UNIVERSITÀ DI PADOVA

Dipartimento di Psicologia Generale

Corso di laurea in Scienze cognitive psicologiche e psicobiologiche

Tesi di laurea triennale

Relazione tra fusione tonale e intervalli consonanti o dissonanti nella cultura occidentale

Relationship between tonal fusion and consonant or dissonant intervals in western culture

Relatore
Prof. Massimo Grassi

Laureando: Emanuele Corvino
Matricola: 2115813

Anno Accademico 2025/2026

INDICE

INTRODUZIONE 1

CAPITOLO 1 – LA CONSONANZA MUSICALE 3

1.1 I suoni e la musica 3
1.2 La percezione uditiva 4
1.3 Consonanza e dissonanza 6
1.4 Il bello come concezione geometrica 8
1.5 L'introduzione del dissonante 9
1.6 L'emancipazione della dissonanza di Schönberg 10
1.7 Il culturalismo e la musica tribale 11
1.8 La fusione tonale di Stumpf 12
1.9 L'ipotesi di ricerca 13

CAPITOLO 2 – IL METODO 14

2.1 I partecipanti 14
2.2 Il piano di ricerca 14
2.3 Gli strumenti 15
2.4 La procedura 17

CAPITOLO 3 – ANALISI DEI DATI 20

CAPITOLO 4 – CONCLUSIONI 21

4.1 Risultati 23
4.2 Limiti e sviluppi futuri 24

BIBLIOGRAFIA 26

INTRODUZIONE

La percezione della musica è un concetto insito nell'essere umano fin dall'alba dei tempi. Negli anni si sono succedute varie teorie riguardo la natura di questo fenomeno. Nello specifico, l'esperienza psicologica e fenomenologica legata all'ascolto di suoni dipende da due concetti distinti e opposti: la consonanza e la dissonanza. Il primo approccio che ha cercato di spiegare la natura di questi fenomeni deriva da Pitagora, con cui inizia il naturalismo, di base innatista. Contrapposto ad esso troviamo invece il culturalismo, che guarda alla storia e cultura delle varie società per spiegare la piacevolezza o meno di un intervallo specifico. Al crocevia tra la dimensione biologica innatista del naturalismo e l'esperienza psicologica soggettiva del culturalismo si trova la teoria della fusione tonale di Carl Stumpf. Stumpf teorizzò che la consonanza non derivasse da meri calcoli matematici, ma dalla qualità dell'esperienza sensoriale. Esiste, per questa teoria, una tendenza psicologica innata a percepire due o più suoni simultanei come un'unica entità sonora integrata. Questa tendenza, per Stumpf, è il grado di fusione per il quale percepiremo un intervallo come più o meno consonante.

Il presente lavoro di tesi si propone di indagare empiricamente la validità della teoria di Stumpf indagando il legame tra fusione tonale, tempi di elaborazione cognitiva e giudizi soggettivi di piacevolezza e fusione all'interno di un campione costituito interamente da partecipanti di cultura occidentale. Il progetto sperimentale adotta la metodologia di studi precedenti (nello specifico DeWitt & Crowder, 1987; McPherson et al., 2020) combinando la misurazione oggettiva di tempi di reazione con valutazioni soggettive espresse su scala Likert a 4 punti.

L'elaborato si sviluppa in quattro capitoli principali.

Il capitolo 1 offre una rassegna del fenomeno partendo dall'anatomia e fisiologia della percezione uditiva per poi percorrere l'evoluzione storica dei concetti di consonanza e dissonanza fino ad arrivare alla teoria della fusione tonale di Carl Stumpf. Il capitolo 2 descrive il metodo adottato e il piano di ricerca, spiegando nei dettagli le caratteristiche del campione utilizzato, gli strumenti adottati, gli stimoli sonori creati per la sperimentazione. Il

capitolo 3 presenta un'analisi dei dati raccolti esplicitando medie e deviazioni standard ottenute nei vari compiti dell'esperimento. Il capitolo 4, infine, discute i risultati emersi sottolineando come questi siano in linea con la teoria di partenza. Vengono inoltre messi in luce anche i limiti metodologici offrendo spunti di riflessione per sviluppi futuri nell'ambito della psicologia della musica.

Capitolo 1 – LA CONSONANZA MUSICALE

1.1 I suoni e la musica

L'enciclopedia Treccani definisce "musica" come "l'arte che consiste nell'ideare e nel produrre successioni strutturate di suoni semplici o complessi, che possono variare per altezza, per intensità e per timbro, per mezzo della voce umana, di strumenti o della combinazione di entrambe queste fonti". Un "suono" consiste in variazioni udibili della pressione dell'aria. Ogni suono può avere altezza e intensità variabili. L'altezza è la frequenza delle vibrazioni del corpo sonoro. L'altezza si misura in Hertz (Hz). L'orecchio umano non percepisce tutte le frequenze sonore. I suoni udibili dall'essere umano vanno dai 20 ai 20.000 Hz. Sotto i 20 Hz parliamo di infrasuoni. Sopra i 20.000 Hz parliamo di ultrasuoni. In base all'ampiezza, quindi a frequenze alte o basse, dividiamo i suoni rispettivamente in suoni alti o acuti e in suoni bassi o gravi. L'intensità determina il volume di un suono. È la forza della vibrazione dell'onda. In base all'intensità dividiamo i suoni in suoni forti, aventi un alto volume, e suoni deboli, aventi un basso volume. L'intensità si misura in Decibel (dB). Un'intensità di 0 dB corrisponde alla soglia dell'udito dell'essere umano. Un'intensità tra i 120 e i 130 dB corrisponde alla soglia del dolore dell'essere umano.

La musica è un concetto che appartiene all'essere umano fin dall'alba dei tempi. Le prime forme di musica erano costituite principalmente da voci e percussioni corporee. Per la voce, si fa riferimento alla *musilanguage*, ovvero l'utilizzo di variazioni di intonazione, ritmo e canti per comunicare un pericolo. In aggiunta i primi ominidi battevano i piedi sul terreno o le mani sul petto per rituali o momenti comunitari. Con i primi ritrovi archeologici, risalenti al paleolitico, abbiamo i primi strumenti musicali, come i flauti d'osso. Questi erano ossa cave di grandi animali in cui venivano praticati dei fori per riprodurre suoni. Charles Darwin, ne "L'origine dell'uomo e la selezione in relazione al sesso" (1871/2011) identifica come ragione principale dell'approccio alla musica, la selezione sessuale e il corteggiamento: "Presumiate pure che gli accordi musicali e i ritmi siano stati acquisiti per la prima volta dai

progenitori maschi o femmine dell'umanità allo scopo di affascinare il sesso opposto." Negli anni, antropologi e psicologi evoluzionisti, hanno collegato l'utilizzo primordiale della musica anche a coesione sociale e ritualità.

In ogni caso troviamo sempre un significato emotivo e sociale associato al fenomeno. Già con il modello tridimensionale delle emozioni di Wundt (1897/1896), i tre assi di esperienza psicologica interna sono correlati a parti diverse della musica. L'asse piacere/dispiacere viene legato ai fenomeni di consonanza/dissonanza. L'asse tensione/rilassamento viene legato al ritmo di un metronomo, lento/veloce. L'asse eccitazione/depressione viene legato all'intensità e alla velocità del suono, alta e incalzante/lenta e monotono. Questi tre assi sono proprio stati studiati utilizzando suoni e melodie, esplicitando quindi anche una forte correlazione tra musica e esperienza emozionale. A livello psicofisiologico questa correlazione è stata resa ancora più chiara. Attraverso l'utilizzo della PET sono state registrate attivazioni neuronali in regioni paralimbiche e neocorticali in risposta all'ascolto di melodie consonanti/dissonanti, piacevoli/spiacevoli (Blood et al., 1999). In uno studio successivo, sempre mediante l'utilizzo di PET, l'ascolto di musica piacevole è stato correlato all'attivazione di regioni coinvolte in: ricompensa/motivazione, emozioni, arousal (Blood & Zatorre, 2001). Tra queste aree troviamo, specificatamente per le emozioni, l'amigdala, fondamentale per la percezione emozionale nell'essere umano.

1.2 La percezione uditiva

L'udito è il senso preposto alla percezione dei suoni, compresa la localizzazione e individuazione di questi. Nell'uomo il sistema uditivo è ciò che ci permette di percepire tali stimoli, trasformando le onde sonore in segnali nervosi, così che il cervello li possa interpretare. L'organo che specificatamente fa percepire vibrazioni sonore è l'orecchio. L'orecchio si divide in orecchio esterno, orecchio medio, orecchio interno. L'orecchio esterno comprende padiglione auricolare e canale uditivo; permette di raccogliere le onde sonore e convogliarle verso l'orecchio medio. L'orecchio medio amplifica le vibrazioni, trasmettendole all'orecchio interno. Una volta arrivati in corrispondenza della membrana

timpanica, infatti, la trasmissione sonora viene trasmessa a un complesso di ossicini composto da martello, incudine, staffa. Questi tre ossicini comunicanti si muovono in base a variazioni pressorie e amplificano la risposta, mandandola verso la finestra ovale, verso l'orecchio interno. I muscoli aventi un ruolo in questo processo sono il muscolo estensore del timpano, tra martello e parete ossea cranica dell'orecchio medio e muscolo stapedio, tra staffa e parete ossea del cranio. Grazie all'azione modulatrice di questi due muscoli, l'orecchio medio è in grado di proteggere l'orecchio da suoni prolungati a volume elevato, ma non da suoni impulsivi con variazioni improvvise di intensità. Questo riflesso protettivo è detto riflesso di attenuazione.

Nell'orecchio interno abbiamo la trasduzione di stimoli sonori in impulsi nervosi. La struttura principale in cui avviene questo processo è la coclea. La coclea, avente forma di spirale avvolta su se stessa, è divisa in tre camere contenenti liquido. Queste sono scala vestibolare, scala media, scala timpanica, separate tra loro da membrane. La membrana basilare, che separa scala media da scala timpanica, comprende l'organo di Corti, all'interno del quale troviamo i recettori uditivi per la trasmissione del segnale. Alla base della coclea la scala vestibolare trova la finestra ovale, mentre la scala timpanica trova la finestra rotonda. I liquidi contenuti all'interno delle scale sono perilinfa e endolinfa. La perilinfa, simile al liquor è contenuta nella scala vestibolare e timpanica e ha potenziale poco minore di 80 mV. L'endolinfa, simile al fluido intercellulare, è contenuta nella scala media e ha potenziale poco maggiore di 80 mV. Il processo di trasferimento delle onde sonore parte con l'arrivo del suono nell'orecchio interno, dove la staffa premerà sulla finestra ovale. La finestra ovale, così, spingerà la perilinfa nella scala vestibolare fino all'elicotrema, piccolo orifizio situato all'apice della coclea che collega scala vestibolare e scala timpanica. La perilinfa passa dunque per la scala timpanica, discendendo fino alla finestra rotonda. Il movimento della perilinfa farà vibrare la membrana basilare, le cui onde determineranno pressioni e decompressioni sulla membrana della finestra rotonda. In base alla frequenza del suono, l'onda si propagherà fino alla fine della coclea. Per frequenze alte sarà la base rigida della coclea a vibrare, dissipando la propria energia meccanica nelle prime zone della struttura cocleare. Per frequenze basse invece la propagazione è totale, arrivando fino all'apice. Ne

consegue la presenza di un codice topografico di frequenza del suono nella coclea in base alla flessione della membrana basilare: più si prosegue verso l'apice della coclea, più ci si avvicina a frequenze basse.

Nella membrana basilare, sarà l'organo di Corti a trasformare l'energia meccanica in arrivo in potenziale di recettore. L'organo di Corti comprende al suo interno cellule ciliate interne e cellule ciliate esterne. Ogni cellula ciliata contiene dalle 30 alle 100 stereocilia. Nella trasduzione, ogni canale del potassio K+ delle sterocilia è sempre parzialmente aperto quando la loro posizione è verticale. Quando però, con l'arrivo dell'energia meccanica, le stereocilia si sposteranno verso la stereocilia più lunga, tutti i canali verranno aperti completamente (entrerà potassio K+), portando a depolarizzazione della cellula ciliata, rilascio del neurotrasmettitore (NT) glutammato e eccitazione delle fibre del ganglio spinale. Quando invece le sterocilia si spostano verso la sterocilia più corta, i canali K+ verranno chiusi e si avrà iperpolarizzazione della cellula ciliata. I neuroni del ganglio spinale costituiscono, con i propri assoni, il nervo uditivo. Attraverso il nervo uditivo il segnale viene mandato dal nucleo cocleare ventrale all'oliva superiore, situata nel tronco encefalico, poco al di sopra del bulbo. Il segnale passa poi al collicolo inferiore nel mesencefalo, fino al nucleo genicolato mediale del talamo (NGM). Infine, il NGM invia l'informazione alla corteccia uditiva primaria (aree 41 e 42 di Broadmann). Nella corteccia uditiva primaria, così come nelle cellule ciliate e nei nuclei che portano alla corteccia, abbiamo suddivisioni specifiche in base a ogni frequenza. Nel caso dei neuroni, questi saranno ognuno specifico per determinate frequenze dette frequenze caratteristiche. Si parla quindi di tonotopia, con mappe tonotopiche in base alla frequenza caratteristica per ogni livello coinvolto nell'elaborazione uditiva (Bear et al., 2016).

1.3 Consonanza e dissonanza

Nella percezione di suoni, alcuni intervalli possono risultare più o meno piacevoli. Ad esempio, accordi come l'ottava o la quinta, fin dai tempi antichi sono sempre stati percepiti come più piacevoli, più stabili. Altri accordi invece, come la seconda minore o la settima

maggiore, sono percepiti più spiacevoli, meno costanti. Normativamente, i primi accordi vengono definiti consonanti; i secondi invece sono dissonanti. Il termine consonanza deriva dal latino *consonare*, ovvero suonare assieme. Per l'enciclopedia Treccani, indica la peculiarità di specifici intervalli e accordi di generare un effetto di affermazione, consenso, riposo, il tutto grazie a un'apparente fusione di più suoni in uno solo. Al contrario, il termine dissonanza indica la capacità di intervalli e accordi musicali di generare un effetto non gradevole. Riguarda una mancanza di armonia, una tensione o instabilità,

Negli anni ci si è chiesti se questa distinzione derivasse da qualità innate del sistema uditivo umano, con una preferenza per determinati intervalli o se invece tali giudizi derivassero dalla propria storia e cultura. Nascono così due approcci diversi nella risposta a questo quesito. Il naturalismo sostiene che consonanza e dissonanza abbiano basi oggettive, dipendenti dalla fisica del suono e dalla biologia del sistema uditivo. Si crede che siano gli intervalli tra armonici, frequenze secondarie generate subito dopo la produzione di una nota, ad essere più piacevoli per l'orecchio. Al contrario vengono etichettati come non piacevoli i battimenti. Un battimento è un fenomeno uditivo che si crea suonando due note molto vicine tra loro, generando una sensazione di attrito tra i suoni. L'idea di base è che l'orecchio interno, specificatamente la membrana basilare, faccia fatica a separare suoni non così distanti tra loro, portando a far percepire, a livello di corteccia, una sensazione di fastidio all'ascolto. Di risposta, il culturalismo sostiene come, seppur possa esistere una fisica dei suoni, è il significato culturale che viene dato a questi fenomeni ad essere alla base dei giudizi di piacevolezza. Esistono costruzioni storiche, geografiche e culturali che governano la nostra valutazione di suoni, per cui uno stesso intervallo potrà essere particolarmente spiacevole per una cultura e particolarmente piacevole per un'altra. Negli anni la visione sull'argomento si è spostata dal naturalismo al culturalismo, arrivando, ad oggi, ad una commistione tra le due ideologie.

1.4 Il bello come concezione geometrica

L'approccio naturalista individua come fondatore del proprio pensiero Pitagora di Samo. Nell'antica Grecia, infatti, il "bello" coincide con la perfezione geometrica, la simmetria. Si parla di *Cosmos*, ordine, da dover rispettare in quanto natura profonda dell'universo. La musica, come conseguenza, aveva il ruolo di riflettere quest'armonia matematica dell'universo. Platone stesso, dai racconti di Boezio (ca. 510/1989) con il mito del monocordo applica la sua simmetria geometrica per ottenere "intervalli musicali perfetti", le nostre consonanze. Pitagora prese una singola corda tesa e, dividendola in segmenti geometrici precisi, vide generarsi "intervalli musicali perfetti". Dividendo la corta a metà (rapporto 2:1) si otteneva l'ottava; dividendo la corda in 2/3 (rapporto 3:2) si otteneva la quinta; dividendo la corda in 3/4 (rapporto 4:3) si otteneva la quarta. I numeri 1, 2, 3, 4, per i pitagorici, formavano la *Tetraktys*, la decade sacra, la cui somma risulta nel numero 10, che rappresentava l'intero universo.

Successivamente, con Aristosseno di Taranto, nel IV a.C. ci si focalizza sulla percezione soggettiva dell'individuo, con l'orecchio "giudice" del bello, guidato dall'intelletto. Nei suoi *Elementi di armonica* (ca. 330 a.C./1954) Aristosseno identifica tre pilastri su cui si articola il bello musicale. Il primo è l'accordo tra senso e ragione: un suono è piacevole non perché corrispondente a frazioni matematiche precise, ma per come l'orecchio lo percepisca come coerente. Il secondo è la memoria e l'attesa: per riconoscere come bello un suono, l'uomo utilizza la memoria per ricordare il suono passato e l'intelletto per anticipare quello in arrivo. Il terzo è il movimento della voce: la voce o lo strumento dovrà produrre suoni che si possono muovere nello spazio secondo intervalli proporzionali che l'essere umano possa percepire come naturali. Aristosseno quindi si allontana dalla concezione pitagorica di rapporto matematico necessario per percepire piacevolezza. Ci si sofferma invece sull'esperienza soggettiva, pur tenendo a mente i concetti di razionalità e un contesto logico di partenza necessario.

In epoca medioevale, tuttavia, Boezio ritorna ad una visione strettamente pitagorica. La musica non è un'arte pratica, derivante da esperienza soggettiva, ma una scienza matematica facente parte del Quadrivio, con aritmetica, geometria, astronomia (Boezio, ca. 510/1989). Il

bello deve riflettere l'ordine di Dio nel cosmo. Come conseguenza la dissonanza viene definita come *Diabolus in Musica*. Specificatamente il tritono, ovvero la distanza di tre toni interi, in quanto non ottenibile con calcoli pitagorici della Tetraktys, viene associato al caos, all'inganno, al diavolo. La conseguenza diretta è un divieto, non formale, ma teorico dell'utilizzo di intervalli dissonanti, che venivano dunque considerati come errori tecnici o addirittura orrori acustici.

1.5 L'introduzione del dissonante

Nella seconda metà del IX secolo compaiono due trattati anonimi sull'argomento: la *Musica enchiriadis* e la *Scolica enchiriadis*. Pur partendo dalle idee di Boezio, questi due trattati affrontano un problema pratico che stava nascendo nei monasteri francesi: i monaci nei propri canti corali di chiesa iniziano ad utilizzare la polifonia. Nella Scolica (ca. 900/1995) vengono definite le regole dell'*organum*, la prima forma strutturata di polifonia occidentale. Troviamo una *vox principalis*, statica, ferma sulla melodia principale del testo sacro, e una *vox organalis*, accompagnamento polifonico che si muove a una distanza fissa. Nell'utilizzo di questa tecnica però era possibile l'utilizzo di intervalli "non perfetti", dissonanti, attraverso l'*occursus*. In questi casi, quando la melodia principale si avvicina al tritono, la vox organalis aveva l'obbligo di rimanere fissa su una nota consonante. Si venivano così a creare intervalli di seconda o terza, per Boezio dissonanti, ma che la Scolica ammette come passaggi temporali, transizione da risolvere in poco per ritornare sul consonante, specificatamente sull'unisono, l'accordo perfetto finale. La dissonanza viene quindi ammessa come tensione da controllare per ritornare verso il consonante.

Solo con il barocco la dissonanza inizia ad avere un'importanza come strumento comunicativo emotivo. Vengono utilizzati accordi dissonanti per esprimere dolore, tensione o pathos e diventano fondamentali per questo ruolo. Tuttavia, è necessario, anche in questo caso, che il tutto trovi una risoluzione sulla tonalità di partenza, così da dare un senso di ordine e stabilità. Un atteggiamento simile viene mantenuto anche nel classicismo. Scompaiono polifonie complesse del barocco e si ritorna ad accordi semplici che trasmettano

la razionalità illuminista. Tra consonanza e dissonanza c'è una rigida gerarchia. La prima regna sovrana, per mantenere stabilità nel pezzo. La seconda viene utilizzata solo per creare brevi momenti di tensione, da risolvere nell'immediato.

Durante il romanticismo, il focus si sposta sull'interiorità, l'espressione del proprio io sotto forma di musica. Il bello diviene quindi qualcosa che può commuovere e sconvolgere e, in questo, la dissonanza risulta essere uno strumento comunicativo fondamentale. Alla triade del classicismo vengono poste note aggiuntive, arrivando ad accordi di settima e di nona nonché accordi alterati alzati o abbassati di un semitono (seste aumentate, quinte diminuite) per creare tensione. Il tritono questa volta è utilizzato appieno, senza più esserne spaventati. La dissonanza non ha più la necessità di scaricarsi, può rimanere ininterrotta, senza la necessità pitagorica di risoluzione.

1.6 L'emancipazione della dissonanza di Schönberg

Arrivati nel 1900, con Schönberg si parla di vera e propria rivoluzione in ambito musicale. Schönberg non guarda ai concetti di dissonanza e consonanza come separati e inconciliabili; al contrario l'idea di base è che siano due versioni diverse di una stessa linea continua. Nel suo *Manuale di armonia* (Schönberg, 1911/2014), Schönberg guarda alla fisica del suono, specificatamente ai suoni armonici. Egli crede che le consonanze, intervalli formati dai primi armonici, siano stati usati maggiormente e considerati più piacevoli nella storia della musica unicamente perché più facili da percepire e da fondere assieme in quanto suoni. Al contrario le dissonanze, intervalli formati da armonici più lontani, saranno semplicemente più complessi, motivo per cui storicamente l'essere umano ha avuto bisogno di più tempo per familiarizzarci. Se però entrambi sono armonici, semplicemente uno più vicino e uno più lontano, allora anche la dissonanza avrà diritto di esistere tanto quanto la consonanza. Questa è quella che Schönberg definisce "emancipazione della dissonanza". La dissonanza ora non dovrà più essere tassativamente risolta. Essa potrà esistere in autonomia senza venire "sottomessa" cercando di tornare verso un accordo perfetto. Per Schönberg il giudizio di consonanza diviene un giudizio puramente storico e culturale, distaccandosi, per la prima

volta, da una visione naturalista, abbracciando al contrario l'approccio culturalista che non era mai stato esplorato così in profondità dai suoi predecessori. A dimostrazione di ciò si parla si un'evoluzione dell'orecchio umano dal Medioevo, specificatamente sul tritono, *Diabolus in Musica*, che ad oggi invece viene considerato tutt'altro che dissonante. Negli anni successivi, Schönberg espanderà questo concetto di emancipazione ad altri aspetti della musica, specificatamente all'atonalità e alla dodecafonia, sistema che pone sullo stesso piano di importanza ogni nota della scala cromatica, eliminando definitivamente le gerarchie armoniche passate. A seguire, grazie al *Manuale di armonia*, sempre più compositori andranno a sperimentare nelle proprie opere, utilizzando addirittura il rumore come strumento di comunicazione emotiva senza porsi alcun limite.

1.7 Il culturalismo e la musica tribale

Come descritto in precedenza, l'approccio culturalista vede i fenomeni di consonanza e dissonanza essere puramente di origine storica e culturale, specificatamente per quanto riguarda la piacevolezza percepita da un accordo. Risulta quindi interessante analizzare la musica e i concetti di consonanza e dissonanza in società non occidentali. I canti tribali balcanici sono uno degli esempi più appropriati in questo contesto. Il "volo della voce" è un fenomeno di canto tradizionale polifonico tipico dei Balcani. Sono presenti due voci: la prima terrà una nota fissa, come se fosse un ronzio di sottofondo; la seconda canterà la nota immediatamente successiva della scala di un tono o un semitono. Tutto ciò risulta in un intervallo di seconda minore, particolarmente dissonante, indicante tensione nella cultura occidentale. Nei Balcani invece, questo intervallo è considerato brillante, pieno di luce. A livello culturale, un canto privo di battimento è visto come debole, privo di valore tecnico. Il dissonante in questo caso viene cercato e elogiato attivamente.

Un secondo esempio arriva dalla musica tribale africana. I Pigmei BaAka dell'Africa Centrale fondano la propria tradizione musicale sul contrasto sonoro. Si parla di *Yelli* o *Limo*, attività di canto legata a caccia e coesione di gruppo. Si basa sulla tecnica dell'*Hocket*, polifonia a incastro con gli altri individui dimostrata essere derivante da rigide regole

compositive implicite (Arom, 1991). Ogni individuo canta un breve complesso di due o tre note, diverse da quelle cantate dagli altri, con un delay di qualche millisecondo. Ne conseguono incastri tra note che si sovrappongono a distanza di un tono o un semitono. L'effetto creato è quello di un complesso di seconde maggiori e minori, un cluster dissonante per la cultura occidentale, ma un tessuto sonoro tridimensionale fondamentale per la cultura dei Pigmei. Lo scopo di questi canti è il mimetismo ambientale. La foresta equatoriale in cui vivono i Pigmei non è mai silenziosa, sono sempre presenti fruscii e versi di uccelli, rane, altri animali. Intervalli così stretti e ronzanti permettono alle varie tribù di fondersi con il paesaggio della foresta. Solo in questo modo potranno essere in grado di dialogare con la natura.

1.8 La fusione tonale di Stumpf

Al fine di comprendere appieno i temi trattati nella ricerca è necessario considerare la teoria della fusione tonale di Carl Stumpf. Pubblicata durante la fine dell'Ottocento, *Psicologia dei toni* si prepone di spiegare i meccanismi puramente psicologici alla base della percezione di consonanza o dissonanza. La fusione tonale è il fenomeno per il quale due o più suoni presentati in contemporanea vengono percepiti come un suono unico anziché separati (Martinelli, 2014). Non si tratta di un'incapacità dell'orecchio di distinguere le note, ma di una qualità psicologica immediata dell'esperienza sensoriale. Maggiore è il grado di fusione, più l'intervallo sarà percepito come consonante. Minore è il grado di fusione, più l'intervallo sarà percepito come dissonante.

Stumpf stabilisce una classifica dei gradi di fusione degli intervalli musicali, basandosi su esperimenti empirici condotti sia su partecipanti musicisti che su partecipanti non musicisti. Questi ultimi soprattutto erano particolarmente rilevanti per Stumpf per dimostrare come la fusione tonale fosse una legge psicologica universale, innata, e non derivante da educazione culturale. L'approccio della teoria quindi è principalmente naturalistico. I gradi di fusione vanno dal grado I, fusione quasi totale, al grado V, nessuna fusione (vedi Tabella 1).

Tabella 1: Gradi di fusione per la teoria della fusione tonale di Stumpf con relativi intervalli musicali associati.

Grado di fusione	Intervallo musicale	Descrizione
I Grado	Ottava	Fusione quasi totale
II Grado	Quinta	Forte fusione
III Grado	Quarta	Fusione moderata
IV Grado	Terze (Maggiori/Minori)	Consonanze imperfette
	Seste (Maggiori/Minori)	
V Grado	Seconde (Maggiori/Minori)	Nessuna fusione
	Settime (Maggiori/Minori)	
	Settima Naturale	
	Tritono	

1.9 L'ipotesi di ricerca

Questa tesi e lo studio che descrive hanno come scopo la conferma o meno della teoria della fusione tonale di Stumpf. La domanda riguarda il possibile legame tra la fusione tonale e i concetti di consonanza e dissonanza. Ci si basa su due ricerche precedenti. Lo studio di DeWitt & Crowder (1987) analizza la correlazione tra la percezione di fusione tra suoni e giudizio uditivo di presenza di uno o più suoni utilizzando tempi di risposta TR. Il secondo studio di McPherson et al. (2020) invece si sofferma sui giudizi di piacevolezza e percezione di fusione tra suoni comparando risultati tra civiltà occidentali e non. L'esperimento descritto in questa tesi unisce i procedimenti di questi due studi su un campione puramente occidentale. L'intento è quello di ottenere giudizi di percezione di fusione tonale non solo attraverso misurazioni dirette con TR, ma anche con giudizi più personali su scala Likert a 4 livelli di fusione tra suoni.

Capitolo 2 – IL METODO

2.1 I Partecipanti

Il campione utilizzato per la ricerca è di 112 partecipanti reclutati per campionamento di convenienza, in questo elaborato tuttavia, ne verranno presi in considerazione solo 61. La partecipazione allo studio era volontaria, aperta a chi fosse interessato, con però dei criteri selettivi. Come criterio di inclusione, i partecipanti dovevano appartenere a un background culturale prevalentemente occidentale. In aggiunta, i partecipanti dovevano aver compiuto almeno 18 anni e non dovevano avere deficit uditivi. L'età dei partecipanti varia dai 20 ai 26 anni ($M = 22.31$, $DS = 1.68$). Per quanto riguarda il genere, non sono state applicate restrizioni per genere di appartenenza con 29 partecipanti maschi e 32 partecipanti femmine.

2.2 Il Piano di Ricerca

Lo studio presenta un approccio correlazionale, con tre variabili d'interesse. Il tempo di reazione (TR), è stato calcolato in millisecondi dalla comparsa dello stimolo uditivo alla pressione dei tasti "Z" e "M". I giudizi di piacevolezza venivano registrati dalla risposta fornita dal partecipante a una scala Likert a 4 punti: dopo aver ascoltato lo stimolo uditivo, il partecipante doveva digitare il tasto "1", "2", "3" o "4" avente il valore della scala Likert corrispondente. Anche per i giudizi di fusione la procedura è stata la medesima. I giudizi di fusione venivano registrati dalla risposta fornita dal partecipante a una scala Likert a 4 punti: dopo aver ascoltato lo stimolo uditivo, il partecipante doveva digitare il tasto "1", "2", "3" o "4" avente il valore della scala Likert corrispondente.

2.3 Gli Strumenti

Gli stimoli sonori sono stati generati utilizzando toni sintetici, così come riportato nello studio di McPherson et al. (2020). Gli stimoli sonori erano composti da due note presentate in contemporanea. Le due note avevano tra di loro un intervallo di altezza (pitch): questo pitch poteva essere un unisono, una seconda maggiore, una terza maggiore, una quarta giusta, un tritono, una quinta giusta, una settima maggiore, un'ottava, una nona minore. Sono stati selezionati questi specifici intervalli al fine da comprendere nella sperimentazione i quattro intervalli canonicamente più consonanti (l'ottava, la quinta, la quarta, la terza) da alternare alle frequenze canonicamente dissonanti. Ogni intervallo poteva essere presentato in tre diversi registri di altezza. Questi registri associati potevano assumere sei frequenze fondamentali: 110 Hz, 220 Hz, 440 Hz e due semitoni al di sopra di queste tre frequenze. Si arrivava in questo modo, a un totale di sei frequenze base. Partendo da ognuna di queste frequenze sono stati generati 9 intervalli di altezza costituiti da due note. Per simulare maggiormente il suono degli strumenti musicali, le ampiezze armoniche sono state ridotte di -14 dB per ottava. Era presente un esponenziale decrescente dello stimolo sonoro, con una costante di decadimento di 4 s^{-1}. Ogni stimolo sonoro durava in tutto 2000 ms. Gli stimoli utilizzati sono riportati nella Tabella 2.

Tabella 2: Intervalli degli stimoli sonori generati per ogni frequenza fondamentale. Sono espressi anche i loro rapporti e i loro rapporti in centesimi utilizzati nell'analisi dei dati.

Unisono	Seconda Maggiore	Terza Maggiore	Quarta Giusta	Tritono	Quinta Giusta	Settima Maggiore	Ottava	Nona Minore
1:1	9:8	5:4	4:3	45:32	3:2	15:18	1:2	32:15
0	204	386	498	590	702	1088	1200	1312

L'esperimento è stato condotto in una camera anecoica, in modo da poter ridurre al massimo la presenza di interferenze e stimoli ambientali esterni. Gli stimoli uditivi sono stati presentati tramite computer Veriton X2640G, con monitor DELLU2715H, Sul computer era presente il programma di calcolo numerico e sviluppo software MATLAB, per la presentazione degli stimoli e la registrazione dei TR. È stata utilizzata una sceda audio FOCUSRITE Scarlet 4i4 il cui segnale era in uscita verso delle cuffie Sennheiser HD-280 Pro il cui volume era poteva

essere regolato secondo le preferenze personali del partecipante. Per la pressione di tasti, al fine di registrare i TR e i dati sulle valutazioni, è stata utilizzata una tastiera USB modello EK-902 di Shenzhen Epoch Technology. Specificatamente dal partecipante sono stati utilizzati i tasti “1”, “2”, “3”, “4”, “M”, Z”.

Sono state utilizzate due scale Likert a 4 punti all’interno dello studio. La prima scala Likert, riguardo il grado di piacevolezza di uno specifico stimolo uditivo, era strutturata nella seguente maniera:

1 = Non mi piace per niente

2 = Non mi piace molto

3 = Mi piace un po’

4 = Mi piace molto

La seconda scala Likert, riguardo il grado di fusione dei vari stimoli uditivi, era strutturata nella seguente maniera:

1 = Per nulla fusi

2 = Poco fusi

3 = Abbastanza fusi

4 = Del tutto fusi

La domanda presentata ai partecipanti riguardo l’esperienza musicale deriva dallo studio di Ollen (2006): *“A criterion-related validity test of selected indicators of musical sophistication using expert ratings”*. È stata tradotta in italiano nella seguente maniera: “Quale titolo ti descrive meglio?”. Le possibili risposte erano le seguenti: “Non musicista”, “Appassionato di musica non musicista”, “Musicista amatoriale”, “Musicista amatoriale esperto”, “Musicista semiprofessionista”, “Musicista professionista”.

2.4 La Procedura

Come prima cosa, ai partecipanti è stato consegnato il foglio informativo riguardo ciò che gli sarebbe stato chiesto di fare nei vari compiti dell'esperimento. Dopo averlo letto e aver dato il consenso informato per la partecipazione alla ricerca, i partecipanti sono stati condotti all'interno della camera anecoica. Lo sperimentatore ha spiegato nuovamente, a voce, i dettagli della ricerca, così da poter fornire completa chiarezza nella procedura sperimentale. Il volume in arrivo alle cuffie è stato regolato secondo le preferenze del partecipante.

Prima di iniziare con la fase sperimentale, sono state poste domande riguardo dati anagrafici basilari (nome e cognome, età, genere, nazionalità) ed una specifica domanda riguardo la propria esperienza musicale (Ollen, J. E., 2006). Riguardo il questionario di Ollen, la risposta veniva registrata come valore da 1 a 6, con 1 "Non musicista" e 6 "Musicista professionale". Ciascun compito veniva somministrato mediante specifiche funzioni MATLAB.

La fase sperimentale era costituita da quattro compiti, ognuno dei quali era rappresentato con una delle seguenti lettere dell'alfabeto: A, B, P, F. L'ordine di presentazione dei diversi compiti è stato randomizzato, facendolo variare a seconda del partecipante. Le quattro combinazioni possibili erano ABPF, ABFP, BAPF, BAFP. Quest'operazione di controbilanciamento ha come scopo quello di controllare eventuali effetti e manipolazioni dovute all'ordine del compito, per garantire maggiore validità allo studio. In caso il controbilanciamento fosse del tipo ABPF o ABFP, tra gli argomenti della funzione del primo compito veniva scritto "AB", mentre tra gli argomenti della funzione del secondo compito veniva scritto "BA". In caso il controbilanciamento fosse del tipo BAPF o BAFP, tra gli argomenti della funzione primo compito veniva scritto "BA", mentre tra gli argomenti della funzione del secondo compito veniva scritto "AB".

Ogni prova dei vari compiti iniziava con lo schermo del monitor nero, con l'assenza di stimoli per 1 secondo. Successivamente andava a comparire una croce di fissazione bianca al centro dello schermo per circa 0.5 s. Dopo un intervallo da 0.5 a 1 s dopo la scomparsa della croce di fissazione, compariva lo stimolo sonoro. Lo stimolo sonoro aveva una durata di circa 2 s. Dopo la pressione del tasto scelto, si poteva passare alla prova successiva, dove il

procedimento appena descritto si ripeteva per ogni stimolo sonoro. Alternativamente, in caso non fosse stata registrata risposta entro 4s dalla comparsa dello stimolo sonoro, il sistema sarebbe passato automaticamente alla prova successiva. Ogni partecipante, all'interno dei primi due compiti, arrivava ad ascoltare lo stesso stimolo sonoro per dieci volte. A metà di ogni compito era presente una pausa per far riposare il partecipante quanto volesse, dandogli completa libertà di quando proseguire con l'esperimento in atto.

Il primo compito (A) prevedeva che ai partecipanti venissero somministrati una serie di stimoli sonori per poi, subito dopo, fare una valutazione percettiva su ciò che avevano appena sentito. La richiesta prevedeva di indicare, nel minor tempo possibile, se il suono appena sentito potesse essere composto da uno o più suoni separati. Nel primo caso, quindi se intuitivamente veniva percepito un unico suono, il partecipante avrebbe dovuto premere il tasto "Z" sulla tastiera. Nel secondo caso invece, quindi se intuitivamente venivano percepiti più suoni, il partecipante avrebbe dovuto premere il tasto "M" sulla tastiera.

Il secondo compito (B) prevedeva una procedura identica al primo; tuttavia, i due tasti da premere venivano invertiti. In caso intuitivamente veniva percepito un unico suono, il partecipante avrebbe dovuto premere il tasto "M" sulla tastiera. Al contrario, in caso intuitivamente venivano percepiti più suoni, il partecipante avrebbe dovuto premere il tasto "Z" sulla tastiera. Sia il primo che il secondo compito sono stati somministrati utilizzando la funzione MATLAB *fusion_rt2*, con in aggiunta come argomento rispetto agli altri compiti, l'ordine di somministrazione randomizzato per il partecipante.

Il terzo compito (P) prevedeva l'esposizione ai partecipanti delle stesse tipologie di stimoli sonori ascoltate nei compiti precedente. Questa volta però, erano presenti solo gli stimoli sonori composti da più suoni; mancava quindi l'unisono. Per ogni stimolo sonoro presentato, al partecipante veniva chiesto di formulare un giudizio di piacevolezza secondo una scala Likert a 4 punti. Per la valutazione, il partecipante doveva premere i tasti da 1 a 4 in corrispondenza con il livello della scala Likert associato. Il terzo compito è stato somministrato utilizzando la funzione MATLAB *pref_*rating2. Al contrario dei primi due compiti, la variabile d'interesse non erano i TR, quindi non era necessario rispondere nel minor tempo possibile.

Il quarto compito (F) prevedeva una procedura identica al terzo; tuttavia, la richiesta di base era diversa. I suoni venivano nuovamente somministrati nella medesima maniera del precedente compito, ma questa volta il giudizio da formulare non era un giudizio di piacevolezza, ma un giudizio di fusione. Il partecipante doveva quindi valutare quanto i suoni che componevano lo stimolo somministrato fossero fusi tra loro. La valutazione avveniva secondo una scala Likert a 4 punti. Anche in questo caso il partecipante doveva premere i tasti da 1 a 4 in corrispondenza con il valore della scala. Il quarto compito è stato somministrato utilizzando la funzione MATLAB *fusion_rating2*.

Con l'ultimo compito lo studio si andava a concludere, arrivando ad avere una durata complessiva di circa un'ora.

Capitolo 3 – ANALISI DEI DATI

Per l'analisi dei dati del primo compito sono state calcolate le medie dei TR dei vari partecipanti in base agli intervalli proposti dalla prova specifica. Insieme alle medie si sono considerate anche le deviazioni standard per ogni intervallo. Sono stati esclusi dall'analisi tutte le prove il cui tempo di risposta superasse i 3s e quelle in cui la risposta fosse minore o uguale a 0, vale a dire i partecipanti che non avevano dato risposta o avevano risposto prima della presentazione dello stimolo sonoro. Inoltre bisogna sottolineare come i dati si riferiscano solo ai TR e non all'accuratezza di tale risposta. I dati analizzati sono riportati nella Tabella 3. Tra i vari intervalli, quelli considerati canonicamente più consonanti risultavano avere TR più brevi, mentre quelli considerati canonicamente dissonanti avevano TR più dilatati. L'intervallo che è risultato avere TR più brevi è stato l'unisono, mentre quello con TR in media più lunghi è stata la quinta giusta, unica eccezione registrata.

Tabella 3: Media e deviazione standard dei TR per ogni intervallo proposto.

Intervalli	Media TR	Dev. Standard TR
0	1345.13	611.74
204	1424.00	667.84
386	1443.04	632.85
498	1424.29	623.18
590	1416.81	637.18
702	1498.44	640.55
1088	1416.18	647.06
1200	1429.12	620.67
1312	1397.03	637.21

Per il secondo compito sui giudizi di piacevolezza sono state considerate media e deviazione standard delle risposte in scala Likert (vedi Tabella 4). Un punteggio più alto in media, quindi, indicava una piacevolezza percepita maggiore nel compito. Anche in questo caso si può notare come l'intervallo che è stato considerato più piacevole sia stata l'ottava, mentre quello

considerato meno piacevole è stata la seconda maggiore. Lo scarto tra il punteggio più alto e quello più basso non è particolarmente ampio, ma comunque significativamente diverso, con il primo avente ricevuto un punteggio medio di 2.79, l'ultimo avente punteggio medio di 1.93. Considerando gli intervalli canonicamente consonanti (unisono, terza maggiore, quarta giusta, quinta giusta, ottava) e quelli canonicamente dissonanti (seconda maggiore, tritono, settima maggiore, nona minore), è possibile osservare come i primi abbiano ottenuto giudizi più alti dei secondi.

Tabella 4: Media e deviazione standard delle risposte alla scala Likert sul grado di piacevolezza percepito per ogni intervallo proposto.

Intervalli	**Media delle risposte**	**Dev. Standard delle risposte**
204	1.93	1.08
386	2.43	1.05
498	2.58	1.06
590	2.20	1.02
702	2.70	1.06
1088	2.03	0.97
1200	2.79	1.02
1312	1.99	0.97

Per il terzo compito sui giudizi di fusione, infine, sono state nuovamente considerate media e deviazione standard delle risposte della scala Likert (vedi Tabella 5). Valori più alti indicavano un grado di fusione percepito maggiore. L'intervallo giudicato più fuso è stata l'ottava, con una media di risposte di 3.22. L'intervallo giudicato meno fuso è stata la nona minore, con una media di risposte di 2.22. Anche in questo caso troviamo gli intervalli canonicamente considerati consonanti aventi ricevuto giudizi di fusione più alti in media, mentre quelli dissonanti sono stati giudicati meno fusi in media.

Tabella 5: Media e deviazione standard delle risposte alla scala Likert sul grado di fusione percepito per ogni intervallo proposto.

Intervalli	Media delle risposte	Dev. Standard delle risposte
204	2.31	1.06
386	2.61	1.03
498	2.72	1.02
590	2.42	1.06
702	2.88	1.00
1088	2.29	1.02
1200	3.22	0.98
1312	2.22	1.03

I dati sono aggiuntivamente mostrati all'interno dei rispettivi grafici (Figura 1).

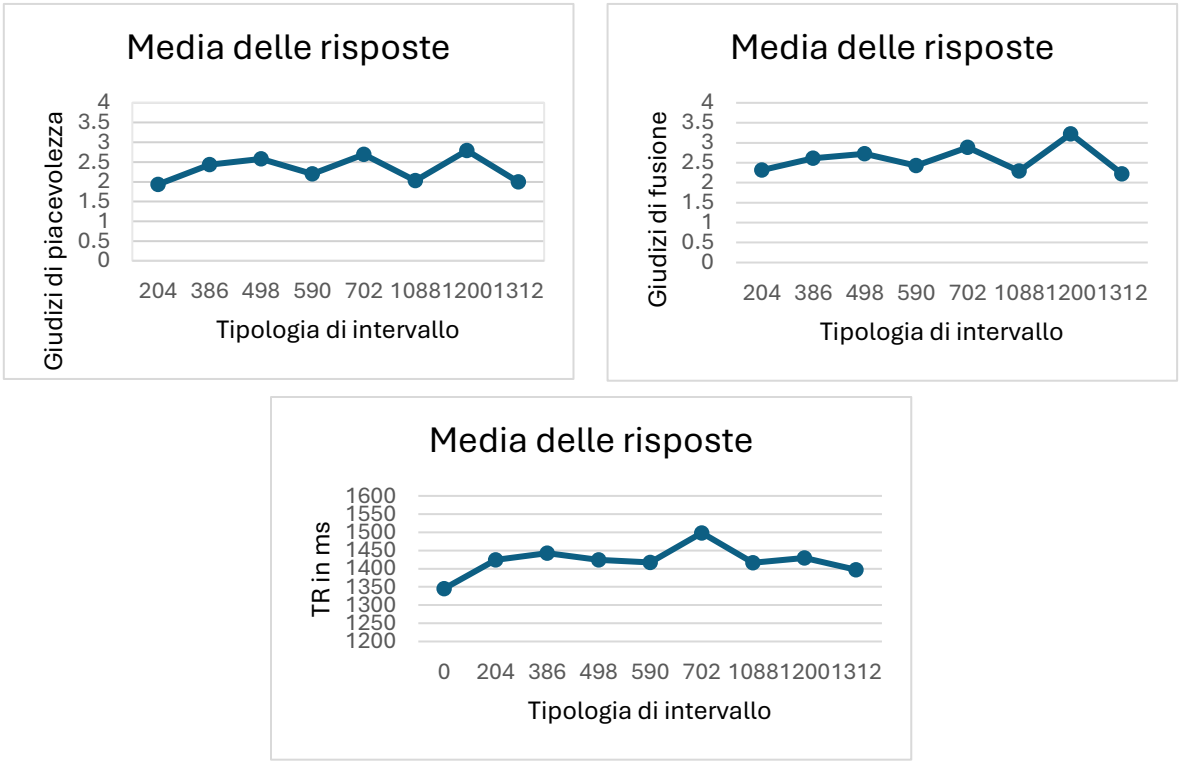

Figura 1: In alto, grafici a confronto di medie di giudizi di piacevolezza (a sinistra) e giudizi di fusione (a destra) per intervalli considerati. In basso, grafico di medie di TR registrati nella discriminazione degli intervalli considerati.

Capitolo 4 – CONCLUSIONI

4.1 Risultati

Lo studio è riuscito a esplorare il rapporto tra fusione tonale e consonanza musicale ottenendo risultati significativi.

Per quanto concerne la seconda parte dell'esperimento, basata sul lavoro di McPherson et al. (2020), emerge una chiara relazione tra i giudizi di piacevolezza e quelli di fusione espressi. Gli intervalli percepiti come più piacevoli risultano essere gli intervalli canonicamente consonanti. Tali intervalli risultano avere anche giudizi di fusione alti. Questa correlazione va a conferma della teoria della fusione tonale di Carl Stumpf, suggerendo come la percezione di fusione e quella di piacevolezza, dunque di consonanza, risultino avere un fondamento psicologico percettivo comune.

Per quanto riguarda l'analisi dei TR, i pattern rilevati vedono i vari intervalli avere per lo più tutti tempi registrati simili. Sicuramente l'intervallo più facile da riconoscere è stato l'unisono, in quanto singolo suono stabile. Di contro lo stimolo sonoro più difficile da riconoscere, che ha richiesto più tempo di elaborazione per la risposta, è stata la quinta giusta, canonicamente consonante. Probabilmente ciò è dovuto alla stretta vicinanza dei due suoni presentati nell'intervallo; dunque, si potrebbe prendere in considerazione come la lontananza di due note tra loro possa rendere più complessa la percezione e discriminazione di suoni. Un dato interessante è quello legato ai TR per la nona minore, dissonante. Dopo l'unisono, questo è l'intervallo a cui sono associati TR minori per la discriminazione. Probabilmente, in linea anche con i risultati della valutazione della fusione percepita, la nona minore risulta un intervallo più semplice da discriminare in quanto avente un indice di fusione particolarmente basso, caratterizzato da forte contrasto tra le due note presentate. Ciò ha facilitato la percezione del suono portando ad una risposta veloce.

4.2 Limiti e sviluppi futuri

Lo studio riscontra dei limiti su cui è possibile lavorare in futuro. Dai risultati specificatamente del primo compito, si possono osservare dei tempi di reazione più dilatati di ciò che ci si poteva aspettare. Ciò potrebbe implicare una minore attendibilità delle risposte date. I vari risultati infatti sembrano essere particolarmente vicini tra loro, con differenze minime tra intervalli consonanti e dissonanti non in linea con gli studi precedenti. Una possibile proposta per sviluppi futuri potrebbe essere quella di implementare "stimoli naturalistici" nella raccolta dati assieme a quelli "sintetici" generati digitalmente. Questo stesso approccio è stato utilizzato proprio negli studi di riferimento di McPherson et al. (2020). Un alternare di diverse tipologie di stimoli, considerati più piacevoli in quanto più consoni all'orecchio umano, permetterebbe al partecipante di rimanere maggiormente vigile durante le varie prove, migliorando i TR. Anche per quanto riguarda giudizi di piacevolezza e fusione, non troviamo grandi differenze, con la maggior parte dei risultati compresi tra 2 e 3. Nello specifico per quello di piacevolezza non troviamo nessun valore oltre il 3. L'utilizzo di stimoli naturalistici aiuterebbe anche nel distinguere maggiormente i vari stimoli sonori, portando a giudizi di gradevolezza più focalizzati e attendibili.

A livello generale lo studio sicuramente apre la strada ad approfondimenti sull'argomento, suggerendo una relazione tra fusione tonale e consonanza. Il rapporto tra i due fenomeni sembrerebbe seguire le teorie di Stumpf, pur rimanendo una relazione complessa che merita ulteriore attenzione per comprenderne al meglio il legame.

BIBLIOGRAFIA

Aristosseno. (1954). *Elementa harmonica* (R. Da Rios, A cura di e Trad.). Istituto Poligrafico dello Stato. (Opera originale pubblicata ca. 330 a.C.).

Arom, S. (1991). *African polyphony and polyrhythm: Musical structure and methodology* (M. Thom, B. Tuckett, & R. Boyd, Trad.). Cambridge University Press.

A.J. Blood, & R.J. Zatorre, Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion, Proc. Natl. Acad. Sci. U.S.A. 98 (20) 11818-11823, https://doi.org/10.1073/pnas.191355898 (2001).

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2016). *Neuroscienze: Esplorare il cervello* (G. A. Carlesimo & M. Oliveri, Trad.; 4ª ed.). Edra. (Opera originale pubblicata 2015).

Blood, A., Zatorre, R., Bermudez, P. *et al.* Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat Neurosci* **2**, 382–387 (1999). https://doi.org/10.1038/7299

Boezio, A. M. T. S. (1989). *De institutione musica* (G. Marzi, A cura di e Trad.). Istituto Italiano per la Storia della Musica. (Opera originale pubblicata ca. 510).

Darwin, C. (2011). *L'origine dell'uomo e la selezione in relazione al sesso* (G. Montalenti, Trad.). Newton Compton Editori. (Opera originale pubblicata 1871).

Dewitt, L.A., Crowder, R.G. Tonal fusion of consonant musical intervals: The oomph in Stumpf. *Perception & Psychophysics* **41**, 73–84 (1987). https://doi.org/10.3758/BF03208216

Martinelli, R. (2014). Melting musics, fusing sounds. Stumpf, Hornbostel and Comparative Musicology in Berlin. Deutsches Jahrbuch für Philosophie, 5, 391–402.

McPherson, M.J., Dolan, S.E., Durango, A. *et al.* Perceptual fusion of musical notes by native Amazonians suggests universal representations of musical intervals. *Nat Commun* **11**, 2786 (2020). https://doi.org/10.1038/s41467-020-16448-6

Musica enchiriadis and Scolica enchiriadis (R. Erickson, Trad.; C. V. Palisca, Cur.). (1995). Yale University Press. (Opera originale pubblicata ca. 900).

Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings* [Doctoral dissertation, Ohio State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=osu1161705351

Schönberg, A. (2014). *Manuale di armonia* (L. Rognoni & G. Manzoni, Trad.). Il Saggiatore. (Opera originale pubblicata nel 1911)

Wundt, W. (1897). *Outlines of psychology* (C. H. Judd, Trans.). Wilhelm Engelmann. (Original work published 1896).