altoparlanti.

Ad ogni modo, anche senza l'inserimento di tali posizioni, l'andamento resta di forma quadratica, il che ci permette di affinare e rafforzare per il sistema uditivo l'ipotesi di un concorso di processi top-down di selezione e bottom-up di monitoraggio e riorientamento, quest'ultimo finalizzato non tanto ad una maggiore precisione, quanto ad una maggiore copertura spaziale e un rapido richiamo delle risorse attentive quando necessario. Nel complesso, i risultati confermano ed estendono la validità delle recenti evidenze, poiché qui è stato coperto l'intero spazio uditivo intorno al soggetto. Infatti, per quanto ne sappiamo, questo è uno dei primi tentativi di studiare l'attenzione spaziale su tutto l'arco dei 360° della scena uditiva, condizione che – auspicabilmente - dovrebbe essere considerata la più ecologica per la ricerca nella modalità uditiva.

Nel complesso, i dati raccolti e discussi in questo primo esperimento hanno sollecitato domande di ricerca ulteriori rispetto a quelle di partenza. Ci chiediamo, infatti se, e fino a che punto, l'informazione visiva costituita dai cue visivi abbia avuto un impatto sull'andamento dei risultati, con particolare riferimento alla maggiore difficoltà di disimpegno dell'attenzione da posizioni nell'emicampo anteriore, rispetto a quanto osservato nell'emicampo posteriore. Ci chiediamo, inoltre, se sia piuttosto il tipo di attenzione richiesto dal compito ad aver influito sui risultati, se cioè il fatto di essere continuamente diretti verso posizioni diverse, anziché verso una posizione costante in un dato tempo, possa aver determinato in tutto o in parte l'andamento dei risultati.

#### **4.4. Esperimento 2: attenzione sostenuta**

Ferme restando le considerazioni fatte nell'introduzione al precedente studio per quanto concerne le scelte sperimentali di fondo, il setting e gli stimoli, i risultati

appena esposti e discussi ci hanno suggerito di esplorare ulteriormente gli effetti dell'attenzione spaziale uditiva in tutto lo spazio circostante.

La letteratura precedente ha evidenziato la mancanza di studi incentrati su un confronto diretto tra l'attenzione spaziale transiente e quella sostenuta, sia nella modalità visiva che uditiva (Spence et al., 2020). La ragionevolezza di questo confronto può essere dimostrata dal fatto che il nostro ambiente uditivo reale è ricco di entrambi i tipi di situazioni, che ora richiedono frequenti spostamenti attentivi (ad es. durante conversazioni con più persone o durante lo svolgimento di più attività contemporanee), ora invece reclamano il mantenimento del focus su una posizione o un oggetto preciso (ad es. durante una conversazione con una sola persona o durante la scrittura). Nessuna delle due tipologie di attenzione può essere quindi sufficientemente rappresentativa di tale varietà, per questo motivo abbiamo cercato di confrontarne gli effetti all'interno di un contesto sperimentale simile.

Una delle nuove domande di ricerca era quindi se un focus attentivo predeterminato e sostenuto su una posizione, anziché continui cambiamenti dovuti a cue casuali, avrebbe suscitato andamenti simili o diversi dei tempi di risposta e delle accuratezze percentuali, in una o in più condizioni sperimentali.

L'adozione di un paradigma di attenzione sostenuta, ci ha anche permesso di non utilizzare i cue visivi, dal momento che il suggerimento spaziale veniva dato ai partecipanti a voce e prima dell'inizio di ogni blocco di prove. Questa scelta sperimentale apriva alla possibilità di valutare se le differenze precedentemente osservate tra gli emicampi anteriore e posteriore fossero dovute all'informazione visiva residua anteriormente o, piuttosto, ad un contributo visivo di alto livello alla rappresentazione dello spazio uditivo.

Infine, eravamo interessati a verificare se il gradiente attentivo di forma quadratica, osservato precedentemente, sarebbe stato replicato o meno in questo nuovo contesto.

#### **4.4.1. Materiali e metodologia**

In questo secondo studio sono stati testati un totale di 30 adulti ( $n=30$ , età media = 22,4; SD = 4,1; 25 femmine; 27 destrorso). Per la scelta della dimensione campionaria, oltre all'esigenza di ottenere due gruppi di egual numero per poterli adeguatamente confrontare a livello statistico, restano ferme le considerazioni e i calcoli effettuati per il primo esperimento. Prima di partecipare all'attività sperimentale, tutti i soggetti hanno letto e sottoscritto il consenso informato e il trattamento dei dati personali a fini di ricerca, secondo i moduli forniti dal Comitato di Bioetica dell'Università di Bologna. Le eventuali patologie a carico dell'udito, il livello di istruzione e il background musicale sono stati sottoposti a screening come nell'esperimento 1. In base ai criteri precedentemente indicati, i dati raccolti non hanno comportato alcuna esclusione di partecipanti.

La procedura sperimentale era simile a quella utilizzata nell'Esperimento 1, con le seguenti eccezioni. Non abbiamo utilizzato cue visivi, poiché la posizione attesa veniva comunicata ai partecipanti prima dell'inizio di ogni blocco di prove. Pertanto, non è stato necessario alcun monitor, e il punto di fissazione per lo sguardo era costituito da una piccola luce led di colore bianco caldo, posizionata di fronte ai partecipanti ad un'altezza corrispondente al punto di fissazione a schermo utilizzato nel primo esperimento. Gli stimoli uditivi, invece, erano esattamente gli stessi.

Come nell'esperimento 1, ogni partecipante ha svolto una fase preliminare di test, per verificare che percepisse correttamente ognuna delle sei possibili posizioni target. Tutti i partecipanti sono stati in grado di percepire i suoni nelle posizioni previste. Sono stati anche istruiti a posizionare i loro indici sui tasti di risposta e sono stati informati sulla mappatura della corrispondenza tasto-suono. Il dispositivo di risposta era lo stesso di quello utilizzato nell'esperimento 1 e abbiamo adottato la stessa procedura per controbilanciare i possibili effetti di compatibilità dovuti alla mano dominante.

Abbiamo iniziato con un breve blocco di pratica, eventualmente ripetuto per coloro che non avevano raggiunto almeno il 70% di accuratezza totale al primo tentativo.

Nelle sessioni di pratica, ai partecipanti è stato detto di fissare la loro attenzione sulla posizione anteriore frontale e di mantenere lì il focus per tutta la durata del blocco di prove, a prescindere dalla direzione dalla quale fosse arrivato il suono target.

L'esperimento era costituito da sei blocchi di prove, uno per ogni possibile posizione degli altoparlanti. Prima di iniziare ogni blocco, i partecipanti ricevevano istruzioni su quale delle sei fosse la posizione attesa per l'intera durata del blocco. L'ordine delle posizioni è stato pseudo-randomizzato, per garantire che ogni partecipante avesse un ordine diverso e, allo stesso tempo, che le posizioni di partenza fossero equamente distribuite tra tutti i soggetti. Inoltre, abbiamo controllato tutte le combinazioni, al fine di evitare che più posizioni adiacenti fossero testate consecutivamente in senso orario o antiorario.

Ogni prova aveva una durata massima di 1800 ms e iniziava con un intervallo pre-stimolo (durata = 400 ms) nel quale i partecipanti erano invitati solamente ad osservare il led di fissazione. Uno dei due suoni target veniva poi presentato in modo casuale su uno dei sei altoparlanti (durata degli stimoli sonori = 500 ms), aprendo contemporaneamente una finestra di risposta massima di 1400 ms. I partecipanti rispondevano a ciascun suono premendo uno dei due tasti sul dispositivo di risposta, terminando così la prova. L'intervallo tra le prove è stato variato in modo casuale da 600 a 1000 ms al fine di evitare possibili *effetti di ritmo* nelle risposte. I partecipanti hanno avuto qualche minuto di riposo prima di iniziare il blocco successivo. Le posizioni delle mani sui tasti di risposta sono state controbilanciate come nell'esperimento 1.

Entrambi gli stimoli sonori potevano essere presentati in ciascuna delle sei posizioni, dando così luogo a 12 diversi target. Ogni blocco consisteva in 120 prove e, per ogni blocco, la maggior parte dei target è stata effettivamente presentata in posizione attesa (prove valide), mentre alcuni di essi sono stati presentati altrove (prove invalide). Le prove valide erano i 2/3 (66,6%) del totale. Di conseguenza, ogni partecipante ha completato un totale di 720 prove (120 prove x 6 blocchi), di cui 480 valide e 240 invalide. Non erano presenti prove neutre, dal momento che

non veniva utilizzato il cue e l'interesse di ricerca era di studiare l'attenzione sostenuta anziché transiente. Alla fine dell'esperimento, abbiamo chiesto a tutti i partecipanti di compilare lo stesso questionario online somministrato nell'esperimento 1. Come già fatto in precedenza, riassumiamo la procedura e i dati principali dell'esperimento nella figura 15.

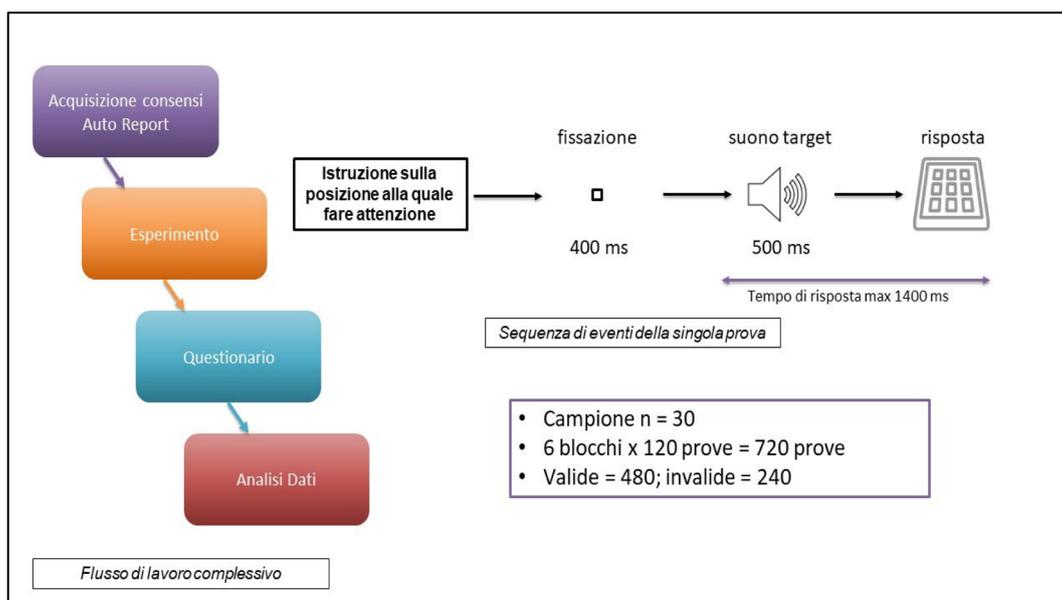


Figura 15: riassunto procedura sperimentale esperimento 2

#### 4.4.2. Analisi dati

Anche per questo secondo esperimento, le variabili dipendenti erano i tempi di risposta e l'accuratezza percentuale. I dati sono stati studiati utilizzando l'analisi della varianza a misure ripetute. Abbiamo corretto il grado di libertà usando l'epsilon di Greenhouse-Geisser ogniqualvolta l'assunzione di sfericità è stata violata. Le prove con risposta errata o mancata e quelle con tempi di risposta eccedenti le 3 deviazioni standard, calcolate singolarmente per ogni partecipante, sono state escluse dall'analisi. Dal totale teorico di 21600 prove, sono quindi state escluse dalle analisi 2088 prove, secondo i criteri appena elencati.

La prima ANOVA generale per misure ripetute è stata eseguita anzitutto per valutare l'efficacia del compito, testimoniata dalla eventuale replica degli risultati tipicamente riportati in letteratura. La stessa analisi mirava a confrontare gli effetti

dell'attenzione nell'emicampo anteriore e posteriore, in tutte le condizioni sperimentali. Ciò ha portato allo studio di due fattori: la *validità della prova* (tre livelli: valida, invalida con spostamento nello stesso emicampo, invalida con spostamento in diverso emicampo) e l'*emicampo del target* (due livelli: anteriore e posteriore). Analisi separate sono state condotte per i tempi di risposta e per le accuratezze percentuali.

Nella seconda analisi della varianza, abbiamo testato le differenze nei costi di spostamento dell'attenzione in funzione della distanza angolare tra la posizione attesa e quella effettiva di presentazione del target. Questo ha portato ad uno studio a fattore singolo, ovvero la *distanza angolare posizione attesa-effettiva* (7 livelli possibili, che vanno da  $-180^\circ$  a  $+180^\circ$  di azimut).

#### **4.4.3. Risultati**

La prima ANOVA generale, sui tempi di risposta, ha mostrato un effetto principale significativo del fattore *validità della prova* ( $F(2, 58) = 38.067, p < 0.001, \eta^2 = 0.568$ ). Abbiamo poi eseguito i confronti tra i dati delle singole condizioni di validità della prova, adottando la correzione di Bonferroni. Le risposte per le prove valide sono state le più rapide (673,5 ms SD 67,5 ms) mentre i tempi di risposta sono stati significativamente più lenti sia per le invalide con spostamento all'interno del medesimo emicampo (702,4 ms SD 75,9 ms,  $p < 0.001$ ) che per le invalide con spostamento tra diversi emicampi (700,9 ms SD 78,5 ms,  $p < 0.001$ ). Guardando poi all'interno delle prove invalide, lo spostamento all'interno dello stesso emicampo o tra diversi emicampi non ha portato a differenze significative ( $p = 1$ ). Inoltre, c'è stato un effetto principale significativo del fattore *emicampo target* ( $F(1, 29) = 7.228, p = 0.012, \eta^2 = 0.200$ ): sia per le prove valide che per le invalide, le risposte sono state più lente quando i target sono stati presentati nell'emicampo posteriore. L'interazione tra i due fattori, invece, non è risultata significativa ( $p = 0.914$ ). I grafici di questi due effetti principali sono riportati, rispettivamente, nelle figure 16 e 17.

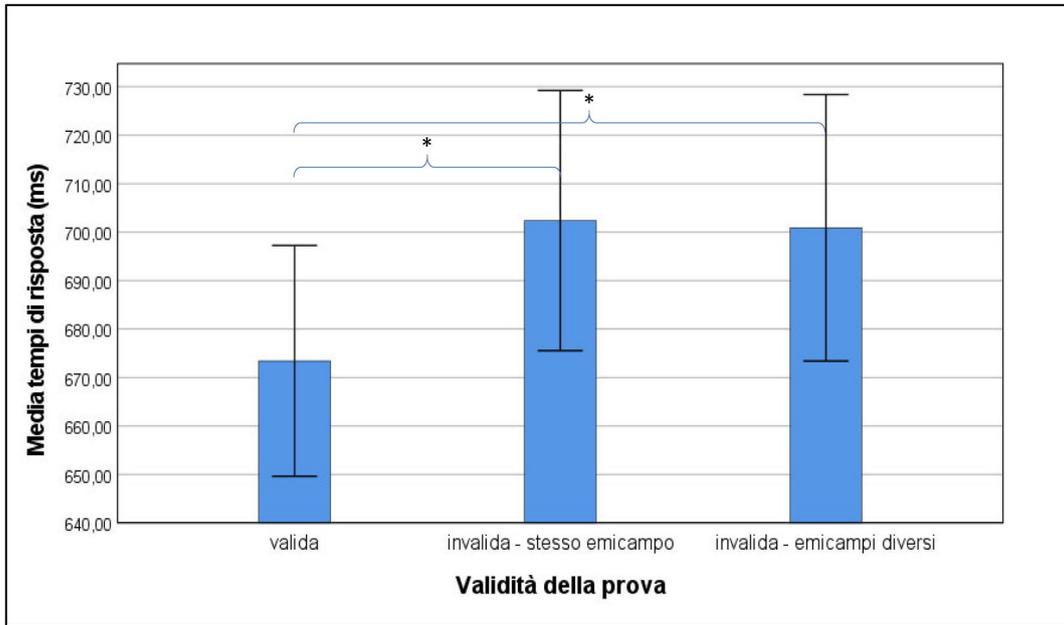


Figura 16: tempi di risposta in funzione della validità della prova

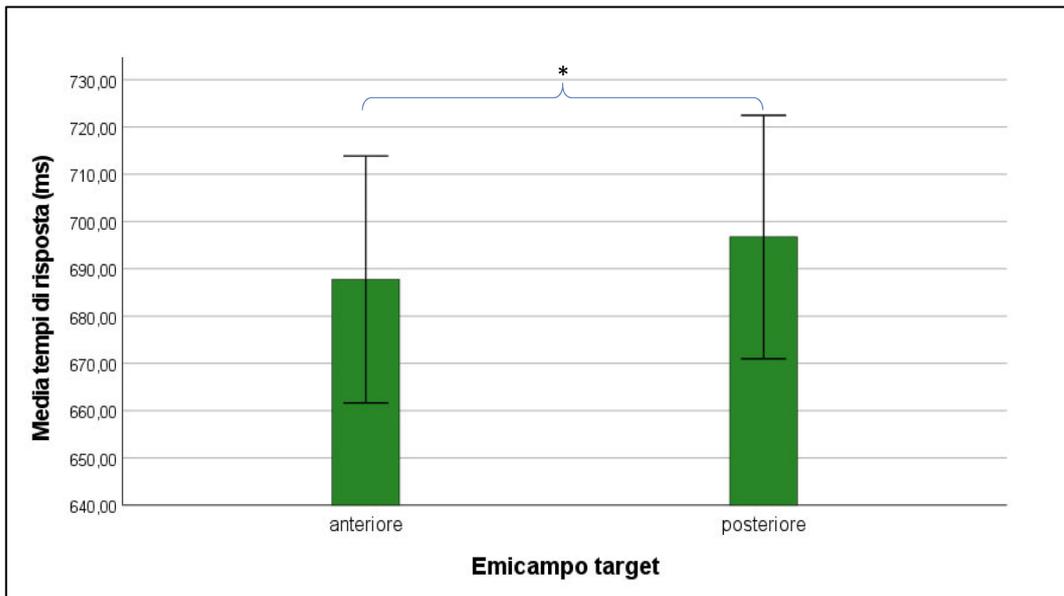


Figura 17: tempi di risposta in funzione dell'emicampo del target

Per quanto riguarda le accuratezze percentuali di risposta, l'analisi dedicata ha mostrato un effetto principale significativo del fattore *emicampo target* ( $F(1, 29) = 5.931, p = 0.021, \eta^2 = 0.170$ ). Le risposte con target presentati nell'emicampo posteriore sono state più accurate di quelle presentate anteriormente (89,9% SD 4,7% vs 88,2% SD 5,9%). I risultati sono riportati nel grafico in figura 18.

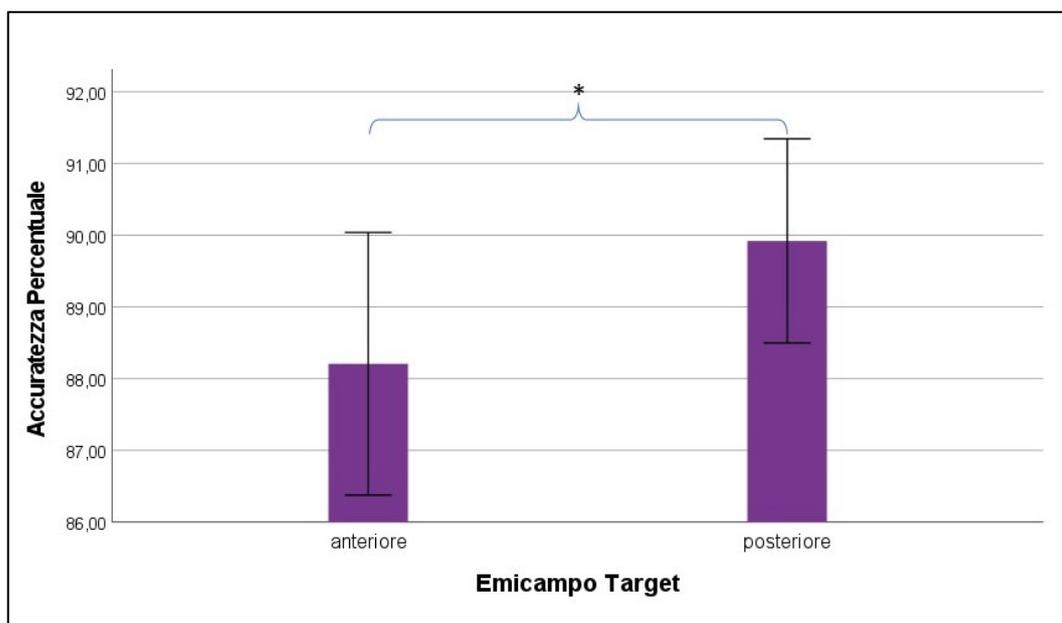


Figura 18: accuratezze percentuali in funzione dell'emicampo del target

Nella stessa analisi, il fattore *validità della prova* risulta molto vicino alla significatività statistica ( $p = 0.053$ ). Le prove valide hanno portato ad una accuratezza media di 89,7% SD 6,3%, mentre le prove invalide con spostamento nello stesso emicampo ad un valore di 89,2% SD 5,7% e le prove invalide con spostamento tra diversi emicampi ad un valore di 88,3% SD 5,7%.

Infine, l'analisi di varianza per i gradienti di attenzione spaziale uditiva, con la quale si volevano indagare i costi dello spostamento dell'attenzione in funzione della distanza angolare tra posizione attesa ed effettiva, ha mostrato un effetto principale significativo del fattore *distanza* ( $F(6, 174) = 13.792, p < 0.001, \eta^2 = 0.322$ ). Come prevedibile, le risposte sono state più rapide nelle prove valide (distanza =  $0^\circ$ ), poi abbiamo osservato una diminuzione della velocità più netta per le distanze più brevi (distanze =  $\pm 60^\circ$ ) seguita da una stabilizzazione e una leggera inversione di tendenza per quelle intermedie ed estreme (distanze =  $\pm 120^\circ$  e  $\pm 180^\circ$ ), come risulta dal grafico in figura 19. Nell'analisi di contrasto, l'approssimazione alla forma quadratica della curva dei tempi di risposta era ancora statisticamente significativa ( $F(1, 29) = 7.479, p = 0.011, \eta^2 = 0.205$ ).

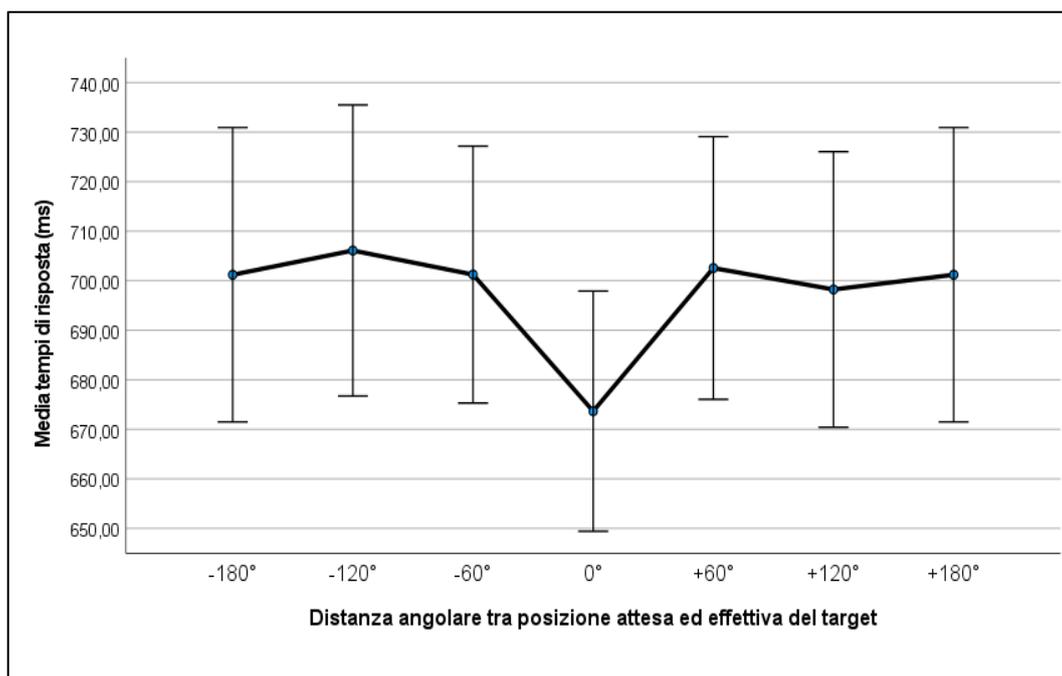


Figura 19: tempi di risposta in funzione della distanza angolare tra posizione attesa ed effettiva del target

#### 4.4.4. Discussione

Una delle domande di ricerca di questo secondo esperimento era se la diversa tipologia di attenzione richiesta avrebbe portato ad osservare andamenti diversi nei risultati di uno stesso compito di discriminazione uditiva.

Sotto questo aspetto, la prima analisi, che mostrava un effetto principale del fattore *validità della prova*, può essere interpretata come un marcatore di efficacia del compito. Infatti, sia per i tempi di risposta che per l'accuratezza, nelle prove valide i partecipanti hanno avuto prestazioni migliori, suggerendo così un vantaggio dovuto alla conoscenza della posizione cui fare attenzione. Di converso, per le prove invalide, la necessità di spostamenti attentivi dalle posizioni attese a quelle effettive di presentazione del target è ben rappresentata da tempi di risposta mediamente più lenti. Inoltre, l'emicampo di appartenenza delle posizioni attese non ha portato ad alcuna differenza sui tempi di risposta, come invece si osserva per l'emicampo di presentazione del target. I partecipanti erano leggermente ma significativamente più veloci quando i suoni venivano presentati anteriormente, mentre i target posteriori portavano a risposte più lente. Questo andamento dei

risultati ci porta a suggerire che, in presenza di una pre-informazione sulla posizione cui prestare attenzione sostenuta, i suoni presentati nell'emicampo anteriore elicitano risposte più rapide, mentre c'è una difficoltà aggiuntiva quando essi provengono da dietro. Le possibili implicazioni derivanti dalle differenze di andamento riscontrate nei compiti di attenzione transiente e sostenuta sono descritte nella successiva discussione generale.

L'analisi delle accuratezze percentuali mostra una maggior precisione nelle risposte con target presentati nell'emicampo posteriore, al contrario di quanto accadeva per i tempi di risposta, più rapidi anteriormente. Questi dati contro-tendenti potrebbero suggerire l'adozione di strategie non univoche per lo svolgimento del compito, da parte di alcuni partecipanti. Ricordiamo infatti che il compito richiedeva sia di rispondere più rapidamente possibile, che di essere più accurati possibile, senza stabilire alcuna priorità tra i due obiettivi. Nell'emicampo anteriore si assiste a risposte più rapide ma meno accurate, mentre nell'emicampo posteriore avviene il contrario, lasciando quindi un interrogativo riguardo a quale aspetto del compito sia stato privilegiato maggiormente nei due casi.

Un'altra motivazione di questo secondo esperimento era verificare se l'andamento di forma quadratica del gradiente di attenzione osservato in condizioni di attenzione transiente sarebbe stato replicato o meno in condizioni di attenzione sostenuta.

La seconda analisi, che metteva a fattore la distanza angolare tra la posizione attesa in modo sostenuto e la posizione effettiva di presentazione del target, ha mostrato lo stesso gradiente di forma quadratica osservato nell'esperimento 1. La discussione fatta in precedenza potrebbe essere considerata valida anche per il presente risultato, con alcune integrazioni. La zona di rallentamento dei TR, generata dall'attenzione sostenuta sulla stessa posizione durante un intero blocco di prove, sembra essere più ristretta, poiché vediamo una significativa diminuzione della velocità solo per distanze angolari di  $\pm 60^\circ$  di azimut. Oltre questi valori, i tempi di risposta tendono a stabilizzarsi e non si osservano variazioni significative. Il vantaggio dovuto alla conoscenza previa della posizione più probabile del target sembra quindi portare non solo a risposte più rapide e tendenzialmente più accurate,

ma anche ad un'area di rallentamento spazialmente più concentrata, al di fuori della quale i costi dello spostamento dell'attenzione tra posizione attesa e posizione reale del target non subiscono ulteriori significativi incrementi all'aumentare della distanza da coprire. Si potrebbe quindi ipotizzare un diverso rapporto reciproco ed una diversa area di influenza tra processi top-down e bottom-up, che porta da un lato ad una selezione spaziale più forte e ristretta ad opera di meccanismi volontari, dall'altro ad un ampliamento dell'area di influenza dei processi di basso livello, che mostra i suoi effetti fino a posizioni assai vicine a quella attesa.

#### **4.5. Confronto tra i due compiti e discussione generale**

Il presente studio di attenzione spaziale intendeva indagare, singolarmente e in un'ottica di confronto, gli effetti dell'attenzione transiente e sostenuta su uno stesso compito di discriminazione uditiva, al fine di comprendere meglio i meccanismi alla base di una competenza chiave per la vita quotidiana. In effetti, siamo costantemente esposti ed inseriti in uno di questi due scenari, ora mantenendo il focus su una sola posizione, ora invece seguendo continuamente suggerimenti spaziali e spostando la nostra attenzione in diverse direzioni (Spence et al., 2020). Dal momento che nessuna di queste due situazioni può essere considerata completamente o anche solo maggiormente rappresentativa del reale, ognuno dei risultati appena riportati non dovrebbe essere letto da solo, bensì come una piccola parte di un modello complesso e sfaccettato.

Per comprendere più da vicino eventuali differenze di andamento tra i due compiti, abbiamo effettuato una serie di analisi. Una prima ANOVA, sui tempi di risposta e sulle accuratezze, metteva quindi a fattore anche il tipo di compito come variabile *between*. I fattori *within*, ovvero la *validità della prova* e l'*emicampo target* sono stati invece mantenuti. Per quanto riguarda le accuratezze percentuali, i risultati non indicano differenze significative tra i due compiti, dal momento che sia il fattore *compito* che l'interazione tra il compito e i diversi fattori intra-soggettivi non raggiungono la significatività (fattore *compito*  $p = 0.93$ , interazione *emicampo target x compito*  $p = 0.46$ , interazione *validità prova x emicampo target x compito*

$p = 0.79$ ). I risultati sui tempi di risposta non hanno mostrato un effetto significativo del fattore *compito* ( $p = 0.18$ ), indicando che, in linea generale, la tipologia di attenzione richiesta non ha influito sulla velocità di esecuzione della discriminazione. Si osserva, invece, una interazione significativa tra i fattori *emicampo target* e *compito* ( $F(1, 58) = 5.733, p = 0.020, \eta^2 = 0.090$ ). Nel compito di attenzione transiente, infatti, i target presentati nell'emicampo posteriore portavano a risposte leggermente più veloci mentre nel compito di attenzione sostenuta le risposte sono state più veloci con target presentati nell'emicampo anteriore.

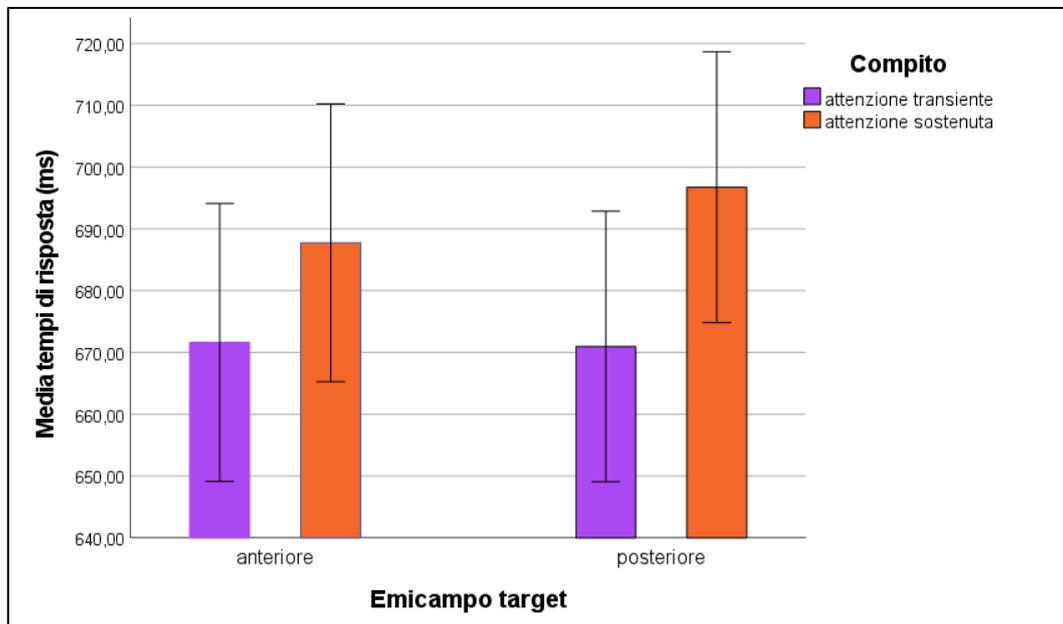


Figura 20: tempi di risposta in funzione di emicampo target e tipo di compito

Ad ogni modo, come si evince dal grafico in figura 20 appena riportato, questo specifico risultato non è accompagnato da differenze statisticamente significative sui confronti separati per singola condizione. Infatti, risulta che non vi siano differenze tra i due compiti né per i tempi di risposta a stimoli presentati nell'emicampo anteriore (attenzione transiente 671,6 ms SD 53,9 ms, attenzione sostenuta 687,8 ms SD 74,1 ms,  $p = 0.31$ ), né per i tempi di risposta a stimoli presentati nell'emicampo posteriore (attenzione transiente 671,0 ms SD 51,0 ms, attenzione sostenuta 696,7 ms SD 73,8 ms,  $p = 0.10$ ). L'effetto di interazione

osservato, quindi, dovrebbe derivare principalmente dalla differenza di andamento dei tempi di risposta tra emicampo anteriore e posteriore, osservata nei due compiti. È possibile che la condizione di attenzione transiente abbia generato maggior allertamento o reattività soprattutto nell'emicampo posteriore dal momento che, prova dopo prova, poteva cambiare sia la direzione del cue che la posizione del target. Al contrario, nel compito di attenzione sostenuta la posizione attesa era conosciuta a priori e mantenuta per l'intera durata di un blocco di prove, e questo può aver contribuito ad un minor grado di reattività e prontezza nelle risposte, soprattutto con target presentati nell'emicampo posteriore. Se questa interpretazione fosse plausibile, ci saremmo dovuti aspettare una velocità generalmente maggiore nel compito di attenzione transiente, cosa che in realtà si verifica, ma la differenza con i dati del compito di attenzione sostenuta non raggiunge la significatività (media TR compito attenzione transiente = 671,3 ms SD 50,2 ms; media TR compito attenzione sostenuta = 692,3 ms SD 68,8 ms;  $p = 0.18$ ). L'interpretazione qui proposta, anche alla luce dei dati contro-tendenti di accuratezza e tempi di risposta per emicampo anteriore e posteriore osservati nel compito di attenzione sostenuta, è da considerarsi con prudenza, poiché i dati a supporto mostrano una tendenza evidente solo in parte.

Un'altra analisi della varianza è stata condotta per confrontare i gradienti di attenzione dei due compiti, prendendo quindi in esame i tempi di risposta al variare della distanza tra posizione attesa e posizione effettiva di presentazione del target. Oltre al fattore *distanza*, abbiamo inserito anche in questo caso il fattore *compito* come variabile tra i soggetti. I risultati confermano che non vi è differenza significativa tra i compiti sulla media generale dei tempi di risposta ( $p = 0.19$ ) e, oltre all'effetto principale del fattore *distanza* ( $F(6, 348) = 20.403, p < 0.001, \eta^2 = 0.260$ ), anche l'andamento di forma quadratica si conferma significativo indipendentemente dal tipo di compito ( $F(1, 58) = 22.314, p < 0.001, \eta^2 = 0.278$ ). Risulta anche un'interazione significativa tra il fattore *distanza* e il fattore *compito* ( $F(6, 348) = 2.342, p = 0.031, \eta^2 = 0.039$ ).

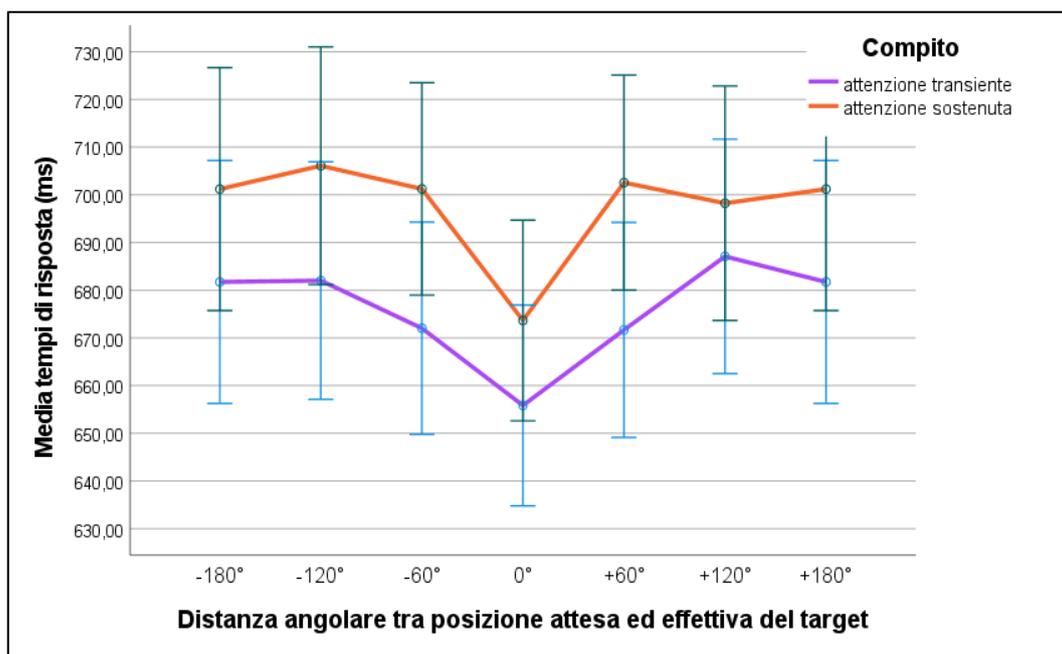


Figura 21: tempi di risposta in funzione della distanza angolare tra posizione attesa ed effettiva del target e del tipo di compito

Come si può osservare dal grafico presentato in figura 21, la distribuzione degli effetti attentivi dovuti allo spostamento dell'attenzione ha un andamento generalmente simile tra i due compiti, come confermato dall'effetto principale della forma di approssimazione quadratica. Tuttavia, nel compito di attenzione transiente essi crescono in maniera chiara fino alle distanze di +/-120° di azimuth tra posizione attesa ed effettiva del target, per poi assestarsi e diminuire leggermente per gli spostamenti più ampi; invece nel compito di attenzione sostenuta gli effetti attentivi sono maggiori per le distanze più brevi (+/-60° di azimuth) mentre rimangono sostanzialmente invariati oltre tale distanza.

Per comprendere più a fondo questa differenza di andamento, abbiamo calcolato la differenza degli effetti attentivi per ogni singolo incremento di distanza da compiere. Volevamo infatti trovare ulteriore conferma al dato sull'interazione tra distanza e compito appena riportato, indagando il diverso andamento di questi effetti attentivi al variare della distanza angolare da coprire con gli spostamenti dell'attenzione. Gli incrementi di distanza erano sei, poiché dalla posizione attesa si potevano dare tre incrementi successivi in senso orario e tre in senso antiorario.

La distanza da coprire, per ognuno di essi, era la medesima, ovvero 60° gradi di azimut. Abbiamo quindi impostato una ANOVA per misure ripetute, mettendo a fattore l'incremento di distanza (con sei livelli) e il fattore compito come variabile inter-soggettiva.

I risultati, pur non mostrando un effetto significativo del fattore compito ( $p = 0.83$ ), indicano un effetto principale del fattore incremento di distanza ( $F(5, 290) = 11.077, p < 0.001, \eta^2 = 0.160$ ) e un'interazione significativa tra il fattore incremento di distanza e il fattore compito ( $F(5, 290) = 3.739, p = 0.003, \eta^2 = 0.061$ ). Il grafico in figura 22 aiuta a chiarire ulteriormente la diversità di distribuzione degli effetti attentivi per i due compiti.

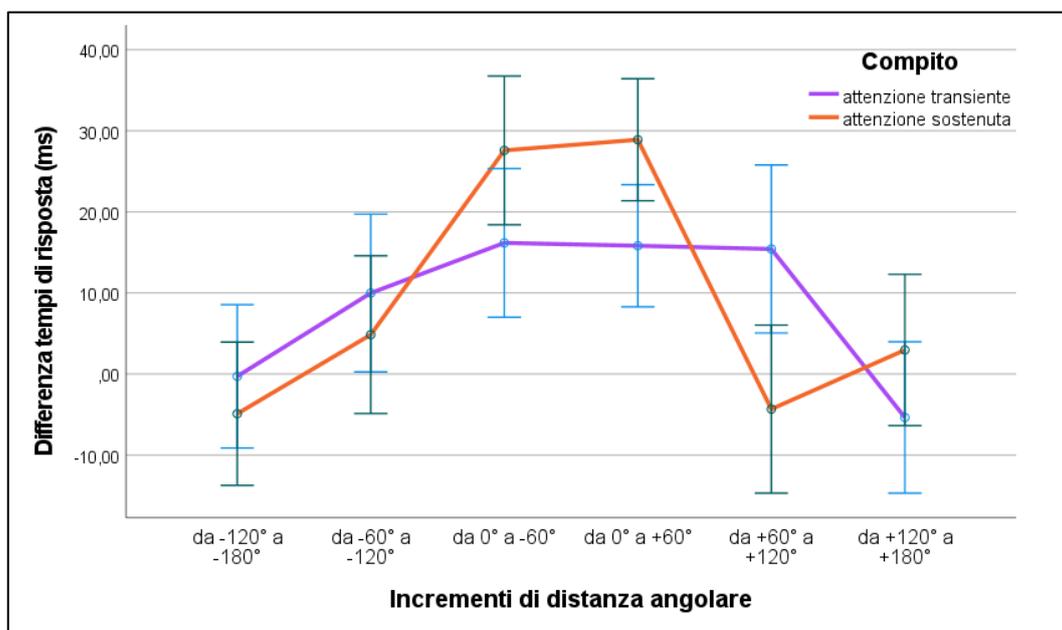


Figura 22: differenza di effetto attentivo al variare degli incrementi di distanza angolare e del tipo di compito

Nel compito di attenzione sostenuta, gli effetti attentivi sono subito massimi per spostamenti attentivi in prossimità della posizione attesa, poi le differenze di effetto con le posizioni più lontane diminuiscono rapidamente, invertendosi anche di segno. Nel compito di attenzione transiente, invece, l'effetto attentivo iniziale, per spostamenti in posizioni vicine a quella attesa, è inferiore ma più persistente all'aumentare della distanza, e diminuisce soltanto per gli spostamenti tra posizioni

più lontane da quella attesa. Questa diversità di andamento sembra riflettere una selettività spaziale più precisa e ristretta quando si esercita attenzione sostenuta su una posizione nota a priori, ciò porta ad una zona di inibizione forte ma contenuta nello spazio. Oltre tale area, la distanza da coprire con lo spostamento attentivo non ha infatti portato ad un aumento dei tempi risposta, bensì ad una sostanziale stabilizzazione. Seguendo l'ipotesi di interazione costante tra processi top-down volontari e bottom-up di monitoraggio e allertamento rapido, già discussa in occasione del primo esperimento (Spence & Driver, 1994; Dalton & Lavie, 2004; Golob et al., 2017), potremmo suggerire che il punto di demarcazione tra le rispettive aree di influenza tra i due processi sia diverso tra i due compiti. Infatti nel compito di attenzione transiente, sembra che la selezione spaziale sia più grossolana ma più ampia, dando così origine ad un'area di inibizione più larga ma anche più debole.

#### **4.6. Conclusioni**

Quanto emerso, dall'analisi dei risultati dei singoli esperimenti e dal confronto tra esperimenti, ci porta a proporre alcune conclusioni riguardo ai fenomeni fin qui osservati e discussi.

Una delle nostre domande di ricerca era se gli effetti dell'attenzione spaziale uditiva, osservati indagando varie porzioni dell'emicampo anteriore, potessero essere replicati anche in un contesto che tiene conto dell'intero spazio circostante rispetto ai partecipanti. I risultati indicano che, anche in questa condizione, si conferma una facilitazione nel completamento di un compito di discriminazione uditiva, dovuta alla correttezza del cue spaziale ricevuto; se invece il cue non è correttamente predittivo quanto alla posizione del target, si osserva un effetto attentivo dovuto alla necessità di riprogrammare lo spostamento verso la posizione effettiva del target e manifestato attraverso tempi di risposta più lunghi in entrambi gli esperimenti e discriminazioni meno accurate nell'esperimento di attenzione sostenuta. Come già evidenziato in altri studi (Sach et al., 2000; Best et al., 2007; Blurton et al., 2015) un cue visivo endogeno risulta anche in questo caso efficace

nell'orientare l'attenzione uditiva verso una determinata posizione, suggerendo che l'informazione spaziale possa avvalersi di un contributo almeno in parte sopra-modale (Spence & Driver, 1994; Banerjee et al., 2011).

La performance sulle prove neutre, nell'esperimento sull'attenzione transiente, differisce in andamento da quanto comunemente mostrato in letteratura, e porta a ipotizzare che il dispiegamento dell'attenzione su tutto lo spazio circostante sia particolarmente gravoso per le risorse attentive. Nel confronto tra i risultati riportati sull'emicampo anteriore (o sue porzioni) e i presenti risultati su tutto lo spazio circostante, questa sembra essere la diversità più evidente, probabilmente dovuta all'unicità della condizione della prova neutra, che richiedeva di esercitare un controllo sull'intero arco dei 360°, vale a dire la selezione spaziale più ampia possibile. Questa selezione totale, sostanzialmente diversa da qualunque altro tipo di selezione spaziale, potrebbe spiegare la differenza osservata.

Per quanto riguarda il confronto tra emicampo anteriore e posteriore sugli effetti dell'attenzione spaziale uditiva possiamo concludere che, complessivamente, i target presentati posteriormente ai partecipanti non hanno portato a risposte più lente o meno accurate di quanto è accaduto con target anteriori, anche se gli andamenti per i due compiti sono stati diversi. Infatti, nel compito di attenzione transiente, un maggior allertamento può aver contribuito a risposte più rapide agli stimoli presentati posteriormente, anche se questa maggior velocità non è accompagnata da maggior accuratezza ed è evidente soprattutto sulle prove neutre. Ciò sarebbe in linea con la prospettiva evoluzionistica che vede il sistema uditivo come il principale dispositivo di allertamento rapido, confermando anche altre evidenze che indicano una selezione spaziale più grossolana nell'emicampo posteriore. Di converso, un minor livello di allerta in condizioni di attenzione sostenuta potrebbe essere responsabile di risposte tendenzialmente più lente, sebbene più rapide con target anteriori.

Tuttavia, alcune differenze tra emicampo anteriore e posteriore sono state trovate sulle prove invalide che, nel compito di attenzione transiente, hanno mostrato una difficoltà maggiore nel disimpegno dell'attenzione da posizioni anteriori, sia in caso

di spostamenti all'interno dello stesso emicampo, sia in caso di spostamenti tra emicampi diversi. Come ipotizzato, l'ancoraggio a posizioni anteriori è risultato più forte ed efficiente, probabilmente a causa del contributo di riferimenti visivi alla rappresentazione spaziale, anche se si tratta di un compito uditivo. Al contrario, cue posteriori hanno elicitato spostamenti attentivi più rapidi, sia all'interno del medesimo emicampo, sia verso l'emicampo opposto.

I risultati delle analisi sui gradienti di attenzione spaziale uditiva confermano le recenti evidenze di una distribuzione di forma quadratica degli effetti attentivi al variare della distanza tra posizione attesa ed effettiva del target, estendendone inoltre la validità all'intero spazio circostante. Infatti, non si ravvisano differenze tra i due emicampi né tra i due compiti su questo punto. I tempi di risposta crescono non linearmente all'aumentare della distanza da coprire con lo spostamento attentivo, e la tendenza si assesta e si inverte per posizioni intermedie ed estreme. L'interazione tra processi top-down (selezione volontaria) e bottom-up (monitoraggio e allertamento) sembra responsabile degli andamenti osservati, in quanto l'area di influenza dei primi appare spazialmente limitata e, oltre tale area, i secondi contribuirebbero a tempi di risposta prima stabili e poi leggermente più veloci per gli spostamenti più ampi. Su quest'ultimo punto, i due compiti hanno mostrato differenze significative dal momento che, in condizioni di attenzione sostenuta, quest'area di selezione appare più ristretta ma più efficace; invece in condizioni di attenzione transiente risulta più ampia ma meno efficace.

#### **4.7. Limitazioni e future direzioni di ricerca**

Nella considerazione dei risultati del presente studio, e delle implicazioni possibili appena discusse, sono da tener presenti alcune possibili limitazioni.

Una prima possibile limitazione emerge dalle analisi dei dati dell'esperimento sull'attenzione transiente, in quanto l'effetto principale del fattore *emicampo target* è significativo ma la differenza tra i tempi di risposta per target anteriori e posteriori è esigua, seppur caratterizzata da scarsa varianza tra i soggetti. Probabilmente un

ampliamento del campione permetterebbe di verificare il consolidarsi o, alternativamente, il vanificarsi di questo effetto.

Per quanto riguarda, invece, il secondo esperimento, i dati contro-tendenti di accuratezza e tempi di risposta suggeriscono una certa cautela nell'interpretazione poiché, come è già stato detto nella sezione dedicata, alcuni partecipanti potrebbero aver messo in atto strategie diverse per completare il compito. Taluni potrebbero aver privilegiato una migliore accuratezza, sacrificando la velocità, talaltri invece potrebbero aver attuato la strategia opposta, privilegiando la velocità a scapito dell'accuratezza. Anche in questo caso, un aumento del campione potrebbe contribuire a chiarire questa eventualità.

Un aspetto comune a entrambi gli esperimenti riguarda, invece, il possibile contributo visivo alla rappresentazione spaziale, che sembra aver rafforzato gli ancoraggi alle posizioni anteriori e reso quindi più difficile il disimpegno dell'attenzione da quelle specifiche posizioni. Il nostro setup sperimentale non ha potuto chiarire se questo contributo dipende dalla direzione dello sguardo oppure, almeno in parte, dalla postura complessiva del corpo. Infatti, sia lo sguardo che il corpo erano orientati verso l'emicampo anteriore. L'informazione visiva residua era assai scarsa, e non risultava possibile per i partecipanti vedere le posizioni target anteriori (è stato esplicitamente chiesto e la risposta è stata sempre negativa), perciò possiamo ritenere che l'effetto osservato non dipenda da un vantaggio delle posizioni anteriori dovuto alla loro visione effettiva. Resta però da chiarire quanto possano incidere la direzione dello sguardo e la postura, e questo aspetto meriterebbe di essere indagato in studi futuri, possibilmente facendo divergere la direzione dello sguardo e la direzione della linea mediana del corpo al fine di valutarne separatamente gli effetti sui risultati eventualmente ottenuti.

Ugualmente interessante e auspicabile in future ricerche sarebbe un approfondimento dell'indagine sul modo in cui si sposta l'attenzione nello spazio, ovvero quali traiettorie essa segue idealmente per raggiungere le posizioni di interesse. I risultati dei nostri studi sembrano suggerire una traiettoria tangenziale, che ricalcherebbe la disposizione circolare degli altoparlanti e quindi delle posizioni

target. Tuttavia, i dati non permettono di escludere altre traiettorie, in particolare per spostamenti tra posizioni diametralmente opposte. I tempi di risposta leggermente più rapidi rispetto a spostamenti tra posizioni meno lontane (ad es. rispetto a spostamenti di  $\pm 120^\circ$  di azimuth) potrebbero anche indicare una traiettoria di spostamento attentivo diversa da quella tangenziale. È possibile, inoltre, che vi siano direttrici preferenziali di spostamento, dipendenti dalla postura o dall'orientamento dello sguardo, così come è possibile immaginare che la differenza sostanziale tra i meccanismi di propagazione dell'informazione visiva e uditiva possano portare l'attenzione a spostarsi con modalità diverse per le due modalità.

Un'altra possibile direzione di ricerca futura riguarda l'aggiunta di nuove posizioni target, che permetterebbero di valutare con maggior precisione gli effetti degli spostamenti attentivi, così come i gradienti di attenzione complessivi e divisi per singolo emicampo. Inoltre, lo stesso accorgimento potrebbe contribuire a comprendere ancora meglio il rapporto tra processi top-down e bottom-up nella definizione dell'ampiezza delle reciproche aree di influenza. È da tener presente, tuttavia, che il disegno sperimentale, già assai complesso per la quantità di condizioni e di dati da raccogliere, ne risulterebbe ulteriormente appesantito. Nondimeno sarebbe opportuno progettare una versione unimodale del presente compito, pur con tutte le difficoltà derivanti dalla scelta e la gestione di cue uditivi, sia per i maggiori tempi di codifica necessari, sia per i possibili effetti di attenzione esogena eventualmente risultanti. Il confronto tra una versione unimodale e la presente versione multimodale dello stesso compito permetterebbe di approfondire il rapporto tra vista e il loro ruolo reciproco udito nella rappresentazione spaziale.



## **V - Studio 2: training musicale e attenzione spaziale uditiva nello spazio circostante**

### **5.1. Introduzione**

La pratica musicale è riconosciuta come una delle attività più impegnative, a motivo della sua natura multidimensionale e della varietà delle competenze coinvolte. L'apprendimento strumentale o vocale, specialmente se unito a esperienze d'insieme, richiede crescenti e concomitanti abilità fisiche, cognitive, emotive e relazionali. Questo training multiforme, specialmente se protratto per molti anni ed iniziato precocemente, ha mostrato di poter indurre cambiamenti strutturali e funzionali in diverse regioni del cervello, come il corpo calloso (vedasi, ad es., Schlaug et al., 1995), la corteccia motoria (Amunts et al., 1997), la corteccia somato-sensoriale (Elbert et al., 1995), la corteccia uditiva (Pantev et al., 1998) e l'ippocampo (Herdener et al., 2010; per una revisione delle differenze cerebrali associate al training musicale, vedasi anche Gaser & Schlaug (2003), Herholz & Zatorre (2012) e Wang & Agius (2018). Questi cambiamenti strutturali e funzionali nel cervello dei musicisti, dovuti al fenomeno della neuro-plasticità, influenzano il funzionamento cognitivo e possono portare a differenze comportamentali rispetto ai non musicisti. Nella maggioranza degli studi menzionati, il confronto con un gruppo di controllo ha permesso di mettere in risalto i cambiamenti potenzialmente dovuti al training e non soltanto al fisiologico processo di maturazione che procede attraverso una continua influenza reciproca tra esercizio di abilità senso-cognitive e modificazioni anatomo-funzionali.

Diversi lavori hanno poi mostrato associazioni positive tra la formazione musicale di tipo formale e il miglioramento delle funzioni cognitive di ordine superiore, come la memoria di lavoro, il controllo inibitorio e l'attenzione (vedasi, in proposito, quanto presentato al capitolo 3 sul rapporto fra training musicale e sviluppo delle funzioni esecutive).

Sebbene ci sia ancora un dibattito aperto sulla possibilità che i benefici del training musicale si generalizzino ad aree di competenza ritenute non specifiche (Swaminathan & Shellenberg, 2018), non pochi studi hanno riportato effetti positivi della pratica musicale su una serie di abilità dominio generali come la memoria visuo-spaziale, la pre-lettura e lettura, la creatività, il successo scolastico e l'intelligenza generale (vedasi quanto riportato al capitolo 3, in particolare Rauscher & Hinton, 2011; Flaugnacco et al., 2015; Costa-Giomi, 2004; Schellenberg, 2004; 2011).

Tra le funzioni cognitive di carattere generale, l'attenzione ricopre un ruolo cruciale in tutte situazioni della vita quotidiana, così come durante le attività musicali. Durante l'ascolto, ma ancor più pienamente durante l'esecuzione, infatti, essi sono tenuti a prestare attenzione contemporaneamente a diverse fonti di informazione come il proprio flusso sonoro, lo spartito, la postura in relazione allo strumento, le dita e le loro specifiche funzioni in ordine al suono, il tempo musicale. Inoltre, nella musica d'insieme, a prescindere dal genere e dall'organico, si aggiungono altri focus attentivi quali i suoni degli altri musicisti e i gesti del direttore, nonché la percezione del risultato sonoro complessivo. Il tutto deve essere utilizzato costantemente dal singolo esecutore per adattare la sua performance a quella collettiva, in vista del raggiungimento del migliore risultato estetico d'insieme (Buma et al., 2015). Quando, ad esempio, i musicisti suonano in orchestra, in un ensemble da camera, in una rock band o in un coro, ognuno dovrà alternare momenti di esecuzione e di attesa, e in entrambi i casi (particolarmente nel secondo) è necessario contare e tenere il tempo della pulsazione interiormente. Inoltre, i musicisti devono eseguire complesse analisi della scena uditiva soprattutto durante i momenti di attacco, transizione e conclusione, poiché è assai probabile che la loro parte sia legata ad alcune note specifiche della parte di un altro esecutore, e queste devono essere isolate e riconosciute all'interno del flusso sonoro complessivo, al fine di sincronizzarsi correttamente con il tempo, il ritmo e il discorso musicale complessivo (Carey et al., 2015). Pertanto, durante l'attività musicale è ragionevole pensare che vengano attivati tutti i diversi processi di attenzione (selettiva, sostenuta e divisa), e che questo training così impegnativo possa sviluppare capacità

più elevate ed evidenti anche al di là del contesto specifico. Tuttavia, anche se si registra un consenso sul fatto che il training musicale sia particolarmente esigente per le risorse attentive - richiedendo il loro dispiegamento nel tempo e nello spazio (Clayton et al., 2016) – risulta ancora aperto il dibattito sui possibili benefici che questo training possa portare allo sviluppo, l'esercizio e il mantenimento delle abilità di attenzione.

Di tutti i lavori presentati al cap. 2, riproponiamo qui una sintesi dei risultati più rilevanti per la comprensione del quadro teorico e delle motivazioni di fondo del presente esperimento.

Le associazioni tra abilità musicali e attenzione sono state per lo più indagate attraverso paradigmi centrati sulla modalità visiva. Questi studi hanno prodotto un corpo eterogeneo di evidenze, spesso caratterizzate da risultati contrastanti (vedasi, ad es., Rodrigues et al., 2013; Patston et al., 2007). Ad ogni modo, la ricerca sui possibili benefici del training musicale sull'attenzione sia visiva che uditiva ha suggerito che le differenze più forti tra gruppi di musicisti e non musicisti possono essere osservate somministrando compiti di attenzione uditiva. Ad esempio, Strait e colleghi (2010) riportano che i musicisti hanno ottenuto performance migliori dei controlli in compiti di attenzione uditiva ma non in quelli di attenzione visiva. Gli autori hanno somministrato la batteria di test IMAP (IHR Multicentre Battery for Auditory Processing, Barry et al., 2010) riscontrando soglie percettive di frequenza e intensità più basse (e quindi più ampie) nei musicisti. Differenze simili sono state osservate anche nei compiti di *backward masking* e di attenzione uditiva, mentre non sono state riscontrate differenze significative nei compiti più direttamente correlati alle abilità di attenzione visiva. Inoltre, Martens e colleghi (2015) hanno studiato gli effetti del training musicale sull'attenzione somministrando un compito di *attentional blink*, in cui ai partecipanti è stato richiesto di identificare due target presentati in rapida sequenza all'interno di un flusso di non-target. L'analisi delle misure (tipicamente l'effetto è quantificato attraverso il ritardo e la minor precisione nella risposta al secondo dei due target) raccolte con stimoli uditivi ha mostrato che i partecipanti musicalmente allenati avevano un effetto attenuato e ritardato rispetto

ai non musicisti, mentre nessuna differenza tra gruppi è emersa con gli stimoli visivi. Inoltre, mentre non è stata trovata alcuna correlazione significativa tra le misure uditive e visive nel gruppo dei musicisti, ciò accadeva per i non musicisti. Questo risultato è stato interpretato come un beneficio modalità-specifico sull'attenzione selettiva per i musicisti. Allo stesso modo, Carey e colleghi (2015) hanno testato due gruppi (musicisti e non musicisti) e hanno trovato forti associazioni tra training specifico e abilità in qualche modo vicine alla musica (come la percezione dell'intonazione) ma deboli correlazioni con funzioni cognitive più generali o abilità non musicali, ad eccezione dell'attenzione sostenuta. Gli autori hanno misurato l'attenzione uditiva sostenuta con il compito SART (Robertson et al., 1997), che testa la capacità di cambiare velocemente il tipo di risposta, e hanno interpretato i risultati suggerendo che la pratica musicale possa avere un effetto modalità-specifico sull'attenzione uditiva, non ravvisando quindi evidenza di un effetto sopra-modale su questa abilità generale. Per di più, l'assenza di differenze significative tra i due gruppi nel compito di analisi della scena uditiva proposto è stata interpretata come una limitazione dei vantaggi dei musicisti alla cornice ristretta dei suoni "musicali".

Quando la nostra attenzione seleziona alcune informazioni rilevanti nell'ambiente circostante, è necessario contemporaneamente non essere distratti da informazioni non utili per il compito da svolgere in quel dato momento. Una selezione efficace non è necessariamente associata ad un efficace disimpegno da informazioni irrilevanti. A questo scopo, Kaganovich et al. (2013), hanno adottato un compito di distrazione uditiva (Schröger & Widmann, 1998) per verificare l'esistenza di possibili differenze tra musicisti e non musicisti in questa particolare abilità. I partecipanti dovevano discriminare tra due possibili durate del suono (la durata, quindi, costituisce la caratteristica rilevante), ignorando il verificarsi di cambiamenti timbrici (caratteristica irrilevante per il compito). I risultati indicavano che i musicisti erano meno distratti dalla caratteristica irrilevante, rispetto alla loro controparte, portando gli autori a ipotizzare una migliore abilità di disimpegno dell'attenzione da stimoli irrilevanti, probabilmente dovuta ad una sottostante abilità di codifica spettrale più accurata, acquisita durante la pratica musicale.

Effettivamente, sono numerosi gli studi che riportano differenze tra musicisti e controlli in abilità percettive di livello sensoriale (vedasi, ad es., Tervaniemi et al., 2009; Strait & Kraus, 2011; Ouimet et al., 2012; Seppänen et al., 2012). Eppure un recente lavoro di Celma-Miralles & Toro (2019) ha dimostrato che i musicisti posseggono anche migliori capacità di selezione top-down di alto livello, utilizzando un particolare compito di discriminazione ritmica spaziale. Ai partecipanti è stato richiesto di riconoscere un metro musicale ternario, presentato con diverse disposizioni spaziali nell'emicampo frontale. Il primo accento forte era presentato in una posizione, mentre i due accenti deboli provenivano da posizioni diverse. Le misure ERP raccolte durante l'esecuzione del compito hanno provato che, durante il riconoscimento del metro, è in atto un processo di alto livello e che i musicisti hanno avuto risposte neurali più forti sia alla pulsazione che al metro, suggerendo effetti positivi del training di lungo periodo. Tuttavia, questo compito era fortemente dominio-specifico della musica, e i risultati dovrebbero essere quindi letti anche alla luce di questa considerazione.

Krzyzak (2021), al fine di indagare possibili differenze nella capacità di selezione uditiva tra musicisti esperti e controlli, ha somministrato un compito di ascolto dicotico nel quale si testavano tre diverse condizioni di attenzione: non focalizzata, focalizzata sull'orecchio destro e focalizzata sull'orecchio sinistro. I risultati, qui riassunti rispetto alla descrizione dettagliata nel capitolo terzo, mostrano una performance più omogenea tra destra e sinistra per i musicisti nella condizione di assenza di focus attentivo, dovuta probabilmente al training, che a sua volta potrebbe aver alterato la tipica risposta neurale agli stimoli uditivi, rendendola meno soggetta alla dominanza emisferica sinistra. In condizioni di attenzione focalizzata, invece, il training musicale non sembra aver influito sull'andamento dei risultati.

Nel complesso, queste evidenze sembrano indicare che i musicisti abbiano capacità attentive superiori e che queste potrebbero essere più evidenti in compiti di attenzione uditiva, dai quali emergono anche chiari benefici del training sull'elaborazione di diversi tipi di informazioni uditive. Ciò potrebbe non

sorprendere, data la peculiarità della pratica musicale e le sue notevoli esigenze in fatto di attenzione. Tuttavia, un aspetto che è stato relativamente trascurato in questi studi, anche considerando la sola modalità uditiva, è l'attenzione spaziale. ovvero la capacità di selezionare e privilegiare le informazioni provenienti da una porzione specifica dello spazio, utilizzandole poi per una ulteriore elaborazione orientata ad un obiettivo contingente (Kidd et al. 2005; Shinn-Cunningham, 2008).

Uno dei primi tentativi di stabilire associazioni dirette tra esperienza musicale e attenzione spaziale uditiva è documentato al paragrafo 2.2 del presente lavoro. Si tratta dello studio di Nager e colleghi (2003) che ha messo a confronto direttori d'orchestra, pianisti e non musicisti su un compito di localizzazione e discriminazione spaziale uditiva. I partecipanti ascoltavano sequenze di stimoli costituiti da brevi emissioni di rumore che, a seconda della diversa frequenza, rappresentavano il target (molto più probabile) o il distrattore. Gli stimoli erano presentati in ordine casuale da uno di sei diversi altoparlanti, tre dei quali situati di fronte e tre a destra dei soggetti. L'angolo di azimut complessivo, coperto dall'esperimento, era quindi di 90 gradi, la parte destra dell'emicampo anteriore. Ai partecipanti veniva data indicazione preventiva sulla posizione cui prestare attenzione sostenuta, o quella di angolo 0° o quella di 90° di azimut. Era loro richiesto di premere un pulsante soltanto nel caso in cui il suono deviante fosse stato presentato in posizione attesa, in tutti gli altri casi era richiesto di non premere alcun pulsante. Venivano raccolti, oltre ai dati comportamentali sulle accuratezze e i tempi di risposta, anche dati elettrofisiologici durante tutto l'esperimento. I tempi di risposta non differivano significativamente tra i gruppi mentre, per quanto riguarda le accuratezze, i due gruppi di musicisti avevano una precisione maggiore dei controlli. Inoltre, le percentuali dei direttori d'orchestra erano le più alte, non tanto nelle due posizioni attese, quanto nelle posizioni devianti.

Questi risultati portano gli autori a ipotizzare che la pratica musicale, specialmente quando accompagnata dalla necessità di selezionare e presidiare ampie porzioni di spazio (come nel caso dei direttori orchestrali e, in misura minore ma apprezzabile, dai pianisti), possa aver contribuito a rafforzare detta selettività, soprattutto in zone

più periferiche, per le quali è ragionevole pensare ad un peggioramento delle performance di rilevamento e discriminazione in condizioni normali. Le misure elettrofisiologiche, invece, indicavano che la differenza più evidente tra i tre gruppi (soprattutto tra i due gruppi di musicisti e i controlli) riguardava la Mismatch Negativity (MMN), marcatore di azione inibitoria per la zone lontane da quella del target. Questo risultato è stato interpretato dagli autori come una possibile conseguenza del training intensivo, che potrebbe aver rafforzato l'elaborazione pre-attentiva dei suoni, soprattutto per quanto riguarda l'informazione spaziale, che sembra essere processata nei primissimi momenti successivi alla ricezione degli stimoli.

Ad oggi e al meglio della nostra conoscenza, non vi sono altre evidenze che abbiano indagato possibili nessi tra la pratica musicale e l'attenzione spaziale uditiva oltre a quelle appena presentate. In generale, poi, abbiamo visto che gli studi su questa particolare abilità, benché numericamente più consistenti, erano limitati principalmente alla valutazione degli effetti dell'attenzione nell'emicampo anteriore (talvolta indagato per intero, talvolta invece solamente una porzione) e non sappiamo quasi niente sugli effetti dell'attenzione nell'emicampo posteriore (vedasi in proposito la revisione della letteratura in Spence et al., 2020). Questa apparente mancanza di interesse verso lo spazio posteriore è alquanto sorprendente, dato che l'udito è uno dei pochi sensi che consente agli esseri umani di monitorare ciò che accade dietro a noi e la sua sensibilità è costantemente e inevitabilmente panoramica (Perrott et al., 2006; Golob & Mock, 2020). Sebbene gli studi sulla percezione e localizzazione uditiva abbiano talvolta considerato l'intero spazio circostante (vedasi, ad es., Wightman & Kistler, 1989; Hofman & Van Opstal, 1998; Macpherson & Middlebrooks, 2000; Zhang & Hartmann, 2010), la ricerca sull'attenzione spaziale uditiva – anche in studi multimodali - si è concentrata principalmente sulle differenze tra il lato sinistro e quello destro dell'emicampo anteriore, trascurando quasi totalmente quello posteriore (Spence & Driver, 2004). Anche lo studio di Popov e colleghi (2023), che indagava ben otto posizioni dislocate sull'intero arco dei 360° attorno ai partecipanti, era primariamente improntato a studiare i meccanismi di allocazione spaziale dell'attenzione nel

tempo, attraverso misure ERP; e il compito proposto non prevedeva spostamenti attentivi di alcun tipo, dal momento che i suggerimenti o cue uditivi erano sempre validi, cioè anticipavano sempre correttamente la posizione verso la quale dirigere l'attenzione per completare il compito di discriminazione richiesto. I dati comportamentali di questo studio, inoltre, non mostrano alcuna differenza significativa tra le otto posizioni esaminate.

Infine, non è chiaro se le recenti evidenze sugli andamenti dei costi di spostamento dell'attenzione nell'emicampo anteriore, già definiti gradienti di attenzione (Golob, et al., 2017, Golob & Mock, 2020), possano o meno essere replicati con stimoli provenienti da dietro. Altrettanto incerta è la loro generalizzazione ad un contesto sperimentale che tenga conto dell'intero spazio circostante.

## **5.2. Domande e ipotesi di ricerca**

Il corpo di evidenze appena presentato, unito ai risultati del nostro primo studio e alle implicazioni suggerite, ci ha portato a formulare alcune domande di ricerca, che ci apprestiamo ad illustrare al termine di questa parte introduttiva.

Una prima questione riguarda le differenze tra gli effetti dell'attenzione negli emicampi anteriore e posteriore, osservati nell'esperimento 1. Abbiamo visto come, in generale, un cue posteriore porti a risposte leggermente più veloci rispetto a quanto accade per i cue che indicano posizioni anteriori. Inoltre, abbiamo osservato differenze significative tra spostamenti attentivi all'interno dello stesso emicampo, con tempi di risposta più veloci per spostamenti tra posizioni posteriori, rispetto a quelli tra posizioni anteriori. Una ulteriore differenza è stata riportata per spostamenti verso l'emicampo opposto, con risposte più rapide a partire da cue posteriori e spostamenti verso l'emicampo anteriore e, viceversa, un disimpegno più difficoltoso e lento dalle posizioni anteriori. Al contempo, alcuni autori riportano performance più omogenee nello spazio per gruppi di musicisti, anche quando lo spazio indagato conteneva posizioni periferiche, e non pochi sono gli studi che riportano una allocazione e un disimpegno dell'attenzione più efficienti. Tuttavia, non sono disponibili evidenze su un confronto diretto tra musicisti e non

musicisti su un compito di suggerimento spaziale, né tantomeno è stato preso in esame l'intero arco dei 360° nei confronti esistenti.

La nostra ipotesi di partenza è che i partecipanti con alle spalle una significativa storia di training musicale possano essere generalmente più veloci e accurati dei controlli nello svolgimento del compito richiesto, anche se gli stimoli uditivi proposti nel nostro studio appartengono alla categoria dei suoni comuni ed è da escludere una maggiore familiarità a priori, dovuta al contenuto dei suoni scelti per il compito. Abbiamo infatti evitato di utilizzare toni puri, toni eseguiti da strumenti musicali o semplicemente modulazioni di rumore bianco, e abbiamo invece optato per lo squillo di un telefono che, a nostro avviso, poteva garantire un grado di familiarità equivalente tra i due gruppi. Inoltre, l'adozione di suoni non specificamente "musicali" potrebbe contribuire a comprendere fino a che punto i benefici del training musicale eventualmente osservati possano estendersi oltre i confini di ciò che si ritiene dominio-specifico.

Ipotizziamo anche che i musicisti potrebbero avere una performance più omogenea nello spazio, con effetti attentivi meno dipendenti dall'emicampo di presentazione del cue e dalla direzione e ampiezza dello spostamento attentivo eventualmente richiesto per raggiungere la posizione del target.

Un confronto che ci pare molto interessante riguarderà le prove neutre, per le quali solitamente la letteratura indica performance di velocità e di accuratezza intermedi rispetto alle prove valide e invalide. Questi valori sono assunti come riferimento per calcolare i cosiddetti *benefici* e *costi* attentivi, cioè la misura della facilitazione data dalla correttezza del cue e, all'opposto, la misura della difficoltà aggiuntiva dovuta alla errata informazione spaziale ricevuta e al conseguente spostamento attentivo necessario. Come abbiamo visto in precedenza, nell'esperimento 1 i partecipanti hanno avuto risultati sorprendenti per le prove neutre, che sono risultate le più lente tra tutte. Questo risultato, forse indice di una particolare difficoltà a dispiegare e mantenere l'attenzione diffusa su tutti i 360°, potrebbe essere diverso per il gruppo dei musicisti, data la loro familiarità con situazioni che richiedono di prestare attenzione contemporaneamente a più posizioni spaziali anche lontane tra

loro. Per di più la pratica musicale d'insieme, a prescindere dal numero e dalla disposizione dei componenti, implica tra l'altro anche l'esercizio di attenzione diffusa, o meglio di una selezione spaziale piuttosto ampia che può estendersi ben oltre l'area del campo visivo e arrivare a coprire anche l'intero spazio circostante, come ad esempio in orchestra, nei cori, nelle grandi band jazz o anche in organici più piccoli ma con disposizione spaziale più ampia come può accadere anche nelle band rock-pop o nei gruppi di musica popolare.

Per quanto riguarda, infine, i gradienti di attenzione, le evidenze più recenti e i risultati dei nostri primi due esperimenti indicano chiaramente una distribuzione di forma quadratica, con risultati stabili e indipendenti dal tipo di attenzione richiesto. Per questo, ci aspettiamo che anche il gruppo dei musicisti possa mostrare un andamento simile, anche se una possibile differenza tra gruppi potrebbe riguardare gli effetti attentivi, dal momento che vi sono evidenze di una più efficiente elaborazione di basso livello per i musicisti, che potrebbe comportare un minor richiamo di risorse attentive di alto livello per l'esecuzione del compito richiesto.

### **5.3. Materiali e metodologia**

Sono stati reclutati un totale di 60 adulti: 30 non musicisti (età media = 28,5; SD = 6,9; 19 femmine; 24 destrimani) e 30 musicisti (età media = 29,1; SD = 6,2; 16 femmine; 28 destrimani). I partecipanti hanno letto e sottoscritto i moduli di consenso informato e di trattamento dei dati personali a fini di ricerca prima dell'esperimento, nell'ambito di un protocollo approvato dal Comitato di Bioetica dell'Università di Bologna. Abbiamo controllato in ingresso la familiarità con patologie uditive attraverso un auto-report compilato dai partecipanti, e nessuno di essi ha riportato disfunzioni pregresse. Pertanto, nessun soggetto è stato scartato a causa di problemi di udito. Un auto-report sul livello di istruzione e sulla dominanza manuale completava la parte di autovalutazione.

L'analisi preliminare della potenza statistica, per la definizione del campione ottimale, è stata eseguita utilizzando il software G-Power (Faul et al., 2007; 2009) e assumendo come valore di riferimento per il calcolo l'eta quadro parziale riportato

da Golob & Mock (2020) nella sua analisi del gradiente di attenzione spaziale uditiva, ovvero 0.15. Il software ha stimato una dimensione totale del campione di 40 partecipanti (20 per ciascun gruppo da testare) per ottenere una potenza di 0.95 con un valore di significatività atteso di 0.05 e una dimensione dell'effetto stimata di 0.42. Ad ogni modo, considerando il fatto che lo studio di riferimento non adottava un paradigma di confronto tra gruppi, abbiamo preso in esame anche uno studio specifico sul paradigma di Posner, già menzionato nel metodo dell'esperimento 1 (Hayward & Ristic, 2013). Gli autori hanno reclutato due gruppi di 22 unità, per un totale di 44 partecipanti. Nonostante questa grandezza campionaria potesse essere ritenuta sufficiente, la novità del compito e la molteplicità delle condizioni sperimentali ci hanno portato a fissare il campione di progetto a 30 soggetti per ogni gruppo.

Per quanto riguarda il gruppo dei musicisti, vale la pena richiamare quanto discusso all'inizio del cap. 2 a proposito della definizione di *musicista*. Essa a tutt'oggi rimane dibattuta, dal momento che il tipo di competenza musicale può essere molto variabile tra i soggetti (Gingras et al., 2015; Slevc et al., 2016; Talamini et al., 2017; Swaminathan & Schellenberg, 2018). Per questo motivo, abbiamo scelto di valutare il background musicale dei partecipanti con la batteria Goldsmith Music Sophistication Index (Müllensiefen et al., 2014) nella versione italiana proposta da Correia e colleghi (2022). Questo strumento, riconosciuto e validato a livello internazionale, permette di tracciare molti aspetti diversi della musicalità e aiuta a definire un profilo musicale assai sfaccettato. Tuttavia, poiché eravamo interessati principalmente a valutare l'impatto della pratica musicale sulle abilità attentive, abbiamo dato un peso maggiore a questo aspetto del background musicale complessivo, introducendo per il gruppo dei musicisti una soglia minima per quanto riguarda gli anni di training strumentale. Inoltre, una ripartizione in gruppi basata semplicemente sul raggiungimento o meno di un punteggio soglia della batteria GOLD-MSI non poteva essere considerata affidabile, nè vi sono evidenze che indicano con precisione dei punteggi di riferimento per musicisti e non musicisti.

Pertanto, tutti i soggetti hanno completato la batteria GOLD-MSI, ma solo coloro che hanno riportato almeno sei anni di formazione strumentale o vocale (non necessariamente accademica) sono stati inclusi nel gruppo dei musicisti, mentre per i non musicisti abbiamo stabilito un valore soglia di massimo due anni di training strumentale. A supporto dell'efficacia dei criteri stabiliti, l'analisi dei punteggi dell'indice di sofisticazione generale mostra una differenza significativa tra i due gruppi ( $F(1, 58) = 77.804, p < 0.001, \eta^2 = 0.573$ ) come mostrato nella figura 23. I musicisti hanno riportato un punteggio medio di 102, SD 11, min 79, max 119, mentre i non musicisti hanno riportato un punteggio medio di 69, SD 18, min 36, max 96.

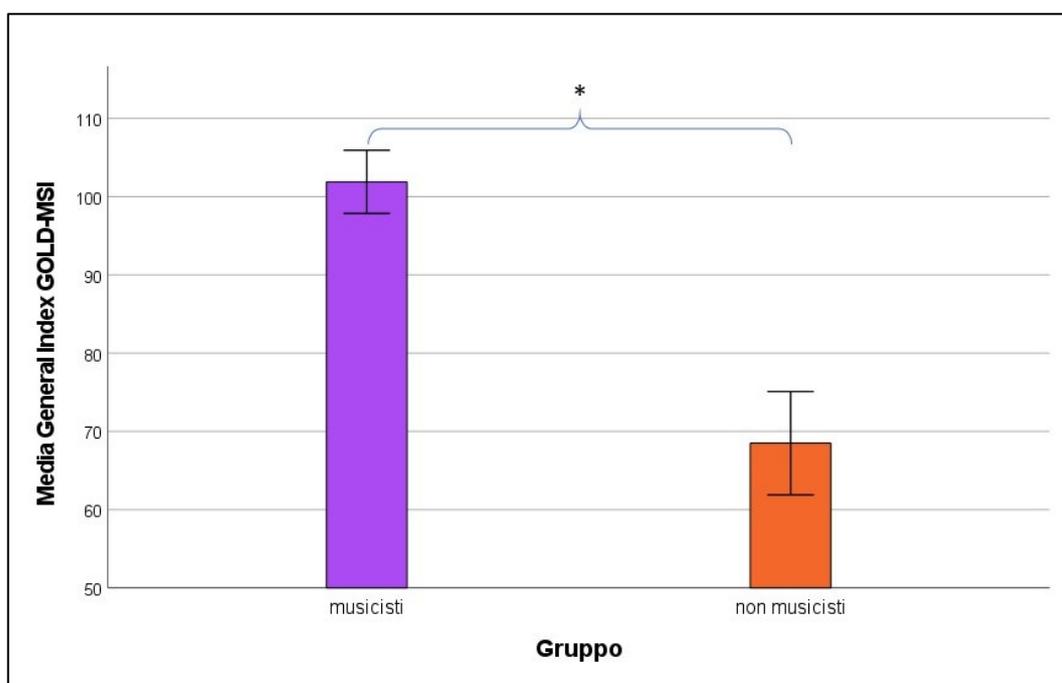


Figura 23: media dei punteggi del General Index della batteria GOLD-MSI per ciascun gruppo

La procedura sperimentale in questo secondo studio è identica a quella adottata per il nostro primo esperimento sull'attenzione transiente, e viene qui riassunta per semplificare la lettura.

La stanza per l'esperimento è stata oscurata per minimizzare l'informazione visiva disponibile. Le sessioni sperimentali si sono svolte in assenza di rumori interferenti, interni o esterni alla stanza, ma non si è proceduto all'insonorizzazione per

preservare un ascolto ecologico dei suoni. Per presentare i suggerimenti o cue visivi, è stato posizionato un monitor esattamente davanti ai partecipanti. La distanza dagli occhi era di 100 cm e la grandezza delle frecce è stata proporzionata alla distanza. Per presentare gli stimoli uditivi, sono stati utilizzati sei altoparlanti cablati, collocati in modo circolare attorno al partecipante (a 120 cm di distanza dalla testa e a 60° di azimut di distanza l'uno dall'altro). La loro altezza è stata regolata per adattarsi al livello medio dell'orecchio. Durante l'esperimento tutti gli altoparlanti sono stati ricoperti da un sottile tessuto nero, per nasconderli alla vista senza tuttavia compromettere la qualità del segnale audio. Gli stimoli uditivi sono stati gestiti da un'interfaccia audio professionale multicanale collegata ad un PC desktop, dotato del software di progettazione sperimentale Presentation. Un dispositivo di risposta dotato di tasti, collegato al PC per la raccolta dei dati, è stato posizionato su un tavolo di fronte al partecipante. I due tasti di risposta sono stati allineati sia con la linea mediana del corpo dei partecipanti che con il centro del monitor, per evitare effetti di compatibilità spaziale e contribuire a controbilanciare gli effetti della mano dominante. Una telecamera a circuito chiuso è stata posizionata di fronte ai partecipanti, per controllare ed evitare qualsiasi movimento della testa o degli occhi. Gli stimoli uditivi target erano costituiti da due squilli prodotti da campanelli all'interno di un telefono d'epoca, con una differenza in frequenza del 10%, tale da farne percepire la diversità da tutti i partecipanti. Il suono più acuto è stato denominato "campanello piccolo", e quello più grave "campanello grande". La pressione sonora a livello delle orecchie era la medesima, da tutte le direzioni e per tutti i suoni. La durata degli stimoli uditivi è stata portata a 500 ms. Gli stimoli visivi erano costituiti da frecce bianche su sfondo nero, con sei diversi orientamenti, presentate al centro del monitor. Una settima immagine, costituita da tutte e sei le frecce precedenti, disposte a raggera, costituiva il cue neutro. Il compito prevedeva di discriminare quale dei due possibili suoni fosse stato presentato in ogni prova, premendo il pulsante corrispondente nel minor tempo possibile e più accuratamente possibile. Ai partecipanti era richiesto di mantenere lo sguardo fisso al centro dello schermo, e di portare implicitamente la propria attenzione verso la direzione indicata dalla freccia. Successivamente alla presentazione del suono

target, essi rispondevano premendo un tasto e concludendo così la singola prova. L'esperimento era composto da 4 blocchi di 192 prove ciascuno, per un totale di 768 prove complessive, di cui 480 valide, 240 invalide e 48 neutre. Ai partecipanti è stato richiesto di invertire la posizione delle mani sul dispositivo di risposta prima dell'inizio di ogni blocco di prove successivo al primo. La mappatura della corrispondenza suono-tasto è rimasta invece costante per tutto l'esperimento. In tal modo, i partecipanti hanno eseguito un numero uguale di prove con entrambe le posizioni, ottenendo così un contro-bilanciamento interno. Prima dell'inizio dell'esperimento, i partecipanti hanno svolto un breve training per familiarizzare con il setting e con il compito. Abbiamo anzitutto verificato che ogni partecipante fosse in grado di percepire e riportare correttamente la posizione degli stimoli uditivi presentati da ciascun altoparlante. Successivamente, hanno posizionato i loro indici sui tasti di risposta e, dopo aver introdotto la mappatura della corrispondenza tasto-suono, si sono esercitati in un breve blocco di pratica costituito da 44 prove, con una soglia di accuratezza minima del 70% per poter procedere con l'esperimento principale. Al termine della procedura sperimentale, abbiamo fornito un link per la compilazione da remoto del questionario sulla musicalità. I dati, anonimizzati attraverso codici alfanumerici univoci, sono stati raccolti sulla piattaforma Qualtrics.

#### **5.4. Analisi dati**

Il disegno sperimentale includeva due variabili indipendenti: la direzione del cue visivo (con 7 livelli, costituiti dalle 6 possibili direzioni più il cue neutro) e la posizione di presentazione del target uditivo (con 6 livelli, costituiti dalle 6 possibili posizioni degli altoparlanti).

Durante l'attività sono stati registrati e raccolti i tempi di risposta (abbreviati, come in precedenza, con "RT") e le accuratezze percentuali nelle risposte. Sono state escluse dall'analisi le prove con errori, le risposte mancate e le prove corrette con tempi di risposta superiori a  $\pm 3$  deviazioni standard dalla media dei tempi di risposta (calcolata separatamente per ciascun partecipante). Tali criteri di

esclusione hanno portato a scartare un totale di 1727 prove su un totale di 23040 prove.

Sono state effettuate analisi separate per i dati dei TR e delle accuratèzze. I dati sono stati studiati utilizzando l'analisi della varianza per misure ripetute (ANOVA). Ogni volta che l'assunto di sfericità veniva violato (la verifica è avvenuta attraverso il test di sfericità di Mauchy) è stata applicata la correzione epsilon di Greenhouse-Geisser al valore dei gradi di libertà.

Una prima ANOVA è stata eseguita per confrontare gli effetti dell'attenzione spaziale uditiva (Posner & Boies, 1971; Posner & Cohen, 1984) nel gruppo dei musicisti e in quello dei non musicisti. Inoltre, eravamo interessati alle potenziali differenze tra gruppi nella distribuzione degli effetti attentivi negli emicampi anteriore e posteriore. Di conseguenza, i fattori *validità della prova* (4 livelli: prova valida, prova invalida e spostamento nello stesso emicampo, prova invalida e spostamento tra diversi emicampi, prova neutra) ed *emicampo target* (2 livelli: anteriore e posteriore) sono stati inclusi come fattori intra-soggettivi, mentre il fattore *gruppo* (2 livelli: musicisti e non musicisti) era quello inter-soggettivo.

La seconda ANOVA, simile a quella proposta nel nostro primo studio, aveva lo scopo di indagare possibili differenze tra gruppi nell'andamento degli effetti dell'attenzione al variare della distanza angolare cue-target (gradiente di attenzione). Gli spostamenti attentivi in senso orario, dalla posizione attesa a quella effettiva di presentazione del target, sono stati codificati con un valore positivo che indica la distanza angolare coperta dall'attenzione (ad esempio +60° per indicare uno spostamento di una posizione, +120° per uno spostamento di due posizioni ecc.), mentre gli spostamenti in senso antiorario sono stati indicati da valori negativi. In questa analisi, i fattori *distanza cue-target* (con 7 livelli possibili, che vanno da -180° a +180°) e *gruppo* (musicisti o non musicisti) sono stati considerati, rispettivamente, come intra-soggettivo e inter-soggettivo.

## 5.5. Risultati

I risultati della prima ANOVA, per quanto riguarda i tempi di risposta, hanno mostrato un effetto principale significativo del fattore *gruppo* ( $F(1, 58) = 12.697$ ,  $p < 0.001$ ,  $\eta^2 = 0.180$ ). Infatti, come si evince anche dal grafico in figura 24, le risposte dei musicisti sono state complessivamente più veloci di quelle dei non musicisti (musicisti: media = 634,5 ms, SD 36,4 ms; non musicisti: media = 676,6 ms, SD 53,5 ms).

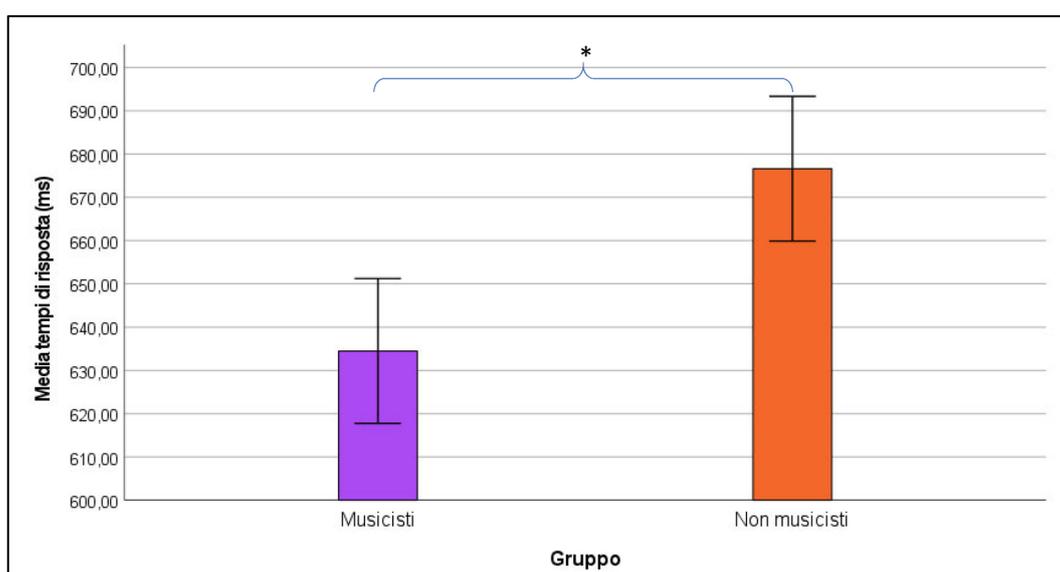


Figura 24: tempi di risposta medi in funzione del gruppo di appartenenza

Anche la *validità della prova* ha avuto un effetto principale significativo ( $F(2.288, 132.704) = 18.442$ ,  $p < 0.001$ ,  $\eta^2 = 0.241$ ). Nelle prove valide, la media dei TR era di 640,1 ms, SD 47,5 ms; nelle prove invalide con spostamenti all'interno dello stesso emicampo, la media è stata di 660,4 ms, SD 49,4 ms; nelle prove invalide con spostamenti tra diversi emicampi, la media è stata di 660,6 ms, SD 47,6 ms e nelle prove neutre, la media è stata di 661,2 ms, SD 65,3 ms.

Anche l'interazione tra i fattori *validità della prova* ed *emicampo target* è risultata complessivamente significativa ( $F(2.427, 140.792) = 11.795$ ,  $p < 0.001$ ,  $\eta^2 = 0.169$ ). Sono stati eseguiti confronti accoppiati, adottando la correzione di Bonferroni per confronti multipli. Nelle prove valide, l'emicampo di presentazione

del target non ha portato a differenze significative sui TR (640,3 ms, SD 50,9 ms emicampo anteriore contro 639,8 ms, SD 44,8 ms emicampo posteriore,  $p = 0.7$ ). Nelle prove invalide con spostamento dell'attenzione all'interno dello stesso emicampo, invece, i target in posizioni posteriori hanno portato a risposte più rapide (655,3 ms, SD 47,8 ms contro i 665,5 ms, SD 53,1 ms con target presentati anteriormente,  $p < 0.001$ ). Anche per le prove invalide con spostamenti di attenzione tra emicampi diversi, l'emicampo target ha portato a differenze significative nei tempi di risposta, che sono risultati più rapidi quando era richiesto di spostarsi da una posizione attesa posteriore ad una posizione effettiva anteriore (655,0 ms, SD 46,1 ms contro 666,2 ms, SD 51,2 ms per spostamenti dalle posizioni anteriori verso quelle posteriori,  $p < 0.001$ ). Nelle prove neutre, infine, non si rilevano differenze nei tempi di risposta, al variare dell'emicampo del target (664,0 ms, SD 72,0 ms emicampo anteriore contro 658,3 ms, SD 68,1 ms emicampo posteriore,  $p = 0.1$ ).

Un'altra interazione significativa, molto importante ai fini del confronto tra musicisti e non musicisti, è quella osservata tra il fattore *validità della prova* e il fattore *gruppo* ( $F(2.208, 132.704) = 8.809, p < 0.001, \eta^2 = 0.132$ ). Come mostrato nella figura 25, i musicisti hanno avuto tempi di risposta significativamente più rapidi in tutte le condizioni sperimentali. I confronti multipli, corretti con l'adattamento di Bonferroni, hanno mostrato che nelle prove valide, la differenza tra i due gruppi è stata di 31,5 ms (musicisti 624,3 ms, SD 40,8 ms, non musicisti 655,8 ms, SD 49,2 ms,  $p = 0.009$ ), nelle prove invalide con spostamenti all'interno dello stesso emicampo è stata di 39,8 ms (musicisti 640,5 ms, SD 37,4 ms, non musicisti 680,3 ms, SD 52,3 ms,  $p = 0.001$ ), nelle prove invalide con spostamenti tra diversi emicampi è stata di 34,2 ms (musicisti 643,5 ms, SD 33,8 ms, non musicisti 677,7 ms, SD 53,4 ms,  $p = 0.004$ ), mentre nelle prove neutre la differenza è stata di 62,9 ms (musicisti 629,7 ms, SD 45,4 ms, non musicisti 692,6 ms, SD 67,6 ms,  $p < 0.001$ ).

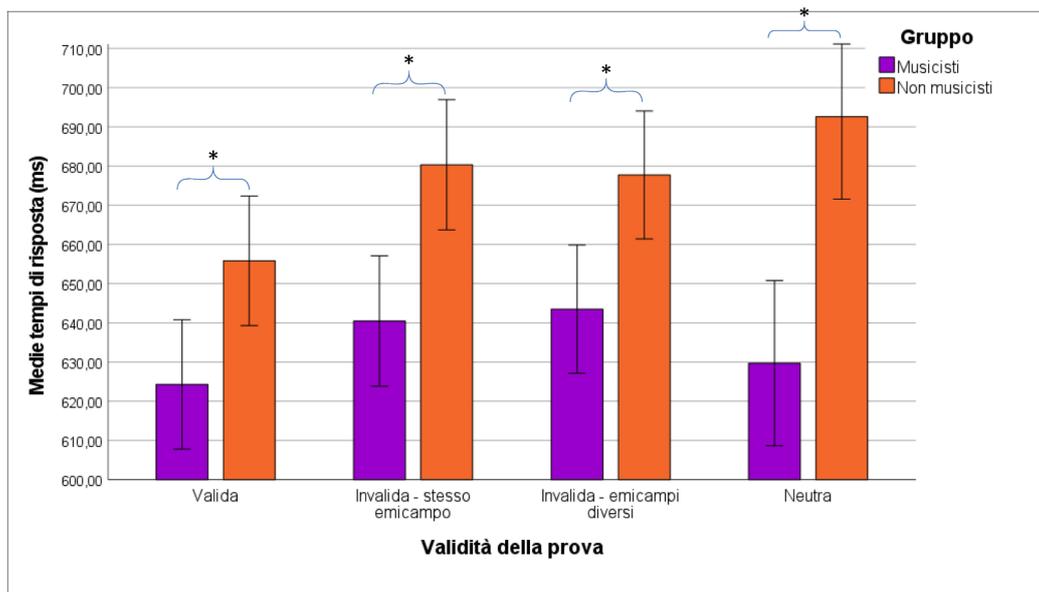


Figura 25: tempi di risposta in funzione di validità prova e gruppo di appartenenza

L'analisi ha mostrato anche un'interazione di ordine superiore tra i fattori *validità della prova*, *emicampo target* e *gruppo* ( $F(2.427, 140.792) = 3.171, p = 0.036, \eta^2 = 0.052$ ). Per indagare su questa interazione, sono state effettuate analisi separate per ciascun gruppo. Per quanto riguarda le analisi a fattore singolo di nostro interesse, ovvero *l'emicampo target*, esso è risultato significativo solo nel gruppo dei non musicisti ( $F(1, 29) = 4.497, p = 0.043, \eta^2 = 0.134$ ), al contrario dei musicisti, per i quali il fattore non ha portato ad effetti significativi ( $p = 0.5$ ). Similmente, nelle analisi a due fattori l'interazione *validità prova x emicampo target* è risultata significativa soltanto per il gruppo dei non musicisti ( $F(3, 87) = 11.823, p < 0.001, \eta^2 = 0.290$ ), mentre per i musicisti non ha raggiunto la significatività ( $p = 0.062$ ).

Nonostante per il gruppo dei musicisti questa interazione non abbia raggiunto la significatività, al solo scopo esplorativo sono stati eseguiti t-test a campioni accoppiati che confrontavano le medie dei tempi di risposta con target anteriori e posteriori, separatamente per ciascuna condizione di validità della prova. Essi hanno rivelato differenze significative solo per le prove invalide con spostamenti attentivi tra emicampi diversi ( $p = 0.040, d \text{ di Cohen} = 0.39$ ), infatti quando i cue

indicavano le posizioni posteriori e i target erano presentati anteriormente, le risposte erano più veloci rispetto alla condizione opposta (639,4 ms, SD 33,7 ms contro 647,6 ms, SD 40,0 ms). Nessun altro confronto è risultato significativo (tutti i valori  $F < 3.654$ , tutti i valori  $p > 0.066$ , tutti i valori  $\eta^2 < 0.112$ ).

Invece, nel gruppo dei non musicisti gli stessi t-test a campioni accoppiati (corretti con l'adattamento di Bonferroni) hanno rivelato differenze statisticamente significative per prove invalide e spostamenti attentivi nello stesso emicampo ( $p = 0.003$ ,  $d$  di Cohen = 0.59), prove invalide e spostamenti tra diversi emicampi ( $p = 0.001$ ,  $d$  di Cohen = 0.65) e prove neutre ( $p = 0.004$ ,  $d$  di Cohen = 0.58). Nel primo caso, gli spostamenti all'interno dell'emicampo anteriore hanno comportato risposte più lente, rispetto agli spostamenti all'interno dell'emicampo posteriore (687,0 ms, SD 56,6 ms contro 673,6 ms, SD 50,3 ms). Per quanto riguarda gli spostamenti tra diversi emicampi, invece, spostare l'attenzione da quello posteriore a quello anteriore portava a risposte più rapide, mentre lo spostamento opposto risultava più lento (670,6 ms, SD 51,9 ms contro 684,8 ms, SD 57,1 ms). Nelle prove neutre, i target presentati nell'emicampo anteriore hanno elicitato tempi di risposta più lenti rispetto a quelli presentati posteriormente (699,8 ms, SD 72,6 ms contro 685,4 ms, SD 64,7 ms).

Ulteriori analisi separate per ciascuna delle quattro condizioni di validità della prova hanno rivelato una significativa interazione tra *emicampo target* e *gruppo* solo per le prove neutre ( $F(1, 58) = 6.083$ ,  $p = 0.017$ ,  $\eta^2 = 0.095$ ). I tempi di risposta medi dei musicisti erano di 628,2 ms, SD 51,3 ms per target presentati nell'emicampo anteriore e di 631,2 ms, SD 43,7 ms per target presentati nell'emicampo posteriore (differenza non significativa,  $p = 0.5$ ), mentre i TR medi dei non musicisti erano di 699,8 ms, SD 72,7 ms per target presentati nell'emicampo anteriore e di 685,4 ms SD, 64,7 ms per target presentati nell'emicampo posteriore (differenza significativa,  $p = 0.005$ ).

L'andamento dei risultati appena descritti è riportato nei grafici in figura 26 e 27.

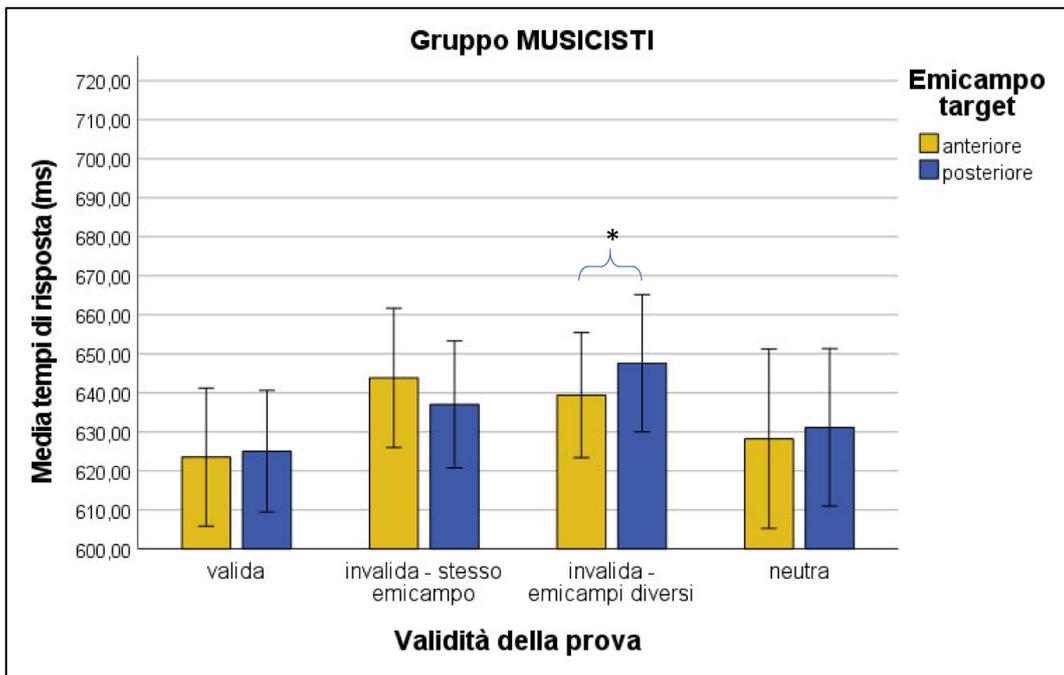


Figura 26: tempi di risposta del gruppo musicisti, in funzione della validità della prova e dell'emicampo target

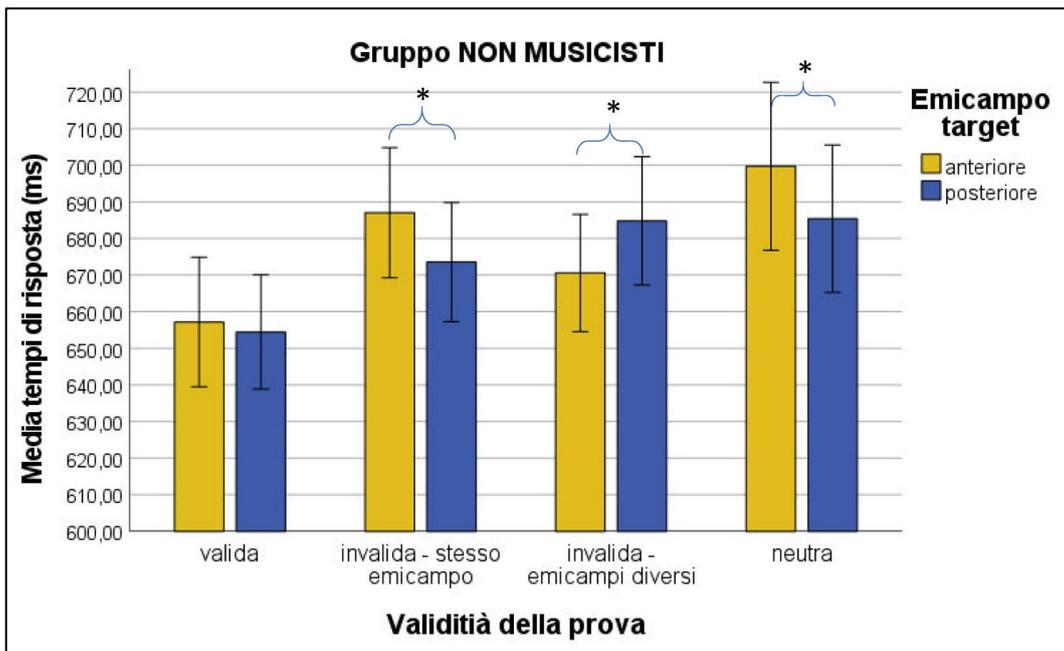


Figura 2726: tempi di risposta del gruppo non musicisti, in funzione della validità della prova e dell'emicampo target

Per una comprensione più approfondita dell'interazione a tre fattori sopra descritta, abbiamo eseguito un'altra ANOVA per misure ripetute, con l'obiettivo di evidenziare i benefici e i costi attentivi. Come già abbiamo avuto modo di vedere al paragrafo 2.1, una letteratura ben consolidata sul paradigma di suggerimento spaziale considera la performance sulle prove neutre come punto di riferimento per calcolare i benefici e i costi dovuti alla correttezza o meno dell'informazione spaziale data dal cue. Il beneficio dell'attenzione infatti, espresso sotto forma di facilitazione nelle risposte, dovrebbe derivare dalla conoscenza previa della posizione del target, assicurata da un cue valido. Al contrario, il costo dell'attenzione, evidenziato da risposte più lente e/o meno accurate, dovrebbe derivare dal tempo aggiuntivo necessario per riorientare l'attenzione dalla posizione attesa a quella effettiva di presentazione del target. Il beneficio attentivo sui TR può essere ottenuto sottraendo i valori medi delle prove valide da quelli delle prove neutre. Se si verifica un beneficio effettivo, questa operazione dovrebbe dare un valore negativo come risultato. Con lo stesso criterio, i costi attentivi sono calcolati sottraendo la media dei TR delle prove invalide dalla media dei TR delle prove neutre. L'evidenza di questo costo è rappresentata da un valore positivo come risultato di questa operazione.

Questa ANOVA ha preso in considerazione un fattore intra-soggettivo ed uno inter-soggettivo. Il fattore interno ai soggetti era il *costo/beneficio* (due livelli, uno per i benefici e l'altro per i costi), e il fattore tra i soggetti era il *gruppo* (due livelli, musicisti o non musicisti).

I risultati hanno mostrato un effetto principale significativo del fattore *gruppo* ( $F(1, 58) = 17.158, p < 0.001, \eta^2 = 0.228$ ). la media costi/benefici dei musicisti è stata di +3,4 ms, SD 21,7 ms, mentre la media dei non musicisti è stata di -25,2 ms, SD 31,0 ms. Come si vede nella figura 28, entrambi i gruppi hanno mostrato un beneficio attentivo per le prove valide, anche se significativamente diverso in valore assoluto. Invece, solo i musicisti hanno avuto un costo attentivo sulle prove

invalide, mentre i non musicisti avevano ancora un beneficio rispetto alle prove neutre.

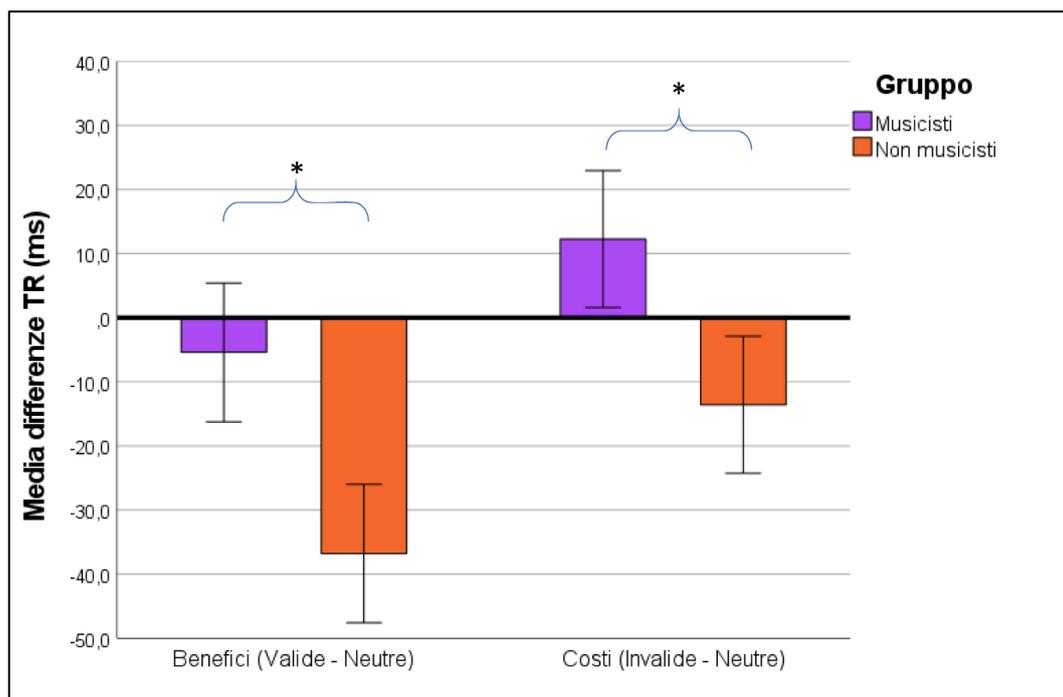


Figura 28: benefici e costi attentivi in funzione del gruppo di appartenenza

L'ANOVA complessiva è stata eseguita anche sui tassi di accuratezza percentuale, e ha rivelato un effetto principale significativo del fattore *gruppo* ( $F(1, 58) = 7.484$ ,  $p = 0.008$ ,  $\eta^2 = 0.114$ ). L'accuratezza media dei musicisti era del 92,4%, SD 3,5%, mentre l'accuratezza dei non musicisti era dell'89,2%, SD 5,4%. In questa analisi non sono stati registrati altri effetti significativi, nè per i singoli fattori intra-soggettivi, nè per le interazioni tra fattori. I risultati sono riportati nel grafico in figura 29.

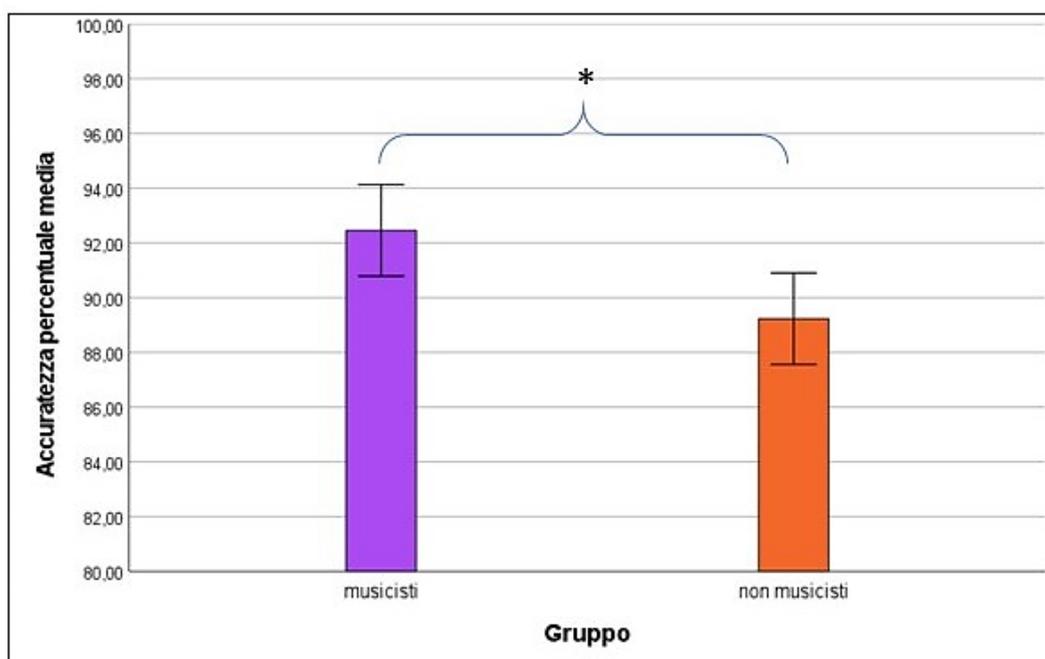


Figura 29: accuratezza percentuale media in funzione del gruppo di appartenenza

I risultati dell'analisi dei gradienti attentivi per i due gruppi hanno evidenziato un effetto principale significativo del fattore *distanza angolare cue-target* ( $F(6, 348) = 16.312, p < 0.001, \eta^2 = 0.220$ ). Come era lecito aspettarsi, visti i risultati delle analisi precedenti, i TR medi complessivi per le prove valide (identificate sul grafico con il valore di  $0^\circ$ , corrispondente ad uno spostamento attentivo nullo) sono stati i più veloci (640,1 ms, SD 47,6 ms), mentre per le distanze angolari più brevi si è assistito ad un incremento dei tempi di risposta, tendenti poi a stabilizzarsi per le distanze più ampie. Le analisi di contrasto hanno mostrato che la forma di approssimazione quadratica era statisticamente significativa ( $F(1, 58) = 28.903, p < 0.001, \eta^2 = 0.333$ ). I costi di spostamento dell'attenzione, in termini di tempi di risposta, sono cresciuti rapidamente per brevi distanze cue-target, per poi ridursi e stabilizzarsi sostanzialmente per le distanze più ampie. I musicisti erano generalmente più veloci, come dimostrato dall'effetto principale significativo del fattore inter-soggettivo *gruppo* ( $F(1, 58) = 9.711, p = 0.003, \eta^2 = 0.143$ ). I tempi di risposta dei musicisti sono stati mediamente di 640,3 ms, SD 35,2 ms, mentre quelli dei non musicisti sono stati mediamente di 676,0 ms, SD 51,9 ms.

Non è stata osservata alcuna differenza tra gruppi nella forma del gradiente attentivo, come dimostrato dal fatto che l'interazione tra i fattori *distanza angolare* e *gruppo* non ha raggiunto la significatività ( $F(6, 348) = 1.249, p = 0.280, \eta^2 = 0.021$ ). La figura 30 rappresenta, in un unico grafico, gli andamenti dei gradienti per entrambi i gruppi.

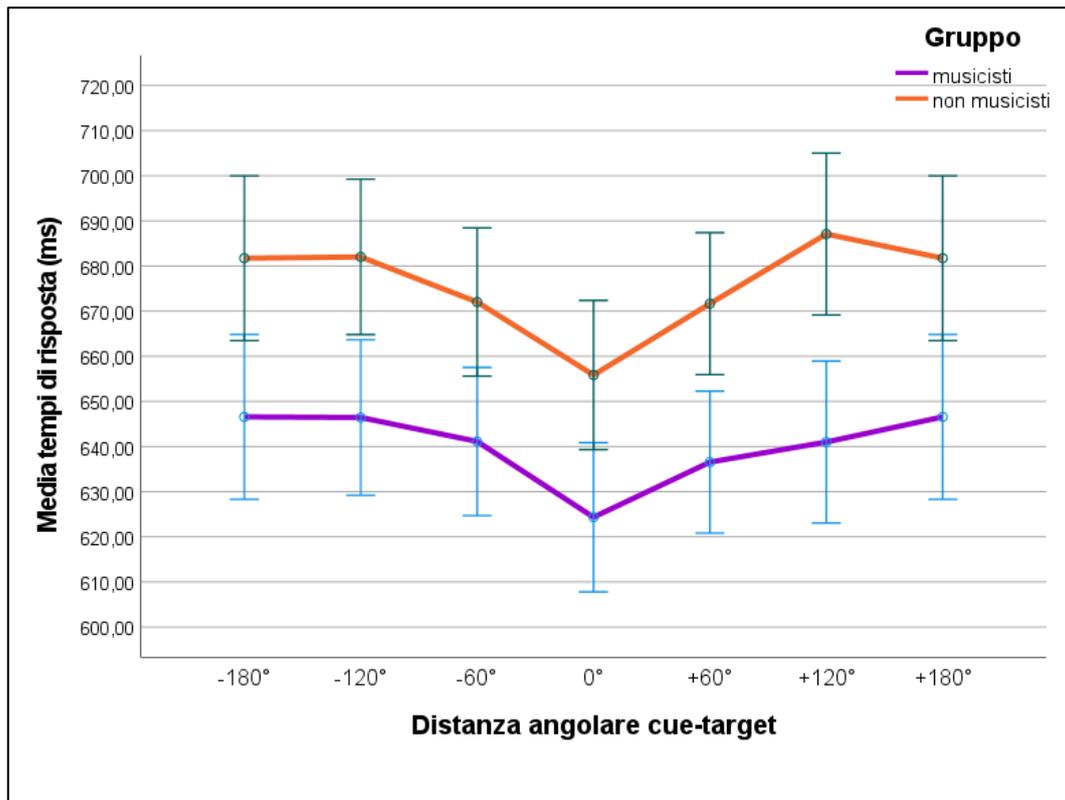


Figura 30: tempi di risposta in funzione della distanza angolare cue-target e del gruppo di appartenenza

## 5.6. Discussione generale

Il presente studio mirava ad indagare se gli effetti dell'attenzione spaziale uditiva in tutto lo spazio circostante, osservati nel nostro primo esperimento, potessero o meno essere replicati prendendo in esame anche un gruppo di musicisti con una significativa storia di training strumentale. Ci siamo chiesti, innanzitutto, se potessero esservi differenze complessive, ascrivibili con buona probabilità

all'esperienza musicale pregressa e, andando più nello specifico, se le differenze tra effetti dell'attenzione negli emicampi anteriore e posteriore, osservate nel gruppo di controllo, potessero variare o meno considerando un campione di musicisti. La stessa domanda di fondo riguardava i gradienti di attenzione evidenziati nel nostro primo esperimento, ovvero se la pratica musicale potesse in qualche modo influire sulla distribuzione degli effetti attentivi al variare della distanza tra posizione attesa ed effettiva del target.

I risultati dell'ANOVA complessiva, che mostrano un effetto principale del fattore *validità della prova*, indicano che tutti i partecipanti, indipendentemente dal gruppo di appartenenza, hanno effettivamente spostato la loro attenzione uditiva in conseguenza dei cue ricevuti e dei target presentati. In effetti, le risposte per le prove valide sono state le più rapide, per effetto della correttezza dell'informazione spaziale data dal cue. Al contrario, in presenza di cue invalidi, i partecipanti hanno avuto un generale rallentamento dei tempi di risposta, a causa della necessità di spostare l'attenzione dalla posizione indicata dal cue a quella effettiva di presentazione del target. Inoltre, questo risultato può essere considerato un marcatore di efficacia del compito, appositamente progettato per studiare gli effetti dell'attenzione *covert* con cue visivi e target uditivi su tutto lo spazio circostante.

L'interazione tra i fattori *validità della prova* ed *emicampo target* replica quanto osservato nel primo studio, anche tenendo conto di una popolazione di musicisti, che si è aggiunta al gruppo di controllo. Le possibili implicazioni, sia per quanto riguarda un contributo di riferimenti visivi alla rappresentazione spaziale anche in compiti uditivi, sia per l'ipotesi di un maggiore effetto di allertamento dato dai cue indicanti posizioni posteriori, sono state discusse al paragrafo 3.2.4, cui si rimanda.

La stessa analisi ha portato anche ad un effetto principale del fattore *gruppo*: i musicisti infatti sono risultati sia più veloci che più accurati dei non musicisti. La lunga attività pratica con i suoni, singoli ma assai più spesso concomitanti, provenienti talvolta da una direzione principale ma assai più spesso da diverse direzioni anche concorrenti, potrebbe aver plasmato la capacità dei musicisti di discriminare le caratteristiche fisiche (Kaganovich et al., 2013; Slater et al., 2015;

Abel et al., 2016), e questa capacità sembra estendersi oltre il dominio specifico dei suoni prodotti da strumenti musicali (Kraus et al., 2014; Sares et al., 2018).

Infatti, il presente studio non ha utilizzato stimoli costituiti da toni puri o suoni di strumenti musicali, poiché eravamo interessati a valutare se le migliori capacità uditive, riportate in precedenza da gruppi di musicisti confrontati con gruppi di controllo, potessero estendersi oltre i suoni specifici del contesto musicale (Martens et al., 2015). L'adozione di suoni prodotti da campanelli di telefoni a disco, oramai in disuso nelle abitazioni civili e adoperati quasi esclusivamente in contesti cinematografici e televisivi, non poteva far presupporre alcun vantaggio del gruppo dei musicisti, dovuto ad una eventuale maggior familiarità. Sebbene i risultati indicanti migliori capacità uditive nei musicisti non siano talvolta considerati un esempio efficace di trasferimento di competenza, poiché la modalità uditiva è considerata una specificità della musica (Carey et al., 2015), alcuni autori hanno proposto una visione più complessa della pratica musicale come attività eminentemente multimodale, che non recluta solamente né prioritariamente il sistema uditivo, ma coinvolge molte altre aree cognitive e sensoriali (Zatorre et al., 2007; Penhune, 2011; Miendlarzewska & Trost, 2014; Wang & Agius, 2018; Van Vugt et al., 2021), così come molti sono i processi di basso e alto livello implicati, ivi compresi quelli attentivi. I presenti risultati, che mostrano migliori abilità di attenzione spaziale uditiva per i musicisti, andrebbero quindi interpretati non soltanto come la conseguenza di un aspetto specifico del training musicale, che riguarda la discriminazione dei suoni. Essi indicano infatti l'applicazione di competenze specifiche ad un materiale sonoro non riconducibile al contesto musicale e ad un compito che richiedeva l'attivazione di meccanismi di attenzione selettiva spaziale, rappresentativo di una molteplicità di situazioni di vita reale.

L'interazione riportata tra il fattore *gruppo* e il fattore *validità della prova* indica che la differenza tra musicisti e non musicisti sui tempi di risposta non era costante in tutte le condizioni sperimentali. Essa è statisticamente significativa in ognuna delle quattro condizioni di validità della prova, ma particolarmente evidente nelle prove neutre, in cui i partecipanti non venivano indirizzati verso una precisa

posizione target e hanno dovuto così mantenere l'attenzione diffusamente su tutto lo spazio circostante. Questa condizione, unica nel suo genere, è risultata la più impegnativa per i non musicisti, anche se confrontata con le prove invalide, per le quali la letteratura riporta tipicamente tempi di risposta più lenti (ad es. Posner, 1980). I musicisti, invece, hanno mostrato l'attesa facilitazione per le prove neutre, rispetto a quelle invalide. Una delle nostre ipotesi, supportata da evidenze precedenti (Nager et al., 2003) era che i musicisti potessero avere un più efficace ed efficiente dispiegamento dell'attenzione su ampie porzioni di spazio. Nel presente studio, la scena uditiva coincideva con l'intero arco dei 360° attorno ai soggetti, rappresentando così la massima ampiezza possibile da presidiare. I risultati delle prove neutre potrebbero confermare questa ipotesi, dal momento che in queste prove i partecipanti erano tenuti a distribuire equamente le risorse attentive in ogni direzione, e proprio questa è risultata essere la condizione in cui i musicisti mostravano la maggior differenza di velocità con il gruppo di controllo. La pratica musicale, come abbiamo avuto modo di vedere, richiede spesso di dispiegare l'attenzione su ampie porzioni di spazio e di elaborare le informazioni – uditive ma anche multimodali - provenienti anche contemporaneamente da posizioni diverse. Soprattutto durante la pratica d'insieme, ogni musicista deve adattare la sua produzione sonora a quella degli altri componenti del gruppo, e spesso le note, le pause, gli attacchi, i gesti e persino i respiri degli altri musicisti sono cruciali per il risultato sonoro complessivo (Carey et al., 2015; Bishop & Goebel, 2018). La suddetta familiarità con la gestione di ambienti sonori multidimensionali e complessi potrebbe aver contribuito ai risultati osservati, soprattutto nelle prove neutre.

Le evidenze precedenti potrebbero essere ulteriormente supportate dai risultati dell'analisi costi/benefici. L'effetto principale del fattore *gruppo* indica che, per i musicisti, i benefici attentivi erano presenti ma significativamente inferiori a quelli dei non musicisti. Per quanto riguarda, invece, i costi attentivi, l'andamento per i due gruppi è opposto. Infatti, i musicisti hanno mostrato il consueto costo per le prove invalide (rapportato a quanto avviene per le prove neutre), mentre i non musicisti hanno avuto un beneficio anche in questo caso. Infatti, i tempi di risposta

nelle prove invalide erano comunque più veloci di quelli nelle prove neutre. Ciò ha comportato complessivamente ad un beneficio attentivo per i non musicisti, mentre i musicisti hanno mostrato più costi che benefici. Il beneficio quasi nullo riportato dai musicisti può essere interpretato come uno scarso effetto di facilitazione dovuto alla conoscenza preliminare della posizione target. Ciò potrebbe indicare che la condizione di attenzione diffusa sui 360° risulti alquanto familiare e di facile gestione, a tal punto che non si ravvisa un evidente vantaggio nel restringere la selezione spaziale, cosa di cui invece i non musicisti hanno notevolmente giovato. Quanto appena affermato rafforza l'ipotesi di un più efficiente dispiegamento delle risorse attentive sull'intera scena uditiva, da parte dei musicisti.

Una ulteriore conferma potrebbe giungere dai dati sull'accuratezza. Infatti, mentre nel complesso i musicisti sono risultati più accurati della loro controparte, i risultati specifici delle prove neutre non hanno mostrato differenze significative tra i gruppi ( $p = 0.088$ ). Ciò significa che i risultati sopra descritti non possono essere spiegati semplicemente in termini di migliore capacità di discriminazione dei suoni per i musicisti. Inoltre, l'evidenza di un costo attentivo per le prove invalide suggerisce che la migliore abilità di discriminazione dei musicisti non ha eliminato il costo dello spostamento dell'attenzione dalle posizioni attese a quelle effettive di presentazione del target. In altre parole, l'evidenza degli effetti dell'attenzione porta ad escludere che le differenze tra i due gruppi siano esclusivamente ascrivibili ad una migliore abilità di discriminazione dei suoni, pur evidente nel gruppo dei musicisti.

Diversamente, i non musicisti hanno mostrato un beneficio attentivo sia in presenza di prove valide che di prove invalide. Il risultato inaspettato per il confronto tra la performance nelle prove neutre e quella nelle prove invalide suggerisce che anche la selezione spaziale di una posizione non correttamente predittiva costituiva in qualche modo un vantaggio per i partecipanti. Infatti, dopo aver orientato in modo implicito la loro attenzione verso la posizione indicata dal cue, se questo era invalido veniva chiesto loro di spostare l'attenzione per la seconda volta verso la posizione effettiva del target. Sembra quindi che il primo punto di ancoraggio,

costituito dalla posizione indicata dal cue, costituisse un punto di partenza in ogni caso migliore rispetto alla situazione di prova neutra, nella quale i partecipanti non ricevevano una indicazione di selezione spaziale specifica, ma dovevano prestare attenzione contemporaneamente a tutto lo spazio circostante.

Un altro risultato significativo riguarda l'interazione a tre fattori tra *validità della prova*, *emicampo target* e *gruppo*. Al fine di far luce sui possibili aspetti alla base di questa differenza, abbiamo effettuato una serie di analisi separate, che abbiamo descritto al paragrafo sui risultati. Dal confronto tra i due gruppi, emergeva una diversità di andamento nelle singole condizioni di validità della prova, dal momento che per i musicisti l'unica differenza significativa è stata osservata nelle prove invalide con spostamenti tra diversi emicampi, dove il passaggio dall'emicampo posteriore a quello anteriore è risultato più facile del suo opposto. La stessa differenza è stata osservata anche nel gruppo dei non musicisti ma, oltre a questo, anche su prove invalide con spostamenti intra-emicampo e su prove neutre il fattore *emicampo target* ha portato a differenze significative. Ulteriore conferma della stessa ipotesi può arrivare dalle analisi separate per ciascuno dei due gruppi, che vede un effetto principale del fattore *emicampo target* e una interazione significativa tra *validità della prova* ed *emicampo target* soltanto per il gruppo dei non musicisti, mentre per i musicisti lo stesso fattore e la stessa interazione non portano a risultati significativi. L'ipotesi di una maggiore omogeneità della performance dei musicisti su tutto lo spazio circostante, desunta da alcune evidenze dello stesso fenomeno riscontrate tuttavia in porzioni di spazio più ristrette (Krzyzak, 2021), trova qui una prima conferma sperimentale. Sugeriamo che la grande quantità di training uditivo cui i musicisti si sottopongono per molti anni, molto spesso caratterizzata dalla gestione di più suoni concomitanti, provenienti dal proprio e dagli altri strumenti o voci, potrebbe aver modellato la capacità di bilanciare l'asimmetria nella percezione del suono, riportata in studi precedenti. Questo risulta valido non più soltanto in un confronto tra destra e sinistra, come precedentemente dimostrato, ma anche nel confronto tra davanti e dietro e, più in generale, sull'intero spazio circostante.

Infine, l'effetto principale del fattore *gruppo* nell'ANOVA che studiava i gradienti di attenzione spaziale uditiva conferma quanto detto sopra riguardo la maggior velocità dei musicisti rispetto ai non musicisti, che risulta qui indipendente dalla distanza angolare cue-target. L'assenza di un'interazione significativa tra *distanza* e *gruppo*, così come l'effetto principale del fattore *distanza* (considerando i dati di tutti i partecipanti, a prescindere dal gruppo di appartenenza) conferma che il gradiente di attenzione di forma quadratica è il miglior modello di approssimazione possibile per descrivere i costi di spostamento dell'attenzione in funzione della distanza cue-target. L'esperienza musicale pregressa, in questo caso, non ha portato ad alcuna eccezione rispetto alla tendenza generale osservata nella letteratura recente (Golob et al., 2017; Golob & Mock, 2020), nel nostro primo studio e nello studio presente.

## **5.7. Conclusioni**

Alla luce dei risultati appena riportati e discussi nelle loro possibili implicazioni, possiamo concludere che il training musicale, se protratto per almeno sei anni, può contribuire a migliorare le abilità di attenzione spaziale uditiva su tutto lo spazio circostante, trasferendo – seppur in una modalità specifica come quella uditiva – le competenze di discriminazione sonora e di elaborazione spaziale verso un dominio più generale e verso suoni e compiti non specifici del contesto musicale.

Il contributo del sistema visivo alla rappresentazione spaziale, anche per un compito di tipo uditivo, sembra essere significativo, a conferma di quanto evidenziato in studi precedenti e nel nostro primo studio. Gli ancoraggi alle posizioni anteriori sono risultati più efficienti, e il conseguente disimpegno dell'attenzione più impegnativo. Il training musicale sembra mitigare solo in parte questo fenomeno.

La familiarità dei musicisti con la gestione contemporanea di stimoli multimodali, spesso concorrenti e provenienti da diverse direzioni nello spazio, può portare ad una maggiore omogeneità della performance di discriminazione uditiva tra emicampo anteriore e posteriore. Essa risulta infatti indipendente sia

dall'emicampo indicato dai cue spaziali, sia dall'emicampo di presentazione del target sul quale operare il compito di discriminazione.

Le stesse competenze, acquisite durante il training, possono inoltre favorire una maggiore efficienza nel dispiegare l'attenzione in maniera diffusa su tutti i 360° dello spazio circostante, una situazione limite di selettività spaziale che risulta alquanto gravosa per le risorse attentive di una popolazione di non musicisti. Al contrario, per i musicisti, questa condizione di attenzione diffusa sembra essere familiare, al punto che una selezione più ristretta, operata grazie ad un cue spaziale valido, non porta un vantaggio significativo in termini di tempo di risposta e di accuratezza nelle risposte.

Infine, il gradiente di attenzione spaziale, che descrive la distribuzione dei costi di spostamento dell'attenzione al variare della distanza angolare tra cue e target, segue un andamento di forma quadratica, a conferma della recente letteratura e dei risultati del nostro primo studio. Il training musicale, in questo caso, non altera l'andamento riscontrato in una popolazione di non musicisti.

### **5.8. Limitazioni e future direzioni di ricerca**

Stanti le considerazioni fatte in precedenza, riguardo ad alcune delle scelte sperimentali adottate sia nel primo che nel secondo studio, dalla lettura e l'analisi dei risultati del presente studio possono emergere alcune limitazioni specifiche.

Una prima questione riguarda il campione di musicisti. Sebbene siano stati applicati criteri di riconosciuta validità, e l'analisi comparata della musicalità ne abbia confermato la bontà, resta una certa variabilità nelle possibili traiettorie di training che i singoli partecipanti hanno percorso nella loro carriera. Infatti, sotto la voce "training", possono confluire diversi tipi di formazione, da quella più formale e classica a quella moderna, jazzistica o popolare, come già abbiamo avuto modo di argomentare nell'introduzione al capitolo dedicato ai musicisti. In secondo luogo, sarebbe importante tenere traccia della quantità e tipologia dell'attività di musica d'insieme cui i partecipanti hanno preso parte durante la loro carriera. Sia la

tipologia di formazione, sia l'ammontare di musica d'insieme potrebbero avere un effetto sulla performance in un compito come quello adottato in questo studio. Infatti, le diverse carriere musicali possono portare ad una più o meno intensa attività d'insieme, in formazioni più o meno numerose. L'esposizione continuativa e la gestione di stimoli provenienti da porzioni più o meno ampie di spazio, potrebbero avere un impatto sulla distribuzione spaziale dell'attenzione. È da notare che, anche senza aver tenuto conto di queste possibili variabili, gli effetti del training musicale sono già evidenti. Sarebbe quindi opportuno, in studi futuri, reclutare campioni diversi di musicisti, caratterizzati da maggiore o minore familiarità con la musica d'insieme, oppure da diversità di generi musicali di formazione e di pratica quotidiana.

Un altro aspetto che meriterebbe approfondimenti futuri è, a nostro avviso, la valutazione separata dell'impatto delle competenze di ascolto e di esecuzione su un compito come quello somministrato in questo studio. Infatti, l'esposizione continuativa alla musica, anche se non accompagnata dalla pratica strumentale o vocale, potrebbe almeno in parte contribuire a sviluppare abilità attentive spaziali.

Abbiamo infine riscontrato nei dati una tendenza verso minori effetti attentivi complessivi per i musicisti. Questa tendenza tuttavia non ha raggiunto la significatività. L'ipotesi di un minor effetto attentivo sarebbe supportata dalla documentata miglior elaborazione di basso livello degli stimoli uditivi da parte dei musicisti. Se ciò fosse vero, sarebbe appunto ipotizzabile un minor carico attentivo richiesto in compiti di discriminazione. La ricerca futura potrebbe contribuire a verificare questa ipotesi.

## VI - Riferimenti

- Abel, M. K., Li, H. C., Russo, F. A., Schlaug, G., & Loui, P. (2016). Audiovisual interval size estimation is associated with early musical training. *PLoS One*, *11*(10), e0163589.
- Addessi, A. R., Baroni, M., Luzzi, C., & Tafuri, J. (1995). The development of musical stylistic competence in children. *Bulletin of the Council for Research in Music Education*, 8-15.
- Aggius-Vella, E., Campus, C., & Gori, M. (2018). Different audio spatial metric representation around the body. *Scientific Reports*, *8*(1), 9383.
- Aggius-Vella, E., Gori, M., Campus, C., Moore, B. C., Pardhan, S., Kolarik, A. J., & Van der Stoep, N. (2022). Auditory distance perception in front and rear space. *Hearing Research*, *417*, 108468.
- Alain, C., & Izenberg, A. (2003). Effects of attentional load on auditory scene analysis. *Journal of cognitive neuroscience*, *15*(7), 1063-1073.
- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proceedings of the national academy of sciences*, *98*(21), 12301-12306.
- Altenmüller, E., Wiesendanger, M., & Kesselring, J. (Eds.). (2006). *Music, motor control and the brain* (p. xi327). Oxford: Oxford University Press.
- Amer, T., Kalender, B., Hasher, L., Trehub, S. E., & Wong, Y. (2013). Do older professional musicians have cognitive advantages?. *PloS one*, *8*(8), e71630.
- Amunts, K., Schlaug, G., Jäncke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., & Zilles, K. (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Human brain mapping*, *5*(3), 206-215.

- Anderson, P., Anderson, V., & Lajoie, G. (1996). The tower of London test: Validation and standardization for pediatric populations. *The Clinical Neuropsychologist*, *10*(1), 54-65.
- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of experimental child psychology*, *83*(2), 111-130.
- Arbogast, T. L., & Kidd Jr, G. (2000). Evidence for spatial tuning in informational masking using the probe-signal method. *The Journal of the Acoustical Society of America*, *108*(4), 1803-1810.
- Arjona, A., Escudero, M., & Gómez, C. M. (2016). Cue validity probability influences neural processing of targets. *Biological Psychology*, *119*, 171-183.
- Asutay, E., & Västfjäll, D. (2015). Attentional and emotional prioritization of the sounds occurring outside the visual field. *Emotion*, *15*(3), 281–286.
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in cognitive sciences*, *16*(8), 437-443.
- Azzopardi, P., & Cowey, A. (1993). Preferential representation of the fovea in the primary visual cortex. *Nature*, *361*(6414), 719-721.
- Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision research*, *39*(1), 71-86.
- Banerjee, S., Snyder, A. C., Molholm, S., & Foxe, J. J. (2011). Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual target locations: supramodal or sensory-specific control mechanisms?. *Journal of Neuroscience*, *31*(27), 9923-9932.
- Barry, J. G., Ferguson, M. A., & Moore, D. R. (2010). Making sense of listening: the IMAP test battery. *JoVE (Journal of Visualized Experiments)*, (44), e2139.

Benetti, T. F. (n.d.). Jingle Bells. <https://www.titofiorenzobenetti.it/-canti-di-natale---canzoni-natalizie.html>

Benso, F., Turatto, M., Mascetti, G. G., & Umiltà, C. (1998). The time course of attentional focusing. *European Journal of Cognitive Psychology, 10*(4), 373–388.

Benz, S., Sellaro, R., Hommel, B., & Colzato, L. S. (2016). Music makes the world go round: The impact of musical training on non-musical cognitive functions—A review. *Frontiers in psychology, 6*, 2023.

Bergman Nutley, S., Darki, F., & Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Frontiers in human neuroscience, 7*, 926.

Berti, S., & Schröger, E. (2003). Working memory controls involuntary attention switching: evidence from an auditory distraction paradigm. *European Journal of Neuroscience, 17*(5), 1119-1122.

Best, V., Ozmeral, E. J., & Shinn-Cunningham, B. G. (2007). Visually-guided attention enhances target identification in a complex auditory scene. *Journal for the Association for Research in Otolaryngology, 8*, 294-304.

Best, V., Ozmeral, E. J., Kopčo, N., & Shinn-Cunningham, B. G. (2008). Object continuity enhances selective auditory attention. *Proceedings of the National Academy of Sciences, 105*(35), 13174–13178.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12*(4), 403.

Bialystok, E., & DePape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 565.

Bidelman, G. M., & Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *Journal of Neuroscience*, 35(3), 1240-1249.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. *PloS one*, 8(4), e60676.

Bigand, E. (2003). More about the musical expertise of musically untrained listeners. *Annals of the New York Academy of Sciences*, 999(1), 304-312.

Bishop, L., & Goebel, W. (2018). Communication for coordination: Gesture kinematics and conventionality affect synchronization success in piano duos. *Psychological Research*, 82(6), 1177-1194.

Blauert, J. (1997). *Spatial hearing: the psychophysics of human sound localization*. MIT press.

Blurton, S. P., Greenlee, M. W., & Gondan, M. (2015). Cross-modal cueing in audiovisual spatial attention. *Attention, Perception, & Psychophysics*, 77, 2356-2376.

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, 137(1), 378-387.

Bowen, J. D., Alforque, C. V., & Silver, M. A. (2023). Effects of involuntary and voluntary attention on critical spacing of visual crowding. *Journal of vision*, 23(3), 2-2.

Bowmer, A., Mason, K., Knight, J., & Welch, G. (2018). Investigating the impact of a musical intervention on preschool children's executive function. *Frontiers in Psychology*, 9, 2389.

- Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1-11.
- Brandler, S., & Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychology of music*, *31*(2), 123-138.
- Brattico, E. (2015). From pleasure to liking and back: Bottom-up and top-down neural routes to the aesthetic enjoyment. *Art, aesthetics, and the brain*, 303-318.
- Bregman, A. S. (1990; 1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.
- Bregman, A. S. (2002). The auditory scene. In D. J. Levitin (Ed.), *Foundations of cognitive psychology: Core readings* (pp. 213–248). MIT Press.
- Bregman, A. S. (2015). Progress in understanding auditory scene analysis. *Music Perception an Interdisciplinary Journal*, *33*(1), 12–19.
- Brickenkamp, R., & Cubero, N. S. (2002). *D2: test de atención*. Madrid: Tea.
- Bright, J. (2006). Factors influencing outstanding band students' choice of music education as a career. *Contributions to Music Education*, 73-88.
- Broadbent, D. (1958). *Perception and Communication*. London: Pergamon Press.
- Broadbent, D. E. (1952). Listening to one of two synchronous messages. *Journal of Experimental Psychology*, *44*(1), 51–55.
- Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of experimental psychology*, *47*(3), 191.
- Broadbent, D. E. (1956). Successive responses to simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, *8*(4), 145-152.
- Brochard, R., Dufour, A., & Despres, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain and cognition*, *54*(2), 103-109.

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics*, *77*(5), 1465-1487.

Brown, T. E., & Landgraf, J. M. (2010). Improvements in executive function correlate with enhanced performance and functioning and health-related quality of life: evidence from 2 large, double-blind, randomized, placebo-controlled trials in ADHD. *Postgraduate medicine*, *122*(5), 42-51.

Buchtel, H. A., & Butter, C. M. (1988). Spatial attentional shifts: Implications for the role of polysensory mechanisms. *Neuropsychologia*, *26*(4), 499-509.

Bugos, J. A. (2019). The effects of bimanual coordination in music interventions on executive functions in aging adults. *Frontiers in Integrative Neuroscience*, *13*, 68.

Bugos, J. A., & DeMarie, D. (2017). The effects of a short-term music program on preschool children's executive functions. *Psychology of Music*, *45*(6), 855-867.

Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., & Bedenbaugh, P. H. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging and mental health*, *11*(4), 464-471.

Buma, L. A., Bakker, F. C., & Oudejans, R. R. (2015). Exploring the thoughts and focus of attention of elite musicians under pressure. *Psychology of Music*, *43*(4), 459-472.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, *315*(5820), 1860-1862.

Caclin, A., McAdams, S., Smith, B. K., & Giard, M. H. (2008). Interactive processing of timbre dimensions: an exploration with event-related potentials. *Journal of Cognitive Neuroscience*, *20*(1), 49-64.

- Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America*, *118*(1), 471-482.
- Calcus, A. (2024). Development of auditory scene analysis: a mini-review. *Frontiers in Human Neuroscience*, *18*, 1352247.
- Caparos, S., & Linnell, K. J. (2010). The spatial focus of attention is controlled at perceptual and cognitive levels. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1080.
- Carey, D., Rosen, S., Krishnan, S., Pearce, M. T., Shepherd, A., Aydelott, J., & Dick, F. (2015). Generality and specificity in the effects of musical expertise on perception and cognition. *Cognition*, *137*, 81-105.
- Carlyon, R. P. (2004). How the brain separates sounds. *Trends in cognitive sciences*, *8*(10), 465-471.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision research*, *51*(13), 1484-1525.
- Cave, K. R. (2015). Spatial attention. *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*, 1-14.
- Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic bulletin & review*, *6*, 204-223.
- Celma-Miralles, A., & Toro, J. M. (2019). Ternary meter from spatial sounds: Differences in neural entrainment between musicians and non-musicians. *Brain and cognition*, *136*, 103594.
- Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in cognitive sciences*, *17*(4), 179-193.

- Chen, J., Scheller, M., Wu, C., Hu, B., Peng, R., Liu, C., ... & Chen, J. (2022). The relationship between early musical training and executive functions: Validation of effects of the sensitive period. *Psychology of Music*, *50*(1), 86-99.
- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979.
- Chieffo, R., Straffi, L., Inuggi, A., Gonzalez-Rosa, J. J., Spagnolo, F., Coppi, E., ... & Leocani, L. (2016). Motor cortical plasticity to training started in childhood: The example of piano players. *PloS one*, *11*(6), e0157952.
- Chobert, J., François, C., Velay, J. L., & Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral cortex*, *24*(4), 956-967.
- Chung, K. K. H. (2015). Socioeconomic status and academic achievement. *International encyclopedia of the social & behavioral sciences*, 924-930.
- Clark, I. N., & Tamplin, J. (2016, April). How music can influence the body: Perspectives from current research. In *Voices: A World Forum for Music Therapy* (Vol. 16, No. 2).
- Clarke, E., DeNora, T., & Vuoskoski, J. (2015). Music, empathy and cultural understanding. *Physics of life reviews*, *15*, 61-88.
- Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., & Kidd Jr, G. (2016). Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PloS one*, *11*(7), e0157638.
- Cogo-Moreira, H., & Lamont, A. (2018). Multidimensional measurement of exposure to music in childhood: Beyond the musician/non-musician dichotomy. *Psychology of Music*, *46*(4), 459-472.

Conway, A. R., & Kane, M. J. (2001). Capacity, control and conflict: An individual differences perspective on attentional capture. *Attention, distraction and action: Multiple perspectives on attention capture*, 133, 349-372.

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, 8, 331-335.

Correia, A. I., Vincenzi, M., Vanzella, P., Pinheiro, A. P., Lima, C. F., & Schellenberg, E. G. (2022). Can musical ability be tested online?. *Behavior Research Methods*, 54(2), 955-969.

Corrigall, K. A., & Schellenberg, E. G. (2015). Predicting who takes music lessons: Parent and child characteristics. *Frontiers in psychology*, 6, 282.

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in psychology*, 4, 222.

Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of music*, 32(2), 139-152.

Cowan, N., & Wood, N. L. (1997). Constraints on awareness, attention, processing, and memory: Some recent investigations with ignored speech. *Consciousness and Cognition*, 6(2-3), 182-203.

Cox, A. (2016). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Indiana University Press.

Cross, I. (2014). Music and communication in music psychology. *Psychology of music*, 42(6), 809-819.

Culling, J. F., & Summerfield, Q. (1995a). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *The Journal of the Acoustical Society of America*, 98(2), 785-797.

- Culling, J. F., & Summerfield, Q. (1995b). The role of frequency modulation in the perceptual segregation of concurrent vowels. *The Journal of the Acoustical Society of America*, 98(2), 837-846.
- D'Ausilio, A., Novembre, G., Fadiga, L., & Keller, P. E. (2015). What can music tell us about social interaction?. *Trends in cognitive sciences*, 19(3), 111-114.
- Dahl, S., Bevilacqua, F., & Bresin, R. (2010). Gestures in performance. In Godøy, R. I., & Leman, M. (Eds.). (2010). *Musical gestures: Sound, movement, and meaning*. Routledge.
- Dalcroze, E. J. (2008). *Il ritmo, la musica e l'educazione* (Vol. 21). EDT srl.
- Dalton, P., & Lavie, N. (2004). Auditory attentional capture: effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 30(1), 180.
- Dalton, P., Santangelo, V., & Spence, C. (2009). The role of working memory in auditory selective attention. *Quarterly Journal of Experimental Psychology*, 62(11), 2126-2132.
- Darwin, C. J. (1997). Auditory grouping. *Trends in cognitive sciences*, 1(9), 327-333.
- Darwin, C. J. (2008). Spatial hearing and perceiving sources. In *Auditory perception of sound sources* (pp. 215-232). Boston, MA: Springer US.
- Darwin, C. J., & Carlyon, R. P. (1995). Auditory grouping. In B. C. J. Moore (Ed.), *Hearing* (pp. 387-424). Academic Press.
- De Coensel, B., & Botteldooren, D. (2010). A model of saliency-based auditory attention to environmental sound. In *20th International Congress on Acoustics (ICA-2010)* (pp. 1-8).

- De Freitas, J., Liverence, B. M., & Scholl, B. J. (2014). Attentional rhythm: A temporal analogue of object-based attention. *Journal of Experimental Psychology: General*, *143*(1), 71–76.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21.
- Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in psychology*, *2*, 124.
- Degé, F., Wehrum, S., Stark, R., & Schwarzer, G. (2014). Music lessons and academic self-concept in 12-to 14-year-old children. *Musicae Scientiae*, *18*(2), 203-215.
- Delalande, F., (2016) *La musica è un gioco da bambini*. Franco Angeli.
- Delalande, F., Guardabasso, G., & Marconi, L. (1993). *Le condotte musicali: comportamenti e motivazioni del fare e ascoltare musica*. Clueb.
- Delalande, F., Vidal, J., & Reibel, G. (1984). *La musique est un jeu d'enfant*. Paris: Buchet/Chastel.
- Delfrati, C. (2008). *Fondamenti di pedagogia musicale*. EDT srl.
- Deliège I., Sloboda J. A. *Naissance et développement du sens musical*. Presses Universitaires de France.
- Delis, D. C., Kramer, J. H., Kaplan, E., & Holdnack, J. (2004). Reliability and validity of the Delis-Kaplan Executive Function System: an update. *Journal of the International Neuropsychological Society*, *10*(2), 301-303.
- Dell'Anna, A., Leman, M., & Berti, A. (2021). Musical interaction reveals music as embodied language. *Frontiers in Neuroscience*, *15*, 667838.
- Deng, Y., Choi, I., & Shinn-Cunningham, B. (2020). Topographic specificity of alpha power during auditory spatial attention. *Neuroimage*, *207*, 116360.

- Deng, Y., Reinhart, R. M., Choi, I., & Shinn-Cunningham, B. G. (2019). Causal links between parietal alpha activity and spatial auditory attention. *elife*, 8, e51184.
- Deouell, L. Y., Heller, A. S., Malach, R., D'Esposito, M., & Knight, R. T. (2007). Cerebral responses to change in spatial location of unattended sounds. *Neuron*, 55(6), 985-996.
- Deriu, R. (2000). Intorno al concetto di competenza musicale. *Musica Domani* 116, 3-11.
- Deutsch, D. (2013). Absolute pitch. *The psychology of music*, 3, 141-182.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review*, 70(1), 80.
- Diamond, A. (2013). Executive functions. *Annual review of psychology*, 64(1), 135-168.
- Diamond, A. (2016). Why improving and assessing executive functions early in life is critical. In J. A. Griffin, P. McCardle, & L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 11–43). American Psychological Association.
- Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental cognitive neuroscience*, 18, 34-48.
- Ding, N., Melloni, L., Tian, X., & Poeppel, D. (2017). Rule-based and word-level statistics-based processing of language: insights from neuroscience. *Language, cognition and neuroscience*, 32(5), 570-575.
- Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition*, 74(1), 1-32.

- Driver, J. (2001). A selective review of selective attention research from the past century. *British journal of psychology*, 92(1), 53-78.
- Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in cognitive sciences*, 2(7), 254-262.
- Du, Y., & Zatorre, R. J. (2017). Musical training sharpens and bonds ears and tongue to hear speech better. *Proceedings of the National Academy of Sciences*, 114(51), 13579-13584.
- Eimer, M., & Van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: Evidence from event-related potentials. *Psychophysiology*, 39(4), 437-449.
- Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, 270(5234), 305-307.
- Elhilali, M., Fritz, J. B., Chi, T. S., & Shamma, S. A. (2007). Auditory cortical receptive fields: stable entities with plastic abilities. *Journal of Neuroscience*, 27(39), 10372-10382.
- Elmer, S., Hänggi, J., Meyer, M., & Jäncke, L. (2013). Increased cortical surface area of the left planum temporale in musicians facilitates the categorization of phonetic and temporal speech sounds. *Cortex*, 49(10), 2812-2821.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, 16(1), 143-149.
- Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & psychophysics*, 40(4), 225-240.

- Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 583.
- Erulkar, S. D. (1972). Comparative aspects of spatial localization of sound. *Physiological reviews*, *52*(1), 237-360.
- Escoffier, N., & Tillmann, B. (2008). The tonal function of a task-irrelevant chord modulates speed of visual processing. *Cognition*, *107*(3), 1070-1083.
- Fairnie, J., Moore, B. C., & Remington, A. (2016). Missing a trick: Auditory load modulates conscious awareness in audition. *Journal of experimental psychology: human perception and performance*, *42*(7), 930.
- Fan, J., & Posner, M. (2004). Human attentional networks. *Psychiatrische Praxis*, *31*(S 2), 210-214.
- Farah, M. J., Wong, A. B., Monheit, M. A., & Morrow, L. A. (1989). Parietal lobe mechanisms of spatial attention: Modality-specific or supramodal?. *Neuropsychologia*, *27*(4), 461-470.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.
- Fiebelkorn, I. C., & Kastner, S. (2019). A rhythmic theory of attention. *Trends in cognitive sciences*, *23*(2), 87-101.
- Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in systems neuroscience*, *7*, 68.
- Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in

developmental dyslexia: A randomized control trial. *PloS one*, *10*(9), e0138715.

Forgeard, M., Schlaug, G., Norton, A., Rosam, C., Iyengar, U., & Winner, E. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music perception*, *25*(4), 383-390.

Fortuna, S., & Nijs, L. (2019). Children's verbal explanations after verbal versus body movement interactions with music. In *Counterpoints of the senses: bodily experiences in musical learning (MERYC19)* (pp. 78-79).

Francès, R. (1984). *La perception de la musique* (Vol. 14). Vrin.

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038-2043.

Franklin, M. S., Sledge Moore, K., Yip, C. Y., Jonides, J., Rattray, K., & Moher, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, *36*(3), 353-365.

Frankowska, N., Parzuchowski, M., Wojciszke, B., Olszanowski, M., & Winkielman, P. (2020). Rear negativity: Verbal messages coming from behind are perceived as more negative. *European Journal of Social Psychology*, *50*(4), 889-902.

Freschi, A. M., & Neulicherl, R. (2012). *Metodologia dell'insegnamento strumentale*. ETS.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of experimental psychology: General*, *133*(1), 101.

Frischen, U., Schwarzer, G., & Degé, F. (2019). Comparing the effects of rhythm-based music training and pitch-based music training on executive functions in preschoolers. *Frontiers in Integrative Neuroscience*, *13*, 41.

- Fritz, J. B., Elhilali, M., & Shamma, S. A. (2007). Adaptive changes in cortical receptive fields induced by attention to complex sounds. *Journal of neurophysiology*, 98(4), 2337-2346.
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—focusing the searchlight on sound. *Current Opinion in Neurobiology*, 17(4), 437–455.
- Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience*, 35(45), 15187-15198.
- Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye–hand span during musical sight–reading. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1436), 2435-2440.
- Gagné, F., & McPherson, G. E. (2016). Analyzing musical prodigiousness using Gagné’s integrative model of talent development. In McPherson, G. E. (Ed.). (2016). *Musical prodigies: Interpretations from psychology, education, musicology, and ethnomusicology*. Oxford University Press.
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, 22(3-4), 455-479.
- Gamble, M. L., & Luck, S. J. (2011). N2ac: An ERP component associated with the focusing of attention within an auditory scene. *Psychophysiology*, 48(8), 1057-1068.
- Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of neuroscience*, 23(27), 9240-9245.
- Gembris, H. (1997). Historical phases in the definition of musicality. *Psychomusicology: A Journal of Research in Music Cognition*, 16(1-2), 17.

- George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083-1094.
- Gibson, C., Folley, B. S., & Park, S. (2009). Enhanced divergent thinking and creativity in musicians: A behavioral and near-infrared spectroscopy study. *Brain and cognition*, *69*(1), 162-169.
- Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2015). Defining the biological bases of individual differences in musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140092.
- Golob, E. J., & Holmes, J. L. (2011). Cortical mechanisms of auditory spatial attention in a target detection task. *Brain research*, *1384*, 128-139.
- Golob, E. J., & Mock, J. R. (2019). Auditory spatial attention capture, disengagement, and response selection in normal aging. *Attention, Perception, & Psychophysics*, *81*, 270-280.
- Golob, E. J., & Mock, J. R. (2020). Dynamics of auditory spatial attention gradients. *Cognition*, *194*, 104058.
- Golob, E. J., Nelson, J. T., Scheuerman, J., Venable, K. B., & Mock, J. R. (2021). Auditory spatial attention gradients and cognitive control as a function of vigilance. *Psychophysiology*, *58*(10), e13903.
- Golob, E. J., Venable, K. B., Anderson, M. T., Benzell, J. A., & Scheuerman, J. (2016). Modelling Auditory Spatial Attention with Constraints. In *AIC* (pp. 59-72).
- Golob, E. J., Winston, J., & Mock, J. R. (2017). Impact of spatial and verbal short-term memory load on auditory spatial attention gradients. *Frontiers in psychology*, *8*, 2028.
- Gomes, H., Barrett, S., Duff, M., Barnhardt, J., & Ritter, W. (2008). The effects of interstimulus interval on event-related indices of attention: An auditory

selective attention test of perceptual load theory. *Clinical Neurophysiology*, 119(3), 542-555.

Good, A., Gordon, K. A., Papsin, B. C., Nespoli, G., Hopyan, T., Peretz, I., & Russo, F. A. (2017). Benefits of music training for perception of emotional speech prosody in deaf children with cochlear implants. *Ear and hearing*, 38(4), 455-464.

Gordon, E. (1965). The Musical Aptitude Profile: A new and unique musical aptitude test battery. *Bulletin of the Council for Research in Music Education*, 12-16.

Gordon, E. E. (1979). Developmental music aptitude as measured by the Primary Measures of Music Audiation. *Psychology of Music*, 7(1), 42-49.

Gordon, E. E. (1989). Audiation, Music Learning Theory, Music Aptitude, and Creativity. In *Suncoast music education forum on creativity* (Vol. 75, p. 81).

Gouzouasis, P., Guhn, M., & Kishor, N. (2007). The predictive relationship between achievement and participation in music and achievement in core Grade 12 academic subjects. *Music Education Research*, 9(1), 81-92.

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: current findings and future perspectives. *Topics in cognitive science*, 4(4), 585-606.

Graziano, M. S., Gross, C. G., Taylor, C. S., & Moore, T. (2004). Multisensory neurons for the control of defensive movements. In Calvert, G. (2004). *The handbook of multisensory processes*. MIT press.

Graziano, M. S., Reiss, L. A., & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, 397(6718), 428-430.

Green, J. J., & Woldorff, M. G. (2012). Arrow-elicited cueing effects at short intervals: Rapid attentional orienting or cue-target stimulus conflict?. *Cognition*, 122(1), 96-101.

- Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (1999). Absolute pitch: prevalence, ethnic variation, and estimation of the genetic component. *The American Journal of Human Genetics*, *65*(3), 911-913.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annu. Rev. Neurosci.*, *27*(1), 649-677.
- Grothe, B., Pecka, M., & McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiological reviews*, *90*(3), 983-1012.
- Gyu, S., & Hae, J. (2017). Statistical Round. *Korean journal of anesthesiology*, *2005*(6419), 2005-7563.
- Habibi, A., Damasio, A., Ilari, B., Veiga, R., Joshi, A. A., Leahy, R. M., Haldar, J. P., Varadarajan, D., Bhushan, C., & Damasio, H. (2017). Childhood Music Training Induces Change in Micro and Macroscopic Brain Structure: Results from a Longitudinal Study. *Cerebral Cortex*, *1*, 12.
- Haftner, E. R., & Saberi, K. (2001). A level of stimulus representation model for auditory detection and attention. *The Journal of the Acoustical Society of America*, *110*(3), 1489-1497.
- Haftner, E. R., Sarampalis, A., & Loui, P. (2007). Auditory attention and filters. In *Springer eBooks* (pp. 115–142).
- Haftner, E. R., Schlauch, R. S., & Tang, J. (1993). Attending to auditory filters that were not stimulated directly. *The Journal of the Acoustical Society of America*, *94*(2), 743-747.
- Hallam, S. (2010). 21st century conceptions of musical ability. *Psychology of music*, *38*(3), 308-330.
- Hallam, S. (2016). Levels of expertise and musical aspirations in young instrumental players. In Stakelum, M. (Ed.). (2013). *Developing the Musician: Contemporary Perspectives on Teaching and Learning* (1st ed.). Routledge.

- Hallam, S., & Prince, V. (2003). Conceptions of musical ability. *Research Studies in Music Education*, 20(1), 2-22.
- Hanna-Pladdy, B., & MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology*, 25(3), 378.
- Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in cognitive sciences*, 11(11), 466-472.
- Hansen, M., Wallentin, M., & Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychology of Music*, 41(6), 779-793.
- Hartman, D. E. (2009). Wechsler Adult Intelligence Scale IV (WAIS IV): return of the gold standard. *Applied neuropsychology*, 16(1), 85-87.
- Hayward, D. A., & Ristic, J. (2013). Measuring attention using the Posner cuing paradigm: the role of across and within trial target probabilities. *Frontiers in Human Neuroscience*, 7, 205.
- Healey, C. G., & Sawant, A. P. (2012). On the limits of resolution and visual angle in visualization. *ACM Transactions on Applied Perception (TAP)*, 9(4), 1-21.
- Heffner, R. S., & Heffner, H. E. (1992). Evolution of sound localization in mammals. In Fay, D. B. W. R. R., & Popper, A. N. *The evolutionary biology of hearing* (pp. 691-715). New York, NY: Springer New York.
- Heine, L., Ferré, F., Lavandier, M., & Perrin, F. (2020, December). Enhanced spatial realism of sound boosts cognition in DOC patients. In *Forum Acusticum* (pp. 1191-1191).
- Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B., ... & Cattapan-Ludewig, K. (2010). Musical training induces functional plasticity in human hippocampus. *Journal of Neuroscience*, 30(4), 1377-1384.

- Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*, 76(3), 486-502.
- Hill, N. I., & Darwin, C. J. (1996). Lateralization of a perturbed harmonic: Effects of onset asynchrony and mistuning. *The Journal of the Acoustical Society of America*, 100(4), 2352-2364.
- Ho, C., & Spence, C. (2009). Using peripersonal warning signals to orient a driver's gaze. *Human Factors*, 51(4), 539-556.
- Hofman, P. M., & Van Opstal, A. J. (1998). Spectro-temporal factors in two-dimensional human sound localization. *The Journal of the Acoustical Society of America*, 103(5), 2634-2648.
- Hofman, P. M., Van Riswick, J. G., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature neuroscience*, 1(5), 417-421.
- Huang, N., & Elhilali, M. (2017). Auditory salience using natural soundscapes. *The Journal of the Acoustical Society of America*, 141(3), 2163-2176.
- Hugdahl, K. (2003). Dichotic Listening in the study of auditory laterality. In K. Hugdahl & R. J. Davidson (Eds.), *The asymmetrical brain* (pp. 441-475). Boston Review.
- Hugdahl, K. (2011). Fifty years of dichotic listening research – Still going and going and. . . *Brain and Cognition*, 76(2), 211-213.
- Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair?. *Developmental neuropsychology*, 28(2), 645-668.
- Hughes, H. C., & Zimba, L. D. (1985). Spatial maps of directed visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 11(4), 409.
- Humes, L. E. (1996). Speech understanding in the elderly. *Journal of the American Academy of Audiology*, 7(3), 161-167.

Imberty, M. (1993). How do we perceive atonal music? Suggestions for a theoretical approach. *Contemporary music review*, 9(1-2), 325-337.

Iwaya, Y., Suzuki, Y., & Kimura, D. (2003). Effects of head movement on front-back error in sound localization. *Acoustical science and technology*, 24(5), 322-324.

Jääskeläinen, I. P., & Ahveninen, J. (2014). Auditory-Cortex Short-Term Plasticity Induced by Selective Attention. *Neural plasticity*, 2014(1), 216731.

James, C. E., Oechslin, M. S., Michel, C. M., & De Preto, M. (2017). Electrical neuroimaging of music processing reveals mid-latency changes with level of musical expertise. *Frontiers in neuroscience*, 11, 613.

Janus, M., Lee, Y., Moreno, S., & Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *Journal of experimental child psychology*, 144, 84-97.

Jaschke, A. C., Honing, H., & Scherder, E. J. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Frontiers in neuroscience*, 12, 103.

Jentsch, I., Mkrtchian, A., & Kansal, N. (2014). Improved effectiveness of performance monitoring in amateur instrumental musicians. *Neuropsychologia*, 52, 117-124.

Jewell, G., & McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia*, 38(1), 93-110.

Johnson, M. (2007). The meaning of the body. In *Developmental perspectives on embodiment and consciousness* (pp. 35-60). Psychology Press.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, 37, 43-75.

- Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological review*, 83(5), 323.
- Jones, M. R. (2018). *Time will tell: A theory of dynamic attending*. Oxford University Press.
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological review*, 96(3), 459.
- Justus, T., & List, A. (2005). Auditory attention to frequency and time: an analogy to visual local–global stimuli. *Cognition*, 98(1), 31-51.
- Kaganovich, N., Kim, J., Herring, C., Schumaker, J., MacPherson, M., & Weber-Fox, C. (2013). Musicians show general enhancement of complex sound encoding and better inhibition of irrelevant auditory change in music: an ERP study. *European Journal of Neuroscience*, 37(8), 1295-1307.
- Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Karhson, D. S., Mock, J. R., & Golob, E. J. (2015). The role of right inferior parietal cortex in auditory spatial attention: a repetitive transcranial magnetic stimulation study. *Plos one*, 10(12), e0144221.
- Karma, K. (2007). Musical aptitude definition and measure validation: ecological validity can endanger the construct validity of musical aptitude tests. *Psychomusicology: A Journal of Research in Music Cognition*, 19(2), 79.
- Kaviani, H., Mirbaha, H., Pournaseh, M., & Sagan, O. (2014). Can music lessons increase the performance of preschool children in IQ tests?. *Cognitive processing*, 15, 77-84.
- Kaya, E. M., & Elhilali, M. (2017). Modelling auditory attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714), 20160101.

- Kayser, C., Petkov, C. I., Lippert, M., & Logothetis, N. K. (2005). Mechanisms for allocating auditory attention: an auditory saliency map. *Current biology, 15*(21), 1943-1947.
- Kean, M., & Crawford, T. J. (2008). Cueing visual attention to spatial locations with auditory cues. *Journal of Eye Movement Research, 2*(3).
- Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception, 28*(1), 27-46.
- Kenny, D. (2011). *The psychology of music performance anxiety*. OUP Oxford.
- Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a “cocktail party”. *Journal of Neuroscience, 30*(2), 620-628.
- Kidd, G., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listen. *The Journal of the Acoustical Society of America, 118*(6), 3804-3815.
- Kimura, D. (1967). Functional asymmetry of the brain in dichotic listening. *Cortex, 3*(2), 163-178.
- Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and human behavior, 31*(5), 354-364.
- Klingberg, T. (1998). Concurrent performance of two working memory tasks: potential mechanisms of interference. *Cerebral cortex (New York, NY: 1991), 8*(7), 593-601.
- Klumpp, R. G., & Eady, H. R. (1956). Some measurements of interaural time difference thresholds. *The Journal of the Acoustical Society of America, 28*(5), 859-860.

- Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in cognitive sciences*, 23(1), 63-77.
- Koffka, K. (1922). Perception: an introduction to the Gestalt-Theorie. *Psychological bulletin*, 19(10), 531.
- Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., & Pardhan, S. (2016). Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics*, 78, 373-395.
- Kolinsky, R., Cuvelier, H., Goetry, V., Peretz, I., & Morais, J. (2009). Music training facilitates lexical stress processing. *Music Perception*, 26(3), 235-246.
- Kong, L., Michalka, S. W., Rosen, M. L., Sheremata, S. L., Swisher, J. D., Shinn-Cunningham, B. G., & Somers, D. C. (2014). Auditory spatial attention representations in the human cerebral cortex. *Cerebral Cortex*, 24(3), 773-784.
- Kopiez, R., & In Lee, J. (2008). Towards a general model of skills involved in sight reading music. *Music education research*, 10(1), 41-62.
- Koreimann, S., Gula, B., & Vitouch, O. (2014). Inattentional deafness in music. *Psychological research*, 78, 304-312.
- Korsakova-Kreyn, M. (2018). Two-level model of embodied cognition in music. *Psychomusicology: Music, Mind, and Brain*, 28(4), 240.
- Koshimori, Y., & Thaut, M. H. (2019). New perspectives on music in rehabilitation of executive and attention functions. *Frontiers in Neuroscience*, 13, 1245.
- Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in cognitive sciences*, 22(10), 896-910.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature reviews neuroscience*, 11(8), 599-605.

- Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *Journal of Neuroscience*, *34*(36), 11913-11918.
- Krzyżak, A. (2021). Different pattern of auditory processing lateralization in musicians and non-musicians. *Acta Neuropsychologica*, *19*, 105-119.
- Kurkowski, Z. M. (2007). Test rozdzielności słyszenia. *Stowarzyszenie Przyjaciół Osób Niestyszających i Niedosłyszących „Człowiek Człowiekowi”*, Warszawa.
- Lamont, A. (2002). Musical identities and the school environment. In MacDonald, R. A., Hargreaves, D. J., & Miell, D. (Eds.). (2002). *Musical identities*. OUP Oxford.
- Laukka, P. (2007). Uses of music and psychological well-being among the elderly. *Journal of happiness studies*, *8*, 215-241.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance*, *21*(3), 451.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences*, *9*(2), 75-82.
- Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current directions in psychological science*, *19*(3), 143-148.
- Lavie, N., & Tsai, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & psychophysics*, *56*, 183-197.
- Law, L. N., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the profile of music perception skills. *PloS one*, *7*(12), e52508.

- Lee, A. K., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M. S., & Shinn-Cunningham, B. G. (2013). Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. *Frontiers in neuroscience*, *6*, 190.
- Lee, L., & Lin, S. C. (2015). The impact of music activities on foreign language, English learning for young children. *Journal of the European Teacher Education Network*, *10*(0), 13-23.
- Leman, M. (2016). *The expressive moment: How interaction (with music) shapes human empowerment*. MIT press.
- Leman, M., Maes, P. J., Nijs, L., & Van Dyck, E. (2018). What is embodied music cognition?. *Springer handbook of systematic musicology*, 747-760.
- Lewald, J., & Getzmann, S. (2015). Electrophysiological correlates of cocktail-party listening. *Behavioural brain research*, *292*, 157-166.
- Liljenström, H. (2003). Neural stability and flexibility: a computational approach. *Neuropsychopharmacology*, *28*(1), S64-S73.
- Lima, C. F., Correia, A. I., Müllensiefen, D., & Castro, S. L. (2020). Goldsmiths Musical Sophistication Index (Gold-MSI): Portuguese version and associations with socio-demographic factors, personality and music preferences. *Psychology of Music*, *48*(3), 376-388.
- List, A., Justus, T., Robertson, L. C., & Bentin, S. (2007). A mismatch negativity study of local–global auditory processing. *Brain Research*, *1153*, 122-133.
- Liu, H., Bai, Y., Zheng, Q., Liu, J., Zhu, J., & Ni, G. (2024). Electrophysiological correlation of auditory selective spatial attention in the “cocktail party” situation. *Human Brain Mapping*, *45*(11).
- Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *Quarterly Journal of Experimental Psychology*, *64*(11), 2153-2167.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception, 11*(2), 115-128.

Lou, H., Lorist, M. M., & Pilz, K. S. (2022). Effects of cue validity on attentional selection. *Journal of Vision, 22*(8), 15-15.

Loui, P., & Guetta, R. E. (2018). Music and attention, executive function, and creativity. In Hodges, D., & Thaut, M. (Eds.). (2019). *The Oxford handbook of music and the brain*. Oxford University Press.

Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception & psychophysics, 69*, 1084-1092.

Loui, P., Grent, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Cognitive Brain Research, 25*(3), 678-687.

Luck, G., & Nte, S. (2008). An investigation of conductors' temporal gestures and conductor—musician synchronization, and a first experiment. *Psychology of Music, 36*(1), 81-99.

Luck, G., & Sloboda, J. A. (2007). An investigation of musicians' synchronization with traditional conducting beat patterns. *Music Performance Research, 1*, 26-46.

Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: An automated feature-extraction analysis. *Music Perception, 24*(2), 189-200.

Luck, S. J. (1995). Multiple mechanisms of visual-spatial attention: recent evidence from human electrophysiology. *Behavioural brain research, 71*(1-2), 113-123.

Luria, A. R., Karpov, B. A., & Yarbuss, A. L. (1966). Disturbances of active visual perception with lesions of the frontal lobes. *Cortex, 2*(2), 202-212.

- Macaluso, E., Eimer, M., Frith, C. D., & Driver, J. (2003). Preparatory states in crossmodal spatial attention: spatial specificity and possible control mechanisms. *Experimental Brain Research*, *149*, 62-74.
- Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*(5482), 1206-1208.
- MacGregor, C., & Müllensiefen, D. (2019). The Musical Emotion Discrimination Task: A new measure for assessing the ability to discriminate emotions in music. *Frontiers in Psychology*, *10*, 1955.
- Macpherson, E. A., & Middlebrooks, J. C. (2000). Localization of brief sounds: effects of level and background noise. *The Journal of the Acoustical Society of America*, *108*(4), 1834-1849.
- Maddox, R. K., & Shinn-Cunningham, B. G. (2012). Influence of task-relevant and task-irrelevant feature continuity on selective auditory attention. *Journal of the Association for Research in Otolaryngology*, *13*(1), 119-129.
- Maddox, R. K., Pospisil, D. A., Stecker, G. C., & Lee, A. K. (2014). Directing eye gaze enhances auditory spatial cue discrimination. *Current Biology*, *24*(7), 748-752.
- Maes, P. J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in psychology*, *4*, 1008.
- Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *Journal of cognitive neuroscience*, *18*(2), 199-211.
- Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, *32*(1), 4-18.
- Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of cognitive neuroscience*, *23*(2), 294-305.

- Marmel, F., Tillmann, B., & Dowling, W. J. (2008). Tonal expectations influence pitch perception. *Perception & Psychophysics*, *70*, 841-852.
- Martens, S., & Wyble, B. (2010). The attentional blink: Past, present, and future of a blind spot in perceptual awareness. *Neuroscience & Biobehavioral Reviews*, *34*(6), 947-957.
- Martens, S., Wierda, S. M., Dun, M., de Vries, M., & Smid, H. G. (2015). Musical minds: attentional blink reveals modality-specific restrictions. *Plos One*, *10*(2), e0118294.
- Martínez-Castilla, P., Gutiérrez-Blasco, I. M., Spitz, D. H., & Granot, R. (2021). The efficacy of music for emotional wellbeing during the COVID-19 lockdown in Spain: an analysis of personal and context-related variables. *Frontiers in psychology*, *12*, 647837.
- Matyja, J. R., & Schiavio, A. (2013). Enactive music cognition: background and research themes. *Constructivist foundations*, *8*(3).
- Mazzucchi, A., Cattelani, R., & Umiltà, C. (1983). Hemispheric prevalence in acoustical attention. *Brain and Cognition*, *2*(1), 1-11.
- McPherson, G. (Ed.). (2015). *The child as musician: A handbook of musical development*. Oxford University Press.
- McPherson, G. E., & Hallam, S. (2009). Musical potential. In Hallam, S., Cross, I., & Thaut, M. (Eds.). (2009). *Oxford handbook of music psychology*. Oxford University Press.
- Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological science*, *21*(7), 914-919.
- Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, *485*(7397), 233-236.

Michalka, S. W., Rosen, M. L., Kong, L., Shinn-Cunningham, B. G., & Somers, D. C. (2016). Auditory spatial coding flexibly recruits anterior, but not posterior, visuotopic parietal cortex. *Cerebral Cortex*, *26*(3), 1302-1308.

Middlebrooks, J. C., & Green, D. M. (1991). Localization by human listeners. *Annu. Rev. Psychol*, *42*, 135-59.

Miendlarzewska, E. A., & Trost, W. J. (2014). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Frontiers in neuroscience*, *7*, 279.

Mills, M. (2010). Being a musician: Musical identity and the adolescent singer. *Bulletin of the Council for Research in Music Education*, (186), 43-54.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100.

Moller, A. (2012). *Auditory physiology*. Elsevier.

Møller, A. R. (1961). Network model of the middle ear. *The Journal of the Acoustical Society of America*, *33*(2), 168-176.

Moncrieff, D. W. (2011). Dichotic listening in children: Age-related changes in direction and magnitude of ear advantage. *Brain and cognition*, *76*(2), 316-322.

Mondor, T. A., & Lacey, T. E. (2001). Facilitative and inhibitory effects of cuing sound duration, intensity, and timbre. *Perception & Psychophysics*, *63*(4), 726-736.

Mondor, T. A., & Terrio, N. A. (1998). Mechanisms of perceptual organization and auditory selective attention: The role of pattern structure. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1628.

Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21(2), 387.

Moore, D. R. (1991). Anatomy and physiology of binaural hearing. *Audiology*, 30(3), 125-134.

Moore, G. P., & Chen, J. (2010). Timings and interactions of skilled musicians. *Biological cybernetics*, 103, 401-414.

Moradzadeh, L. (2014). *The Association Between Musical Training, Bilingualism, And Executive Function* (Doctoral Dissertation, York University Toronto).

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology*, 11(1), 56-60.

Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing research*, 308, 84-97.

Moreno, S., & Farzan, F. (2015). Music training and inhibitory control: a multidimensional model. *Annals of the New York Academy of Sciences*, 1337(1), 147-152.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological science*, 22(11), 1425-1433.

Moreno, S., Lee, Y., Janus, M., & Bialystok, E. (2015). Short-term second language and music training induces lasting functional brain changes in early childhood. *Child development*, 86(2), 394-406.

Moreno, S., Lee, Y., Janus, M., & Bialystok, E. (2015). Short-term second language and music training induces lasting functional brain changes in early childhood. *Child development, 86*(2), 394-406.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cerebral cortex, 19*(3), 712-723.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences, 114*(42), E8913-E8921.

Mortari, L. (2010). Cercare il rigore metodologico per una ricerca pedagogica scientificamente fondata. *Education Sciences & Society, 1*(1), 143-156.

Mortari, L., & Tarozzi, M. (2010). Phenomenology as philosophy of research: An introductory essay. Zeta books.

Mortari, L., Ghirotto, L., (2019). *Metodi per la ricerca educativa*. Carocci.

Mosing, M. A., Madison, G., Pedersen, N. L., & Ullén, F. (2016). Investigating cognitive transfer within the framework of music practice: Genetic pleiotropy rather than causality. *Developmental Science, 19*(3), 504-512.

Mukari, S. Z., Keith, R. W., Tharpe, A. M., & Johnson, C. D. (2006). Development and standardization of single and double dichotic digit tests in the Malay language: Desarrollo y estandarización de pruebas de dígitos dicóticos sencillos y dobles en lengua malaya. *International Journal of Audiology, 45*(6), 344-352.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PloS one, 9*(2), e89642.

- Müller, J. R., Philiastides, M. G., & Newsome, W. T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences*, *102*(3), 524-529.
- Münste, T. F., Kohlmetz, C., Nager, W., & Altenmüller, E. (2001). Superior auditory spatial tuning in conductors. *Nature*, *409*(6820), 580-580.
- Münste, T. F., Nager, W., Beiss, T., Schroeder, C., & Altenmüller, E. (2003). Specialization of the specialized: Electrophysiological investigations in professional musicians. *Annals of the New York academy of sciences*, *999*(1), 131-139.
- Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next?. *Psychonomic bulletin & review*, *23*, 1316-1340.
- Murphy, S., & Dalton, P. (2016). Out of touch? Visual load induces inattentional numbness. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 761–765.
- Murphy, S., Fraenkel, N., & Dalton, P. (2013). Perceptual load does not modulate auditory distractor processing. *Cognition*, *129*(2), 345-355.
- Murphy, S., Spence, C., & Dalton, P. (2017). Auditory perceptual load: A review. *Hearing Research*, *352*, 40-48.
- Nager, W., Kohlmetz, C., Altenmüller, E., Rodriguez-Fornells, A., & Münste, T. F. (2003). The fate of sounds in conductors' brains: an ERP study. *Cognitive Brain Research*, *17*(1), 83-93.
- Navon, D. (1989). Attentional selection: Early, late, or neither?. *European Journal of Cognitive Psychology*, *1*(1), 47-68.
- Neelon, M. F., Williams, J., & Garell, P. C. (2011). Elastic attention: enhanced, then sharpened response to auditory input as attentional load increases. *Frontiers in Human Neuroscience*, *5*, 41.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, *13*(2), 87–110.

Niarchou, M., Gustavson, D. E., Sathirapongsasuti, J. F., Anglada-Tort, M., Eising, E., Bell, E., McArthur, E., Straub, P., Aslibekyan, S., Auton, A., Bell, R. K., Bryc, K., Clark, S. K., Elson, S. L., Fletez-Brant, K., Fontanillas, P., Furlotte, N. A., Gandhi, P. M., Heilbron, K., . . . Gordon, R. L. (2022). Genome-wide association study of musical beat synchronization demonstrates high polygenicity. *Nature Human Behaviour*, *6*(9), 1292–1309.

Nicholls, M. E., Ellis, B. E., Clement, J. G., & Yoshino, M. (2004). Detecting hemifacial asymmetries in emotional expression with three-dimensional computerized image analysis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1540), 663-668.

Nicholls, M., Orr, C., & Lindell, A. (2005). Magical ideation and its relation to lateral preference. *Laterality: Asymmetries of Body, Brain, and Cognition*, *10*(6), 503-515.

Nozaradan, S. (2013). *Exploring the neural entrainment to musical rhythms and meter: A steady-state evoked potential approach*. Universite de Montreal (Canada).

O'Neill, S. A. (2002). The self-identity of young musicians, in MacDonald, R. A., Hargreaves, D. J., & Miell, D. (Eds.). (2002). *Musical identities*. OUP Oxford. p. 79-96.

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., & Lalor, E. C. (2014). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cerebral Cortex*, *25*(7), 1697–1706.

Ocelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in front and rear space. *Neuroscience & Biobehavioral Reviews*, *35*(3), 589-598.

- Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C. A., & James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cerebral cortex*, *23*(9), 2213-2224.
- Okada, B. M. (2016). *Musical training and executive functions* (Master's thesis, University of Maryland, College Park).
- Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings* (Doctoral dissertation, The Ohio State University).
- Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), B13-B23.
- Ouimet, T., Foster, N. E., & Hyde, K. L. (2012). Auditory global-local processing: Effects of attention and musical experience. *The Journal of the Acoustical Society of America*, *132*(4), 2536-2544.
- Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music perception*, *26*(5), 489-504.
- Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PloS one*, *5*(6), e11120.
- Palmer, C., Spidle, F., Koopmans, E., & Schubert, P. (2013). Temporal coordination in vocal duet performances of musical rounds. In *Proceedings of the Stockholm music acoustics conference* (pp. 678-682). Stockholm: KTH Royal Institute of Technology.
- Palmer, S. E. (1992). Modern theories of Gestalt perception. In G. W. Humphreys (Ed.), *Understanding vision: An interdisciplinary perspective* (pp. 39-70). Blackwell Publishing.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392(6678), 811-814.

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience*, 29(45), 14100-14107.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, 6(7), 674-681.

Pato, M. V., Jones, S. J., Perez, N., & Sprague, L. (2002). Mismatch negativity to single and multiple pitch-deviant tones in regular and pseudo-random complex tone sequences. *Clinical neurophysiology*, 113(4), 519-527.

Patston, L. L., Corballis, M. C., Hogg, S. L., & Tippett, L. J. (2006). The neglect of musicians: Line bisection reveals an opposite bias. *Psychological Science*, 17(12), 1029-1031.

Patston, L. L., Hogg, S. L., & Tippett, L. J. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality*, 12(3), 262-272.

Pearsons, K. S., Bennett, R. L., & Fidell, S. A. (1977). *Speech levels in various noise environments*. Office of Health and Ecological Effects, Office of Research and Development, US EPA.

Pecenka, N., & Keller, P. E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, 1169(1), 282-286.

Penhune, V. B. (2011). Sensitive periods in human development: evidence from musical training. *cortex*, 47(9), 1126-1137.

Perlovsky, L. (2015). Origin of music and embodied cognition. *Frontiers in Psychology*, 6, 538.

Perrenoud, M., & Bataille, P. (2017). Artist, craftsman, teacher: “being a musician” in France and Switzerland. *Popular Music and Society*, 40(5), 592-604.

Perrot, X., Ryvlin, P., Isnard, J., GuÉnot, M., Catenoix, H., Fischer, C., Maugulere, F. & Collet, L. (2006). Evidence for corticofugal modulation of peripheral auditory activity in humans. *Cerebral Cortex*, 16(7), 941-948.

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual review of neuroscience*, 35(1), 73-89.

Pfordresher, P. Q. (2012). Musical training and the role of auditory feedback during performance. *Annals of the New York Academy of Sciences*, 1252(1), 171-178.

Pinker, S., & Downing, C. J. (1985). The spatial structure of visual attention. In M. Posner and O. Marin (Eds.), *Attention and Performance XI: Mechanisms of attention and visual search*. Hillsdale, NJ: Erlbaum.

Popov, T., Gips, B., Weisz, N., & Jensen, O. (2023). Brain areas associated with visual spatial attention display topographic organization during auditory spatial attention. *Cerebral Cortex*, 33(7), 3478-3489.

Porzionato, G. (1980). *Psicobiologia della musica*. Pàtron.

Posedel, J., Emery, L., Souza, B., & Fountain, C. (2012). Pitch perception, working memory, and second-language phonological production. *Psychology of Music*, 40(4), 508-517.

Posner, M. I. (1978). *Chronometric explorations of mind*. Lawrence Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, 32(1), 3-25.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, 78, 391-408.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bowhuis (Eds.), *Attention and performance X* (pp. 531–556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: an information-processing account of its origins and significance. *Psychological review*, 83(2), 157.

Poulton, E. C. (1953). Two-channel listening. *Journal of Experimental Psychology*, 46(2), 91.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music perception*, 2(4), 411-440.

Previc, F. H., Liotti, M., Blakemore, C., Beer, J., & Fox, P. (2000). Functional imaging of brain areas involved in the processing of coherent and incoherent wide field-of-view visual motion. *Experimental Brain Research*, 131, 393-405.

Prinzmetal, W., Zvinyatskovskiy, A., Gutierrez, P., & Dilem, L. (2009). Voluntary and involuntary attention have different consequences: The effect of perceptual difficulty. *Quarterly Journal of Experimental Psychology*, 62(2), 352-369.

Proverbio, A. M. (2019). *Neuroscienze cognitive della musica. Il cervello musicale tra arte e scienza*. Zanichelli.

Rauscher, F. H., & Hinton, S. C. (2011). Music instruction and its diverse extra-musical benefits. *Music perception*, 29(2), 215-226.

Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, 1(1), 0007.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink?. *Journal of experimental psychology: Human perception and performance*, 18(3), 849.

Rayner, K., & Pollatsek, A. (1997). Eye movements, the eye-hand span, and the perceptual span during sight-reading of music. *Current Directions in Psychological Science*, 6(2), 49-53.

Remington, R., & Pierce, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception & Psychophysics*, 35, 393-399.

Rhodes, G. (1987). Auditory attention and the representation of spatial information. *Perception & Psychophysics*, 42(1), 1-14.

Ribner, A. D., Willoughby, M. T., Blair, C. B., & Family Life Project Key Investigators. (2017). Executive function buffers the association between early math and later academic skills. *Frontiers in psychology*, 8, 869.

Rickard, N. S., & Chin, T. (2017). Defining the musical identity of “non-musicians.” In MacDonald, R., Hargreaves, D. J., & Miell, D. (Eds.). (2017). *Handbook of musical identities*. Oxford University Press.

Rivenez, M., Darwin, C. J., & Guillaume, A. (2006). Processing unattended speech. *The journal of the Acoustical Society of America*, 119(6), 4027-4040.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25(1), 31-40.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 35(6), 747-758.

Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychology of Music*, 42(2), 284-298.

- Rodrigues, A. C., Guerra, L., & Loureiro, M. (2007, August). Visual attention in musicians and non-musicians: a comparative study. In *Proceedings of the 3rd International Conference on Interdisciplinary Musicology* (pp. 15-19).
- Rodrigues, A. C., Loureiro, M. A., & Caramelli, P. (2013). Long-term musical training may improve different forms of visual attention ability. *Brain and cognition*, 82(3), 229-235.
- Rogers, W.L., Bregman, A.S. An experimental evaluation of three theories of auditory stream segregation. *Perception & Psychophysics* 53, 179–189 (1993)
- Romero-Díaz, A., Peñaloza-López, Y., García-Pedroza, F., Pérez, S. J., & Camacho, W. C. (2011). Central auditory processes evaluated with psychoacoustic tests in normal children. *Acta Otorrinolaringologica (English Edition)*, 62(6), 418-424.
- Rorden, C., & Driver, J. (2001). Spatial deployment of attention within and across hemifields in an auditory task. *Experimental Brain Research*, 137, 487-496.
- Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, 109(4), 444–474
- Rosenbaum, D. A., & Kornblum, S. (1982). A priming method for investigating the selection of motor responses. *Acta Psychologica*, 51(3), 223-243.
- Rosenbaum, D. A., Inhoff, A. W., & Gordon, A. M. (1984). Choosing between movement sequences: a hierarchical editor model. *Journal of Experimental Psychology: General*, 113(3), 372.
- Rosenblum, L. D., Wuestefeld, A. P., & Saldaña, H. M. (1993). Auditory Looming Perception: Influences on anticipatory judgments. *Perception*, 22(12), 1467–1482.

- Rotaru, I., Geirnaert, S., Heintz, N., Van de Ryck, I., Bertrand, A., & Francart, T. (2024). What are we really decoding? Unveiling biases in EEG-based decoding of the spatial focus of auditory attention. *Journal of Neural Engineering*, *21*(1), 016017.
- Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PloS one*, *9*(1), e86980.
- Saarikivi, K., Putkinen, V., Tervaniemi, M., & Huotilainen, M. (2016). Cognitive flexibility modulates maturation and music-training-related changes in neural sound discrimination. *European Journal of Neuroscience*, *44*(2), 1815-1825.
- Sach, A. J., Hill, N. I., & Bailey, P. J. (2000). Auditory spatial attention using interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 717.
- Salimpoor, V. N., & Zatorre, R. J. (2013). Neural interactions that give rise to musical pleasure. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 62.
- Santangelo, V., Fagioli, S., & Macaluso, E. (2010). The costs of monitoring simultaneously two sensory modalities decrease when dividing attention in space. *Neuroimage*, *49*(3), 2717-2727.
- Sares, A. G., Foster, N. E., Allen, K., & Hyde, K. L. (2018). Pitch and time processing in speech and tones: The effects of musical training and attention. *Journal of Speech, Language, and Hearing Research*, *61*(3), 496-509.
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, *112*(29), 8987-8992.
- Schafer, R. M. (1977). *The Tuning of the World*. New York: Knopf.

Scharf, B. (1998). Auditory attention: The psychoacoustical approach. In H. Pashler (Ed.), *Attention* (pp. 75–117). East Sussex, UK: Psychology Press.

Scharf, B., Quigley, S., Aoki, C., Peachey, N., & Reeves, A. (1987). Focused auditory attention and frequency selectivity. *Perception & psychophysics*, *42*, 215-223.

Scheich, H., Baumgart, F., Gaschler-Markefski, B., Tegeler, C., Tempelmann, C., Heinze, H. J., Schindler, F., & Stiller, D. (1998). Functional magnetic resonance imaging of a human auditory cortex area involved in foreground–background decomposition. *European Journal of Neuroscience*, *10*(2), 803–809.

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological science*, *15*(8), 511-514.

Schellenberg, E. G. (2011a). Examining the association between music lessons and intelligence. *British journal of psychology*, *102*(3), 283-302.

Schellenberg, E., G. (2011b). Music lessons, emotional intelligence, and IQ. *Music Perception*, *29*(2), 185-194.

Schellenberg, E. G. (2015). Music training and speech perception: a gene–environment interaction. *Annals of the New York Academy of Sciences*, *1337*(1), 170-177.

Scheuerman, J., Venable, K. B., Anderson, M. T., & Golob, E. J. (2018). Modeling spatial auditory attention in ACT-R: a constraint-based approach. *Procedia Computer Science*, *145*, 797-804.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*(8), 1047-1055.

Schnupp, J., Nelken, I., & King, A. (2011). *Auditory neuroscience: Making sense of sound*. MIT press.

- Schröger, E., & Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, *35*(6), 755-759.
- Seashore, C. E., Lewis, D., & Saetveit, J. G. (1956). *Seashore measures of musical talents*. Psychological Corp.
- Seinfeld, S., Figueroa, H., Ortiz-Gil, J., & Sanchez-Vives, M. V. (2013). Effects of music learning and piano practice on cognitive function, mood and quality of life in older adults. *Frontiers in psychology*, *4*, 810.
- Seppänen, M., Hämäläinen, J., Pesonen, A. K., & Tervaniemi, M. (2012). Music training enhances rapid neural plasticity of N1 and P2 source activation for unattended sounds. *Frontiers in human Neuroscience*, *6*, 43.
- Serafine, M. L., Davidson, J., Crowder, R. G., & Repp, B. H. (1986). On the nature of melody-text integration in memory for songs. *Journal of memory and language*, *25*(2), 123-135.
- Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport*, *16*(16), 1781-1785.
- Shaw, B. K., McGowan, R. S., & Turvey, M. T. (1991). An acoustic variable specifying time-to-contact. *Ecological Psychology*, *3*(3), 253–261.
- Shen, Y., Lin, Y., Liu, S., Fang, L., & Liu, G. (2019). Sustained effect of music training on the enhancement of executive function in preschool children. *Frontiers in Psychology*, *10*, 1910.
- Sheppard, A., & Broughton, M. C. (2020). Promoting wellbeing and health through active participation in music and dance: a systematic review. *International journal of qualitative studies on health and well-being*, *15*(1), 1732526.
- Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in cognitive sciences*, *12*(5), 182-186.

Shomstein, S., & Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *Journal of Neuroscience*, *26*(2), 435-439.

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural brain research*, *291*, 244-252.

Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: A case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, *22*, 637-652.

Slevc, L. R., Davey, N. S., Buschkuehl, M., & Jaeggi, S. M. (2016). Tuning the mind: Exploring the connections between musical ability and executive functions. *Cognition*, *152*, 199-211.

Sloboda, J. (2005). *Exploring the musical mind: Cognition, emotion, ability, function*. Oxford University Press.

Sloboda, J. A. (1988). *Generative processes in music: The psychology of performance, improvisation, and composition*. Clarendon Press/Oxford University Press.

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of experimental psychology: Human perception and performance*, *20*(3), 555.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of experimental psychology: Human perception and performance*, *22*(4), 1005.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & psychophysics*, *59*(1), 1-22.

- Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford University Press.
- Spence, C., & Santangelo, V. (2009). Capturing spatial attention with multisensory cues: A review. *Hearing research*, 258(1-2), 134-142.
- Spence, C., Lee, J., & Van der Stoep, N. (2020). Responding to sounds from unseen locations: crossmodal attentional orienting in response to sounds presented from the rear. *European Journal of Neuroscience*, 51(5), 1137-1150.
- Sridharan, D., Levitin, D. J., Chafe, C. H., Berger, J., & Menon, V. (2007). Neural Dynamics of Event Segmentation in Music: Converging evidence for dissociable ventral and dorsal networks. *Neuron*, 55(3), 521–532.
- Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: evidence for a sensitive period. *Journal of Neuroscience*, 33(3), 1282-1290.
- Stevens, S. S., & Davis, H. (1938). *Hearing: its psychology and physiology*. Wiley.
- Störmer, V. S., McDonald, J. J., & Hillyard, S. A. (2009). Cross-modal cueing of attention alters appearance and early cortical processing of visual stimuli. *Proceedings of the National Academy of Sciences*, 106(52), 22456-22461.
- Strait, D. L., Chan, K., Ashley, R., & Kraus, N. (2012). Specialization among the specialized: auditory brainstem function is tuned in to timbre. *cortex*, 48(3), 360-362.
- Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hearing research*, 261(1-2), 22-29.
- Strait, D., & Kraus, N. (2011). Playing music for a smarter ear: cognitive, perceptual and neurobiological evidence. *Music perception*, 29(2), 133-146.

Strasburger, H., & Pöppel, E. (2002). Visual field. *Encyclopedia of neuroscience*, 2127-2129.

Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme perception in a foreign language. *Psychonomic Bulletin & Review*, 24, 1929-1934.

Swaminathan, S., & Schellenberg, E. G., (2018). Music Training and Cognitive Abilities: Associations, Causes, and Consequences, in Michael H. Thaut, and Donald A. Hodges (eds), *The Oxford Handbook of Music and the Brain*, Oxford Library of Psychology.

Symonds, R. M., Zhou, J. W., Cole, S. L., Brace, K. M., & Sussman, E. S. (2020). Cognitive resources are distributed among the entire auditory landscape in auditory scene analysis. *Psychophysiology*, 57(2), e13487.

Tafuri, J. (2007). *Nascere musicali: percorsi per educatori e genitori* (Vol. 19). EDT srl.

Tafuri, J., & Villa, D. (2002). Musical elements in the vocalisations of infants aged 2–8 months. *British Journal of Music Education*, 19(1), 73-88.

Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. (2010). Embodied auditory perception: the emotional impact of approaching and receding sound sources. *Emotion*, 10(2), 216.

Talamini, F., Altoè, G., Carretti, B., & Grassi, M. (2017). Musicians have better memory than nonmusicians: A meta-analysis. *PloS one*, 12(10), e0186773.

Talamini, F., Grassi, M., Toffalini, E., Santoni, R., & Carretti, B. (2018). Learning a second language: Can music aptitude or music training have a role?. *Learning and Individual Differences*, 64, 1-7.

Talamini, F., Vigl, J., Doerr, E., Grassi, M., & Carretti, B. (2023). Auditory and visual mental imagery in musicians and non-musicians. *Musicae Scientiae*, 27(2), 428-441.

- Tan, Y. T., McPherson, G. E., Peretz, I., Berkovic, S. F., & Wilson, S. J. (2014). The genetic basis of music ability. *Frontiers in psychology*, *5*, 658.
- Teder-Sälejärvi, W. A., Hillyard, S. A., Röder, B., & Neville, H. J. (1999). Spatial attention to central and peripheral auditory stimuli as indexed by event-related potentials. *Cognitive Brain Research*, *8*(3), 213-227.
- Teraoka, R., Hayashida, Y., & Teramoto, W. (2023). Difference in auditory time-to-contact estimation between the rear and other directions. *Acoustical Science and Technology*, *44*(2), 77-83.
- Teraoka, R., Sakamoto, S., Cui, Z., Suzuki, Y., & Shioiri, S. (2020). Temporal characteristics of auditory spatial attention on word intelligibility. *Acoustical Science and Technology*, *41*(1), 394-395.
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental brain research*, *161*, 1-10.
- Tervaniemi, M., Kruck, S., De Baene, W., Schröger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience*, *30*(8), 1636-1642.
- Tervaniemi, M., Tao, S., & Huotilainen, M. (2018, October). Promises of music in education?. In *Frontiers in Education* (Vol. 3, p. 74). Frontiers Media SA.
- Theeuwes, J. (2014). Spatial orienting and attentional capture. In A. C. Nobre & S. Kastner (Eds.), *The Oxford handbook of attention* (pp. 231–252). Oxford University Press.
- Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help?. *Emotion*, *4*(1), 46.

Thomson, J. M., Leong, V., & Goswami, U. (2013). Auditory processing interventions and developmental dyslexia: A comparison of phonemic and rhythmic approaches. *Reading and Writing, 26*, 139-161.

Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus, N. (2013). High school music classes enhance the neural processing of speech. *Frontiers in psychology, 4*, 855.

Timmers, R., & Li, S. (2016). Representation of pitch in horizontal space and its dependence on musical and instrumental experience. *Psychomusicology: Music, Mind, and Brain, 26*(2), 139.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology Section A, 37*(4), 571-590.

Tipper, S. P., & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. *The Quarterly Journal of Experimental Psychology, 37*(4), 591-611.

Tipper, S. P., & Driver, J. (1988). Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. *Memory & Cognition, 16*, 64-70.

Tomatis, A. (2015). *L'orecchio e la vita*. Xenia.

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1664), 20140089.

Travis, F., Harung, H. S., & Lagrosen, Y. (2011). Moral development, executive functioning, peak experiences and brain patterns in professional and amateur classical musicians: Interpreted in light of a Unified Theory of Performance. *Consciousness and cognition, 20*(4), 1256-1264.

- Treisman, A. M. (1960). Contextual Cues in Selective Listening. *Quarterly Journal of Experimental Psychology*, 12(4), 242-248.
- Treisman, A. M. (1964). The effect of irrelevant material on the efficiency of selective listening. *The American journal of psychology*, 77(4), 533-546.
- Tucker, I. (2010). The potentiality of bodies. *Theory & Psychology*, 20(4), 511-527.
- Tune, S., Wöstmann, M., & Obleser, J. (2018). Probing the limits of alpha power lateralisation as a neural marker of selective attention in middle-aged and older listeners. *European Journal of Neuroscience*, 48(7), 2537-2550.
- Ucchino Produzioni Elettroniche.(n.d.). *Sirena bitonale italiana mod. 2s2 de luxe, scheda tecnica*. <https://www.ucchino.net/wp-content/uploads/2018/06/Manuale-ITA.pdf>
- van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *Journal of Neuroscience*, 32(30), 10408-10412.
- van Vugt, F. T., Hartmann, K., Altenmüller, E., Mohammadi, B., & Margulies, D. S. (2021). The impact of early musical training on striatal functional connectivity. *NeuroImage*, 238, 118251.
- van't Hullenaar, C. D., Hermans, B., & Broeders, I. A. (2017). Ergonomic assessment of the da Vinci console in robot-assisted surgery. *Innovative surgical sciences*, 2(2), 97-104.
- VandenBos, G. R. (Ed.). (2007). *APA Dictionary of Psychology*. American Psychological Association.
- Váradi, J. (2022). A review of the literature on the relationship of music education to the development of socio-emotional learning. *Sage Open*, 12(1), 21582440211068501.

Vaughn, K., & Winner, E. (2000). SAT scores of students who study the arts: What we can and cannot conclude about the association. *Journal of Aesthetic Education*, 34(3/4), 77-89.

Vecera, S. P., & Rizzo, M. (2003). Spatial attention: Normal processes and their breakdown. *Neurologic Clinics*, 21(3), 575–607.

Ventura, J. A. (2018). *Predicting Working Memory and Fluid Intelligence from Measures of Musicality*. LSU Master's Theses. 4755.

Vernon, P. A. (1993). Der Zahlen-Verbindungs-Test and other trail-making correlates of general intelligence. *Personality and Individual Differences*, 14(1), 35-40.

Vuilleumier, P., & Trost, W. (2015). Music and emotions: from enchantment to entrainment. *Annals of the New York Academy of Sciences*, 1337(1), 212-222.

Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., & Tervaniemi, M. (2012). The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia*, 50(7), 1432-1443.

Vuust, P., Wallentin, M., Mouridsen, K., Østergaard, L., & Roepstorff, A. (2011). Tapping polyrhythms in music activates language areas. *Neuroscience letters*, 494(3), 211-216.

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, 20(3), 188-196.

Wan, C. Y., & Schlaug, G. (2010). Neural pathways for language in autism: the potential for music-based treatments. *Future neurology*, 5(6), 797-805.

Wang, S., & Agius, M. (2018). The neuroscience of music; a review and summary. *Psychiatria Danubina*, 30(suppl. 7), 588-594.

Wertheimer, M. (1938). Gestalt theory. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 1–11). Kegan Paul, Trench, Trubner & Company.

White, E. J., Hutka, S. A., Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Frontiers in systems neuroscience*, 7, 90.

Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free-field listening. I: Stimulus synthesis. *The Journal of the Acoustical Society of America*, 85(2), 858-867.

Wilsch, A., Mercier, M. R., Obleser, J., Schroeder, C. E., & Haegens, S. (2020). Spatial attention and temporal expectation exert differential effects on visual and auditory discrimination. *Journal of Cognitive Neuroscience*, 32(8), 1562-1576.

Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: merging auditory and visual worlds. *Neuron*, 48(3), 489-496.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and clinical neurophysiology*, 79(3), 170-191.

Woldorff, M. G., Hillyard, S. A., Gallen, C. C., Hampson, S. R., & Bloom, F. E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology*, 35(3), 283-292.

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature neuroscience*, 10(4), 420-422.

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1), 255.

Woods, W. S., & Colburn, H. S. (1992). Test of a model of auditory object formation using intensity and interaural time difference discrimination. *The Journal of the Acoustical Society of America*, *91*(5), 2894-2902.

Wöstmann, M., Waschke, L., & Obleser, J. (2019). Prestimulus neural alpha power predicts confidence in discriminating identical auditory stimuli. *European Journal of Neuroscience*, *49*(1), 94-105.

Wright, R. D., & Ward, L. M. (2008). *Orienting of attention*. Oxford university press.

Xiang, J. (2008). *Hearing vs. listening: Attention changes the neural representations of auditory percepts*. University of Maryland, College Park.

Yost, W. A., & Killion, M. C. (1997). Hearing thresholds. *Encyclopedia of acoustics*, *3*, 1545-1554.

Yurgil, K. A., & Golob, E. J. (2010). Neural activity before and after conscious perception in dichotic listening. *Neuropsychologia*, *48*(10), 2952-2958.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature reviews neuroscience*, *8*(7), 547-558.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature protocols*, *1*(1), 297-301.

Zerbetto, R. (1998). *La gestalt. Terapia della consapevolezza*, Pavia, Xenia.

Zhang J. D., Susino M., McPherson G., Schubert E. (2018). Using the three-component model of the musician definition, a musician is someone who has six years of musical expertise. In Parncutt R., Sattmann S. (Eds.), *Proceedings of ICMPC15/ESCOM10* (pp. 503-505). Graz, Austria: Centre for Systematic Musicology, University of Graz.

Zhang, C., Tao, R., & Zhao, H. (2019). Auditory spatial attention modulates the unmasking effect of perceptual separation in a “cocktail party” environment. *Neuropsychologia*, *124*, 108-116.

Zhang, J. D., & Schubert, E. (2019). A single item measure for identifying musician and nonmusician categories based on measures of musical sophistication. *Music Perception: An Interdisciplinary Journal*, *36*(5), 457-467.

Zhang, J. D., Susino, M., McPherson, G. E., & Schubert, E. (2020). The definition of a musician in music psychology: A literature review and the six-year rule. *Psychology of Music*, *48*(3), 389-409.

Zhang, P. X., & Hartmann, W. M. (2010). On the ability of human listeners to distinguish between front and back. *Hearing research*, *260*(1-2), 30-46.

Zuanazzi, A., & Noppeney, U. (2020). The intricate interplay of spatial attention and expectation: a multisensory perspective. *Multisensory Research*, *33*(4-5), 383-416.

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PloS one*, *9*(6), e99868.