



UNIVERSITÀ DEGLI STUDI DI PADOVA

**Dipartimento di Filosofia, Sociologia,
Pedagogia, e Psicologia Applicata (FISPPA)**

Corso di laurea in Scienze Psicologiche Sociali e del Lavoro

Elaborato finale

**La relazione tra fusione tonale e
consonanza musicale**

***The relationship between tonal fusion and musical
consonance***

Relatore

Prof. Massimo Grassi

Laureando: Thomas Righi

Matricola: 2123980

Anno accademico 2025/2026

SOMMARIO

SOMMARIO	3
INTRODUZIONE	4
CAPITOLO 1- CONSONANZA E DISSONANZA TONALE	6
1.1 Concetti introduttivi.....	6
1.2 Dalla speculazione matematica al razionalismo	6
1.3 L'approccio psicoacustico e fisiologico.....	7
1.4 L'approccio psicologico e fenomenologico	9
1.5 Il culturalismo e i modelli empirici contemporanei	10
CAPITOLO 2 – METODO E PROCEDURA SPERIMENTALE.....	12
2.1 Partecipanti	12
2.2 Gli stimoli e l'attrezzatura.....	13
2.3 Procedura e disegno sperimentale.....	14
CAPITOLO 3 – ANALISI DEI DATI	17
CAPITOLO 4 – CONCLUSIONI	23
4.1 Discussione dei risultati	23
4.2 Limiti dello studio e sviluppi futuri	23
BIBLIOGRAFIA.....	25

INTRODUZIONE

L'ascolto della musica e la percezione dei suoni rappresentano un'esperienza complessa che va ben oltre la semplice ricezione di onde acustiche. Quando ascoltiamo un brano o sentiamo un singolo accordo, il nostro sistema uditivo elabora continuamente stimoli fisici per trasformarli in significati e risposte emotive, capaci di evocare ricordi, alterare l'umore o suscitare sensazioni immediate di piacere o fastidio.

Spiegare la reazione dell'ascoltatore basandosi unicamente sulle proprietà acustiche dei suoni risulterebbe riduttivo. Sebbene l'armonia si fondi su precise leggi fisiche e matematiche, l'esperienza percettiva ed edonica è il prodotto di una complessa interazione tra vincoli neurofisiologici, processi cognitivi di organizzazione del segnale e fattori di natura culturale.

Un interrogativo che da oltre un secolo affascina la psicologia della percezione e la psicoacustica riguarda il legame tra la fusione tonale (*Tonverschmelzung*) e la percezione di consonanza e dissonanza. Storicamente introdotta da Carl Stumpf tra la fine dell'Ottocento e l'inizio del Novecento, la fusione tonale è stata a lungo considerata uno dei costrutti chiave per comprendere la consonanza degli intervalli musicali. Stumpf ipotizzò infatti che il grado di fusione, ovvero la tendenza di due o più toni simultanei a perdere la propria individualità per integrarsi in una percezione sonora unitaria, fosse una proprietà intrinseca del sistema uditivo, strettamente legata ai rapporti matematici tra le frequenze dei toni e indipendente dalle componenti edoniche o culturali.

Tuttavia, la ricerca contemporanea ha ampliato questo dibattito, indagando come la percezione di fusione sia influenzata non solo dalla struttura acustica dello stimolo, come l'ampiezza dell'intervallo armonico e la frequenza fondamentale (f_0), ma anche da come tali processi percettivi si riflettano sul giudizio estetico ed edonico dell'ascoltatore.

In questa prospettiva risulta cruciale distinguere tra meccanismi di elaborazione ad alta velocità, guidati dalle caratteristiche fisiche del segnale, e i giudizi espliciti formulati dall'ascoltatore. L'analisi combinata delle risposte implicite e delle valutazioni consapevoli consente di chiarire se la tendenza dei toni a fondersi

costituisca un vincolo percettivo primario o se sia piuttosto un fenomeno modulabile del gradimento estetico.

L'obiettivo centrale della presente ricerca è comprendere in che modo la facilità di elaborazione percettiva guidi l'apprezzamento estetico. A questo scopo, è stato condotto un esperimento in ambiente controllato all'interno di una stanza insonorizzata, esponendo i partecipanti all'ascolto di una batteria di suoni sintetici complessi. Il protocollo ha previsto una valutazione integrata del fenomeno: da un lato, attraverso compiti di decisione rapida mirati a misurare i tempi di reazione e la tendenza immediata a percepire un suono unico o doppio; dall'altro, mediante scale di valutazione soggettiva volte a quantificare sia il grado di fusione percepito sia la gradevolezza edonica attribuita a ciascun intervallo armonico.

Il presente elaborato si articola in quattro capitoli. Il primo capitolo ripercorre l'evoluzione teorica dei concetti di consonanza e dissonanza, dai primi modelli matematici e psicoacustici fino all'approccio fenomenologico di Stumpf e ai modelli culturalisti contemporanei. Il secondo capitolo descrive dettagliatamente la metodologia, il campione di partecipanti e la procedura sperimentale adottata. Il terzo capitolo espone l'analisi descrittiva dei dati raccolti nei diversi compiti. Infine, il quarto capitolo discute i risultati emersi alla luce dell'ipotesi di Stumpf, delineando i limiti metodologici della ricerca e le prospettive per gli sviluppi futuri.

CAPITOLO 1- CONSONANZA E DISSONANZA TONALE

1.1 Concetti introduttivi

Dal punto di vista acustico, il fenomeno musicale si descrive come la disposizione organizzata di sequenze di suoni, semplici o complessi, distribuiti lungo l'asse temporale. Tuttavia, limitarsi a una visione puramente tecnica non basta a chiarire la varietà delle reazioni psicologiche ed emotive che la musica provoca nell'ascoltatore. Numerose indagini psicofisiologiche e di neuroimaging hanno infatti evidenziato come l'ascolto di strutture armoniche stimoli circuiti neurali profondi legati alla ricompensa, all'arousal e all'elaborazione emotiva, coinvolgendo regioni sottocorticali come l'amigdala in risposta al grado di stabilità o tensione del segnale sonoro.

Nello studio psicofisico della percezione uditiva, uno degli argomenti più discussi è rappresentato dai concetti di consonanza e dissonanza tonale. Si tratta della tendenza, comune all'esperienza umana, a percepire alcune combinazioni di suoni simultanei (dette intervalli armonici) come stabili, armonici e piacevoli, rispetto ad altre che invece trasmettono tensione, instabilità e una certa asprezza acustica.

Nel corso della storia, i tentativi di spiegare questa dicotomia si sono sviluppati lungo due principali orientamenti discordanti: quello naturalista e quello culturalista (Di Stefano & Bertolaso, 2014). In questa prospettiva, la questione centrale è capire se la consonanza sia un fenomeno innato e biologicamente determinato, oppure il risultato malleabile dell'apprendimento e dell'esperienza, influenzato dal contesto socio-culturale e dalla familiarità statistica con determinati stimoli sonori. Il primo orientamento considera la consonanza governata da precise leggi fisiche o da limiti anatomo-fisiologici dell'apparato uditivo periferico (DeWitt & Crowder, 1987), mentre il secondo interpreta l'apprezzamento musicale in termini di esperienza soggettiva e apprendimento (Cazden, 1945).

1.2 Dalla speculazione matematica al razionalismo

Le prime formulazioni teoriche volte a comprendere la natura della consonanza affondano le proprie radici nell'antica Grecia, all'interno della tradizione pitagorica. Pitagora osservò come la gradevolezza e la stabilità percettiva di un intervallo acustico fossero direttamente proporzionali alla semplicità matematica del rapporto numerico che legava le lunghezze

delle corde vibranti. Relazioni geometriche elementari come rapporti 2:1, 3:2 e 4:3 – corrispondenti rispettivamente agli intervalli ottava, quinta e quarta – venivano interpretate come espressione di un ordine razionale e universale intrinseco alla natura stessa del cosmo (Di Stefano & Bertolaso, 2014). All'opposto, rapporti numerici più complessi venivano associati a una progressiva instabilità uditiva, relegando quindi la dissonanza a una deviazione da tale armonia geometrica. In epoca medievale, tale rigidità portò persino a bandire intervalli complessi come il tritono, etichettandolo come *Diabolus in Musica* e considerandolo un errore acustico da evitare rigorosamente.

Questo rigido riduzionismo incontrò evidenti paradossi empirici nel corso della sua evoluzione, come dimostrato dal dibattito intorno all'intervallo di undicesima¹. Sebbene dal punto di vista del calcolo pitagorico questo stimolo dovesse essere catalogato come dissonante, i teorici dell'epoca e gli esecutori pratici ne riportavano una chiara percezione di dolcezza e consonanza acustica. Tale discrepanza mise in crisi l'approccio matematico puro, favorendo una transizione verso l'empirismo guidata da pensatori come Aristosseno, il quale ricondusse la qualificazione dei toni al giudizio dinamico e soggettivo dell'orecchio e della memoria uditiva (Aristosseno, 1955; cit. in DeWitt e Crowder, 1987).

Nel XVII secolo, il passaggio verso una concezione più meccanicistica e psicofisica trovò un punto di svolta nel pensiero di Cartesio. All'interno del suo *Compendium Musicae*, egli cercò di spiegare il gradimento uditivo non più come una legge cosmica, ma in funzione della frequenza con cui le onde sonore colpiscono contemporaneamente i sensi (Di Stefano & Bertolaso, 2014). Pur mantenendo un orientamento basato su relazioni fisiche, Cartesio riconobbe il ruolo della soggettività, dell'esperienza pregressa e del contesto individuale nel determinare il giudizio estetico dell'ascoltatore. L'armonia musicale cessò così di essere una mera caratteristica strutturale dello stimolo, trasformandosi progressivamente in un oggetto di indagine legato ai processi percettivi e cognitivi dell'essere umano.

1.3 L'approccio psicoacustico e fisiologico

La prima transizione verso una spiegazione scientifica e fisiologica della consonanza si verifica nella seconda metà del XIX secolo grazie agli studi del medico e fisico tedesco

¹ L'intervallo di undicesima è un intervallo composto (ovvero superiore all'ottava) costituito dalla somma di un'ottava e di una quarta. Nel sistema proporzionale pitagorico, esso corrisponde al rapporto di frequenza 8:3

Hermann von Helmholtz. All'interno del suo modello teorico, egli propone di ridurre la dicotomia tra gli intervalli eufonici e quelli disfonici a un fenomeno di natura prevalentemente sensoriale e periferica, le cui cause risiedono nei vincoli anatomici della coclea e della membrana basilare.

A livello meccanico, quindi, il comportamento del sistema uditivo periferico può essere descritto come un banco di filtri passa-banda. La coclea agisce come un analizzatore di frequenze: ogni punto specifico lungo la membrana basilare si caratterizza per una frequenza preferenziale d'oscillazione, detta frequenza caratteristica, che risponde in modo selettivo a una ristretta gamma di stimolazioni pressorie. In altre parole, questa membrana funziona come la tastiera di un pianoforte al contrario, dove zone diverse entrano in risonanza solo quando vengono colpite da note specifiche. Quando due toni puri (suoni composti da una sola frequenza isolata), o le parziali superiori che costituiscono lo spettro di due toni complessi armonici (frequenze secondarie più acute che compongono naturalmente ogni suono reale, come quello di uno strumento o una voce), vengono eseguiti contemporaneamente, le rispettive onde interagiscono nello spazio generando un effetto oscillatorio di sommazione noto come battimento².

La spiegazione fisiologica si focalizza sul concetto di banda critica, ovvero un limite spaziale d'interferenza che all'interno della coclea equivale approssimativamente all'ampiezza di un quarto d'ottava (Plomp & Levelt, 1965). Se la distanza in frequenza tra due parziali acustiche è sufficientemente ampia da superare i confini della banda critica, i filtri uditivi riescono a risolvere lo stimolo, permettendo all'ascoltatore di separare nettamente le diverse componenti e di percepire un suono in modo nitido e continuo. Al contrario, se le armoniche cadono contemporaneamente nell'area d'azione del medesimo filtro, la membrana basilare non è in grado di discriminarle; in questo stato di mancata risoluzione, i battimenti rapidi generano una percezione intermittente del tono, descritta psicologicamente in termini di rugosità, asprezza e sgradevolezza.

Nell'approccio psicoacustico classico, di conseguenza, la dissonanza si configura come il mero riflesso cosciente della rugosità prodotta dai battimenti irrisolti a livello periferico. La consonanza, al contrario, viene interpretata in chiave negativa, ovvero

² Questo fenomeno si manifesta come un tremolio o fluttuazione periodica dell'intensità sonora. Ciò accade quando due onde caratterizzate da frequenza molto prossime si sovrappongono nello spazio: l'apparato uditivo li fonde in un unico tono instabile che, al crescere della velocità delle fluttuazioni, tende a essere percepito come sgradevole

come una condizione acustica ottimale definita dalla totale assenza di fluttuazioni e battimenti udibili all'interno della banda critica.

1.4 L'approccio psicologico e fenomenologico

In netta contrapposizione con l'impostazione fisiologica di Helmholtz, lo psicologo e filosofo tedesco Carl Stumpf propose un modello interpretativo alternativo, spostando l'indagine della consonanza dal livello periferico a quello cognitivo e centrale (DeWitt & Crowder, 1987). Stumpf (1890) infatti riteneva insufficiente la spiegazione basata sui battimenti e sulla banda critica: l'esperienza della consonanza non poteva essere ridotta alla semplice assenza di un disturbo sensoriale periferico, ma doveva configurarsi come un processo psicologico attivo e unitario.

Il nucleo di questa teoria risiede nel concetto di fusione tonale (*Tonverschmelzung*), definita come la tendenza per cui due suoni eseguiti simultaneamente tendono a interpenetrarsi e a fondersi nella coscienza dell'ascoltatore, fino ad essere percepiti come un'unica entità coerente (DeWitt & Crowder, 1987). Secondo Stumpf, il grado di consonanza di un intervallo armonico è direttamente proporzionale al suo livello di fusione: più l'intervallo è consonante, più è difficile per il soggetto isolare le singole note che lo compongono.

Questa relazione psicofisica permise a Stumpf di tracciare una gerarchia articolata su cinque gradi decrescenti di fusione:

- I Grado (fusione quasi totale): ottava
- II Grado (forte fusione): quinta giusta
- III Grado (fusione moderata): quarta giusta
- IV Grado (consonanze imperfette): terze e seste (maggiori e minori)
- V Grado (nessuna fusione): seconde, settime e tritono

Per verificare sperimentalmente la validità di questa gerarchia e dimostrare che la fusione fosse un fenomeno cognitivo universale, Stumpf strutturò un paradigma di indagine basato sulla discriminazione uditiva. Presentando a soggetti sia esperti che non, stimoli composti da una sola nota o da intervalli di due note, egli osservò che gli errori di valutazione aumentavano sistematicamente in presenza di intervalli a elevata fusione. Davanti a un'ottava o a una quinta, anche gli ascoltatori privi di un *background*

musicale faticavano a comprendere se il suono fosse singolo o doppio, mentre l'identificazione di due altezze distinte risultava immediata e priva di esitazioni in presenza di intervalli dissonanti come la seconda.

Sebbene la sovrapposizione tra consonanza e fusione tonale sia stata successivamente ridimensionata da autori che evidenziarono la necessità di mantenere distinti il giudizio di piacevolezza estetica e la capacità di analisi acustica (Huron, 1991), l'eredità metodologica di Stumpf rimane cruciale. Il suo lavoro ha dimostrato che lo studio della percezione musicale non può prescindere dall'analisi dei tempi di reazione e dei tassi d'errore nei compiti di discriminazione, ponendo le basi per i moderni protocolli di indagine psicofisica sul funzionamento cognitivo dell'udito.

1.5 Il culturalismo e i modelli empirici contemporanei

Parallelamente all'evoluzione dei modelli naturalisti di matrice biologica e cognitiva, la ricerca psicologica ha sviluppato un orientamento teorico di stampo culturalista ed empirista. I sostenitori di questo approccio, tra i quali spiccano i contributi di Cazden (1945) e Lundin (1947), contestano l'idea che la consonanza sia il riflesso automatico di leggi universali o di limiti anatomo-fisiologici dell'orecchio. Da questa prospettiva, l'apprezzamento o il rifiuto di determinate combinazioni sonore viene interpretato come un comportamento appreso, strettamente legato all'esposizione statistica, alla familiarità con gli stimoli e all'interiorizzazione delle regole del sistema musicale di appartenenza. Il giudizio di consonanza si configura quindi come un costrutto plastico e malleabile, fortemente influenzato dal contesto socio-culturale e soggetto a mutamenti nel corso delle diverse epoche storiche.

A sostegno di tale ipotesi, alcuni autori contemporanei hanno ipotizzato che i criteri di gradevolezza acustica derivino dall'apprendimento implicito delle caratteristiche spettrali della voce umana. L'esposizione quotidiana alla vocalizzazione linguistica, i cui spettri presentano una marcata stabilità e precise ricorrenze intervallari, fungerebbe da matrice per la costruzione di aspettative uditive stabili, portando l'ascoltatore a valutare eufonici gli intervalli musicali più comuni e rappresentati nell'ambiente vocale. Sebbene questo approccio offra una spiegazione dell'universalità di alcuni tratti percettivi, il culturalismo radicale incontra un limite metodologico nella

complessa verificabilità sperimentale dei suoi assunti, oltre a dover giustificare le ampie variazioni individuali riscontrate all'interno delle medesime cornici culturali.

Il dibattito scientifico moderno ha tentato di superare la storica polarizzazione tra naturalismo e culturalismo integrando entrambi gli orientamenti all'interno di modelli multifattoriali. Un punto di svolta in questa direzione è rappresentato dalle ricerche empiriche di McDermott, Lehr e Oxenham (2010), i quali sono riusciti ad isolare sperimentalmente l'impatto dei battimenti periferici dal ruolo esercitato dalla somiglianza armonica delle parziali. I loro risultati hanno evidenziato che la preferenza estetica per la consonanza è legata in modo significativo alla selettività per gli spettri armonici e tale sensibilità risulta fortemente amplificata dal livello di *training* ed esperienza musicale del soggetto.

L'architettura delle evidenze contemporanee suggerisce pertanto la necessità di scindere il giudizio estetico soggettivo (legato alla piacevolezza e fortemente mediato dall'esposizione culturale) dai processi di fusione e risoluzione uditiva operati dal sistema cognitivo centrale. Questa distinzione epistemologica ridefinisce lo studio della percezione uditiva non più come la ricerca di una singola causa meccanica, ma come l'analisi di un sistema integrato in cui l'accuratezza e la velocità nella discriminazione degli intervalli riflettono complessi meccanismi di elaborazione delle informazioni acustiche.

CAPITOLO 2 – METODO E PROCEDURA SPERIMENTALE

Per comprendere a fondo come percepiamo la consonanza, è utile affiancare alle misurazioni comportamentali dirette anche le valutazioni soggettive degli ascoltatori. Alla luce del dibattito teorico presentato nel primo capitolo, questo studio esamina la fusione tonale attraverso un doppio livello di indagine: da un lato la misurazione oggettiva delle capacità di discriminazione uditiva, dall'altro la raccolta dei giudizi espliciti sulla qualità e sulla piacevolezza degli intervalli.

Il disegno sperimentale unisce due impostazioni ben consolidate in letteratura. Per la rilevazione dei tempi di reazione e dell'accuratezza nelle risposte si fa riferimento al paradigma di DeWitt e Crowder (1987), che permette di stimare la fusione tonale in base alla difficoltà incontrata nel separare i due suoni. A questa prima valutazione dicotomica (suono singolo o suono doppio) si aggiunge una misurazione graduata su scala Likert a 4 punti, ripresa dai lavori di McPherson e collaboratori (2020)³, pensata per cogliere le sfumature nei giudizi di fusione e gradimento estetico.

Nelle sezioni che seguono vengono descritti il campione dei partecipanti coinvolti, le caratteristiche degli stimoli sonori impiegati e lo svolgimento dei quattro compiti che hanno composto la procedura sperimentale.

2.1 Partecipanti

Allo studio ha preso parte un campione di 28 partecipanti, composto da 13 femmine e 15 maschi, di età compresa tra i 20 e i 26 anni, con un'età media di 22,2 anni e una deviazione standard di 1,23. Tutti i soggetti sono stati reclutati su base volontaria.

I partecipanti hanno riferito di avere un udito nella norma e di non presentare deficit sensoriali o cognitivi che potessero ostacolare l'esecuzione dei compiti di discriminazione uditiva. Per valutare il livello di formazione e competenza musicale del campione è stato utilizzato l'indice Ollen (*Ollen Musical Sophistication Index*)⁴,

³ In questo studio gli autori hanno mostrato come la valutazione della consonanza e della fusione non si limiti a un meccanismo "tutto o niente", ma presenti gradazioni continue. L'impiego di una scala Likert a 4 punti permette quindi di catturare la variabilità individuale e l'intensità percepita di tali fenomeni senza forzare una scelta dicotomica

⁴ La competenza musicale è stata misurata tramite la scala validata da Ollen (2006), la quale richiede al partecipante di scegliere quale titolo descriva meglio il profilo tra sei categorie crescenti: 1) Non musicista, 2) Appassionato di musica non musicista, 3) Musicista amatoriale, 4) Musicista amatoriale esperto, 5) Musicista semiprofessionista, 6) Musicista professionista

ottenendo un punteggio medio di 2,11 con una deviazione standard di 1,10, su una scala da 1 a 6. Questo dato conferma che il campione era composto in prevalenza da ascoltatori non professionisti, con una formazione musicale assente o di livello amatoriale.

2.2 Gli stimoli e l'attrezzatura

L'esperimento si è svolto all'interno della cabina silente (Laboratorio C-09) del Dipartimento di Psicologia Generale (DPG) dell'università di Padova. Si tratta di un ambiente insonorizzato progettato per isolare l'ascoltatore dai rumori di fondo e dalle distrazioni ambientali.

I test sono stati somministrati tramite un personal computer Acer Veriton X2640G collegato a un monitor DELL U2715H da 27 pollici. La gestione della presentazione degli stimoli e la registrazione dei tempi di reazione sono state programmate in ambiente MATLAB, sfruttando la libreria Psychtoolbox.

Gli stimoli visivi comprendevano esclusivamente una croce bianca di fissazione (+) presentata al centro dello schermo su sfondo nero prima di ogni prova. La riproduzione dei suoni è avvenuta tramite una scheda audio esterna Focusrite Scarlett 4i4, con frequenza di campionamento impostata a 48.000 Hz, collegata a cuffie stereofoniche professionali chiuse Sennheiser HD-280 Pro. Il volume d'ascolto è stato regolato prima dell'avvio del test a un livello confortevole e costante per ciascun partecipante.

Le risposte comportamentali e i giudizi valutativi sono stati registrati mediante una tastiera USB dedicata, utilizzando nello specifico i tasti M, Z (per i compiti di discriminazione rapida BA ed AB) e i tasti numerici da 1 a 4 (per le valutazioni su scala Likert nei compiti F e P).

Tutti gli stimoli acustici utilizzati erano suoni sintetici generati al computer della durata di 2 secondi ciascuno. Ciascun suono (sia tono inferiore che superiore dell'intervallo) è stato sintetizzato come un tono complesso composto da 12 armoniche caratterizzate da un inviluppo con decadimento esponenziale (rampa di -24 dB sull'intera durata dello stimolo).

La batteria sperimentale comprendeva 14 intervalli armonici differenti, espressi in *cents*, la cui ampiezza spaziava da un minimo di 111 cents (seconda minore) fino a un

massimo di 1404 cents (nona maggiore), con le relative ampiezze intermedie riportate nella Tabella 1. A questi stimoli veniva inoltre affiancata la condizione di controllo dell'unisono (0 cents, rapporto 1:1), impiegata nelle prove di discriminazione rapida per presentare uno stimolo a frequenza singola.

Tabella 1: Intervalli armonici generati per ciascuna frequenza fondamentale, con i relativi rapporti di frequenza e ampiezze in centesimi

Intervallo	Rapporto	Cents	Intervallo	Rapporto	Cents
Seconda Minore	16:15	111	Sesta Minore	8:5	814
Seconda Maggiore	9:8	204	Sesta Maggiore	5:3	884
Terza Minore	6:5	316	Settima Minore	9:5	1018
Terza Maggiore	5:4	386	Settima Maggiore	15:8	1088
Quarta Giusta	4:3	498	Ottava	2:1	1200
Tritono	45:32	590	Nona Minore	32:15	1312
Quinta Giusta	3:2	702	Nona Maggiore	9:4	1404

I suoni sono stati generati a partire da 6 frequenze fondamentali (f_0) di base: 110 Hz, 123.5 Hz, 220 Hz, 246.9 Hz, 440 Hz e 493.9 Hz. L'intervallo di tempo silenzioso tra la presentazione della croce di fissazione sullo schermo e l'inizio effettivo del suono (intervallo ISI) è stato fissato a 0,5 secondi.

2.3 Procedura e disegno sperimentale

Dopo la raccolta del consenso informato, i partecipanti hanno ricevuto una descrizione scritta delle istruzioni relative al compito sperimentale, le quali sono state successivamente spiegate a voce dallo sperimentatore per assicurare una comprensione chiara dei compiti richiesti. Una volta completata questa fase, i partecipanti sono stati condotti individualmente all'interno della cabina silente e fatti accomodare davanti alla postazione di ascolto indossando correttamente le cuffie prima dell'inizio del test.

L'esperimento non prevedeva alcuna fase preliminare e si avviava direttamente con la sessione sperimentale. L'intero protocollo si articolava in quattro compiti distinti –

definiti convenzionalmente AB/BA, F e P – la cui esecuzione veniva avviata manualmente dallo sperimentatore al termine di ciascun blocco. Per il campione esaminato in questo studio, la sequenza dei quattro compiti è stata mantenuta fissa secondo l'ordine BAFP.

I primi due compiti (BA ed AB) corrispondevano alla condizione di risposta rapida sulla fusione tonale (*fusion RT*). In queste due sessioni venivano somministrati 180 stimoli per blocco. Ai partecipanti veniva chiesto di indicare nel minor tempo possibile se lo stimolo acustico ascoltato corrispondeva a un unico suono oppure a due suoni distinti sovrapposti, premendo uno dei due tasti dedicati nella tastiera (ad esempio il tasto M per indicare un suono unico e Z per due suoni distinti). La seconda sessione presentava la medesima struttura della prima, ma con l'associazione dei tasti di risposta invertita, al fine di controbilanciare l'associazione motoria della risposta tra le due mani. In entrambe le sessioni, il sistema registrava sia la scelta effettuata sia il tempo di reazione (RT).

Completati i primi due blocchi di risposta, lo sperimentatore avviava la terza fase (F), denominata *fusion rating*, composta da 84 stimoli. In questo compito si chiedeva ai partecipanti di esprimere un giudizio soggettivo sul grado di fusione percepito tra i suoni, fornendo una valutazione su una scala Likert a 4 punti strutturata nella seguente maniera:

- 1 = Per nulla fusi
- 2 = Poco fusi
- 3 = Abbastanza fusi
- 4 = Del tutto fusi

Infine, il compito (P) di valutazione della gradevolezza (*preference rating*), prevedeva la somministrazione di ulteriori 84 stimoli. In questa fase i soggetti dovevano esprimere un giudizio qualitativo ed edonico avvalendosi di una scala Likert a 4 punti strutturata nel seguente modo:

- 1 = Non mi piace per niente
- 2 = Non mi piace molto
- 3 = Mi piace un po'
- 4 = Mi piace molto

In tutti i compiti, la sequenza temporale della singola prova prevedeva la presentazione della croce di fissazione, l'intervallo di silenzio ISI, la riproduzione dello stimolo sonoro di 2 secondi e la registrazione della risposta tramite tastiera, seguita da una pausa di 3 secondi (*Inter-Trial Interval*) prima della prova successiva. Inoltre, a metà di ciascuna sessione il software interrompeva temporaneamente la sequenza presentando il messaggio "Premere la barra spaziatrice per continuare", permettendo al partecipante di fare una breve pausa e di riprendere il test solo quando si sentiva nuovamente pronto.

L'esperimento ha adottato un disegno fattoriale bilanciato entro i soggetti, in cui ogni partecipante è stato esposto all'incrocio completo tra i 14 livelli di ampiezza dell'intervallo armonico e i 6 livelli di frequenza fondamentale (f_0). L'ordine di presentazione delle prove all'interno di ciascun blocco veniva completamente randomizzato per ogni soggetto. Le variabili dipendenti considerate nell'analisi comprendevano la percentuale di fusione percepita e l'accuratezza nei compiti di reazione rapida, i punteggi medi di fusione soggettiva (compito F), i punteggi medi di gradevolezza edonica (compito P) e i relativi tempi di reazione registrati in millisecondi.

CAPITOLO 3 – ANALISI DEI DATI

Per l'analisi delle prove di discriminazione rapida (compiti BA ed AB) sono state calcolate le medie dei tempi di reazione (TR) dei partecipanti in base agli intervalli proposti della prova specifica, unitamente alle relative deviazioni standard. Sono stati esclusi dall'analisi tutti i trial non validi (*outlier*), ovvero le prove in cui il tempo di risposta superava i 2000 ms e quelle in cui la risposta era minore o uguale a 0 ms (risposte anticipate o nulle). Su un totale di 10.080 prove somministrate nei due blocchi di risposta rapida (360 per partecipante), i trial analizzati sono stati 9417. I dati si riferiscono ai TR di risposta registrati in millisecondi (vedi Tabella 2)

Tabella 2: Media e deviazione standard dei TR per ogni intervallo proposto

Intervalli	Media TR	Dev. Standard TR
0	1183.42	544.61
111	1196.21	602.49
204	1216.32	619.73
316	1210.51	590.17
386	1254.34	593.43
498	1232.91	585.42
590	1252.70	586.74
702	1310.97	592.14
814	1259.24	585.71
884	1238.30	593.75
1018	1213.29	577.07
1088	1230.68	605.91
1200	1291.24	579.35
1312	1211.67	579.35
1404	1194.92	538.07

Tra i vari intervalli, quello in cui i partecipanti hanno risposto più rapidamente è stato l'unisone (0 cents), con una media di 1183,42 ms. Al contrario, l'intervallo con TR in media più lunghi è stato la quinta giusta (702 cents) con 1310,97 ms, seguito dall'ottava (1200 cents) con 1291,24 ms. Raggruppando gli stimoli, gli intervalli considerati canonicamente consonanti (unisone, terze, quarta giusta, quinta giusta, seste e ottava) hanno mostrato tempi di reazione mediamente più lunghi con una media di 1247,51 ms (DS = 583,98) rispetto agli intervalli dissonanti (seconde, tritono, settime e none), i quali hanno fatto registrare risposte più rapide (M = 1216,59 ms, DS = 587,41).

Per il secondo compito sui giudizi di fusione (compito F, *fusion rating*) sono state considerate media e deviazione standard delle risposte espresse sulla scala Likert a 4 punti, in cui valori più elevati indicano un grado di fusione percepito maggiore tra le due componenti sonore (vedi Tabella 3).

Tabella 3: 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
111	2.13	1.06
204	2.29	1.00
316	2.49	0.89
386	2.58	0.93
498	2.77	0.92
590	2.55	0.98
702	2.98	0.94
814	2.47	0.98
884	2.64	0.92
1018	2.41	0.96
1088	2.32	0.94
1200	3.40	0.82
1312	2.38	0.96
1404	2.33	0.97

L'intervallo giudicato più fuso è risultato l'ottava (1200 cents), con una media di risposte pari a 3,40, seguito dalla quinta giusta (702 cents) con 2,98 e dalla quarta giusta (498 cents) con 2,77. Gli intervalli valutati come meno fusi sono stati la seconda minore (111 cents), con una media di 2,13, e la seconda maggiore (204 cents) con 2,29. Anche per la fusione tonale, gli intervalli canonicamente considerati consonanti hanno ricevuto giudizi di fusione mediamente più alti con una media di 2,76 (DS = 0,96) rispetto a quelli dissonanti che vedono una media di 2,34 (DS = 0,99).

Per il terzo compito sui giudizi di piacevolezza (compito P, *preference rating*) sono state considerate media e deviazione standard delle risposte fornite sulla scala likert a 4 punti, dove un punteggio più alto indicava una maggiore gradevolezza percepita nel compito (vedi Tabella 4).

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
111	1.79	0.95
204	1.86	0.93
316	2.07	0.96
386	2.38	1.01
498	2.62	1.05
590	2.25	0.97
702	2.64	1.03
814	2.24	1.00
884	2.51	1.10
1018	2.29	1.01
1088	2.12	1.04
1200	3.02	1.05
1312	2.07	0.94
1404	2.51	1.13

Anche sul piano della gradevolezza percepita, l'intervallo che è stato considerato più piacevole in assoluto è risultato l'ottava (1200 cents), con una media di 3,02, seguito dalla quinta giusta (702 cents) con 2,64 e dalla quarta giusta (498 cents) con 2,62. Dall'altro lato gli intervalli giudicati meno piacevoli sono stati la seconda minore (111 cents) con un punteggio medio di 1,79 e la seconda maggiore (204 cents) con 1,86. Considerando gli intervalli canonicamente consonanti e quelli canonicamente dissonanti, è possibile osservare come i primi abbiano ottenuto giudizi mediamente più alti ($M = 2,50$, $DS = 1,07$) rispetto ai secondi ($M = 2,13$, $DS = 1,02$).

I dati complessivi dei tre compiti sperimentali sono illustrati all'interno dei rispettivi grafici nella Figura 1, consentendo un confronto visivo diretto tra le tre misure analizzate.

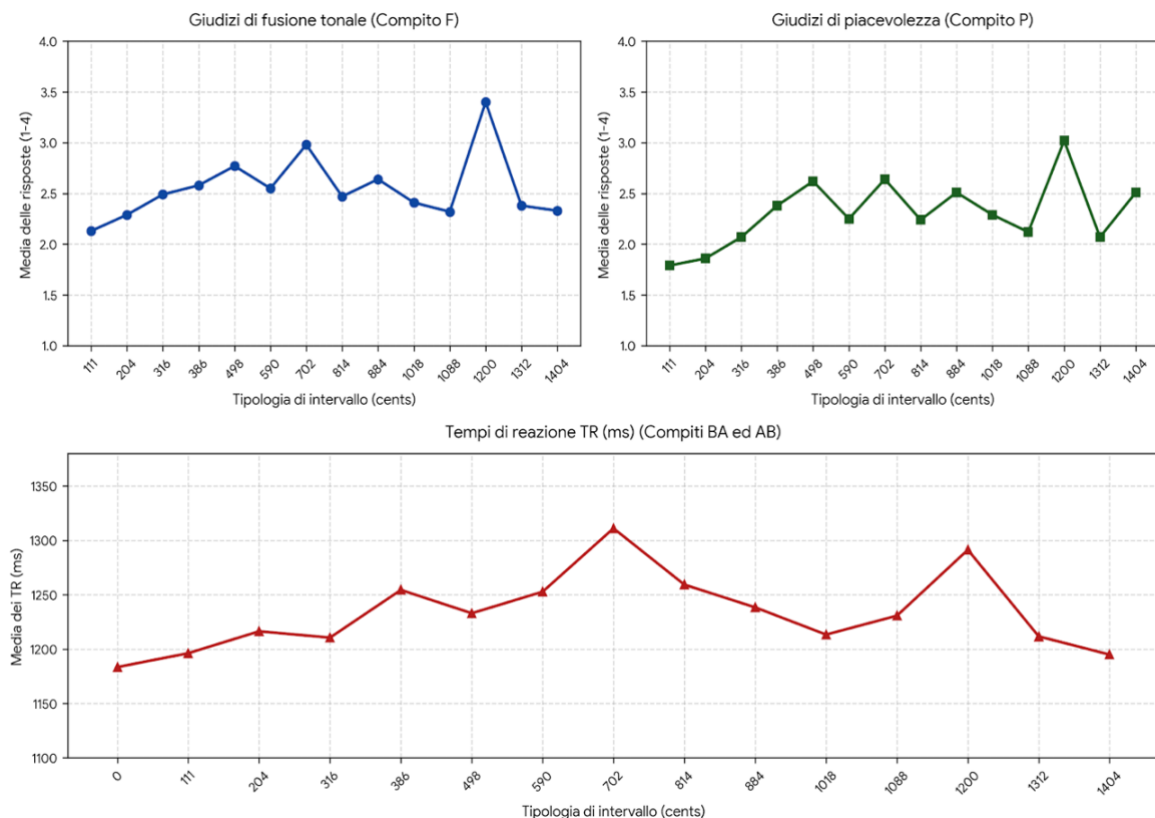


Figura 1: In alto i grafici delle medie dei giudizi di fusione tonale (a sinistra) e dei giudizi di piacevolezza (a destra) per ciascun intervallo considerato. In basso, grafico delle medie dei TR registrati nella discriminazione degli intervalli considerati

Oltre all'ampiezza dell'intervallo, l'analisi ha evidenziato un chiaro effetto della frequenza fondamentale (f_0) sulla piacevolezza dell'ascolto (vedi Tabella 5). Dai dati emerge una netta preferenza per i registri gravi e intermedi (110 Hz, 123,5 Hz, 220 Hz

e 246,9 Hz) con medie comprese tra i 2,50 e i 2,58. Quando viceversa gli stimoli vengono presentati nel registro acuto (440 Hz e 493,9 Hz), il gradimento scende fino a un minimo di 1,75. Questo andamento indica che le frequenze fondamentali elevate tendono ad accentuare l'asprezza del suono, riducendo la gradevolezza complessiva indipendentemente dalla consonanza dell'intervallo.

Tabella 5: Media e deviazione standard delle risposte alla scala Likert sul grado di piacevolezza in funzione della frequenza fondamentale (f_0)

F0	Media delle risposte	Dev. Standard delle risposte
110.0	2.58	1.07
123.5	2.50	1.07
220.0	2.57	1.01
246.9	2.53	1.01
440.0	1.94	0.95
493.9	1.75	0.93

Infine, sovrapponendo i giudizi medi ottenuti nei compiti di fusione tonale e di piacevolezza (Figura 2), emerge un parallelismo percettivo evidente.

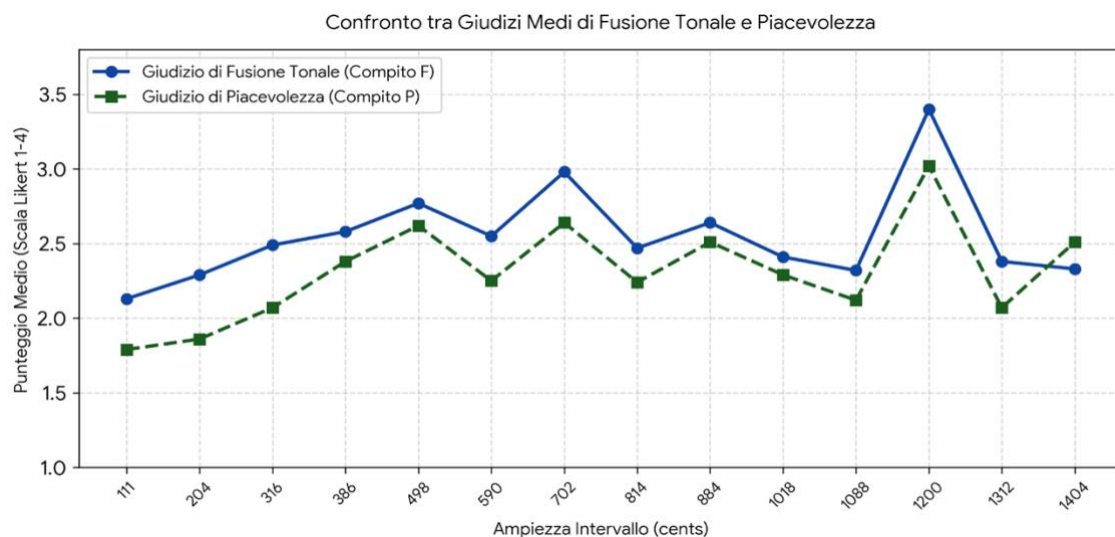


Figura 2: Confronto tra i giudizi medi di fusione tonale (compito F) e i giudizi di piacevolezza (compito P) per ciascun intervallo armonico.

Come illustrato nella Figura 2, l'andamento della curva di fusione ricalca abbastanza accuratamente le oscillazioni della curva di gradevolezza lungo l'intero asse. I picchi di massima fusione (ottava, quinta giusta, quarta giusta) corrispondono esattamente agli intervalli valutati come più gradevoli. Al contrario, gli intervalli a bassa fusione (seconde minori e maggiori) registrano i minimi di gradimento. Questa convergenza supporta l'ipotesi teorica di Stumpf, suggerendo che la sensazione soggettiva di consonanza musicale e il gradimento estetico siano radicati nei meccanismi di fusione e integrazione tonale del sistema uditivo.

CAPITOLO 4 – CONCLUSIONI

4.1 Discussione dei risultati

L'obiettivo di questa ricerca era verificare empiricamente l'ipotesi di Carl Stumpf, secondo cui la percezione di consonanza musicale è strettamente legata al fenomeno della fusione tonale. I risultati ottenuti confermano questa relazione attraverso due evidenze principali.

La prima riguarda la forte corrispondenza tra i giudizi di fusione e quelli di piacevolezza: gli intervalli valutati come maggiormente fusi coincidono con quelli giudicati più gradevoli, mentre gli intervalli a bassa fusione, come le seconde, si collocano ai livelli minimi di preferenza. Questo parallelismo supporta l'idea che la tendenza dei suoni a integrarsi in un'unità percettiva non sia un fenomeno secondario, ma una componente centrale nella valutazione estetica della consonanza.

La seconda evidenza emerge dall'analisi dei tempi di reazione nei compiti di risposta rapida. I partecipanti hanno impiegato più tempo a rispondere di fronte agli intervalli consonanti rispetto a quelli dissonanti, poiché l'elevata coesione acustica negli intervalli consonanti rende più difficile la separazione dei due toni e rallenta la risposta.

L'analisi dell'effetto del registro (f_0) aggiunge un ulteriore elemento: il netto calo di gradimento osservato per le frequenze fondamentali più acute (440 Hz e 493.9 Hz) indica che la piacevolezza d'ascolto è modulata anche dalle caratteristiche spettrali dello stimolo acustico, in linea con i modelli multifattoriali più recenti (McDermott et al., 2010).

4.2 Limiti dello studio e sviluppi futuri

Nonostante la coerenza delle evidenze raccolte, questo studio presenta alcuni limiti metodologici. Il primo riguarda la composizione del campione, costituito prevalentemente da ascoltatori non musicisti con un indice di Ollen medio-basso ($M = 2,11$). Indagini future potrebbero confrontare sistematicamente gruppi di musicisti professionisti e soggetti non formati, così da verificare se una specifica competenza musicale accresca la capacità di separare analiticamente le singole note senza alterare la preferenza estetica per gli intervalli con maggiore fusione.

Un altro limite è l'impiego esclusivo di toni sintetici complessi che, pur consentendo un controllo rigoroso delle proprietà acustiche, riducono la naturalezza dell'ascolto rispetto alla ricchezza timbrica degli strumenti reali o della voce umana. Estendere questo paradigma sperimentale a timbri naturali o a strutture armoniche più complesse permetterebbe di valutare il modello sviluppato da Stumpf in contesti musicali più realistici.

Dal punto di vista procedurale, l'adozione di una sequenza fissa dei compiti (BA ed AB, F, P) ha garantito la standardizzazione della somministrazione per tutti i partecipanti; tuttavia, l'introduzione di un controbilanciamento tra i blocchi consentirebbe di escludere eventuali effetti di abitudine o fatica cognitiva legati alla durata della sessione sperimentale.

Infine, sul piano degli sviluppi metodologici, l'affiancamento di misure comportamentali a tecniche di neuroimaging funzionale potrebbe chiarire con maggiore precisione la dinamica temporale con cui il sistema nervoso centrale integra il segnale sonoro, distinguendo l'elaborazione periferica della rugosità acustica dalla rappresentazione unitaria della fusione tonale.

BIBLIOGRAFIA

- Aristosseno. (1955). *Aristoxeni Elementa Harmonica* (R. da Rios, a cura di). Istituto Poligrafico dello Stato.
- Cazden, N. (1945). Musical Consonance and Dissonance: A Cultural Criterion. *The Journal of Aesthetics and Art Criticism*, 4(1), 3-11.
- DeWitt, L. A., & Crowder, R. G. (1987). Tonal fusion of consonant musical intervals: The oomph in Stumpf. *Perception & Psychophysics*, 41(1), 73-84.
- Di Stefano, N., & Bertolaso, M. (2014). Understanding musical consonance and dissonance: Epistemological considerations from a systemic perspective. *Systems*, 2(4), 566-575.
- Huron, D. (1991). Tonal consonance versus tonal fusion in single chordal sonances. *Music Perception*, 9(2), 135-154.
- Lundin, R. W. (1947). Toward a cultural theory of consonance. *The Journal of Psychology*, 23(1), 45-49.
- McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance preferences. *Current Biology*, 20(11), 1035-1041.
- McPherson, M. J., Grace, S. E., & McDermott, J. H. (2020). Harmonicity aids hearing in noise by facilitating perceptual grouping. *Nature Human Behaviour*, 4(11), 1146-1158.
- Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings* [Doctoral dissertation, Ohio State University]. The Ohio State University.
- Plomp, R., & Levelt, W. M. (1965). Tonal Consonance and Critical Bandwidth. *The Journal of the Acoustical Society of America*, 38(4), 548-560.
- Stumpf, C. (1890). *Tonpsychologie* (Vol. 2). S. Hirzel.
- von Helmholtz, H. (1954). *On the sensations of tone as a psychological basis for the theory of music* (A. J. Ellis, Trad.; opera originale pubblicata nel 1863). Dover Publications.