

ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA
SECONDA FACOLTÀ DI INGEGNERIA CON SEDE A CESENA
CORSO DI LAUREA MAGISTRALE IN INGEGNERIA BIOMEDICA

**INTERAZIONE VISUO-ACUSTICA E
FENOMENI DI PLASTICITÀ SINAPTICA:
STUDIO MEDIANTE UN MODELLO DI
RETE NEURALE APPLICATO AL
VENTRILOQUISMO**

Tesi in
Sistemi neurali LM

Relatore
Prof.ssa Elisa Magosso

Presentata da
Domenico Fiorentino

Sessione Terza
Anno Accademico 2011-2012

*Mamma, Papà,
questa tesi la dedico a voi.
Spero che siate orgogliosi di me.*

Sommario

Introduzione	1
---------------------------	----------

Capitolo 1

L'integrazione multisensoriale	5
---	----------

1.1 Aspetti generali.....	5
1.2 Meccanismi di integrazione multisensoriale	6
1.2.1 Convergenza anatomica di informazioni unimodali in aree multimodali ..	7
1.2.2 Effetti multisensoriali in regioni unimodali del cervello	13

Capitolo 2

Il fenomeno del ventriloquismo	17
---	-----------

2.1 Il ventriloquismo	17
2.2 L'aftereffect del ventriloquismo	22
2.3 Trasferimento attraverso le frequenze dell'aftereffect del ventriloquismo	25

Capitolo 3

Modello di rete neurale per il fenomeno del ventriloquismo	30
---	-----------

3.1 Aspetti generali.....	30
3.2 Struttura del modello	31
3.3 Caratteristiche dei neuroni della rete	36
3.4 Input incidente su un neurone della rete	37
3.4.1 Input esterno	37
3.4.2 Input laterale	39
3.4.3 Input cross-modale	42

3.5 Regole di apprendimento di Hebb	44
3.6 Calcolo della posizione percepita dello stimolo	47
3.7 Addestramento della rete	48
3.8 Assegnazione dei valori dei parametri	48

Capitolo 4

Risultati delle configurazioni per la non generalizzazione e la generalizzazione in frequenza dell'aftereffect 54

4.1 Aspetti generali.....	54
4.2 Risultati della configurazione per la non generalizzazione in frequenza dell'aftereffect	55
4.2.1 Pre-addestramento	55
4.2.2 Post-addestramento – Primo paradigma di training	67
4.2.3 Post-addestramento – Secondo paradigma di training	74
4.3 Risultati della configurazione per la generalizzazione in frequenza dell'aftereffect	85
4.3.1 Pre-addestramento	85
4.3.2 Post-addestramento – Primo paradigma di training	94
4.3.3 Post-addestramento – Secondo paradigma di training	101

Capitolo 5

Risultati della configurazione per l'influenza dell'intensità dell'input acustico sull'aftereffect 110

5.1 Aspetti generali.....	110
5.2 Pre-addestramento	110
5.2.1 Costruzione della FRA	110
5.2.2 Stimolazione cross-modale e ventriloquismo	113
5.3 Post-addestramento.....	115
5.3.1 Valutazione dell'aftereffect	116
5.3.2 Verifica del trasferimento dell'aftereffect attraverso le frequenze	118

Capitolo 6

Conclusioni 120

Bibliografia 127

Introduzione

Una delle caratteristiche fondamentali dei sistemi biologici è la capacità di combinare informazioni multisensoriali per ottenere una percezione completa della realtà. Infatti, oggetti ed eventi della vita quotidiana producono spesso segnali che coinvolgono diverse modalità percettive; di conseguenza, il sistema nervoso dei vertebrati si è evoluto al fine di trasdurre i differenti tipi di energia ambientale, per poi integrarli e produrre una rappresentazione unificata del mondo reale.

Un esempio di integrazione multisensoriale è quella che interessa il sistema visivo e il sistema uditivo. Tra i vari aspetti dell'interazione visuo-acustica, uno particolarmente analizzato nel corso degli anni è l'effetto di "cattura" che la vista esercita nei confronti dell'udito per quel che concerne la localizzazione di oggetti nello spazio. Quest'azione di cattura si verifica in quanto il sistema visivo ha una risoluzione spaziale migliore rispetto a quello acustico, e porta di solito alla generazione di illusioni nella percezione. L'esempio classico di questa influenza intersensoriale è rappresentato dal fenomeno del *ventriloquismo*. Esso si manifesta quando a un soggetto vengono presentati uno stimolo visivo e uno sonoro temporalmente coincidenti, ma spazialmente separati. A causa del maggior potere risolutivo del sistema visivo in condizioni normali, il soggetto percepisce la posizione dell'input acustico come se fosse spostata verso quella da cui proviene il segnale visivo. Questo fenomeno è definito appunto con il termine di ventriloquismo, in quanto una delle sue principali manifestazioni è l'illusione creata dal ventriloquo: quando quest'ultimo muove la bocca di un pupazzo e contemporaneamente emette dei suoni senza movimenti visibili della labbra, tali suoni vengono percepiti come se provenissero dalla bocca in movimento del pupazzo.

Il fenomeno del ventriloquismo, oltre all'effetto on-line appena descritto, ne produce uno off-line, definito *aftereffect del ventriloquismo*. Tale effetto si ottiene nel caso in cui un soggetto viene ripetutamente esposto alla situazione di conflitto audio-visiva: se gli vengono presentati costantemente un segnale acustico e uno visivo simultanei, ma in due distinte posizioni dello spazio, quando, dopo il periodo di esposizione al conflitto, viene applicato uno stimolo unicamente sonoro, il soggetto lo avvertirà come se fosse ancora spostato verso la posizione dell'input visivo. Ciò indica che a

causa dell'esposizione al conflitto cross-modale, avviene una ricalibrazione a lunga durata nella percezione dello spazio acustico.

Un aspetto significativo e non ancora chiarito dell'aftereffect del ventriloquismo è la sua generalizzazione attraverso le frequenze sonore. Infatti, non è ancora stato pienamente compreso se tale effetto off-line si verifichi anche presentando un input acustico a una frequenza diversa da quella utilizzata durante il periodo di esposizione al conflitto. In particolare, risultati contrastanti sono presenti nella letteratura scientifica: secondo alcune ricerche l'aftereffect rimane confinato alla frequenza di adattamento [12,18,19], mentre in base ai risultati di altri studi, esso si generalizza per un range di frequenza più ampio [20,21].

Il presente lavoro di tesi è nato perciò con l'obiettivo principale di esaminare questo aspetto dell'aftereffect, in particolare i meccanismi sinaptici alla base del fenomeno e i possibili fattori che influenzano la sua eventuale generalizzazione in frequenza. Una migliore comprensione di questi aspetti, infatti, può essere utile per fare maggiore chiarezza su come avviene l'interazione tra stimoli visivi e stimoli acustici nella corteccia cerebrale.

Lo studio è stato condotto mediante un modello di rete neurale realizzato per indagare le caratteristiche del trasferimento in frequenza o meno dell'aftereffect.

Tale modello prende spunto da un modello di rete neurale precedentemente sviluppato dai ricercatori dell'università di Bologna [22]. La rete costruita nel precedente lavoro era composta da due strati di neuroni, uno rappresentativo dell'area acustica e l'altro dell'area visiva. Ognuno dei due strati era costituito da una catena monodimensionale di neuroni. Ciascun neurone del modello codificava per una certa posizione dello spazio; i neuroni erano collegati da sinapsi laterali se appartenenti allo stesso strato, da connessioni inter-area se situati in aree diverse. Il modello così realizzato era in grado di riprodurre l'effetti on-line del ventriloquismo (spostamento della percezione dello stimolo acustico verso la posizione del contemporaneo input visivo) e il suo effetto off-line (shift spaziale dello stimolo acustico senza input visivo a seguito di un periodo di esposizione alla situazione di conflitto cross-modale). Per riprodurre gli effetti off-line del ventriloquismo, il modello utilizzava regole di apprendimento sinaptico, a cui erano sottoposte le sinapsi laterali durante l'esposizione alla situazione di conflitto. Questo modello, tuttavia, presentava il limite di non tenere in considerazione le caratteristiche in

frequenza dei neuroni uditivi; di conseguenza non esaminava la possibilità di generalizzazione dell'aftereffect per le frequenze sonore.

Il precedente modello è stato dunque modificato e arricchito in modo tale da considerare anche la risposta in frequenza dei neuroni acustici e comprendere meglio i risultati riguardanti il trasferimento in frequenza dell'aftereffect presenti in letteratura [12,18-21]. In particolare, lo strato acustico, da monodimensionale, è stato reso bidimensionale, in maniera tale che i neuroni uditivi potessero codificare sia per la posizione spaziale che per la frequenza sonora; in seguito a tale modifica, sono state apportate significative variazioni anche alle connessioni sinaptiche.

Alla luce di queste modifiche, la rete è stata quindi valutata per la sua capacità di simulare sia l'effetto on-line del ventriloquismo sia l'aftereffect di tale fenomeno. Per considerare il secondo aspetto, è stato considerato un paradigma di addestramento (ispirato da quelli utilizzati negli studi di letteratura [12,18-21]), tramite il quale la rete è stata sottoposta a ripetute situazioni di conflitto spaziale visuo-acustico. Durante la fase di addestramento, le sinapsi laterali nelle aree acustica e visiva sono state modificate sulla base di regole di apprendimento simili a quelle utilizzate nel precedente lavoro [22].

Il modello è stato testato in tre diverse configurazioni, ciascuna delle quali corrisponde a un certo set di parametri, con lo scopo di simulare la generalizzazione e la non generalizzazione in frequenza dell'aftereffect del ventriloquismo, e interpretare le discordanze presenti nei risultati riportati in letteratura, relativi a esperimenti su soggetti umani.

La trattazione della presente tesi è suddivisa in sei capitoli.

Il Capitolo 1 fornisce informazioni generali sui principali aspetti dell'integrazione multisensoriale; vengono infatti discussi i possibili meccanismi alla base di tale processo ed analizzate le aree anatomiche probabilmente coinvolte nell'interazione cross-modale, in riferimento ai risultati riportati in letteratura [1-11].

Il Capitolo 2 affronta gli aspetti caratteristici del fenomeno del ventriloquismo: effetto on-line, aftereffect, eventuale trasferimento di quest'ultimo attraverso le frequenze sonore. Questi aspetti vengono trattati con riferimento alle evidenze sperimentali della letteratura fisiologica e psicofisica [12-21].

Nel Capitolo 3 viene descritto dettagliatamente il modello neurale utilizzato in questa tesi, mettendo in evidenza le modifiche e gli ampliamenti apportati rispetto al

modello originale [22]. In particolare, viene dapprima descritta la sua architettura, si passa quindi alle equazioni matematiche che lo costituiscono, per terminare infine con una spiegazione sulla scelta dei valori dei parametri adottati.

Il Capitolo 4 riporta i risultati prodotti dal modello per quanto riguarda due delle tre configurazioni di parametri che sono state esaminate. Tali risultati vengono distinti in quelli ottenuti prima di addestrare la rete e quelli ottenuti post-adattamento, per analizzare le proprietà dell'aftereffect.

Il Capitolo 5 fornisce i risultati relativi al terzo set di parametri, con la stessa distinzione fatta nel precedente capitolo.

Il Capitolo 6 è di tipo conclusivo: vengono sintetizzati i risultati ottenuti, sono esposti i contributi innovativi del modello, ne vengono discussi i limiti e infine sono delineati i possibili sviluppi futuri.

Capitolo 1

L'integrazione multisensoriale

1.1 Aspetti generali

Gli eventi che avvengono nella vita di tutti i giorni spesso forniscono contemporaneamente informazioni che coinvolgono vie sensoriali differenti. Si pensi ad esempio alla rottura del vetro di una finestra: sia l'immagine del vetro che va in frantumi sia il rumore provocato dalla sua rottura indicano il verificarsi dello stesso evento. Questa è una chiara dimostrazione di come stimoli che interessano due diverse modalità sensoriali agevolino la percezione di ciò che accade nell'ambiente circostante. Ci sono situazioni in cui il coinvolgimento sincrono di due sensi diventa fondamentale per la giusta comprensione di un certo evento. Si faccia infatti riferimento al caso banale di guardare una partita di tennis in televisione. Se con il telecomando viene tolto l'audio, ci si accorge che diventa più difficile seguire le varie fasi di gioco, dato che non è più possibile percepire l'impatto della palla con la racchetta e con il suolo. La situazione diviene ancora più complicata se, per assurdo, si prova a giocare a tennis con le orecchie tappate; questa volta non solo non si sente l'impatto della palla, ma diventa arduo coordinare le proprie azioni in funzione di ciò che sta accadendo. Ecco perché è importante disporre di informazioni provenienti da più modalità percettive ed ecco perché il sistema nervoso degli animali superiori si è evoluto al fine di integrare gli input derivanti da differenti sorgenti e produrre così una percezione unificata della realtà.

Sebbene oggi sia ampiamente riconosciuto che il cervello umano combina informazioni di diverse modalità sensoriale, tuttavia ci sono voluti molti anni prima che si iniziasse a parlare di *integrazione multisensoriale*. Infatti, per molto tempo i diversi canali sensoriali sono stati studiati separatamente e considerati indipendenti in quanto deputati al processamento di forme di energia differenti. Tradizionalmente si ipotizzava che l'integrazione tra le informazioni elaborate indipendentemente dalle diverse modalità avvenisse solo in aree corticali associative di alto livello, per lo

svolgimento di compiti cognitivi complessi. Negli ultimi decenni ha invece preso piede la convinzione secondo cui esistono delle interazioni tra differenti modalità percettive, con conseguente interesse, in numerosi settori di ricerca, nel definire i processi neurali implicati nella sintesi dell'informazione proveniente da stimoli cross-modali. Grazie a tutta questa serie di studi si può quindi affermare che l'essere umano è un organismo essenzialmente ed intrinsecamente multimodale: le sue relazioni con il mondo e le sue azioni in esso si basano sulla possibilità di cogliere i fenomeni integrativi della realtà.

1.2 Meccanismi di integrazione multisensoriale

I segnali provenienti da differenti modalità sensoriali vengono inizialmente elaborati in regioni del cervello separate. Tuttavia, poiché segnali diversi derivano spesso da eventi comuni, il cervello mette in atto una serie di meccanismi che costituiscono quella che viene generalmente definita integrazione multisensoriale. Diversi studi psicologici hanno mostrato che stimoli che coinvolgono modalità percettive distinte possono interagire potentemente in alcune circostanze, in modo tale da determinare la percezione e il conseguente comportamento. Esistono infatti vari esempi di illusioni cross-modali, come l'effetto del ventriloquo (argomento centrale di questa tesi), per cui la posizione di un suono è percepita come se fosse spostata verso la posizione di un input visivo concomitante, oppure l'effetto McGurk, per cui se si guardano delle labbra che scandiscono una certa sillaba mentre se ne ascolta un'altra, il suono percepito è una via di mezzo delle due. Sebbene queste illusioni percettive siano provocate da combinazioni di stimoli anomale e incongruenti, molti altri effetti mostrano l'impatto dell'integrazione multisensoriale sulla percezione. Ad esempio, la capacità di rilevazione di un oggetto da parte degli occhi migliora se quell'oggetto emette qualche suono. Inoltre nel mondo reale, segnali che interessano sensi diversi ma scaturiti dallo stesso oggetto o evento, spesso sono temporalmente e spazialmente coerenti.

Le basi neurali delle interazioni multisensoriali sono state studiate utilizzando registrazioni intracraniche negli animali e misure elettrofisiologiche o emodinamiche negli esseri umani. Sono stati così trovati neuroni che rispondono a stimoli di natura diversa in regioni corticali e subcorticali. Le regioni multisensoriali corticali includono numerose aree della corteccia parietale, come l'area intraparietale ventrale, della corteccia temporale, come la regione polisensoriale temporale superiore, e della corteccia frontale, come la corteccia premotoria ventrale [1]. Le regioni multisensoriali subcorticali comprendono il collicolo superiore, i gangli della base e il putamen [2]. Recenti studi indicano non solo che l'integrazione multisensoriale riguarda la "convergenza anatomica" di cortecce sensori-specifiche (unimodali) in aree del cervello multisensoriali (eteromodali), ma anche che le interazioni multisensoriali spaziali possono influenzare le regioni unimodali del cervello [3].

1.2.1 Convergenza anatomica di informazioni unimodali in aree multimodali

Tra le diverse tipologie di neuroni, ne esistono alcuni che sono i principali protagonisti dell'integrazione cross-modale, i "neuroni multisensoriali": essi infatti rispondono a stimoli provenienti da più di una modalità percettiva. Queste cellule nervose sono presenti in tutti i mammiferi in molte aree cerebrali, corticali e sottocorticali, che ricevono afferenze da aree sensoriali unimodali.

1.2.1.1 Il collicolo superiore

I neuroni multisensoriali sono particolarmente abbondanti nel collicolo superiore del mesencefalo (Figura 1.1). Tale struttura è largamente studiata in neurofisiologia, essendo una fonte di informazione estremamente ricca circa le proprietà dei neuroni multisensoriali [4]. Il collicolo superiore è una struttura sottocorticale mesencefalica

coinvolta nell'orientamento, nell'attenzione, nel controllo dei movimenti oculari e i processi di selezione del target [5]. Ha un ruolo specifico nell'iniziazione e nel controllo dei movimenti di occhi e testa (per esempio, gli spostamenti dello sguardo), funzione che svolge in virtù delle proiezioni ai neuroni del tronco cerebrale e del cordone spinale che governano l'attività muscolare. I suoi input visivi, acustici e somatosensoriali sono derivati da fibre sensoriali ascendenti e da proiezioni discendenti dalla corteccia, che convergono in vari modi sui neuroni del collicolo superiore.

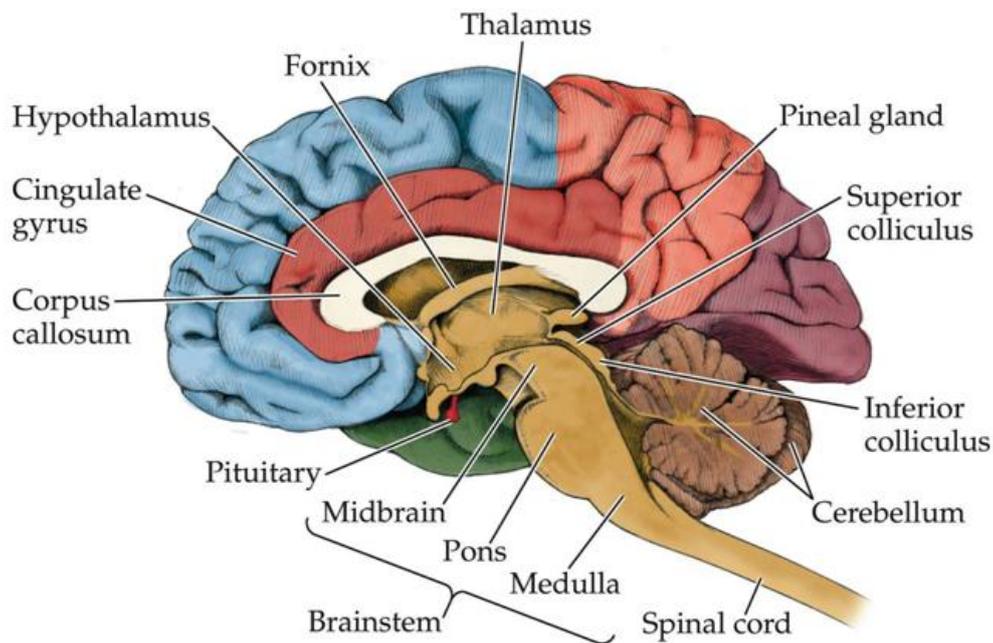


Figura 1.1. Vista sagittale del cervello umano.

Il principale vantaggio che deriva dalla presenza nel sistema nervoso di neuroni multisensoriali è la possibilità di integrare informazioni provenienti da fonti diverse. Affinché si realizzi una vera sintesi, la risposta ad uno stimolo multisensoriale deve differire da tutte quelle suscitate dalle sue componenti modalità-specifiche. Pertanto, l'integrazione multisensoriale viene definita a livello operativo come “la differenza statisticamente significativa tra il numero di impulsi evocati da una combinazione cross-modale di stimoli e il numero di impulsi evocati dal più efficace di questi stimoli preso individualmente” [5]. Descritta così, l'integrazione multisensoriale è una misura dell'efficacia relativa di una stimolazione

multisensoriale rispetto ad una stimolazione modalità-specifica, la cui conseguenza è un incremento della probabilità di rilevazione ed iniziazione della risposta verso la fonte del segnale multisensoriale. L' aumento della risposta multisensoriale è rappresentato dalla formula:

$$(CM-SM_{max}) \times 100 / SM_{max}$$

dove:

CM: numero di impulsi evocati dalla combinazione di stimoli sensoriali differenti

SM_{max}: risposta indotta dallo stimolo unimodale più efficace

1.2.1.2 Aree corticali multisensoriali

L'approccio più semplice per capire quali siano le aree multisensoriali del cervello umano utilizzando tecniche di neuroimaging è misurare l'attività cerebrale durante la stimolazione di particolari modalità sensoriali e quindi valutare se qualche regione risponde alla stimolazione di più vie percettive. Usando stimoli visivi, uditivi e tattili, Bremmer e colleghi [6] hanno così identificato le risposte multisensoriali per il solco intraparietale, il lobulo parietale inferiore e la corteccia premotoria ventrale (Figura 1.2 [5]).



Figura 1.2. Aree unimodali e multimodali del cervello.

Un altro studio ha sfruttato la risonanza magnetica funzionale per indagare le attivazioni posizione-dipendenti per tatto e vista [7]. In tale studio alcuni soggetti ricevevano passivamente degli stimoli tattili sulla mano sinistra o su quella destra, o brevi flash luminosi in stretta corrispondenza spaziale per entrambe le mani. Dai dati ottenuti tramite risonanza magnetica è risultato che le aree occipitali hanno espresso un'attivazione spazialmente specifica nell'emisfero controlaterale alla stimolazione visiva, ma nessuna risposta durante la stimolazione tattile.; il solco intraparietale invece ha mostrato risposte nell'emisfero controlaterale spazialmente specifiche, indipendentemente dalla modalità stimolata (cioè vista o tatto); inoltre regioni delle cortecce frontale, parietale e temporale posteriore hanno risposto in maniera multimodale, indipendentemente dalla posizione dello stimolo. Questi risultati indicano una convergenza di tipo feedforward delle informazioni provenienti da cortecce sensori-specifiche (ovvero quelle occipitale visiva e post-centrale somatosensoriale) in regioni corticali eteromodali (cioè il solco intraparietale).

1.2.1.3 Leggi dell'integrazione multisensoriale

Gli studi neurofisiologici in animali (grazie soprattutto all'analisi delle proprietà dei neuroni del collicolo superiore) hanno evidenziato come, a livello cellulare, gli input che arrivano dalle diverse modalità vengono integrati dai neuroni multisensoriali sulla base di tre principi generali, noti come “leggi integrative” [5], che sono:

- legge spaziale
- legge temporale
- legge dell'efficacia inversa

Legge spaziale

Per la prima legge, solo stimoli spazialmente coincidenti vengono integrati, aumentando in tal modo la risposta neuronale; al contrario, stimoli spazialmente distanti suscitano risposte di depressione, o non determinano integrazione, non producendo pertanto alcuna integrazione a livello cellulare. La proprietà spaziale

dipende dall'organizzazione dei campi recettivi dei neuroni multisensoriali in zone centrali, eccitatorie, e zone periferiche, inibitorie (il campo recettivo è una regione dello spazio nella quale deve essere localizzato uno stimolo sensoriale affinché un neurone possa rispondere). I campi recettoriali acustici presentano regioni eccitatorie più ampie rispetto a quelli visivi, con confini laterali che possono estendersi ben oltre i corrispondenti campi recettivi visivi, mentre la discrepanza del campo recettivo visuo-acustico è minore all'altezza dei bordi mediali. L'estensione asimmetrica del bordo laterale è giustificata dal fatto che la rappresentazione dello spazio acustico si estende ben oltre la rappresentazione dello spazio visivo (il campo visivo si ferma a circa 100° , mentre lo spazio acustico corre tutto intorno alla testa); di conseguenza, uno stimolo visivo a 45° nello spazio visivo periferico può essere potenziato da uno stimolo acustico a 90° perché entrambi gli stimoli cadono all'interno dei rispettivi campi recettivi eccitatori degli stessi neuroni multisensoriali. Al contrario, se l'input acustico è in posizione centrale rispetto al target visivo (ad esempio, lo stimolo acustico è posizionato a 0° e il visivo a 45°), è probabile che esso cada oltre il confine mediale dei campi recettivi acustici di quei neuroni multisensoriali che sono attivati dallo stimolo visivo; è inoltre probabile che lo stimolo acustico cada entro la regione inibitoria che delimita molti di questi neuroni. Il risultato sarà una depressione dell'attività del collicolo superiore evocata visivamente, ed una minore probabilità di evocare chiare risposte. Poiché i campi recettivi acustici e visivi dei neuroni bimodali si sovrappongono in parte, stimoli visuo-acustici spazialmente coincidenti vengono a cadere all'interno delle zone eccitatorie, producendo l'"enhancement" (incremento) della risposta multisensoriale. Se gli stimoli sono spazialmente distanti, uno potrebbe cadere all'interno della regione recettiva inibitoria e deprimere gli effetti dell'altro, oppure potrebbe essere processato come un evento separato. Precisamente, quando uno dei due stimoli cade all'interno della regione inibitoria del suo campo recettivo, la risposta del neurone bimodale viene depressa, mentre quando entrambi gli stimoli cadono dentro la regione inibitoria dei loro rispettivi campi recettivi essi non vengono integrati, e la risposta bimodale del neurone non risulta differente da quella che verrebbe determinata da stimoli unimodali presentati isolatamente.

Legge temporale

Secondo la legge temporale, la massima interazione multisensoriale si ottiene a seguito della presentazione di stimoli temporalmente coincidenti, mentre stimoli separati da un intervallo temporale sono processati come eventi indipendenti. In base alla regola temporale, livelli massimi di enhancement sono ottenuti qualora due input siano presentati simultaneamente, o nel caso in cui la loro disparità temporale sia ridotta al minimo. Sebbene sia stato rilevato che la coincidenza temporale rappresenti il periodo interattivo ottimale per la maggior parte dei neuroni multisensoriali, ciò non è da considerarsi una regola; in alcuni neuroni, ad esempio, la combinazione di stimoli unimodali acustici e visivi a determinati intervalli (50 e 150 ms) produce ugualmente un enhancement della risposta. Sembrerebbe che l'intervallo temporale ottimale per l'integrazione si aggiri attorno ai 100 ms. Ciò che induce l'aumento della risposta multisensoriale è la sovrapposizione dei picchi dei periodi di scarica indotti da ciascuno stimolo unimodale sulla membrana neuronale. Al contrario, per intervalli più lunghi (200 e 300 ms), la stimolazione visuo-acustica provoca una depressione della risposta dell'attività del neurone, o l'assenza di interazione. Il sistema multisensoriale è dunque organizzato in modo tale che le risposte di eccitazione e depressione siano generate solo in virtù delle relazioni spaziali e temporali tra gli stimoli, e non della loro modalità di appartenenza.

Legge dell'efficacia inversa

Per la legge dell'efficacia inversa, la risposta multisensoriale è massima combinando stimoli unimodali deboli, rispetto alla combinazione di stimoli unimodali potenti. In accordo con la terza legge si stabilisce una relazione inversa tra il livello di efficacia degli stimoli e la risposta neurale da loro evocata. La combinazione di stimoli unimodali deboli produce una risposta di enhancement maggiore rispetto alla combinazione di stimoli unimodali potenti; ciò significa che la combinazione di due stimoli unimodali, nessuno dei quali, singolarmente preso, è in grado di evocare un significativo effetto sull'attività del neurone, può aumentare potentemente la risposta nei neuroni multisensoriali. La conseguenza adattiva della terza legge è che l'integrazione multisensoriale risulta maggiormente vantaggiosa per la rilevazione di fonti di stimolazione multimodale costituite da stimoli unimodali deboli. Nelle

condizioni in cui è possibile rilevare la presenza di un evento sensoriale attraverso una sola modalità sensoriale, la componente informativa multisensoriale è minima, in quanto superflua. Qualora, al contrario, l'informazione sensoriale proveniente da un sistema sensoriale risulta ambigua o degradata, la sua salienza è aumentata dalla presenza di informazioni provenienti da altri sistemi sensoriali.

1.2.2 Effetti multisensoriali in regioni unimodali del cervello

Recenti studi nelle scimmie e in soggetti umani hanno evidenziato l'esistenza di una "convergenza multisensoriale" nelle aree a basso livello unisensoriali del pathway corticale.

1.2.2.1 Interazioni visuo-tattili

In recenti studi basati sempre sulla risonanza magnetica funzionale [3], sono stati effettuati degli esperimenti per valutare l'azione di regioni eteromodali rispetto ad aree unimodali. In un primo esperimento, i soggetti vedevano brevi flash luminosi combinati con stimoli tattili (vibrazioni). I risultati hanno mostrato che aggiungendo un tocco spazialmente congruente cresceva l'attività occipitale controlaterale al flash. Questo fatto è degno di nota, in quanto mostra che la corteccia occipitale, tradizionalmente considerata sensibile solo agli stimoli visivi, può essere influenzata anche da stimoli di diversa modalità sensoriale. In un altro esperimento, le risposte tattili controlaterali nell'operculum parietale (che include la corteccia somatosensoriale secondaria, la quale risponde agli stimoli tattili ma non a quelli visivi) sono state analogamente incrementate aggiungendo uno stimolo visivo nella stessa posizione del tocco. L'attività massima nell'operculum parietale è stata osservata per la stimolazione tattile controlaterale quando questa era accoppiata con uno stimolo visivo alla stessa posizione. Questi risultati indicano che gli effetti multisensoriali non incidono solamente sulle regioni eteromodali, come per esempio

il solco intraparietale, ma anche su aree apparentemente unimodali come la corteccia visiva primaria.

1.2.2.2 Influenze visive e tattili sulla corteccia uditiva

Risposte somatosensoriali sono state osservate nell'area belt della corteccia uditiva, oppure è stato rilevato che la posizione degli occhi è in grado di modulare le risposte uditive a livello della corteccia uditiva primaria. RegISTRAZIONI intracraniche in diverse specie di scimmie confermano la convergenza multisensoriale nella corteccia uditiva, mostrando input somatosensoriali e visivi nelle regioni posteriori alla corteccia uditiva primaria. Sono diversi i casi di risposte somatosensoriali nella corteccia uditiva. Per esempio Leinonen et al. [8] hanno riportato una corappresentazione uditiva-somatosensoriale nell'area temporo-parietale dei macachi. Sempre nei macachi, Robinson e Burton [9] hanno descritto una mappa corporea in una regione retro insulare mediale del piano temporale superiore. Tutti questi risultati indicano che in qualche modo la corteccia uditiva riceve degli input che non sono solo acustici [10].

1.2.2.3 Meccanismi anatomici di convergenza multisensoriale nella corteccia uditiva

I pattern di connettività delle cortecce uditive posteriori forniscono diversi percorsi, sia corticali che subcorticali, attraverso cui gli input visivi e somatosensoriali possono convergere nella corteccia uditiva [10]. Molte aree sono potenziali sorgenti corticali di input somatosensoriali, come la corteccia somatosensoriale secondaria, l'area parietale ventrale, il solco intraparietale e la corteccia prefrontale. Si ritiene invece che gli input visivi per le interazioni audio-visive nascano nella corteccia uditiva e che siano mediati da proiezioni a feedback provenienti dall'area polisensoriale temporale superiore. Inoltre recenti studi condotti in diversi laboratori indicano proiezioni che partono da regioni uditive, come la corteccia uditiva primaria, e arrivano alle cortecce visive primaria e secondaria. Ci sono anche regioni

subcorticali che vengono considerate potenziali sorgenti di input somatosensoriali e visivi per la corteccia uditiva, in particolare strutture talamiche, come i nuclei posteriore, posteriore mediale e sopragenicolato del talamo [10].

1.2.2.4 Implicazioni funzionali della convergenza multisensoriale nell'elaborazione corticale a basso livello

Una probabile funzione della convergenza degli input visivi e somatosensoriali è il miglioramento dell'analisi effettuata dal sistema uditivo per un certo stimolo. Ma in che modo l'elaborazione acustica trae vantaggio da un input che non è acustico? Si possono considerare due ipotesi in proposito. La prima ipotesi è che gli input somatosensoriali e visivi, grazie alla loro precisione spaziale maggiore, aiutino la localizzazione spaziale uditiva. Questo è in accordo con la proposta di Rauschecker [11], secondo cui le regioni corticali uditive caudali sono specializzate per la localizzazione dello stimolo e con gli studi di Leiononen et al. [8] che riportano una corrispondenza spaziale tra i campi recettivi uditivi e somatosensoriali dei neuroni bimodali nell'area temporo-parietale. La seconda ipotesi consiste nel fatto che gli input visivi e somatosensoriali per la corteccia uditiva resettino l'attività corticale uditiva in corso, così da potenziare la risposta locale al successivo input acustico. Questo meccanismo è confermato dal fatto che, in una miriade di situazioni quotidiane, stimoli non acustici sono generati prima del segnale uditivo, dato che per produrre un suono è spesso necessaria un'azione visibile o palpabile. Per esempio, è necessario muovere il martello prima che esso possa colpire il chiodo per produrre un rumore, oppure molto spesso l'emissione di un certo suono da parte di un soggetto è preceduta da movimenti labiali o facciali.

1.2.2.5 Osservazioni conclusive

Anche se i segnali che coinvolgono diverse vie sensoriali sono inizialmente elaborati in diverse regioni del cervello in una sorta di "segregazione funzionale", tali segnali vengono poi combinati fino a interagire in maniera cross-modale. Recenti studi di

neuroimaging sulle interazioni cross- modali spaziali indicano che esiste una sorta di convergenza feedforward, cioè un link tra aree unimodali ed eteromodali, con le prime che si proiettano verso le seconde [3]. Tuttavia questo legame non è l'unico meccanismo alla base dell'integrazione funzionale, infatti alcune ricerche neurofisiologiche hanno messo in evidenza che stimolare una modalità sensoriale può avere un'influenza spaziale su zone del cervello responsabili di altre modalità [10]. Questo comportamento può essere dovuto ad un duplice meccanismo:

- a) regioni multimodali del cervello possono influenzare, tramite "retro-proiezioni" (sinapsi feedback) le aree sensori-specifiche;
- b) regioni unimodali, tradizionalmente considerate separate e indipendenti, possono essere collegate direttamente attraverso sinapsi reciproche, sicché la risposta ad uno stimolo in una modalità può essere direttamente modulata da un stimolo concomitante in una differente modalità percettiva.

Capitolo 2

Il fenomeno del ventriloquismo

2.1 Il ventriloquismo

La maggior parte della conoscenza sull'interazione cross-modale deriva da esperimenti che generano una situazione di conflitto tra due modalità sensoriali per indagare come l'evento è percepito da un osservatore [12]. Tali conflitti si traducono spesso in illusioni della percezione. Un classico esempio, nel caso dell'interazione visuo-acustica, è rappresentato dal fenomeno del "ventriloquismo". Quest'ultimo mette in evidenza un'azione di "cattura" che il senso della vista esercita nei confronti del senso dell'udito. Infatti esso si manifesta quando uno stimolo acustico e uno visivo vengono presentati simultaneamente, ma in posizioni leggermente diverse; in questo caso l'osservatore percepisce i due segnali come se provenissero entrambi dalla posizione dell'input visivo. L'effetto appena descritto è stato definito "ventriloquismo" in quanto la sua manifestazione più evidente è rappresentata proprio dall'azione del ventriloquo, che muovendo con una mano la bocca di un pupazzo mentre emette dei suoni senza quasi muovere le labbra, produce l'illusione per cui i suoni emessi provengono dalla bocca del pupazzo anziché dalla sua stessa bocca. Quanto appena descritto è lo stesso effetto che si sperimenta ogni volta che si guarda un film, in cui le parole sembrano provenire dagli attori sullo schermo e non dagli altoparlanti o dalle casse audio.

Alcune ricerche hanno evidenziato come ci siano diversi fattori che possono rendere l'illusione sopra citata più forte o più lieve [13]. In particolare esistono fattori cognitivi e non-cognitivi che influenzano l'incidenza e l'intensità del fenomeno. Il principale fattore cognitivo è rappresentato dall' "ipotesi di unità", in base alla quale per un soggetto vale l'ipotesi a priori per cui due stimoli concomitanti fanno parte dello stesso evento. Per quanto riguarda invece i fattori non-cognitivi, essi sono sostanzialmente tre: il convincimento, la tempistica e l'influenza spaziale. Il convincimento consiste nel fatto che è importante che il suono corrisponda a quello

che l'ascoltatore si aspetta che provenga dall'oggetto osservato (cioè, toni profondi da oggetti larghi e toni alti da oggetti piccoli). Per la tempistica, deve esserci una stretta sincronia temporale tra stimolo acustico e stimolo visivo (per esempio, tornando al caso del ventriloquo, i movimenti della bocca del pupazzo devono essere coerenti con ciò che si sente). Per l'influenza spaziale, se i due segnali sono troppo distanti spazialmente, l'illusione viene meno (quindi la distanza tra ventriloquo e pupazzo deve essere relativamente piccola).

Il fenomeno del ventriloquismo avviene perché, in condizioni normali, l'acuità di localizzazione spaziale del sistema visivo è maggiore rispetto a tutti gli altri sensi. Diverse ricerche hanno infatti evidenziato che la modalità sensoriale che domina la percezione di un oggetto è quella che risulta la più affidabile per una data discriminazione [12]. Perciò, per soggetti con un'acutezza visiva normale, la stima della posizione di un oggetto da parte della vista è più affidabile rispetto alla stima ottenuta tramite l'udito. Se però il soggetto presenta un deficit importante relativo al sistema visivo, allora la stima della localizzazione di un input è determinata primariamente dal sistema uditivo. Quanto detto è stato dimostrato in un esperimento, che coinvolgeva un soggetto con lesioni bilaterali del lobo parietale [14]. Anche se il paziente (d'ora in poi indicato con RM) aveva 10 /10 di vista per entrambi gli occhi, era incapace di vedere più di un oggetto alla volta. Le prove effettuate consistevano nel presentare due stimoli visivi distanti temporalmente un secondo l'uno dall'altro e nel chiedere al soggetto se essi provenivano dalla stessa posizione oppure no. I risultati sono mostrati in Figura 2.1 [12], dove le risposte di RM (cerchi neri) a input visivi sono state confrontate con quelle di soggetti sani (cerchi bianchi), che rappresentano il campione di controllo. Si può notare come i soggetti sani non avessero mai dato una risposta errata, mentre RM mostrava un chiaro deficit per il campo visivo sinistro e un deficit ancora più grave per il campo visivo destro. L'accuratezza della localizzazione acustica era invece molto migliore, come indicato dalla Figura 2.2 [12]. In questo caso, c'era comunque un deficit di RM rispetto al controllo, ma la capacità di localizzare stimoli acustici a destra era migliore rispetto a quella di localizzare di stimoli visivi nella stessa regione dello spazio. In prove successive, uno stimolo acustico era presentato nella stessa posizione e nello stesso istante di quello visivo. Quando a RM veniva chiesto di localizzare gli input visivi ignorando quelli acustici, egli mostrava un chiaro

miglioramento per il lato destro, mentre i soggetti di controllo non presentavano differenze rispetto alle prove precedenti (Figura 2.3 [12]). Da quanto è stato descritto emerge che la dominazione nella percezione da parte di una modalità sensoriale non è una caratteristica assoluta, ma dipende dal grado di affidabilità che un senso garantisce nella valutazione di oggetti ed eventi. Mentre nel fenomeno del ventriloquismo è la vista a catturare l'udito, nell'esperimento sopra citato succede il contrario, dato che l'udito presentava un'acuità spaziale superiore rispetto alla vista.

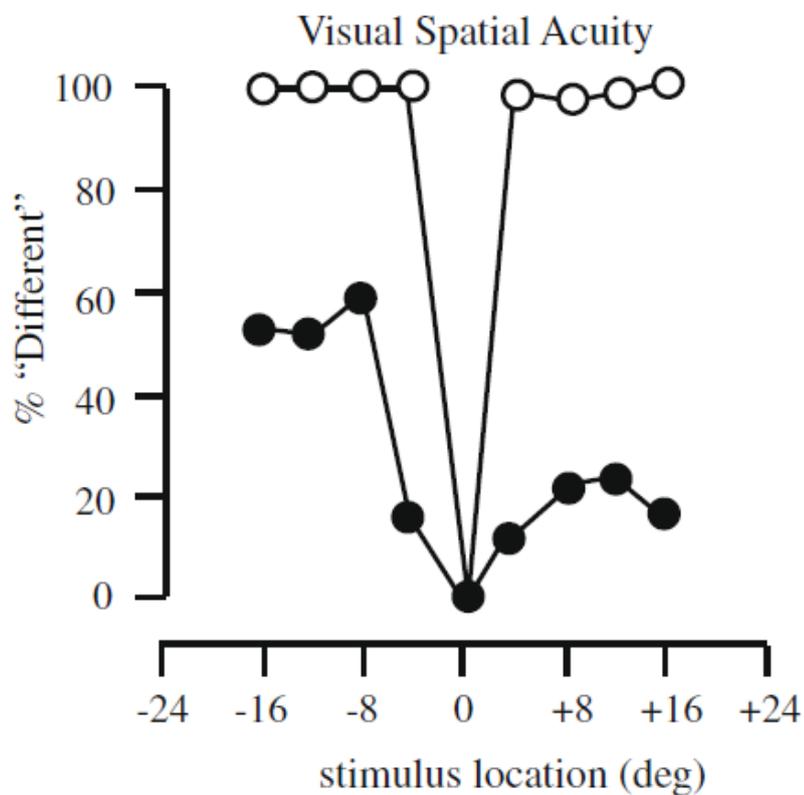


Figura 2.1. Misura dell'acuità spaziale visiva. L'asse y mostra la percentuale del tempo in cui il soggetto rispondeva che il secondo stimolo era presentato in una posizione diversa rispetto al primo.

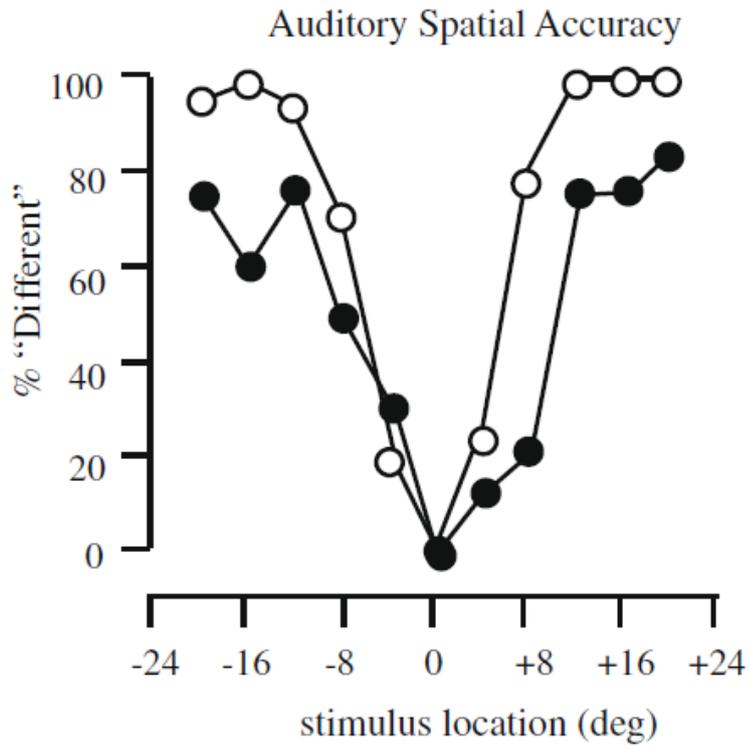


Figura 2.2. Misura dell'acuità spaziale acustica.

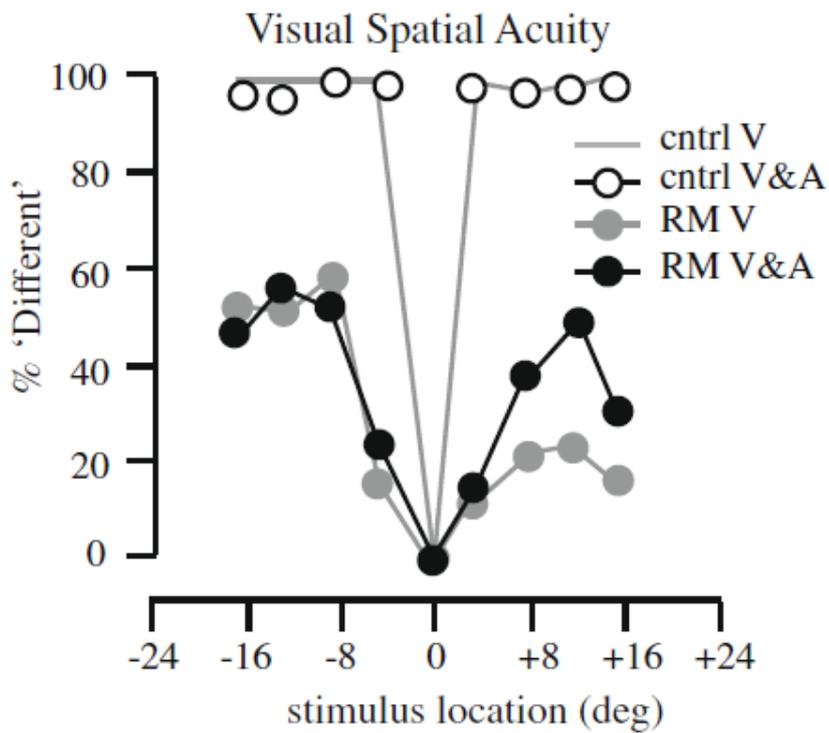


Figura 2.3. Influenza di stimoli acustici e visivi combinati sull'acuità spaziale visiva.

Un numero molto elevato di studi ha dimostrato come, nei soggetti sani (in cui la risoluzione spaziale visiva è nettamente migliore rispetto alla risoluzione spaziale acustica) l'effetto ventriloquismo sia molto robusto e si manifesti anche con stimoli neutri (senza implicazioni cognitive), quali flash luminosi e beep acustici [15,16,17]. In Figura 2.4 vengono riportati i risultati di alcuni studi; in particolare in essa viene mostrato il *bias di localizzazione acustica* (ovvero la differenza tra posizione percepita e posizione reale dello stimolo acustico) in funzione della disparità spaziale tra stimolo visivo e stimolo acustico, calcolata come la posizione dello stimolo visivo (P^v) meno quella dello stimolo acustico (P^a).

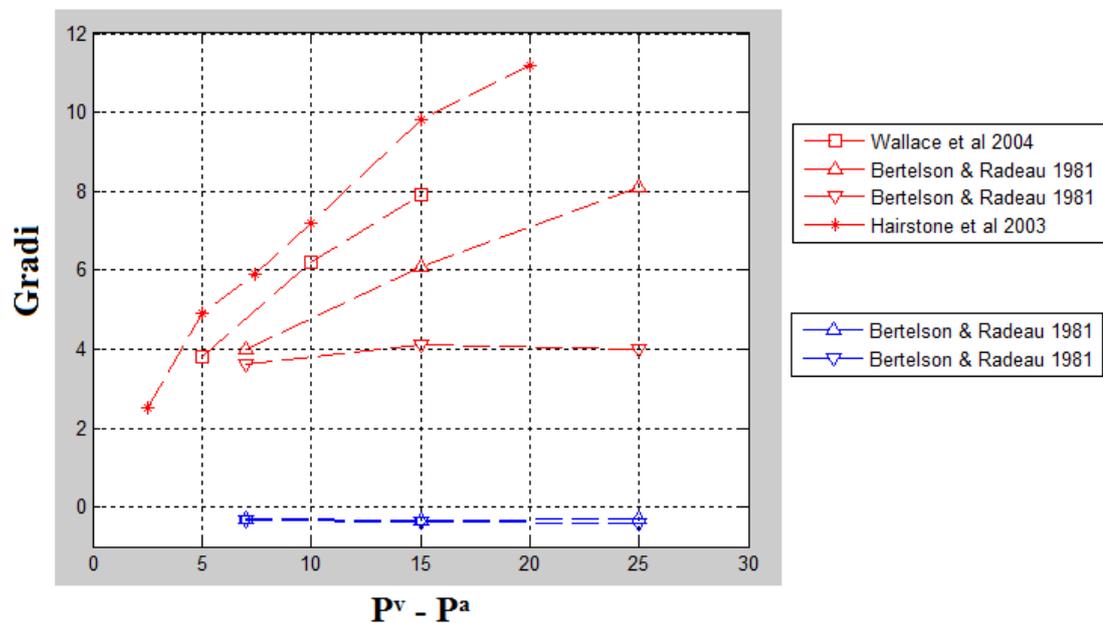


Figura 2.4 Fenomeno del ventriloquismo in soggetti sani.

Le curve in rosso rappresentano lo shift della posizione dell'input acustico verso quella dell'input visivo, mentre le curve in blu indicano lo spostamento dello stimolo visivo verso quello acustico. Si può notare che a seguito dell'esposizione alla situazione di conflitto audio-visivo, c'è un chiaro spostamento nella percezione della localizzazione dell'input uditivo; al contrario la percezione dell'input visivo non subisce alterazioni significative.

2.2 L'aftereffect del ventriloquismo

Una caratteristica peculiare del fenomeno del ventriloquismo è che esso può generare un effetto a lunga durata [18]. Infatti è stato dimostrato in un gran numero di esperimenti e con diversi paradigmi di addestramento, che la percezione della posizione dello stimolo acustico continua ad essere spostata verso la posizione dello stimolo visivo anche off-line, cioè quando l'input uditivo è presentato da solo. Questo risultato può essere ottenuto sottoponendo i soggetti a un periodo di costante esposizione a coppie di input acustici e visivi temporalmente coincidenti, ma spazialmente separati. Un esempio di esperimento riguardante l'aftereffect del ventriloquismo è quello condotto da Recanzone qualche anno fa [18]. Nel corso della prova, i soggetti stavano in una stanza buia e veniva chiesto loro di ascoltare una serie di toni accompagnati da flash luminosi; i due tipi di input percettivo erano mostrati allo stesso istante, ma provenivano da posizioni spaziali distinte. Durante la fase di training la posizione degli input veniva fatta variare. Dopo un'esposizione di 20 minuti, ai soggetti erano presentati solo segnali acustici ed essi dovevano indicare la locazione degli stessi. I risultati mostravano che se la separazione spaziale tra i due tipi di stimoli era relativamente piccola, la posizione dello stimolo acustico era percepita shiftata verso la posizione da cui era stato originato l'input visivo. La Figura 2.5 [18] mostra l'aftereffect in seguito a un addestramento in cui lo stimolo visivo era presentato 8° gradi a destra di quello acustico. I dati sono stati raccolti da una prova in cui la posizione dell'input è stata variata per cinque posizioni; le linee grigie verticali mostrano la posizione dello stimolo. Si può notare che prima dell'addestramento, le stime della posizione dell'input sono centrate vicino alla reale posizione dello stimolo, eccetto che per quelle più periferiche. Dopo una fase di training della durata di 20 minuti, le stime della posizione dell'input si spostano nettamente verso destra. La Figura 2.6 [18] mostra la retta di regressione per le stime della posizione, ottenute da tre soggetti in tre diverse prove ed evidenzia di nuovo lo shift della percezione a causa del fenomeno del ventriloquismo. Infatti è possibile notare come la retta di regressione abbia una pendenza di circa uno e un'intercetta sull'asse y di quasi 8, coerentemente con la differenza spaziale adottata durante la fase di addestramento.

Ventriloquism aftereffect for +8 deg training

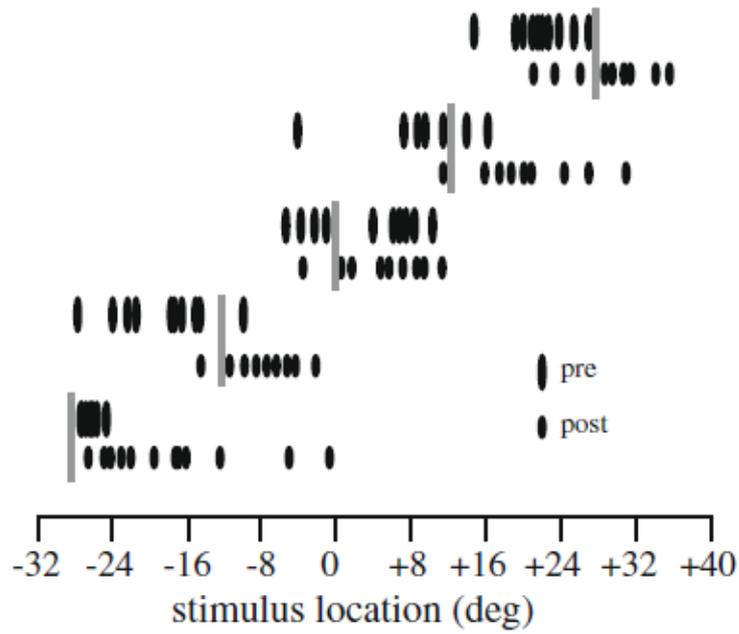


Figura 2.5. Aftereffect del ventriloquismo in un soggetto umano.

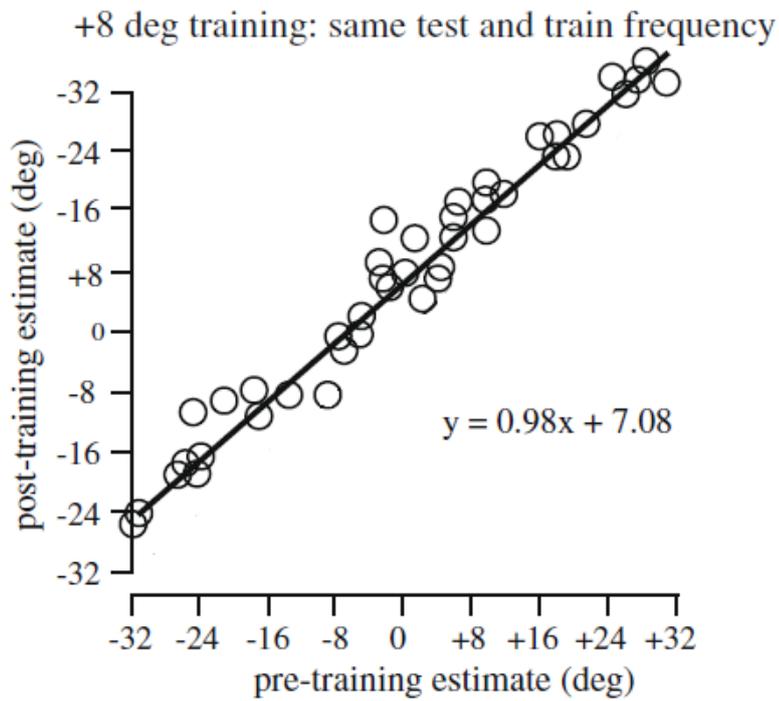


Figura 2.6. Retta di regressione per le stime di posizione.

I dati fin qui mostrati indicano quindi che c'è uno spostamento nella rappresentazione dello spazio acustico dopo un periodo di esposizione relativamente breve a un mismatch nella posizione tra le modalità uditiva e visiva.

L'aftereffect del ventriloquismo è un'illusione a lunga durata che si differenzia da altre, come per esempio l' "illusione a cascata" o "aftereffect del movimento" [12]. Tale illusione nasce se si guarda per un certo tempo un oggetto in movimento mantenendo gli occhi fissi e dopo si prova a fissare un oggetto fermo: quest'ultimo sembrerà muoversi in direzione opposta rispetto allo stimolo in movimento. Per esempio, se si osserva una cascata per qualche minuto e poi si concentra lo sguardo sulle rocce circostanti, queste appariranno muoversi lentamente verso l'alto. Alla base dell'aftereffect del movimento c'è il fenomeno dell'adattamento neurale: i neuroni che codificano un particolare movimento riducono la loro risposta con il passare del tempo di esposizione a uno stimolo costantemente in movimento. L'adattamento neurale fa abbassare anche l'attività spontanea di base di questi stessi neuroni quando essi rispondono a uno stimolo stazionario. Secondo una certa teoria, la percezione di oggetti fermi (come per esempio le rocce ai lati della cascata) è codificata come il bilancio tra le risposte di base dei neuroni che codificano per tutte le possibili direzioni del movimento. Perciò l'adattamento neurale di neuroni stimolati da un moto verso il basso fa diminuire la loro attività spontanea, a favore della percezione del movimento verso l'alto. L'aftereffect del ventriloquismo si differenzia da quello del movimento per due caratteristiche. Innanzitutto esso si traduce in uno spostamento della percezione nella stessa direzione che interessava la situazione di conflitto, mentre l'aftereffect del movimento va nella direzione opposta a quella percepita durante la fissazione di un oggetto in movimento. In secondo luogo, l'illusione dovuta al ventriloquismo può durare decine di minuti, mentre quella a cascata persiste per diversi secondi [18].

2.3 Trasferimento attraverso le frequenze dell'aftereffect del ventriloquismo

Un aspetto interessante e non ancora del tutto chiarito del fenomeno del ventriloquismo è la sua generalizzazione in frequenza: se dopo il periodo di training si utilizzano stimoli acustici con una frequenza diversa rispetto a quella adottata in fase di addestramento, l'aftereffect si verifica comunque? Diverse ricerche sono state compiute per cercare di rispondere a questa domanda, ma i risultati ottenuti sono stati molto spesso contrastanti. Per esempio dagli studi condotti da Recanzone [12], si evince che l'effetto off-line del ventriloquismo non si trasferisce attraverso le frequenze, come indicato in Figura 2.7 [12].

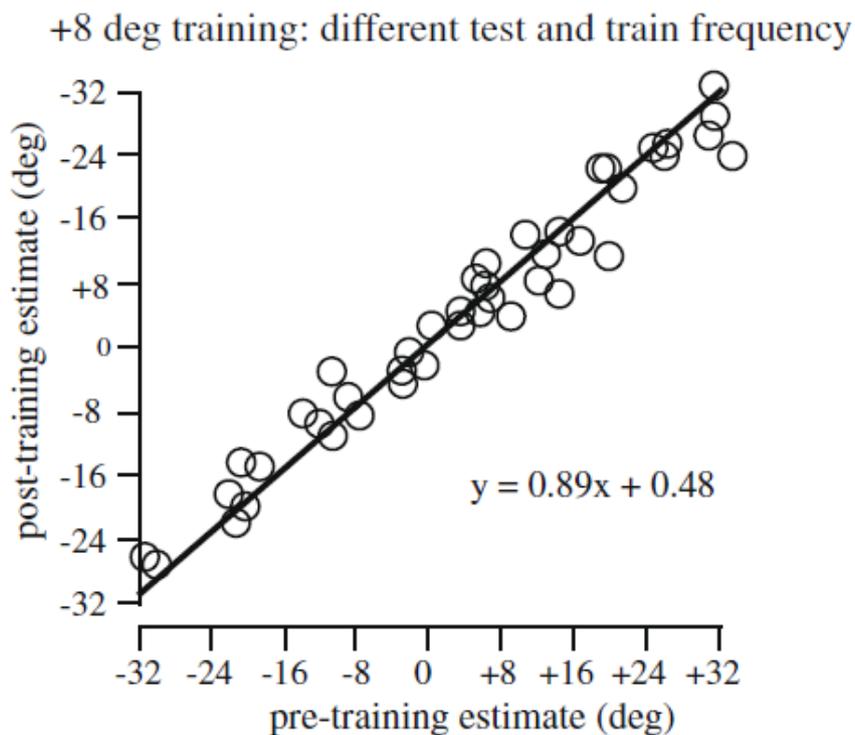


Figura 2.7. Mancato trasferimento dell'aftereffect per le frequenze (Recanzone).

I dati provengono da due esperimenti simili: in uno, dopo aver sottoposto i soggetti a un periodo di training in cui erano esposti a input acustici di 1500 Hz, sono stati presentati stimoli a 750 Hz per testare l'aftereffect; nell'altro, la procedura è stata

identica, ma con frequenze di addestramento e di test invertite rispetto a prima. In nessuno dei due casi si è manifestato l'aftereffect, infatti la retta di regressione indica che non c'è nessuna differenza tra le stime pre- e post- addestramento. Alle stesse conclusioni Recanzone è arrivato con altre prove, uguali a quelle appena descritte, ma usando una frequenza di 3000 Hz anziché di 1500 Hz [18]. Risultati simili a quelli di Recanzone sono stati ottenuti da Lewald [19]. In una serie di esperimenti, i partecipanti erano sottoposti a condizioni di stimolazione visuo-acustica, con una differenza spaziale di $+20^\circ$ o -20° tra i due tipi di input. In una prova sono stati usati suoni a 1 KHz per l'addestramento, in un'altra segnali a 4 KHz. Successivamente, per verificare la generazione dell'aftereffect, ai soggetti è stato chiesto di localizzare la posizione di stimoli visivi e acustici, con questi ultimi aventi una frequenza di 1 KHz.. La Figura 2.8 [19] mostra le risposte ricavate dal test.

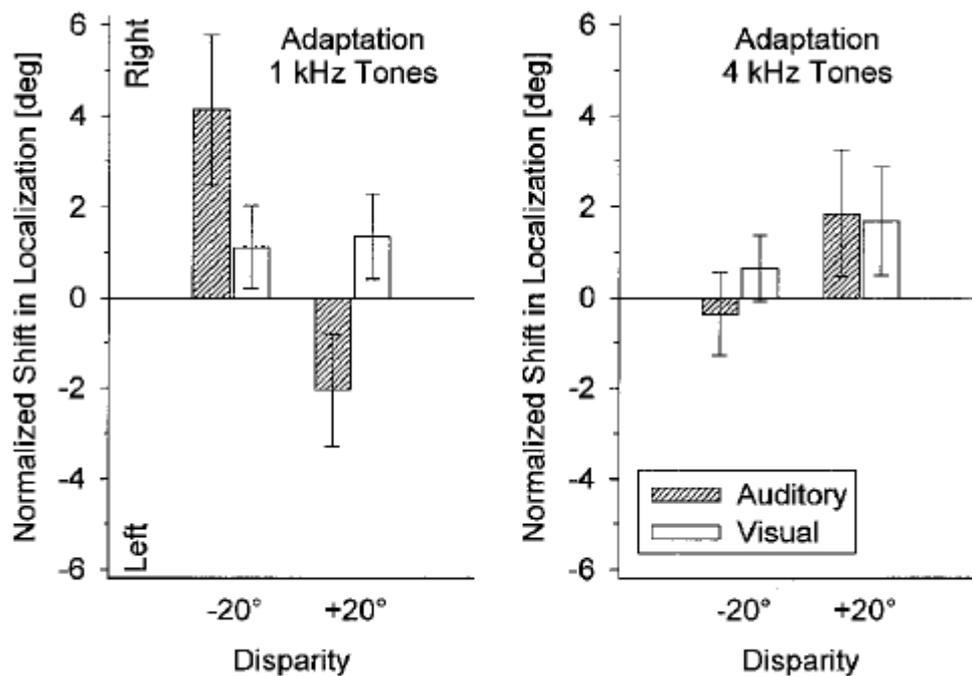


Figura 2.8. Mancato trasferimento dell'aftereffect per le frequenze (Lewald).

E' possibile osservare che quando la frequenza sonora del segnale di post-adattamento era uguale a quella adoperata nella fase di training, c'era uno

spostamento significativo nella localizzazione uditiva; se invece le due frequenze erano differenti, non c'erano alterazioni significative sulla percezione uditiva.

Sulla base di questi risultati sia Recanzone che Lewald hanno ipotizzato che siano i neuroni della corteccia uditiva primaria a essere coinvolti nell'aftereffect del ventriloquismo. Infatti essi mostrano - almeno per certi valori di intensità di stimolazione acustica - stretti campi recettivi in frequenza, che conferiscono loro una capacità di selettività rispetto alla frequenza sonora.

Dalle ricerche condotte da Frissen et al. [20,21], risulta invece una generalizzazione in frequenza dell'aftereffect del ventriloquismo. In uno degli esperimenti compiuti da Frissen et al. [20], i partecipanti venivano esposti a situazioni di conflitto audiovisivo caratterizzate da una discordanza spaziale di 9° tra stimolo acustico e stimolo visivo. L'adattamento è stato fatto usando suoni di 750 Hz e 3000 Hz e i risultati sono indicati in Figura 2.9 [20]. Gli aftereffect sono stati calcolati sottraendo le medie delle posizioni stimate prima della fase di training a quelle dopo la fase di training.

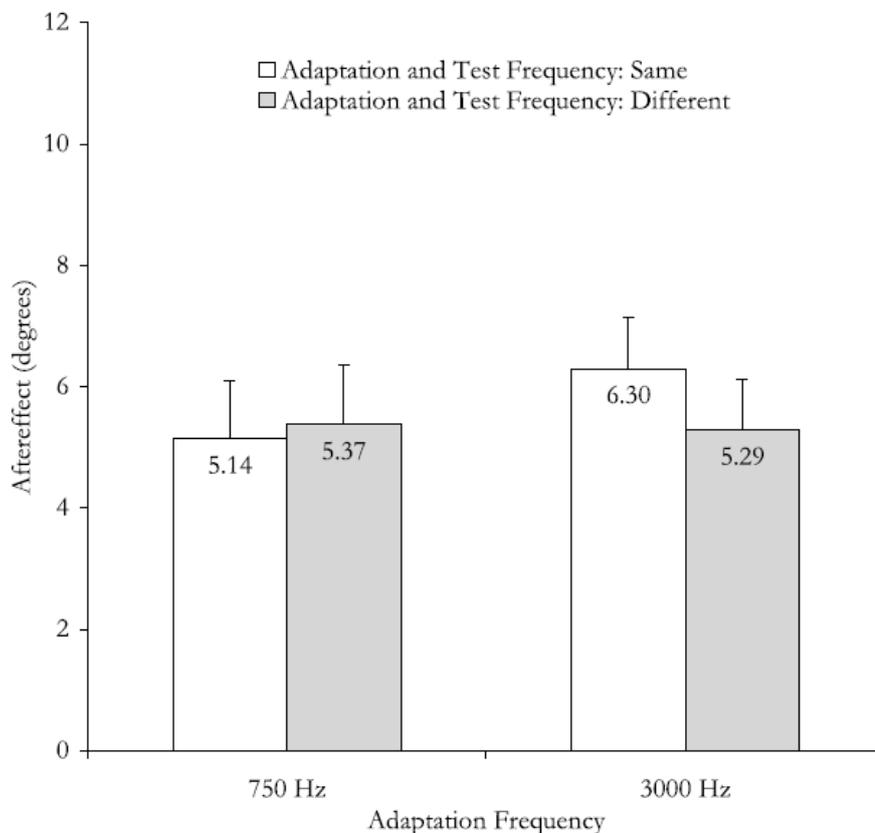


Figura 2.9. Trasferimento dell'aftereffect per le frequenze (Frissen et al. 2003).

Si può notare come testando l'aftereffect con un segnale a 750 Hz dopo aver usato durante l'adattamento suoni a 3000 Hz e viceversa, ci sia una lieve riduzione dello shift nella percezione. Le prove appena descritte hanno perciò evidenziato un trasferimento dell'aftereffect per due ottave di frequenza. In un altro esperimento eseguito da Frissen et al. [21], i soggetti erano sottoposti a periodi di adattamento con una discordanza spaziale di 18° tra i due tipi di input. Le prove sono state effettuate con toni di 400 Hz e 6400 Hz e i risultati sono mostrati in Figura 2.10 [21]. Gli aftereffect sono stati calcolati come nel precedente lavoro.

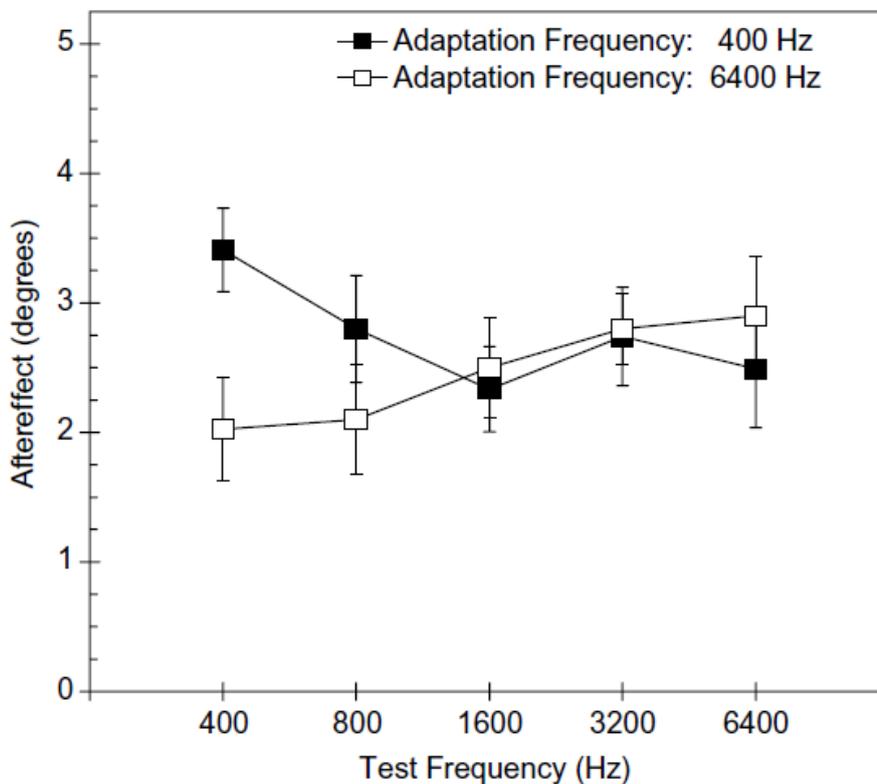


Figura 2.10. Trasferimento dell'aftereffect per le frequenze (Frissen et al. 2005).

In questo caso che c'è stata una generalizzazione dell'aftereffect per un range di frequenza di quattro ottave; si può comunque notare che l'aftereffect si riduce allontanandosi dalla frequenza usata durante il periodo di training. Gli esiti della ricerca di Frissen et al. mostrano quindi che la ricalibrazione visiva della posizione

del suono non è limitata alla frequenza sonora adottata durante la fase di esposizione al conflitto, ma riguarda un intervallo di frequenze più ampio.

Capitolo 3

Modello di rete neurale per il fenomeno del ventriloquismo

3.1 Aspetti generali

Finora sono stati descritti gli aspetti caratteristici del fenomeno del ventriloquismo e le sue manifestazioni in vivo in soggetti umani. A partire dai risultati sperimentali sopra citati, nel presente lavoro di tesi è stato realizzato un modello di rete neurale per il ventriloquismo, con l'obiettivo specifico di studiare un aspetto non ancora chiarito di tale fenomeno, ovvero la generalizzazione in frequenza o meno dell'aftereffect del ventriloquismo.

Il modello qui proposto prende spunto da un modello di rete neurale precedentemente sviluppato dai ricercatori dell'università di Bologna [22]. Tale modello simula gli effetti on-line (spostamento della percezione dello stimolo acustico verso la posizione del concomitante input visivo) e off-line (shift spaziale dello stimolo acustico senza input visivo a seguito di un periodo di esposizione alla situazione di conflitto audio-visiva) del ventriloquismo, indagando i meccanismi alla base del fenomeno, in particolare le modificazioni sinaptiche alla base dell'aftereffect del ventriloquismo.

Tuttavia il precedente studio non prendeva in esame il trasferimento dell'aftereffect attraverso le frequenze sonore. Per esaminare questo ulteriore aspetto, il precedente modello è stato opportunamente ampliato e modificato allo scopo di simulare e interpretare i risultati relativi alla generalizzazione o meno dell'aftereffect nel dominio delle frequenze. In particolare, è stata modificata la struttura della rete e conseguentemente la disposizione delle connessioni sinaptiche. Inoltre sono state realizzate due configurazioni del modello, una che simulasse la non generalizzazione in frequenza dell'effetto off-line e un'altra che ne riproducesse la generalizzazione. Infine è stata realizzata un'ulteriore configurazione per valutare l'influenza

dell'input, fornito per stimolare la rete, sul trasferimento dell'aftereffect attraverso le frequenze sonore.

Tutte le configurazioni create sono state prima valutate nelle loro condizioni di base, cioè quando le sinapsi non hanno ancora subito modifiche. Successivamente la rete neurale è stata sottoposta a un periodo di addestramento tramite delle regole di Hebb e sono state analizzate le variazioni dell'output nelle diverse configurazioni. L'uscita della rete è rappresentata dalla posizione spaziale in cui viene percepito lo stimolo acustico. La posizione percepita dello stimolo uditivo è stata determinata decodificando l'attivazione dell'intera popolazione di neuroni acustici mediante la metrica del baricentro, per cui la posizione percepita coincide con il baricentro dell'attivazione dell'intera popolazione.

Per costruire il modello e risolvere le equazioni che lo caratterizzano è stato utilizzato il programma MATLAB®.

3.2 Struttura del modello

Il modello di rete neurale realizzato in questo lavoro di tesi prende spunto da uno studio portato avanti dal gruppo di Bioingegneria dell'Università di Bologna [22], che ha sviluppato un modello riguardante il fenomeno del ventriloquismo. Esso è formato da due strati monodimensionali di neuroni: uno strato rappresenta l'area visiva e l'altro quella acustica. I neuroni dello stesso layer sono collegati da sinapsi laterali, mentre quelli di aree differenti sono connessi da sinapsi inter-area (Figura 3.1 [22]). Le sinapsi laterali possono essere eccitatorie (quelle che terminano con una freccia) o inibitorie (quelle che terminano con punto), in particolare un neurone eccita i suoi vicini ed inibisce le cellule più distanti. Una cellula di uno strato manda inoltre una sinapsi eccitatoria alla cellula dell'altra modalità sensoriale che si trova nella stessa posizione spaziale. La rete così realizzata è in grado di simulare gli effetti on-line e off-line del ventriloquismo. Uno dei limiti di questo modello è rappresentato dal fatto che sono state trascurate le caratteristiche in frequenza della risposta dei neuroni acustici, infatti questi ultimi, come quelli visivi, codificano

ciascuno per una certa posizione dello spazio, ma non tengono conto della frequenza dell'input sonoro. Perciò, nel presente lavoro di tesi, a partire dalla rete neurale creata dal gruppo di Bioingegneria dell'Università di Bologna [22], ne è stata costruita un'altra che includesse le proprietà dei neuroni acustici relative alla frequenza sonora.

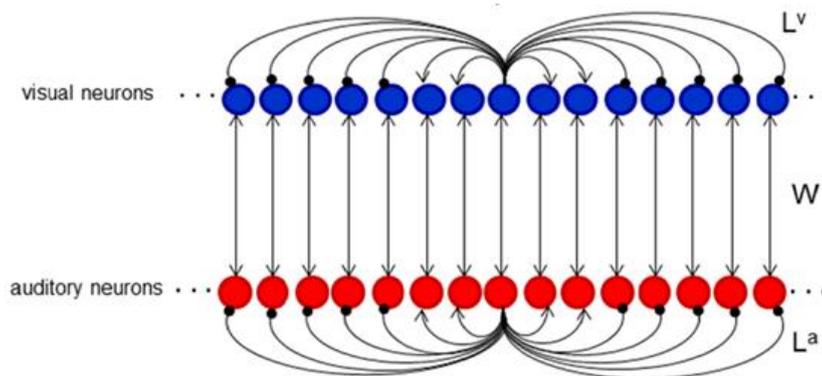


Figura 3.1. Schema della rete neurale del modello di riferimento.

La rete neurale costruita in questa tesi è composta da neuroni visivi e acustici. Le due tipologie di neuroni fanno parte di due strati neurali distinti, ciascuno dei quali rappresenta un'area unisensoriale, quella visiva e quella uditiva. I neuroni visivi compongono uno strato monodimensionale di cellule, ognuno di essi codifica per una particolare posizione dello spazio e sono in numero di N_p ; quelli acustici costituiscono invece uno strato bidimensionale di cellule, codificano per una specifica posizione dello spazio e per una determinata frequenza del suono, e sono in numero di $N_p \times N_f$. Il parametro N_p indica il numero di neuroni (acustici o visivi) che codificano per le posizioni dello spazio, mentre N_f rappresenta il numero di neuroni (solo acustici) che rispondono alle diverse frequenze sonore. Quindi la rete nel complesso codifica per N_p posizioni ed N_f frequenze.

Per quanto riguarda la distanza spaziale tra i neuroni dello stesso layer, è stato ipotizzato che ci fosse una distanza di 1° tra di essi; dato che $N_p = 180$, ogni strato copre un'area di 180° dello spazio visivo in un caso e acustico nell'altro. È stato evidenziato in precedenza che l'area acustica risponde anche alle frequenze. N_f è pari a 40 e il range di frequenze scelto copre un intervallo che va da 200 Hz a 6000 Hz. Poiché per valutare la presenza di generalizzazione in frequenza si è deciso di

ragionare in termini di ottave, ogni 8 neuroni la distanza in frequenza tra una cellula e l'altra è stata fatta raddoppiare, in particolare:

200	225	250	275	300	325	350	375
400	450	500	550	600	650	700	750
800	900	1000	1100	1200	1300	1400	1500
1600	1800	2000	2200	2400	2600	2800	3000
3200	3600	4000	4400	4800	5200	5600	6000

In questo modo dopo 8 neuroni la frequenza raddoppia (ad esempio, il neurone di posto 8 risponde a una frequenza sonora di 375 Hz, mentre quello di posto 16 corrisponde a 750 Hz) e ci si è mossi così di un'ottava.

I neuroni che compongono lo stesso strato sono collegati tra loro da sinapsi laterali (L^m , dove m indica la generica modalità sensoriale; d'ora in avanti si useranno l'apice a per i neuroni acustici e v per quelli visivi), mentre quelli situati in aree unisensoriali differenti sono connessi da sinapsi inter-area (W). La Figura 3.2 schematizza quanto appena detto.

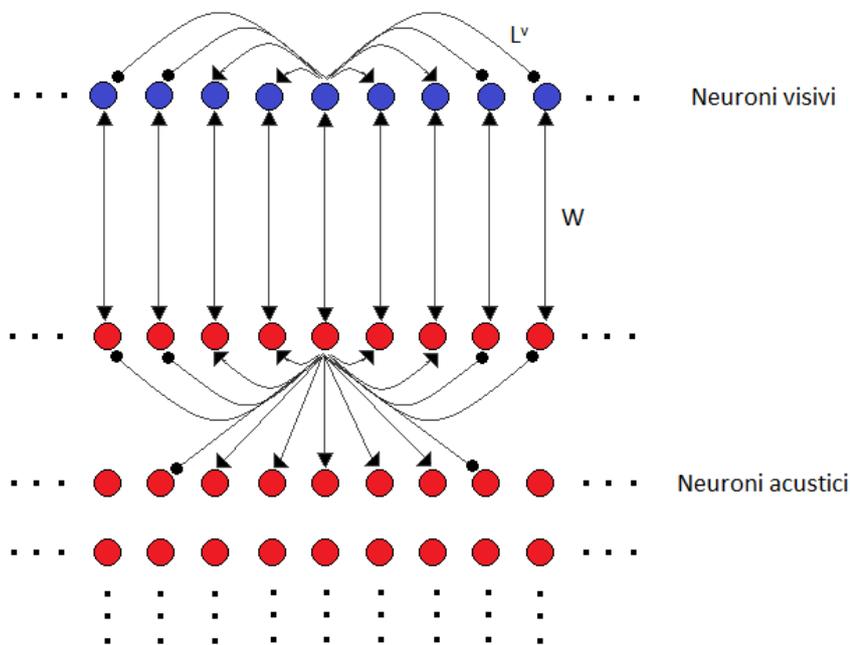


Figura 3.2. Architettura della rete neurale (per semplicità, vengono mostrate le sinapsi laterali provenienti da un solo neurone per ogni strato).

Ogni cerchio rappresenta un neurone (in blu sono riportati i neuroni visivi e in rosso quelli acustici). Ciascuna riga di neuroni nell'area acustica codifica per diverse posizioni spaziali e per un'unica frequenza, mentre ciascuna colonna di neuroni codifica per frequenze diverse e per la medesima posizione spaziale. Le linee che collegano i cerchi sono le sinapsi: quelle che terminano con una freccia sono eccitatorie, mentre quelle che finiscono con un punto sono inibitorie. Si capisce quindi che le sinapsi laterali possono essere eccitatorie o inibitorie, mentre quelle inter-area sono solo eccitatorie. Per quanto concerne le connessioni intra-area, esse hanno una disposizione tale per cui un neurone eccita i suoi vicini e inibisce le cellule più lontane (Figura 3.3). Questo tipo di organizzazione delle sinapsi è definito "cappello messicano": si tratta in sostanza di una zona centrale eccitatoria circondata da un'area inibitoria. Il cappello messicano è ottenuto realizzando separatamente le sinapsi laterali eccitatorie ed inibitorie con una disposizione gaussiana (la gaussiana inibitoria ha una deviazione standard maggiore e un'ampiezza minore rispetto alla gaussiana eccitatoria) e sottraendo alle connessioni eccitatorie quelle inibitorie, come indicato in Figura 3.4.

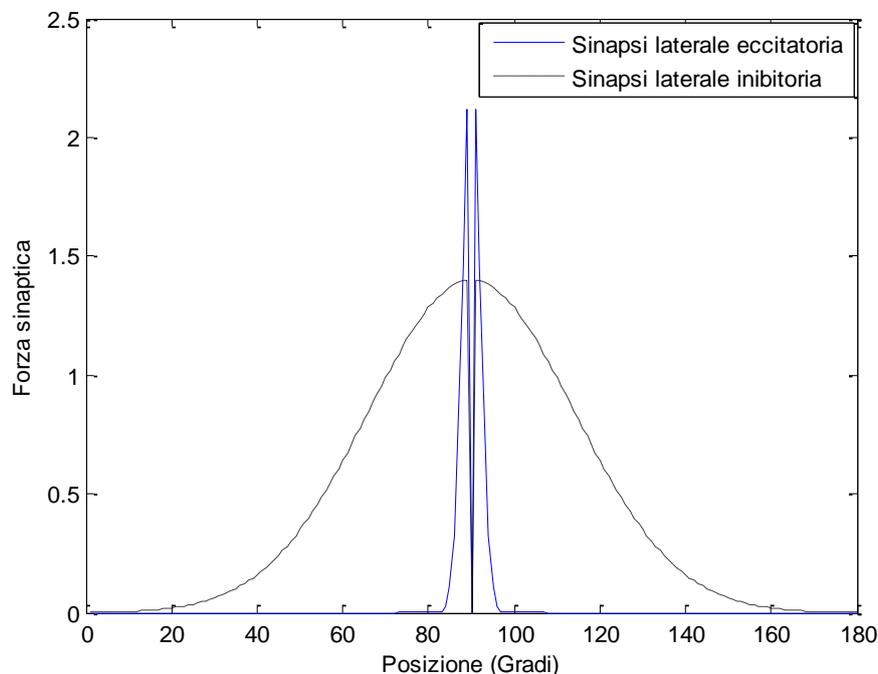


Figura 3.3. Sinapsi laterali eccitatoria e inibitoria.

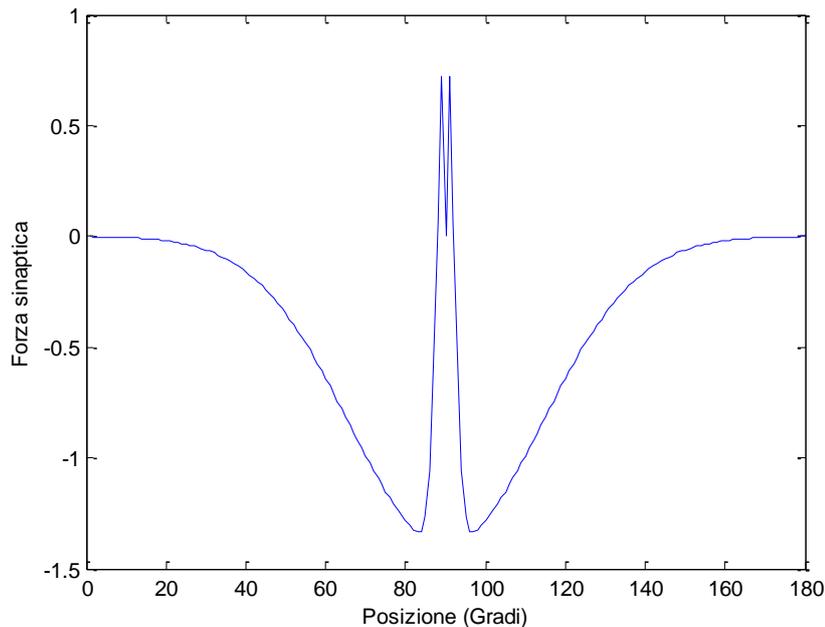


Figura 3.4. Cappello messicano.

La rete neurale così realizzata è stimolata mediante input visivi e acustici. Essi sono stati pensati come delle funzioni Gaussiane: sono centrati in un certo punto dello spazio (e in un certo valore di frequenza nel caso di stimolo acustico), interessando maggiormente la cellula centrata nella loro stessa posizione, ma coinvolgendo anche quelle vicine, con un'influenza che diminuisce con la distanza dalla posizione centrale dello stimolo. Considerare un input del genere vuol dire immaginare che esso venga filtrato dal campo recettivo del neurone su cui incide. Infatti le deviazioni standard delle funzioni Gaussiane rappresentative degli stimoli esterni possono essere considerate come l'ampiezza dei campi recettivi dei neuroni stessi. Un'assunzione fondamentale del modello è che i neuroni acustici hanno campi recettivi più ampi rispetto a quelli visivi; di conseguenza la deviazione standard della Gaussiana rappresentativa dell'input acustico è maggiore rispetto a quella dell'input visivo. Questa ipotesi rispecchia il fatto che la risoluzione spaziale visiva è migliore di quella acustica.

Per quel che concerne l'input di ciascun neurone, nel complesso ogni neurone riceve un input che è la somma di tre contributi: un input esterno, un input laterale

proveniente dai neuroni dello stesso layer e un input cross-modale derivante dall'altra area sensoriale.

3.3 Caratteristiche dei neuroni della rete

Come indicato precedentemente, è stato scelto $N_p = 180$ e $N_f = 40$. Quindi il layer visivo è formato da 180 cellule, mentre quello acustico da $180 \times 40 = 7200$ cellule. Questi valori sono da considerarsi fittizi e sono stati scelti per non appesantire troppo il modello dal punto di vista computazionale. Perciò un singolo neurone della rete è da supporre invece rappresentativo di una popolazione di cellule che codificano per la stessa posizione spaziale e, nel caso dell'area uditiva, per la stessa frequenza sonora. Per riferirsi a un certo neurone della rete, si adopererà l'apice m per indicare la generica modalità percettiva, v per la modalità visiva ed a per quella acustica, e il pedice j per rappresentare la sua posizione nello strato visivo oppure il doppio pedice ij per lo strato uditivo. Con $u(t)$ e $y(t)$ ci si riferisce rispettivamente all'ingresso e all'uscita di un certo neurone della rete. La risposta dei neuroni agli stimoli che essi ricevono è descritta mediante un'equazione differenziale del primo ordine e da una relazione sigmoidale. L'equazione differenziale simula le proprietà integrative della membrana cellulare, mentre la caratteristica sigmoidale indica la presenza di una soglia inferiore e di una saturazione superiore. Perciò considerando una cellula appartenente alla modalità percettiva m , si può scrivere un'equazione differenziale del tipo:

$$\tau_y \frac{dy^m(t)}{dt} = -y^m(t) + F(u^m(t)) \quad \text{Eq. 3.1}$$

dove τ_y è la costante di tempo ed $F(u)$ è la caratteristica sigmoidale, che non è altro che una relazione statica ingresso-uscita della forma:

$$F(u^m) = \frac{1}{1 + e^{-s(u^m - \theta)}} \quad \text{Eq. 3.2}$$

dove s rappresenta la pendenza della sigmoide e θ la sua posizione centrale. Come si nota dall'Eq. 3.2, l'attività di saturazione del neurone è stata impostata a 1; in questo modo tutte le attività sono normalizzate al valore massimo.

3.4 Input incidente su un neurone della rete

È stato già accennato nel paragrafo 3.1 che ciascun neurone della rete riceve un ingresso $u^m(t)$ dato dalla somma di tre termini: uno stimolo esterno $e^m(t)$, un input laterale $l^m(t)$ mandato da una cellula dello stesso strato e un contributo cross-modale $c^m(t)$ proveniente dai neuroni presenti nell'altra area unisensoriale. Perciò si può scrivere:

$$u^m(t) = e^m(t) + l^m(t) + c^m(t) \quad \text{Eq. 3.3}$$

Vengono ora mostrate in dettaglio le tre tipologie di input.

3.4.1 Input esterno

L'input esterno fornito alla rete è dato da una funzione Gaussiana, che rappresenta il risultato dell'azione di filtraggio dello stimolo ambientale da parte del campo recettivo dei neuroni. Considerando uno stimolo visivo esterno centrato nella posizione p^v , l'input della rete può essere definito come:

$$e_j^v(t) = E_0^v e^{-\frac{(d_j^v)^2}{2(\sigma^v)^2}} \quad \text{Eq. 3.4}$$

dove E_0^v indica la forza dell'input esterno, d_j^v rappresenta la distanza tra un neurone di posto j e la posizione p^v dello stimolo, e σ^v è la deviazione standard della Gaussiana ed è rappresentativa dell'ampiezza del campo recettivo della cellula neurale. I neuroni che si trovano ai bordi dello strato potrebbero non ricevere lo stesso input di quelli che si trovano in posizione centrale. Questo fatto potrebbe produrre degli effetti di bordo indesiderati, perciò lo strato è stato immaginato come se avesse una struttura circolare e la distanza d_j^v è stata così definita:

$$d_j^v = \begin{cases} |j - p^v| & \text{se } |j - p^v| \leq N_p / 2 \\ N_p - |j - p^v| & \text{se } |j - p^v| > N_p / 2 \end{cases} \quad \text{Eq. 3.5}$$

Per quel che concerne lo stimolo acustico esterno, è stato spiegato in precedenza che l'area sensoriale uditiva è rappresentata da uno strato bidimensionale, in quanto le cellule acustiche rispondono sia alla posizione che alla frequenza dello stimolo acustico. Per scandire i neuroni lungo la posizione si adopera il pedice j , lungo la frequenza il pedice i . Quindi l'input esterno per le cellule di questo strato è rappresentato da una funzione Gaussiana bidimensionale, che può essere espressa con la seguente formula:

$$e_{ij}^a(t) = E_0^a e^{-\left[\frac{(d_j^a)^2}{2(\sigma^{ap})^2} + \frac{(d_i^a)^2}{2(\sigma^{af})^2} \right]} \quad \text{Eq. 3.6}$$

dove E_0^a indica la forza dell'input esterno, d_j^a rappresenta la distanza in posizione tra un neurone di posto j e lo stimolo, d_i^a indica la distanza in frequenza tra un neurone di posto i e lo stimolo, σ^{ap} è la deviazione standard relativa alla posizione e σ^{af} è quella relativa alla frequenza.

3.4.2 Input laterale

L'input laterale che interessa un certo neurone rappresenta il contributo che le cellule situate nella stessa area unisensoriale inviano al neurone considerato tramite le sinapsi laterali. Come per l'input esterno, anche in questo caso bisogna fare una distinzione tra input laterale visivo e acustico, dato che i due strati percettivi hanno dimensioni diverse. Lo stimolo laterale che coinvolge le cellule dell'area visiva può essere definito nel seguente modo:

$$l_j^v(t) = \sum_k L_{jk}^v \cdot y_k^v(t) \quad \text{Eq. 3.7}$$

dove y_k^v rappresenta l'attività di un neurone presinaptico situato nella posizione k , mentre L_{jk}^v è la forza delle sinapsi laterali che si originano dal neurone presinaptico di posto k e raggiungono un neurone postsinaptico che si trova in posizione j . Come affermato nel paragrafo 3.2, le sinapsi laterali possono essere eccitatorie o inibitorie. La forza delle sinapsi laterali è data dalla differenza tra il contributo di eccitazione e quello di inibizione:

$$L_{jk}^v = L_{ex,jk}^v - L_{in,jk}^v \quad \text{Eq. 3.8}$$

Il contributo di eccitazione e quello di inibizione vengono espressi mediante delle funzioni Gaussiane di questo tipo:

$$L_{ex,jk}^v = \begin{cases} L_{ex0}^v \cdot e^{-\frac{(d_{jk})^2}{2\sigma_{ex,v}^2}} & \text{se } d_{jk} \neq 0 \\ 0 & \text{se } d_{jk} = 0 \end{cases} \quad \text{Eq. 3.9}$$

$$L_{in,jk}^v = \begin{cases} L_{in0}^v \cdot e^{-\frac{(d_{jk})^2}{2\sigma_{in,v}^2}} & \text{se } d_{jk} \neq 0 \\ 0 & \text{se } d_{jk} = 0 \end{cases} \quad \text{Eq. 3.10}$$

dove L_{ex0}^v è una costante che indica la forza delle sinapsi eccitatorie, mentre L_{in0}^v è una costante che corrisponde alla forza delle connessioni inibitorie; $\sigma_{ex,v}$ e $\sigma_{in,v}$ sono deviazioni standard che indicano la velocità con cui la forza sinaptica decresce con la distanza; infine d_{jk} è la distanza tra un neurone alla posizione j e uno alla posizione k , e come per l'input esterno, è stata fatta l'ipotesi di circolarità della struttura, quindi:

$$d_{jk} = \begin{cases} |j-k| & \text{se } |j-k| \leq N_p / 2 \\ N_p - |j-k| & \text{se } |j-k| > N_p / 2 \end{cases} \quad \text{Eq. 3.11}$$

La differenza tra i due termini porta a un'organizzazione sinaptica definita a "cappello messicano", come mostrato in Figura 3.5. Affinché si ottenga la forma di un cappello messicano è necessario che $L_{ex0} > L_{in0}$ e $\sigma_{ex} < \sigma_{in}$. Per quanto riguarda l'input laterale che coinvolge i neuroni dello strato acustico, esso è esprimibile come:

$$l_{ij}^a(t) = \sum_h \sum_k L_{ij,hk}^a \cdot y_{hk}^a(t) \quad \text{Eq. 3.12}$$

dove y_{hk}^a rappresenta l'attività di un neurone presinaptico situato nella posizione hk , mentre $L_{ij,hk}^a$ è la forza delle sinapsi laterali che si originano dal neurone presinaptico di posto hk e raggiungono un neurone postsinaptico che si trova in posizione ij . Anche in questo caso la forza delle sinapsi laterali è data dalla parte eccitatoria meno quella inibitoria:

$$L_{ij,hk}^a = L_{ex,ij,hk}^a - L_{in,ij,hk}^a \quad \text{Eq. 3.13}$$

Come prima, i due contributi sono rappresentati da funzione Gaussiane:

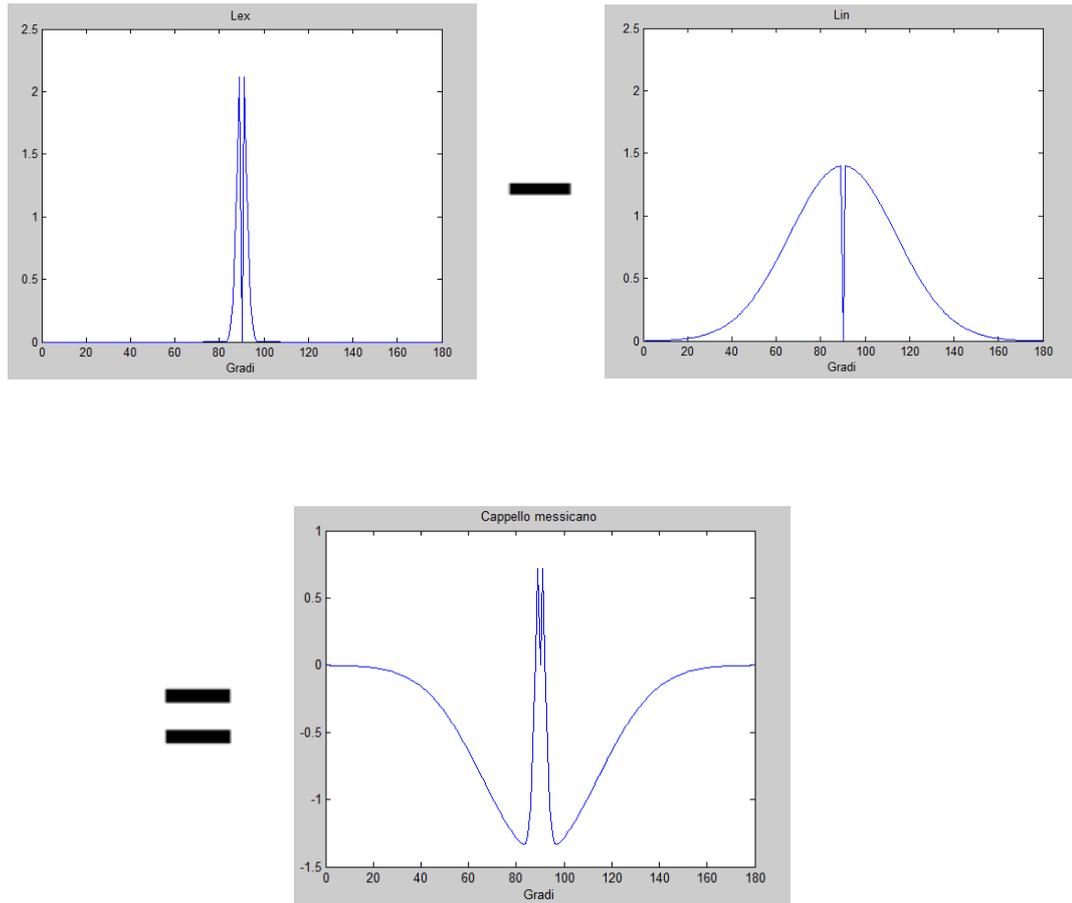


Figura 3.5. Formazione del cappello messicano per le sinapsi laterali visive.

$$L_{ex,ij,hk}^a = \begin{cases} L_{ex0}^a \cdot e^{-\left[\frac{(d_{jk})^2}{2\sigma_{ex,ap}^2} + \frac{(d_{ih})^2}{2\sigma_{ex,af}^2} \right]} & \text{se } d_{jk} = 0 \text{ e } d_{ih} = 0 \\ 0 & \text{altrimenti} \end{cases} \quad \text{Eq. 3.14}$$

$$L_{in,ij,hk}^a = \begin{cases} L_{in0}^a \cdot e^{-\left[\frac{(d_{jk})^2}{2\sigma_{in,ap}^2} + \frac{(d_{ih})^2}{2\sigma_{in,af}^2} \right]} & \text{se } d_{jk} = 0 \text{ e } d_{ih} = 0 \\ 0 & \text{altrimenti} \end{cases} \quad \text{Eq. 3.15}$$

dove L_{ex0}^a ed L_{in0}^a sono delle costanti che indicano rispettivamente la forza delle sinapsi eccitatorie e quella delle sinapsi inibitorie; $\sigma_{ex,ap}$ e $\sigma_{in,ap}$ sono delle deviazioni standard che indicano la velocità con cui la forza sinaptica decresce con la distanza

riferita alla posizione; $\sigma_{ex,af}$ e $\sigma_{in,af}$ sono delle deviazioni standard che indicano la velocità con cui la forza sinaptica diminuisce con la distanza riferita alla frequenza; infine d_{jk} è la distanza tra un neurone alla posizione j e uno alla posizione k , mentre d_{ih} è la distanza tra un neurone alla frequenza i e uno alla frequenza h . Come per l'area visiva, anche per quella uditiva la disposizione delle sinapsi è del tipo a cappello messicano (Figura 3.6). Per entrambe le modalità percettive è stata esclusa la possibilità che un neurone possa auto-eccitarsi o auto-inibirsi. Sia le sinapsi visive che quelle acustiche non sono fisse, ma possono subire delle modifiche a seguito di un periodo di addestramento a stimoli cross-modali, in particolare esse possono rinforzarsi o indebolirsi; queste variazioni avvengono secondo le regole di Hebb, come verrà discusso nel paragrafo successivo.

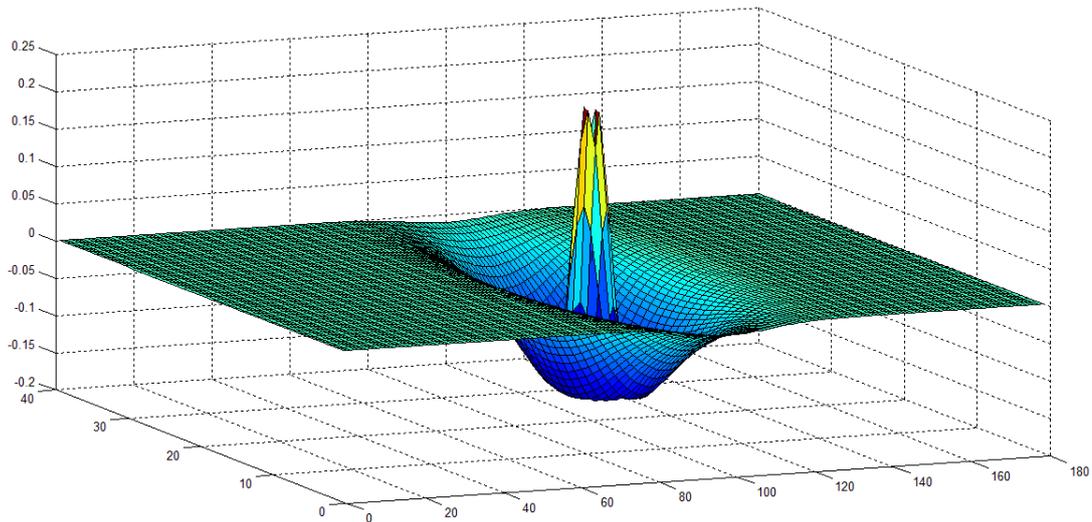


Figura 3.6. Cappello messicano per le sinapsi laterali acustiche.

3.4.3 Input cross-modale

L'input cross-modale è quello che un neurone riceve da una cellula appartenente all'altro layer. La rete è stata costruita in maniera tale che un neurone possa mandare una sinapsi inter-area solo al neurone (nel caso in cui quest'ultimo appartenga all'area visiva) o ai neuroni (se questi si trovano nell'area acustica) appartenenti all'altro strato che si trovano nella stessa posizione spaziale, come indicato in Figura

3.7. Inoltre le connessioni inter-area sono solo eccitatorie. Dalla Figura 3.7 si nota che un neurone visivo manda sinapsi a tutti i neuroni acustici che si trovano nella sua stessa posizione spaziale ma che rispondono a frequenze diverse, mentre una cellula uditiva manda sinapsi solo a una cellula visiva, cioè quella che occupa la sua stessa locazione nello spazio.

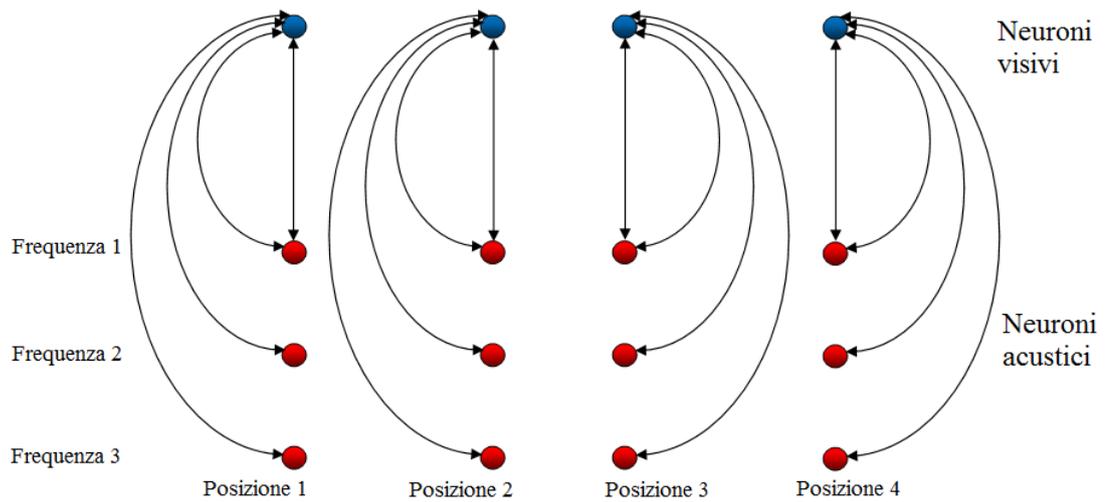


Figura 3.7. Sinapsi inter-area.

Come sarà mostrato in seguito, le sinapsi che vanno dall'area acustica a quella visiva e quelle che fanno il percorso contrario non hanno lo stesso peso, dato che i neuroni acustici sono molti di più di quelli visivi.

L'espressione relativa all'input cross-modale che arriva a un neurone visivo è la seguente:

$$c_j^v(t) = \sum_i W_{va} \cdot y_{ij}^a(t) \quad \text{Eq. 3.16}$$

dove W_{va} rappresenta le sinapsi che i neuroni acustici in posizione j mandano al neurone visivo nella medesima posizione spaziale e y_{ij}^a è l'attività di un neurone acustico di posto ij . Lo stimolo cross-modale che riceve una cellula acustica è dato dalla seguente relazione:

$$c_{ij}^a(t) = W_{av} \cdot y_j^v(t) \quad \text{Eq. 3.17}$$

dove W_{av} rappresenta la sinapsi che un neurone dell'area visiva in posizione j manda al neurone acustico nella medesima posizione spaziale indipendentemente dalla posizione in frequenza, e y_j^v è l'attività di un neurone visivo di posto j .

3.5 Regole di apprendimento di Hebb

Uno degli obiettivi del presente lavoro di tesi è stato quello di riprodurre l'aftereffect del ventriloquismo, ovvero uno spostamento duraturo nella percezione di uno stimolo acustico dopo una fase di esposizione prolungata a input acustici e visivi simultanei ma spazialmente separati. Per raggiungere tale scopo, il modello neurale è stato sottoposto a una fase di addestramento, durante la quale alcune sinapsi della rete subiscono delle modifiche. Le sinapsi coinvolte nel training sono solo quelle laterali. Come legge di addestramento è stata scelta la regola di apprendimento di Hebb; essa rappresenta un modo per riprodurre il fenomeno della plasticità sinaptica dal punto di vista computazionale e, nella sua forma più generale, stabilisce che due neuroni attivi contemporaneamente rafforzano il peso delle loro connessioni sinaptiche. Le regole di apprendimento adottate sono state tratte dal precedente lavoro sviluppato dai ricercatori dell'Università di Bologna [22].

In particolare, nel modello realizzato le espressioni della regola di Hebb per le sinapsi eccitatorie e inibitorie visive sono le seguenti:

$$\tau_L \frac{dL_{ex,jk}^v(t)}{dt} = \alpha_{ex,jk}^v(t) \cdot y_k^v(t) \cdot [y_j^v(t) - \theta_{post}]^+ \quad \text{Eq. 3.18}$$

$$\tau_L \frac{dL_{in,jk}^v(t)}{dt} = -\alpha_{in,jk}^v(t) \cdot y_k^v(t) \cdot [y_j^v(t) - \theta_{post}]^+ \quad \text{Eq. 3.19}$$

dove τ_L è la costante di tempo, $\alpha^{v_{ex,jk}}$ è il fattore di apprendimento per le sinapsi eccitatorie, $\alpha^{v_{in,jk}}$ è il fattore di apprendimento relativo alle sinapsi inibitorie, θ_{post} indica un valore di soglia con cui viene confrontata l'attività y_j^v relativa al neurone postsinaptico e y_k^v è l'attività del neurone presinaptico. La notazione $[]^+$ rappresenta una funzione che è uguale al suo argomento quando questo è positivo, uguale a zero altrimenti. Le Eq. 3.18 e 3.19 costituiscono quindi delle regole di apprendimento con una soglia per l'attività postsinaptica, nel senso che la sinapsi tra due neuroni si modifica solo se l'attività della cellula postsinaptica supera il valore θ_{post} . In base alle Eq. 3.18 e 3.19, le sinapsi eccitatorie possono crescere e quelle inibitorie diminuire. Inoltre sono stati imposti due vincoli fisiologici alle modificazioni delle sinapsi. Il primo è rappresentato da una legge di saturazione per le connessioni eccitatorie: le sinapsi eccitatorie non possono superare un valore massimo indicato con L_{max} . Questo viene realizzato assumendo che il rateo di apprendimento delle sinapsi eccitatorie rispetti la seguente regola:

$$\alpha^{v_{ex,jk}}(t) = \alpha_{ex0} \cdot [L_{max} - L^{v_{ex,jk}}(t)] \quad \text{Eq. 3.20}$$

dove $\alpha_{ex0} \cdot L_{max}$ è il fattore di apprendimento massimo, che si ottiene nel caso in cui la forza della sinapsi è nulla. Secondo l'Eq. 3.20, il fattore di apprendimento $\alpha^{v_{ex,jk}}$ diminuisce al crescere della forza sinaptica. Analogamente per le sinapsi inibitorie si assume che non possano scendere al di sotto del valore nullo; a tale scopo il rateo di apprendimento segue la seguente legge:

$$\alpha^{v_{in,jk}}(t) = \alpha_{in0} \cdot L^{v_{in,jk}}(t) \quad \text{Eq. 3.21}$$

In base alla precedente espressione, il fattore di apprendimento decresce con il decrescere della forza della sinapsi. Il secondo vincolo è rappresentato dal fatto che la somma delle sinapsi eccitatorie e inibitorie che entrano in un certo neurone deve rimanere costante: se una sinapsi eccitatoria cresce, allora un'altra connessione eccitatoria deve indebolirsi; allo stesso modo se una sinapsi inibitoria decresce, allora un'altra connessione inibitoria deve aumentare. Per realizzare questo vincolo, a ogni

step di simulazione tutte le sinapsi eccitatorie ed inibitorie che arrivano su un neurone postsinaptico sono state normalizzate secondo queste equazioni:

$$L_{ex,jk}^v(t + \Delta t) = \frac{L_{ex,jk}^v(t) + \Delta L_{ex,jk}^v(t)}{\sum_{k=1}^{N_p} (L_{ex,jk}^v(t) + \Delta L_{ex,jk}^v(t))} \cdot \sum_{k=1}^{N_p} L_{ex,jk}^v(0) \quad \text{Eq. 3.22}$$

$$L_{in,jk}^v(t + \Delta t) = \frac{L_{in,jk}^v(t) + \Delta L_{in,jk}^v(t)}{\sum_{k=1}^{N_p} (L_{in,jk}^v(t) + \Delta L_{in,jk}^v(t))} \cdot \sum_{k=1}^{N_p} L_{in,jk}^v(0) \quad \text{Eq. 3.23}$$

dove Δt è il passo di integrazione e ΔL indica la variazione delle sinapsi calcolata con le Eq. 3.18-3.21 durante un singolo passo di integrazione.

Anche le sinapsi laterali acustiche sono state addestrate mediante analoghe regole di Hebb:

$$\tau_L \frac{dL_{ex,ij,hk}^a(t)}{dt} = \alpha_{ex,ij,hk}^a(t) \cdot y_{hk}^a(t) \cdot [y_{ij}^a(t) - \theta_{post}]^+ \quad \text{Eq. 3.24}$$

$$\tau_L \frac{dL_{in,ij,hk}^a(t)}{dt} = -\alpha_{in,ij,hk}^a(t) \cdot y_{hk}^a(t) \cdot [y_{ij}^a(t) - \theta_{post}]^+ \quad \text{Eq. 3.25}$$

Valgono gli stessi vincoli fisiologici descritti per le sinapsi laterali visive. Perciò ciascuna sinapsi eccitatoria non può superare un valore massimo di saturazione e ogni sinapsi inibitoria non può scendere al di sotto dello zero; tutto ciò è stato ottenuto usando dei fattori di apprendimento di questo tipo:

$$\alpha_{ex,ij,hk}^a(t) = \alpha_{ex0} \cdot [L_{max} - L_{ex,ij,hk}^a(t)] \quad \text{Eq. 3.26}$$

$$\alpha_{in,ij,hk}^a(t) = \alpha_{in0} \cdot L_{in,ij,hk}^a(t) \quad \text{Eq. 3.27}$$

Inoltre l'input sinaptico per ciascun neurone è mantenuto costante tramite delle espressioni che normalizzano le sinapsi eccitatorie e inibitorie:

$$L_{ex,ij,hk}^a(t + \Delta t) = \frac{L_{ex,ij,hk}^a(t) + \Delta L_{ex,ij,hk}^a(t)}{\sum_{h=1}^{N_f} \sum_{k=1}^{N_p} (L_{ex,ij,hk}^a(t) + \Delta L_{ex,ij,hk}^a(t))} \cdot \sum_{h=1}^{N_f} \sum_{k=1}^{N_p} L_{ex,jk}^a(0) \quad \text{Eq. 3.28}$$

$$L_{in,ij,hk}^a(t + \Delta t) = \frac{L_{in,ij,hk}^a(t) + \Delta L_{in,ij,hk}^a(t)}{\sum_{h=1}^{N_f} \sum_{k=1}^{N_p} (L_{in,ij,hk}^a(t) + \Delta L_{in,ij,hk}^a(t))} \cdot \sum_{h=1}^{N_f} \sum_{k=1}^{N_p} L_{in,jk}^a(0) \quad \text{Eq. 3.29}$$

3.6 Calcolo della posizione percepita dello stimolo

L'aspetto fondamentale del fenomeno del ventriloquismo e del suo aftereffect è che la posizione di un input acustico è percepita come se fosse spostata verso la posizione di un contemporaneo stimolo visivo. Perciò è necessario calcolare una qualche grandezza che sia rappresentativa della posizione percepita di uno stimolo, per capire se il modello è in grado di riprodurre l'effetto on-line e quello off-line del ventriloquismo. La posizione percepita dello stimolo esterno (visivo o acustico) viene determinata sulla base dell'attivazione complessiva della corrispondente area unimodale. In particolare è stata scelta la metrica del baricentro per rappresentare la posizione percepita. Il baricentro è il valore medio delle posizioni codificate da ciascun neurone pesate dalle rispettive attività neuronali. Secondo la metrica del baricentro, la posizione dello stimolo fornito al modello viene percepita dai neuroni visivi secondo questa espressione:

$$z_{bar}^v = \frac{\sum_{k=1}^{N_p} y_k^v \cdot k}{\sum_{k=1}^{N_p} y_k^v} \quad \text{Eq. 3.30}$$

dove y_k^v indica l'attività di un neurone visivo che si trova nella posizione k . Analogamente, il baricentro relativo ai neuroni dell'area acustica è dato dalla seguente relazione:

$$z_{bar}^a = \frac{\sum_{h=1}^{N_f} \sum_{k=1}^{N_p} y_{hk}^a \cdot k}{\sum_{h=1}^{N_f} \sum_{k=1}^{N_p} y_{hk}^a} \quad \text{Eq.}$$

3.31

3.7 Addestramento della rete

Il paradigma di addestramento utilizzato per ricreare la situazione di esposizione al conflitto audio-visivo è stato impostato in maniera tale da realizzare una modalità di training simile a quelle adoperate in molti studi in vivo. In particolare, per ottenere l'aftereffect del ventriloquismo, la rete neurale è stata sottoposta a esposizioni ripetute di stimoli acustici e visivi simultanei e spazialmente separati. La procedura di adattamento per ciascuna delle tre configurazioni in cui il modello è stato testato verrà spiegata in dettaglio nei capitoli successivi.

3.8 Assegnazione dei valori dei parametri

Il modello è stato testato in tre differenti configurazioni, corrispondenti a tre diversi set di parametri, allo scopo di simulare la generalizzazione e la non generalizzazione in frequenza dell'aftereffect del ventriloquismo.

In particolare, dapprima sono stati individuati i principali parametri dell'area acustica che influenzano il comportamento in frequenza dell'aftereffect; questi sono la

deviazione standard in frequenza dello stimolo esterno e il bilancio tra eccitazione e inibizione delle sinapsi laterali. Sono stati quindi individuati due diversi insiemi di valori per questi parametri, così che utilizzando un set di parametri si ottiene la generalizzazione in frequenza, mentre con l'altro l'aftereffect rimane localizzato alla frequenza di training. Questi differenti set di parametri potrebbero indicare una variabilità tra individui o anche tra neuroni della stessa popolazione.

In alternativa, si è cercato di capire se, a parità di set di parametri, differenze nell'intensità degli input acustici possono, almeno parzialmente, spiegare il differente comportamento in frequenza dell'aftereffect.

Alcuni valori dei parametri sono comuni per tutte e tre le configurazioni, altri invece sono differenti. L'assegnazione dei valori numerici ai diversi parametri è stata fatta cercando di seguire alcuni criteri. Innanzitutto un'assunzione fondamentale riguarda le deviazioni standard dell'input esterno (σ^v , σ^{ap} e σ^{af}): il sistema visivo ha una risoluzione spaziale migliore rispetto a quello acustico e questo significa che i campi recettivi dei neuroni visivi sono più stretti di quelli dei neuroni uditivi. La scelta dei valori delle deviazioni standard dello stimolo esterno è stata fatta in base a questa ipotesi. Per quel che concerne le caratteristiche dei neuroni, i parametri sono stati impostati agli stessi valori per tutti i neuroni, in maniera tale da semplificare il modello. In particolare, il valore della posizione centrale della caratteristica sigmoideale (θ) consente di ottenere un'attività trascurabile in assenza di stimoli esterni, mentre la pendenza della relazione sigmoideale (s) ha un valore tale per cui si ha una transizione dolce dalla condizione di inattività a quella di saturazione in risposta a stimoli esterni. I valori delle sinapsi inter-area (W_{va} e W_{av}) sono stati scelti in modo tale che uno stimolo in una modalità non induca un effetto fantasma nell'altra area, cioè un'attività apprezzabile senza che la seconda area venga stimolata. Se i pesi delle connessioni inter-area fossero stati impostati allo stesso valore, ci sarebbe stato uno sbilanciamento nella forza di attrazione da parte dell'area acustica, dato che essa presenta un numero maggiore di neuroni; perciò i valori dei pesi sono stati scelti anche in base a questa considerazione.

Di seguito sono indicati prima i parametri che hanno valori comuni per tutte le configurazioni, poi quelli caratteristici di ciascuna di esse. Dato che ciò che differenzia una configurazione dall'altra è rappresentato dai parametri relativi all'input acustico esterno e alle sinapsi laterali acustiche, per ogni set di parametri

sono mostrati lo stimolo uditivo esterno applicato alla posizione 80 e alla frequenza 20 (che rappresenta 1100 Hz) e le connessioni laterali acustiche prima dell'addestramento entranti in un neurone uditivo di posto $i=20$ e $j=100$ (una spiegazione più dettagliata di quanto illustrato è riportata nei capitoli successivi).

Parametri con valori comuni

Numero di neuroni della rete: $N_p = 180$ $N_f = 40$

Stimolo esterno visivo: $E_0^v = 15$ $\sigma^v = 4$

Stimolo esterno acustico: $\sigma^{ap} = 30$

Caratteristiche dei neuroni: $\tau_y = 3$ $s = 0.6$ $\theta = 12$

Sinapsi laterali visive: $L_{ex0}^v = 2.4$ $L_{in0}^v = 1.4$ $\sigma_{ex,v} = 2$ $\sigma_{in,v} = 24$

Sinapsi inter-area: $W_{va} = 0.22$ $W_{av} = 8.5$

Regole di Hebb: $\tau_L = 3$ $\theta_{post} = 0.5$ $\alpha_{ex0} = 0.03$ $\alpha_{in0} = 0.05$

Parametri della configurazione per la non generalizzazione in frequenza

Stimolo esterno acustico: $E_0^a = 15$ $\sigma^{af} = 8$

Sinapsi laterali acustiche: $L_{ex0}^a = 0.47$ $L_{in0}^a = 0.2$ $\sigma_{ex,ap} = 2$ $\sigma_{in,ap} = 10.5$

$\sigma_{ex,af} = 2$ $\sigma_{in,af} = 8.2$

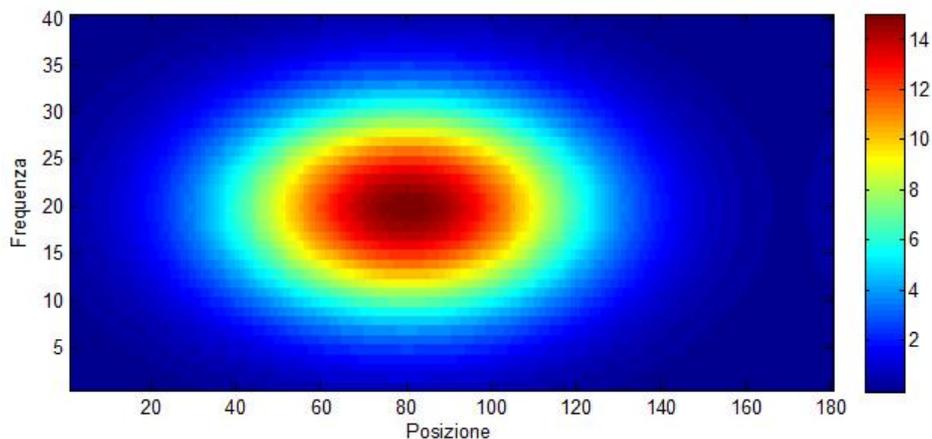


Figura 3.8. Input acustico per la prima configurazione.

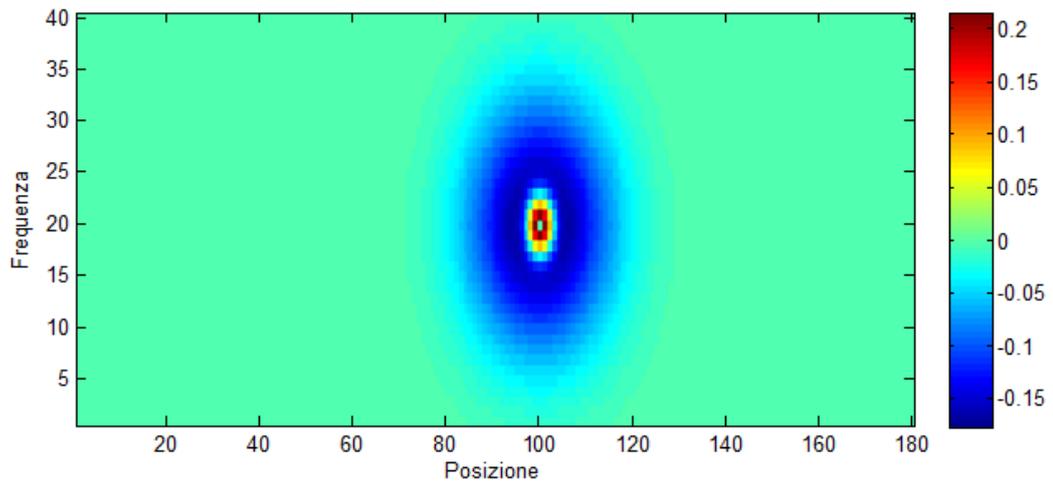


Figura 3.9. Sinapsi laterali acustiche per la prima configurazione.

Parametri della configurazione per la generalizzazione in frequenza

Stimolo esterno acustico: $E_0^a = 15$ $\sigma^{af} = 18$

Sinapsi laterali acustiche: $L_{ex0}^a = 0.4$ $L_{in0}^a = 0.21$ $\sigma_{ex,ap} = 2$ $\sigma_{in,ap} = 10$

$\sigma_{ex,af} = 2.45$ $\sigma_{in,af} = 3$

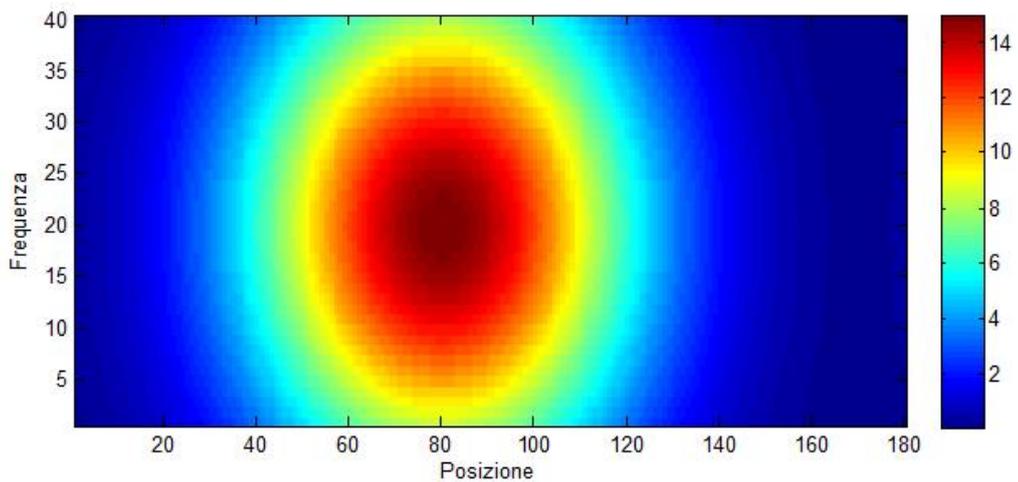


Figura 3.10. Input acustico per la seconda configurazione.

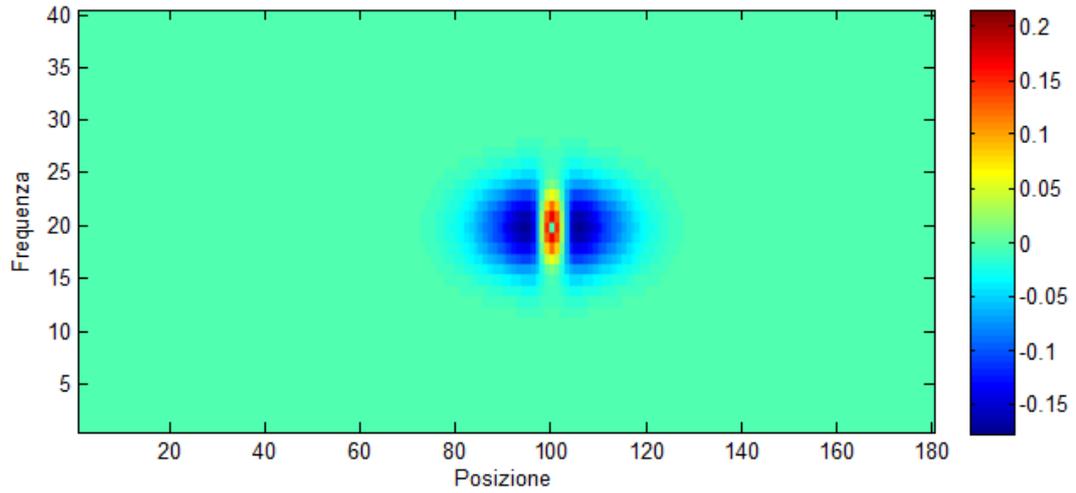


Figura 3.11. sinapsi laterali acustiche per la seconda configurazione.

Parametri della configurazione per l'influenza dell'intensità dell'input esterno

Stimolo esterno acustico: $E_0^a = \text{variabile (13-15-18-20-22)}$ $\sigma^{af} = 14$

Sinapsi laterali acustiche: $L_{ex0}^a = 0.4$ $L_{in0}^a = 0.21$ $\sigma_{ex,ap} = 1.45$ $\sigma_{in,ap} = 8$

$\sigma_{ex,af} = 1.45$ $\sigma_{in,af} = 5$

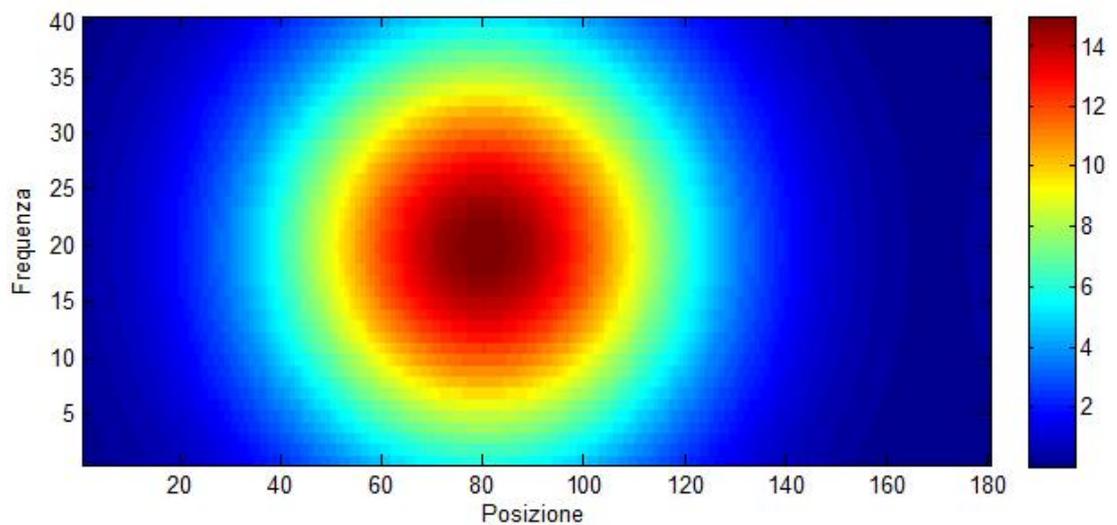


Figura 3.12. Input acustico per la terza configurazione ($E_{0x}=15$).

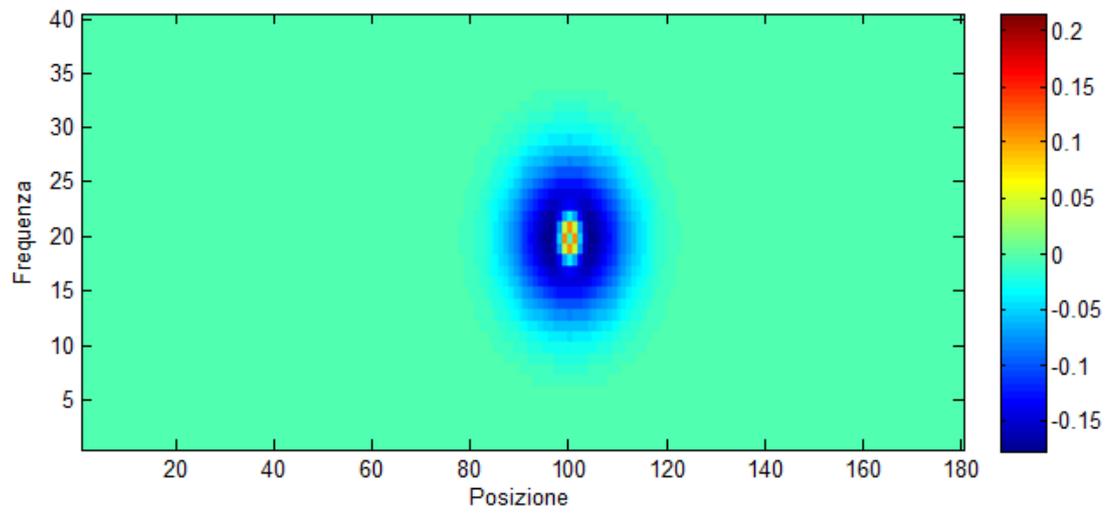


Figura 3.13. Sinapsi laterali acustiche per la terza configurazione.

Capitolo 4

Risultati delle configurazioni per la non generalizzazione e la generalizzazione in frequenza dell'aftereffect

4.1 Aspetti generali

In questo capitolo vengono riportati i risultati forniti dal modello quando quest'ultimo è stato testato con i parametri che danno luogo alla generalizzazione in frequenza o meno dell'aftereffect del ventriloquismo.

Gli esiti ottenuti dalle simulazioni in MATLAB®, per ciascuna delle due configurazioni utilizzate, possono essere distinti in due tipologie, ovvero quelli ricavati prima di addestrare la rete neurale (pre-addestramento) e quelli ottenuti dopo la fase di adattamento, compiuta tramite esposizione ripetuta a input acustici e visivi concomitanti ma spazialmente discrepanti (post-addestramento). I risultati pre-addestramento riproducono l'effetto on-line del ventriloquismo, quelli post-addestramento il suo aftereffect.

Per ottenere i dati di pre-addestramento, sono state compiute diverse prove di simulazione. Innanzitutto il modello è stato testato mediante una stimolazione unisensoriale, cioè ha ricevuto in ingresso solo l'input visivo o solo quello acustico. Questo tipo di simulazione serve a verificare che non ci siano effetti fantasma, ovvero che lo stimolo in una modalità non induca una risposta significativa nell'altra modalità. Un secondo tipo di simulazione ha riguardato l'applicazione di un segnale acustico e un segnale uditivo simultanei e spazialmente coincidenti. Infine la rete è stata stimolata con segnali acustici e visivi concomitanti ma distanti dal punto di vista spaziale.

Per ricavare i risultati post-addestramento il modello è stato sottoposto a una fase di training tramite due diversi paradigmi, replicati per le due le configurazioni in esame; dopodiché, per valutare l'aftereffect, sono stati presentati stimoli unicamente

acustici. Vengono di seguito mostrati separatamente gli esiti delle simulazioni per ognuna delle due configurazioni.

4.2 Risultati della configurazione per la non generalizzazione in frequenza dell'aftereffect

4.2.1 Pre-addestramento

4.2.1.1 Stimolazione unimodale visiva

In questo tipo di test l'ingresso del modello è rappresentato da un input esterno acustico centrato nella posizione 100, generato in base all'Eq. 3.4, che, per una maggiore chiarezza, viene di seguito riportata:

$$e_j^v(t) = E_0^v e^{-\frac{(d_j^v)^2}{2(\sigma^v)^2}}$$

Sulla base di questa espressione viene creato un input visivo della forma di quello mostrato in Figura 4.1.

In Figura 4.2 viene riportata l'attivazione corrispondente nell'area visiva monodimensionale. Si può osservare che il profilo di attivazione nell'area visiva rispecchia la forma del campo recettivo dei neuroni di quest'area e cioè un campo molto stretto dato dall'elevata risoluzione spaziale del sistema visivo.

La Figura 4.3 mostra invece l'attivazione nell'area acustica. A differenza dello strato visivo, quello uditivo è bidimensionale, dato che i neuroni acustici codificano sia per la posizione che per la frequenza. Per maggiore chiarezza in Figura 4.4 viene mostrato il profilo di attività acustica rispetto alla posizione per una generica frequenza.

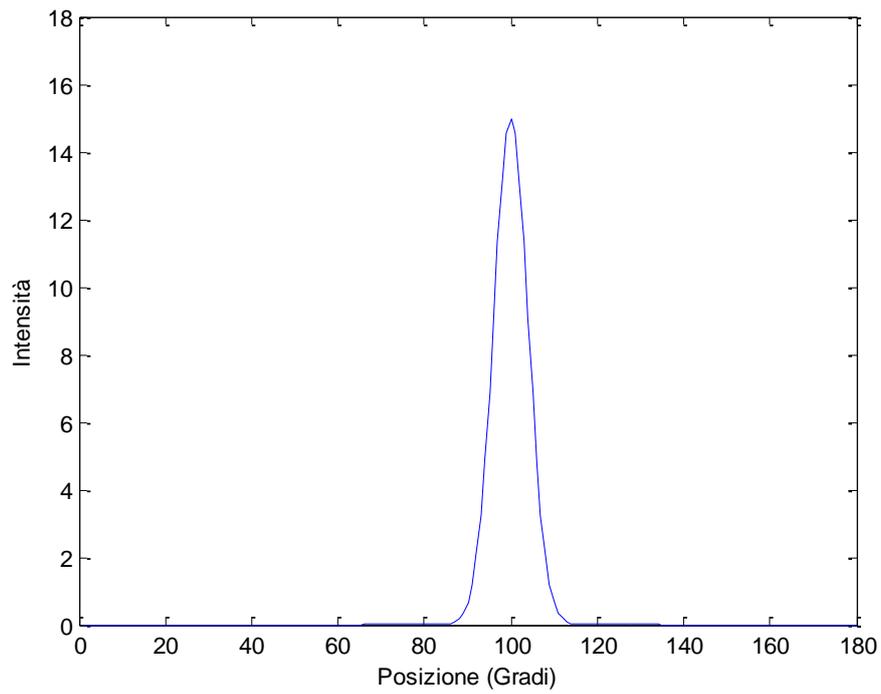


Figura 4.1. Input visivo.

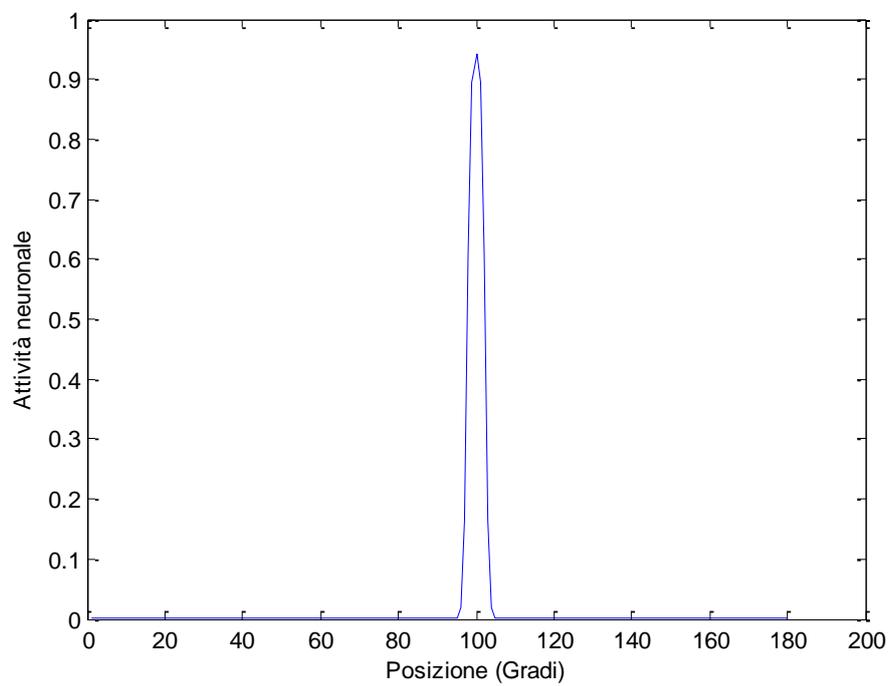


Figura 4.2. Profilo di attività nell'area visiva rispetto alla posizione in seguito a stimolazione unimodale visiva.

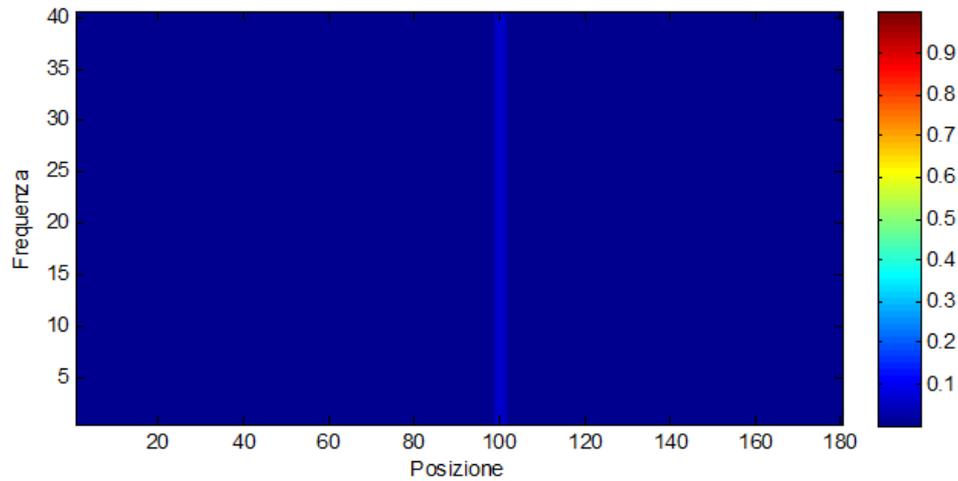


Figura 4.3. Attivazione nell'area acustica in seguito a stimolazione unimodale visiva.

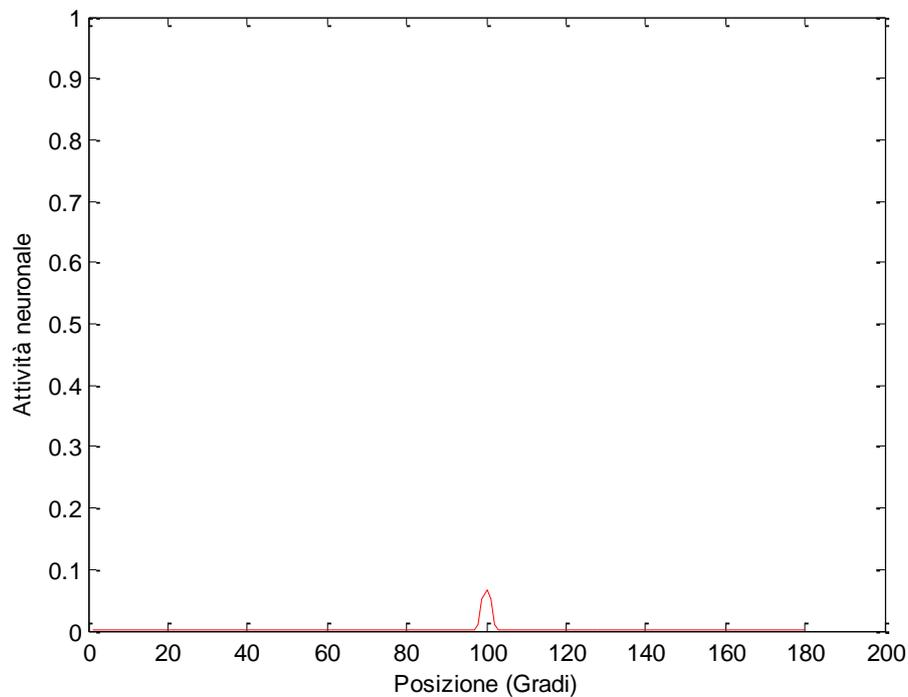


Figura 4.4. Profilo di attività nell'area acustica rispetto alla posizione in seguito a stimolazione unimodale visiva.

E' possibile notare che i neuroni dello strato uditivo presentano un'attivazione trascurabile in corrispondenza della posizione in cui è centrato l'input esterno e un'attivazione nulla altrove. Si può quindi concludere che non si genera effetto fantasma.

4.2.1.2 Stimolazione unimodale acustica

Alla rete è stato dato in ingresso uno stimolo esterno acustico centrato alla posizione 80 e alla frequenza 20 (che in base a quanto detto nel paragrafo 3.2 corrisponde a una frequenza di 1100 Hz) e generato in base all'Eq. 3.6:

$$e_{ij}^a(t) = E_0^a e^{-\left[\frac{(d_j^a)^2}{2(\sigma^{ap})^2} + \frac{(d_i^a)^2}{2(\sigma^{af})^2} \right]}$$

Sulla base di tale espressione, viene creato l'input acustico di Figura 4.5. A differenza dello stimolo visivo, quello acustico è bidimensionale, poiché i neuroni dell'area uditiva sono definiti sia in frequenza che in posizione.

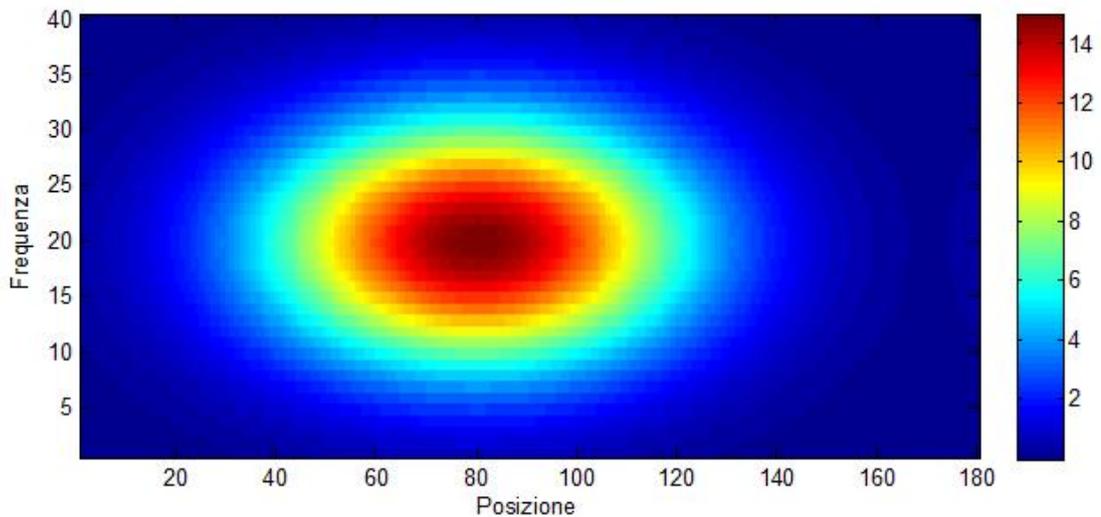


Fig. 4.5. Input acustico.

In risposta a tale stimolo si avrà una mappa bidimensionale di attivazione acustica, come mostrato in Figura 4.6. Per meglio comprendere la morfologia di tale attivazione, in Figura 4.7 viene illustrato il profilo di attività acustica rispetto alla posizione in corrispondenza della frequenza di 1100 Hz (frequenza centrale dello stimolo). Si può osservare che il profilo di attivazione acustica è più largo rispetto a quello dell'area visiva di Figura 4.2 e questo è coerente con il fatto che la risoluzione spaziale uditiva è inferiore a quella visiva.

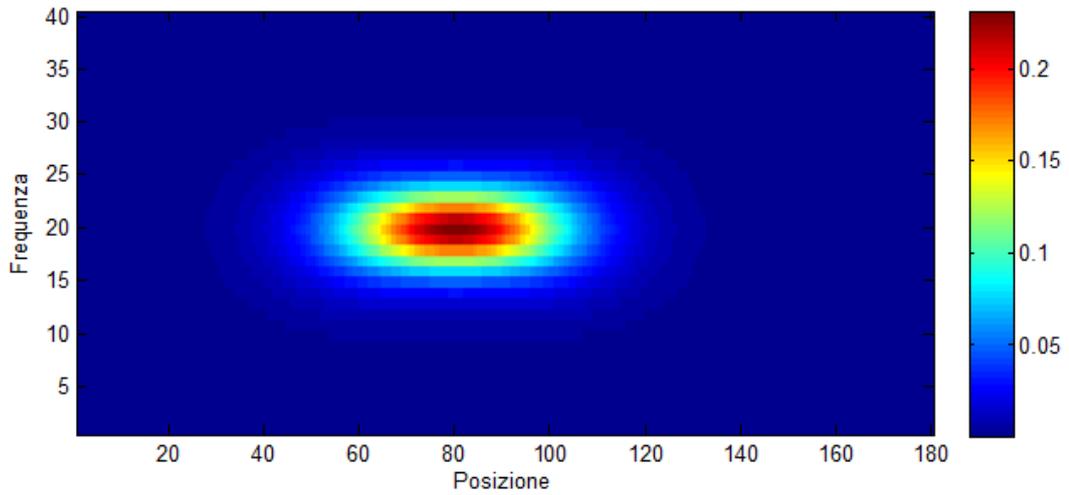


Figura 4.6. Attivazione nell'area uditiva in seguito a stimolazione unimodale acustica.

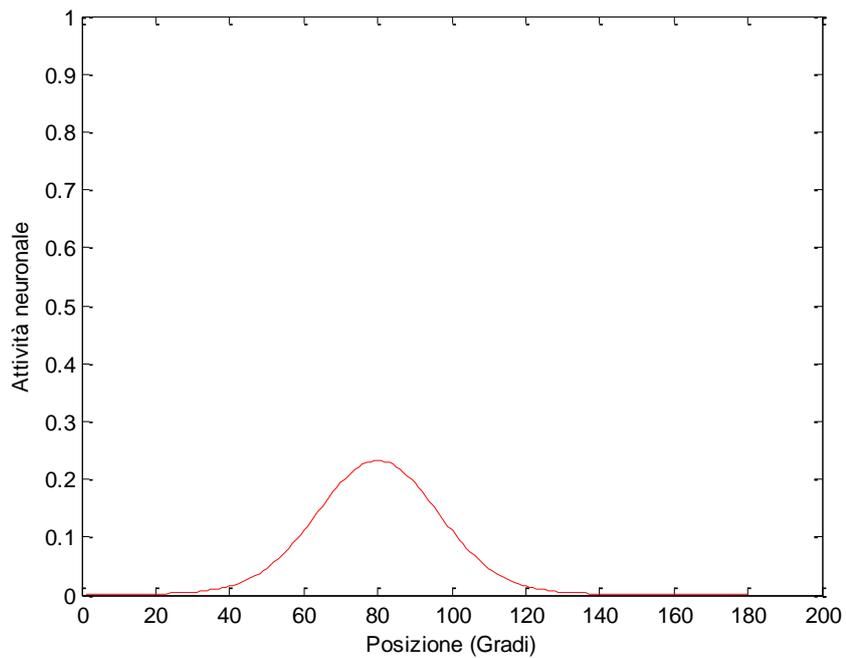


Figura 4.7. Profilo di attività nell'area acustica rispetto alla posizione in seguito a stimolazione unimodale acustica.

Inoltre in Figura 4.8 viene riportato il profilo di attività acustica rispetto alla frequenza in corrispondenza della posizione 80 (posizione centrale dello stimolo).

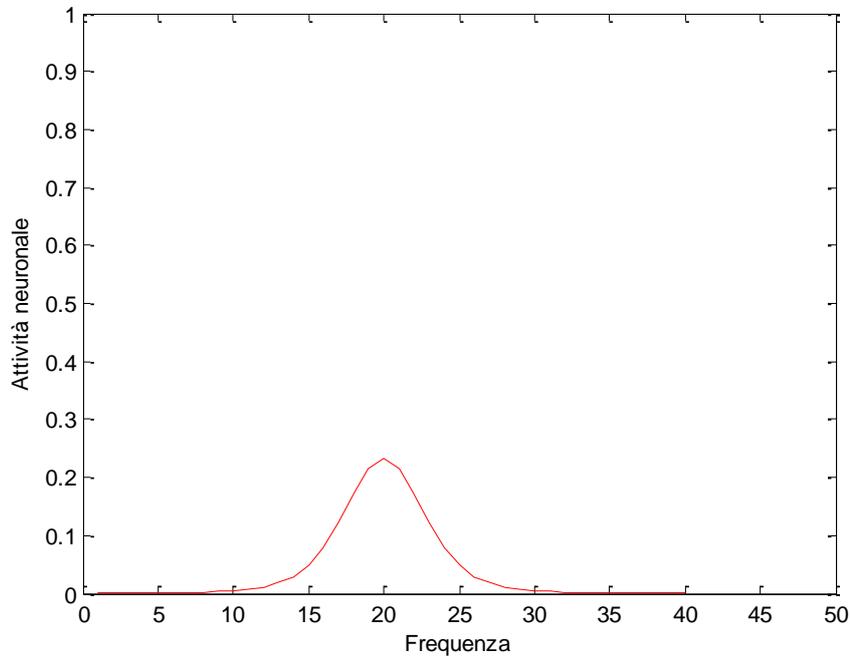


Figura 4.8. Profilo di attività rispetto alla frequenza nell'area uditiva in seguito a stimolazione unimodale acustica.

Si osserva un'attivazione lungo la frequenza abbastanza stretta (quindi molto selettiva) e questo è in accordo con la caratteristica della configurazione in esame per cui l'aftereffect del ventriloquismo non si dovrà trasferire attraverso le frequenze sonore.

Viene infine illustrata l'attivazione nell'area visiva in presenza del solo input acustico (Figura 4.9). Analogamente al caso di stimolazione unimodale visiva, si osserva che non c'è effetto fantasma nella modalità non stimolata.

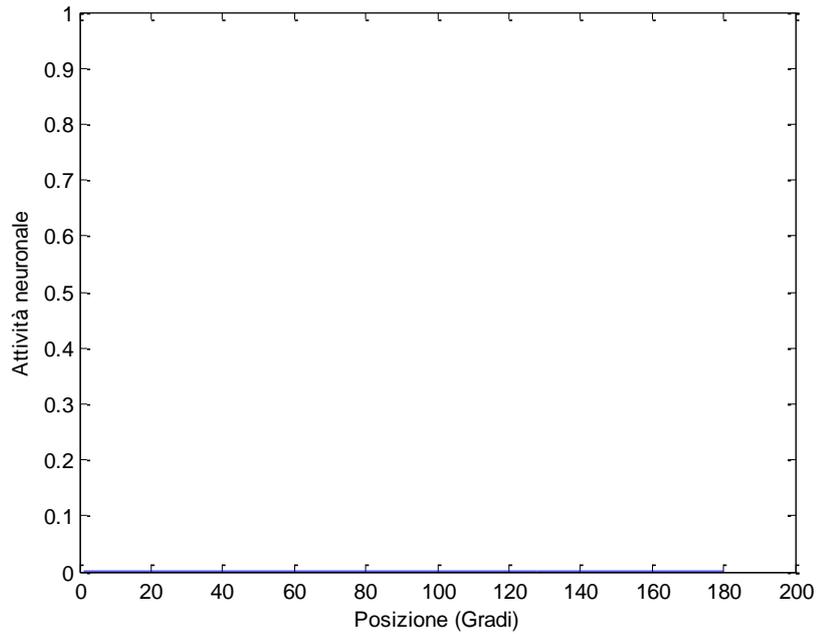


Figura 4.9. Profilo di attività nell'area visiva rispetto alla posizione in seguito a stimolazione unimodale acustica.

4.2.1.3 Stimoli spazialmente coincidenti

Il modello ha ricevuto in ingresso contemporaneamente uno segnale uditivo e uno visivo posizionati nella medesima locazione spaziale. La Figura 4.10 mostra i profili di attività per le due modalità dopo che sono stati applicati due input in posizione 80 (frequenza di applicazione dell'input acustico sempre pari a 1100 Hz). La Figura 4.11 mostra invece l'attivazione nell'intera area uditiva. In questo caso, le attività in entrambe le aree risultano molto alte nel punto di applicazione dei due input; questo si verifica grazie alla presenza delle sinapsi inter-area che collegano i due strati e che favoriscono un rinforzo reciproco delle due attivazioni. I risultati di questa simulazione sono coerenti con gli esperimenti condotti in vivo [23], in base ai quali la capacità di localizzazione del sistema uditivo migliora se viene presentato un segnale visivo nella stessa posizione di quello acustico.

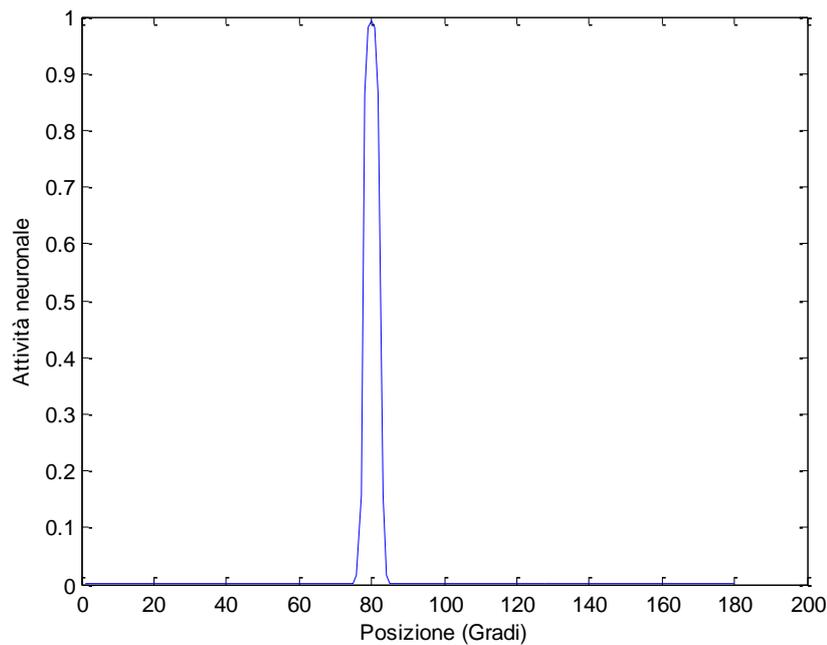
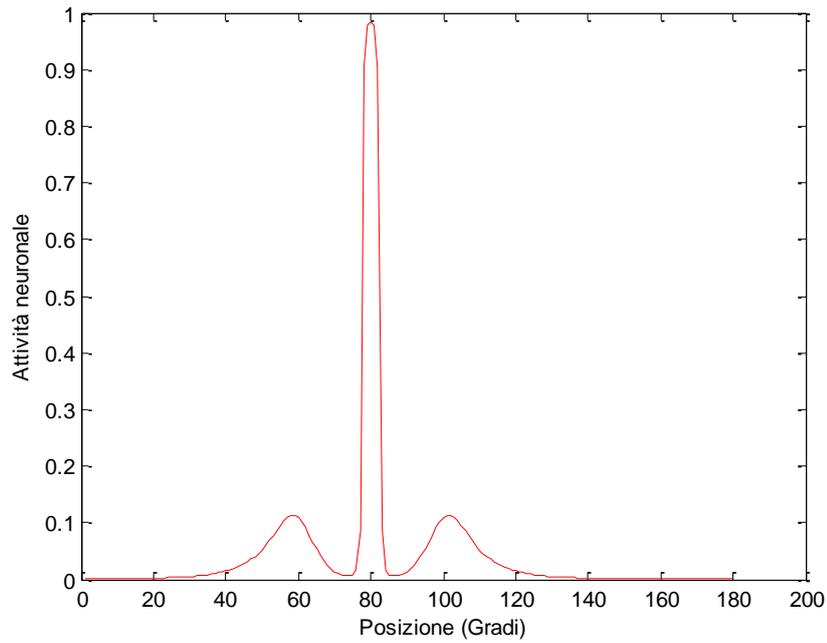


Figura 4.10. Profili di attività rispetto alla posizione nelle aree acustica (sopra) e visiva (sotto) dopo la presentazione di stimoli spazialmente coincidenti. Il profilo di attivazione acustica è mostrato in corrispondenza della frequenza di 1100 Hz.

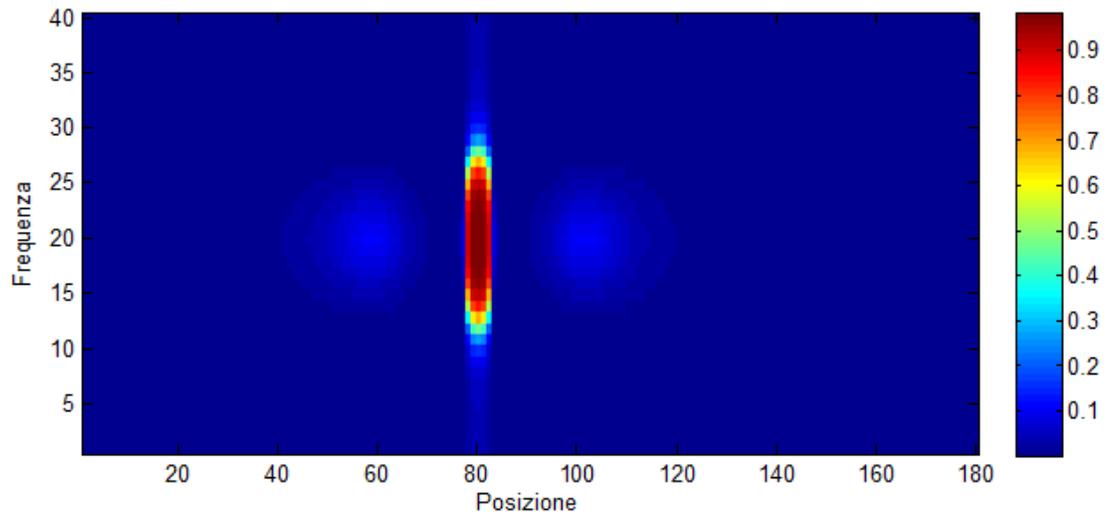


Figura 4.11. Attivazione nell'area uditiva dopo la presentazione di stimoli spazialmente coincidenti.

4.2.1.4 Stimoli spazialmente separati

Sono stati applicati al modello uno stimolo acustico e uno visivo simultanei ma spazialmente separati. In Figura 4.12 viene mostrata la conseguente attivazione nell'intera area acustica.

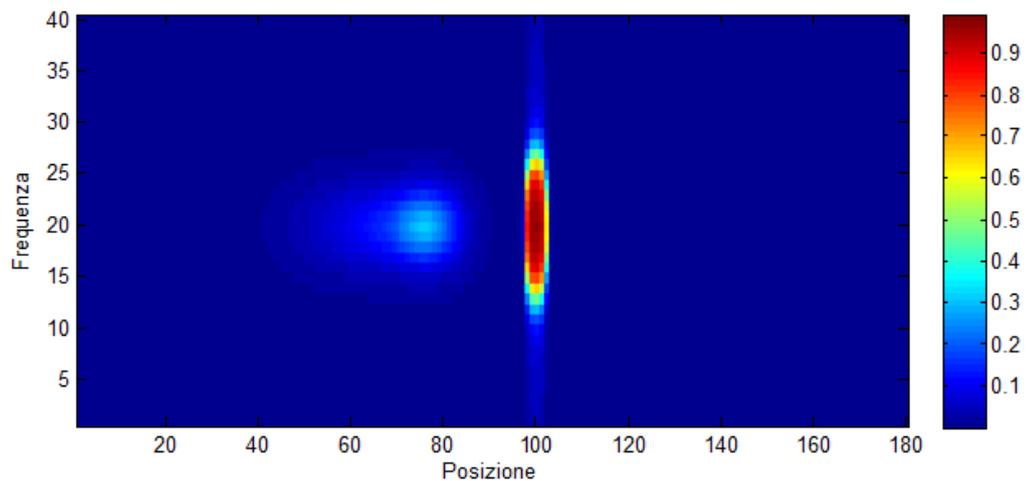


Figura 4.12. Attivazione nell'area acustica dopo la presentazione di due stimoli contemporanei e spazialmente separati.

L'input acustico è stato dato in posizione 80 e alla frequenza 20 (1100 Hz) e quello visivo in posizione 100. La bolla di attivazione coinvolge pochi neuroni lungo la frequenza (dal neurone di posto 15, che corrisponde a 700 Hz, fino a quello di posto 25, che codifica per 1600 Hz). Questo tipo di attività è in accordo con il fatto che la configurazione in esame è stata pensata per non trasferire l'aftereffect attraverso le frequenze sonore. Si nota inoltre che l'attività dei neuroni acustici è fortemente spostata verso la posizione in cui è stato applicato l'input visivo. Per comprendere meglio ciò che avviene, in Figura 4.13 è mostrato il profilo di attività acustica rispetto alla posizione (in corrispondenza della frequenza centrale di 1100 Hz) confrontato con il profilo di attività visiva. La posizione percepita dell'input acustico subisce uno shift: in base alla metrica del baricentro, essa risulta pari a 86.08° , perciò l'alterazione nella percezione è uguale a 6.08° , un valore compatibile con i risultati riportati in letteratura [15,16,18]. La posizione del segnale visivo non è influenzata dal concomitante segnale acustico, come dimostra il fatto che il baricentro visivo sia pari a 99.95° . Ciò indica che il modello è in grado di riprodurre bene l'effetto on-line del ventriloquismo.

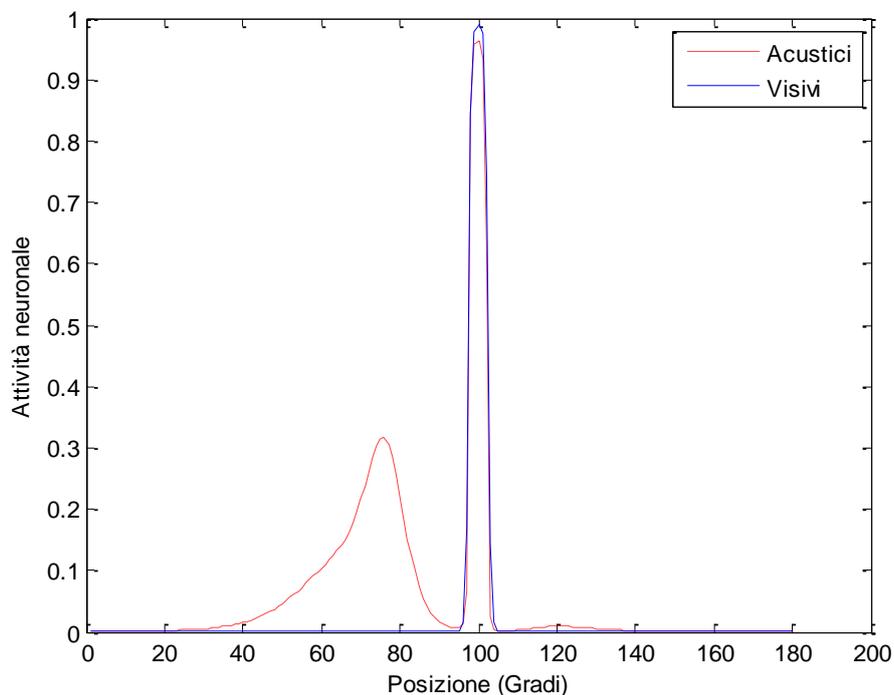


Figura 4.13. Profili di attività rispetto alla posizione nelle aree acustica e visiva dopo la presentazione di stimoli spazialmente separati. Il profilo di attivazione acustica è mostrato in corrispondenza della frequenza di 1100 Hz.

La Figura 4.14 mostra il fenomeno appena descritto a diversi istanti di simulazione, mettendo in evidenza come l'attività acustica venga progressivamente attratta da quella visiva, mentre quest'ultima non subisca alcuno shift.

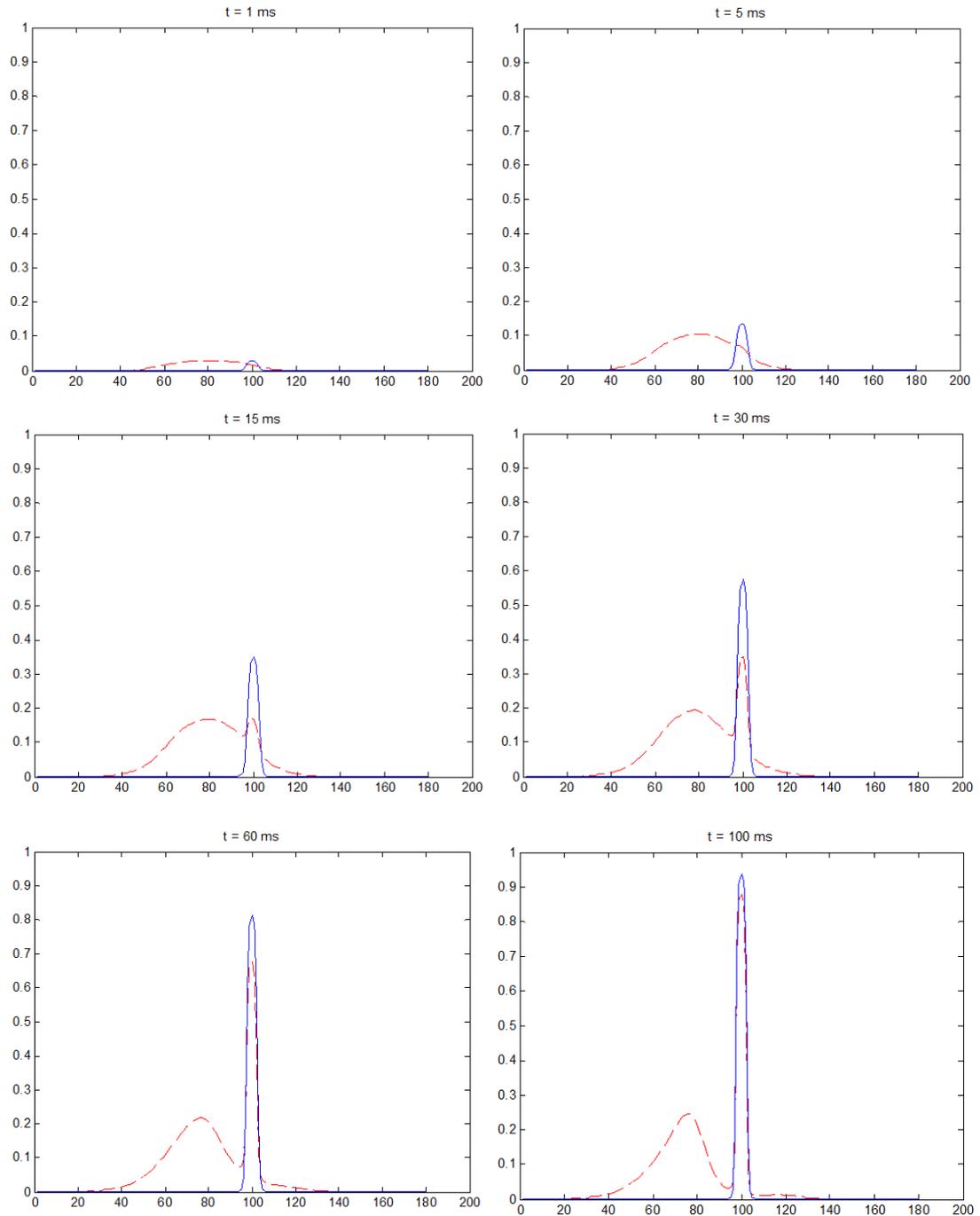


Figura 4.14. Risposta della rete a stimoli spazialmente separati per diversi istanti di tempo. In ascissa è riportata la posizione in gradi e in ordinata l'entità dell'attivazione dei neuroni.

Dopo 1 ms dalla presentazione della stimolazione cross-modale, l'attività nelle due aree sensoriali è ancora molto bassa, tuttavia si può osservare che si attiva un gran numero di neuroni acustici, mentre poche cellule visive rispondono all'input. Questo comportamento è da imputare alla forma dei campi recettivi dei neuroni dei due strati: ampio quello acustico e stretto (quindi molto selettivo) quello visivo. Dopo 15 ms dall'inizio della simulazione, nasce, grazie alla presenza delle sinapsi inter-area, un feedback positivo di rinforzo tra le attività uditive e visive in corrispondenza della posizione dell'input visivo p^v . In questo modo, i neuroni acustici attorno alla posizione p^v eccitano i loro vicini ed inibiscono sempre più le cellule lontane. Alla fine l'attivazione in corrispondenza di p^v diventa così forte da lasciare solo un'attività residua nell'intorno della posizione in cui è stato applicato l'input acustico.

Un'ulteriore prova di simulazione è stata quella in cui la posizione dell'input visivo è stata mantenuta fissa ($p^v=100$), mentre quella dell'input acustico è stata fatta variare ($p^a=70, 75, 80, 85, 90, 95, 98$). Gli esiti della prova, confrontati con alcuni risultati della letteratura [15,16,17], sono illustrati in Figura 4.15.

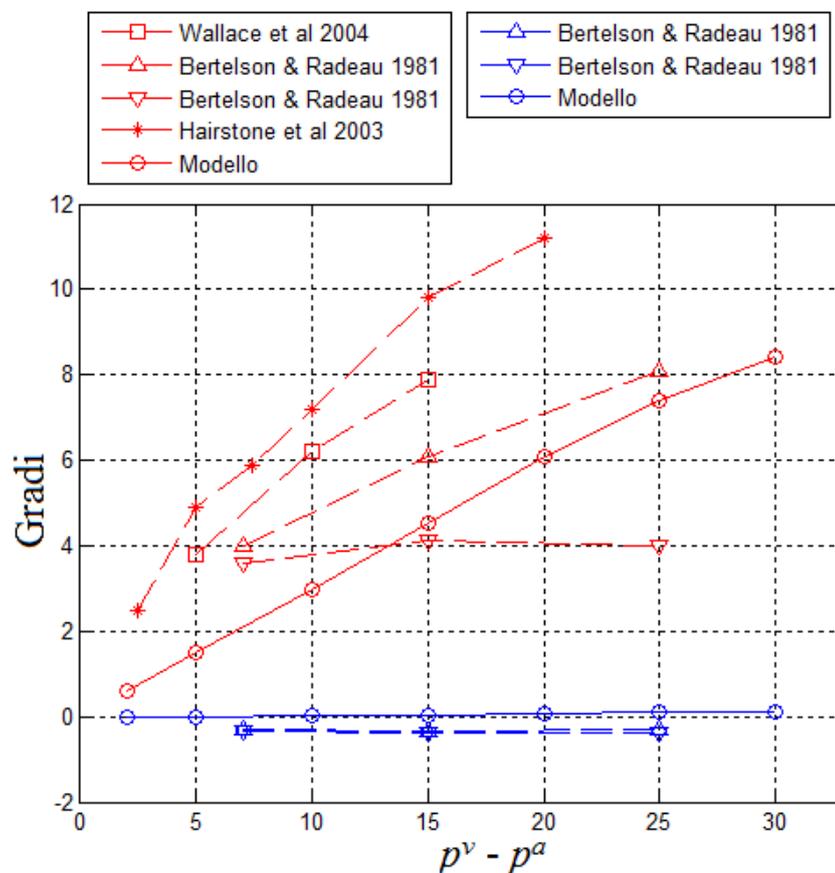


Figura 4.15. Confronto dei risultati del modello con quelli della letteratura.

Nella Figura sopra riportata, viene mostrato lo shift nella percezione acustica (in rosso) e in quella visiva (in blu), calcolato come posizione percepita meno posizione reale dello stimolo acustico, in funzione della differenza tra la posizione in cui è centrato l'input visivo (p^v) e quella in cui è applicato lo stimolo acustico (p^a). Si può notare come i valori forniti dal modello siano all'interno del range dei dati raccolti in esperimenti in vivo [15,16,17].

4.2.2 Post-addestramento – Primo paradigma di training

I risultati di seguito riportati sono stati ricavati dopo che il modello è stato addestrato tramite input acustici e visivi contemporanei e spazialmente separati, mantenendo fisse le posizioni dei due stimoli. L'input sonoro è stato presentato alla posizione 80 e alla frequenza 20 (1100 Hz), quello visivo alla posizione 100. La fase di addestramento si è composta di una prova della durata di 1000 ms. Alla fine dell'adattamento, per verificare la generazione dell'aftereffect, alla rete è stato applicato un input solo uditivo centrato nella stessa posizione spaziale di quelli usati nel training. La valutazione della generalizzazione o meno dell'aftereffect è stata compiuta presentando stimoli sonori a frequenze diverse da quella di training, in modo da coprire tutto l'intervallo di frequenze codificato dallo strato acustico del modello (200 - 6000 Hz).

4.2.2.1 Modifiche delle sinapsi

Una prima analisi è quella relativa alle modifiche a cui vanno incontro le sinapsi laterali a causa del periodo di training. La Figura 4.16 mostra tali variazioni per quanto riguarda l'area acustica. Le immagini superiore e centrale mostrano le sinapsi che entrano nel neurone di posto $i=20$ e $j=100$ nell'area acustica rispettivamente pre- e post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi, considerate però soltanto lungo la posizione per una frequenza fissa ($i=20$).

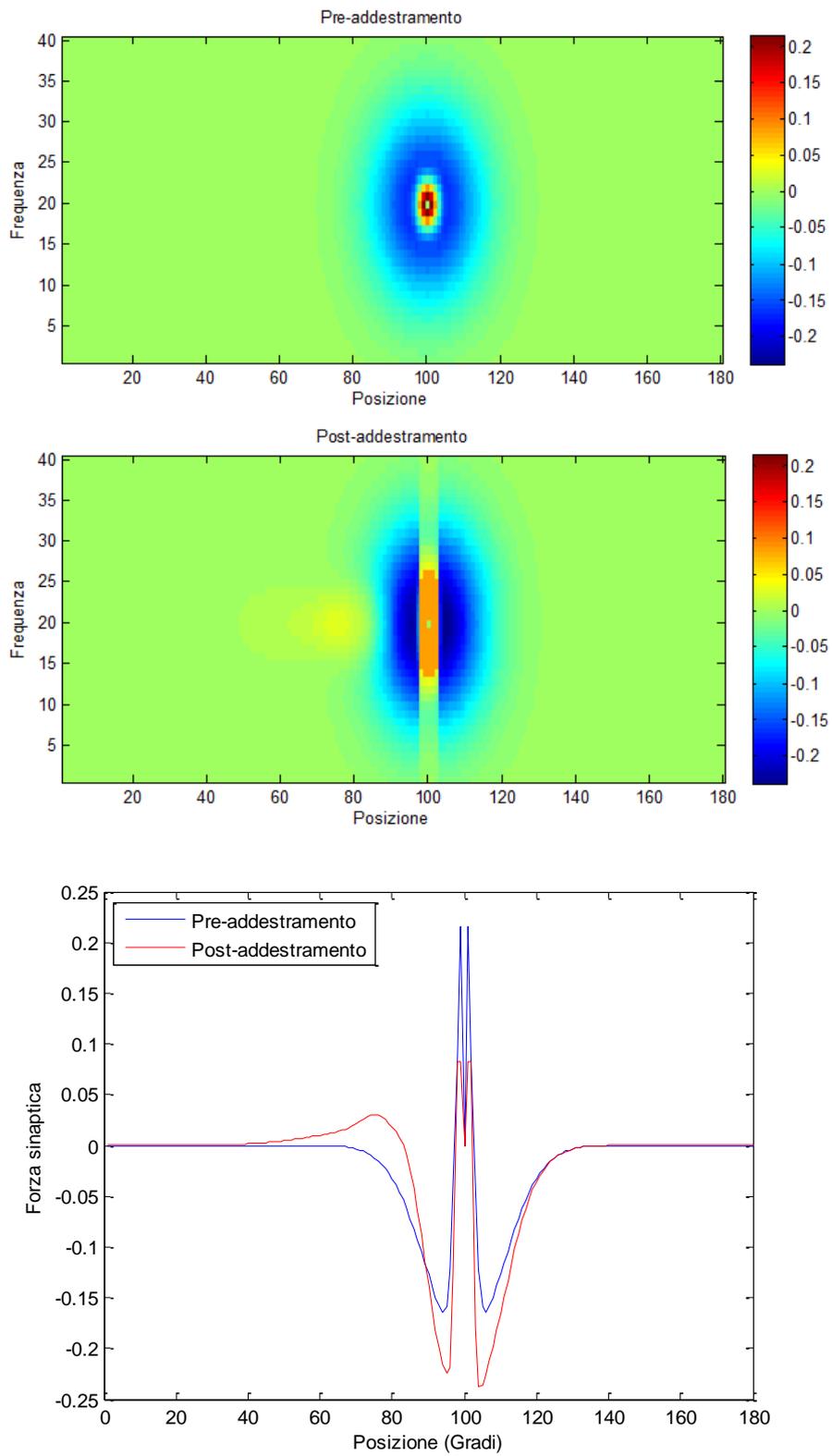


Figura 4.16. Modifiche delle sinapsi laterali acustiche.

Si può osservare che rispetto alla condizione di pre-addestramento, dopo il training la forma del cappello messicano non è più simmetrica: il neurone considerato riceve infatti più eccitazione dalle cellule alla sua sinistra e una maggiore inibizione dai neuroni alla sua destra. L'asimmetria delle sinapsi porterà, come descritto più avanti, a un'alterazione nella percezione acustica.

Anche le sinapsi laterali visive vanno incontro a delle modifiche, come mostrato in Figura 4.17, che illustra le connessioni entranti nel neurone visivo di posto $j=100$. Si può notare che dopo l'adattamento, sia le connessioni eccitatorie provenienti dai neuroni vicini sia quelle inibitorie che partono da cellule più distanti si sono potenziate, senza però alterare la simmetria del cappello messicano. Questo tipo di plasticità farà in modo che non ci saranno shift duraturi nella percezione dello spazio visivo.

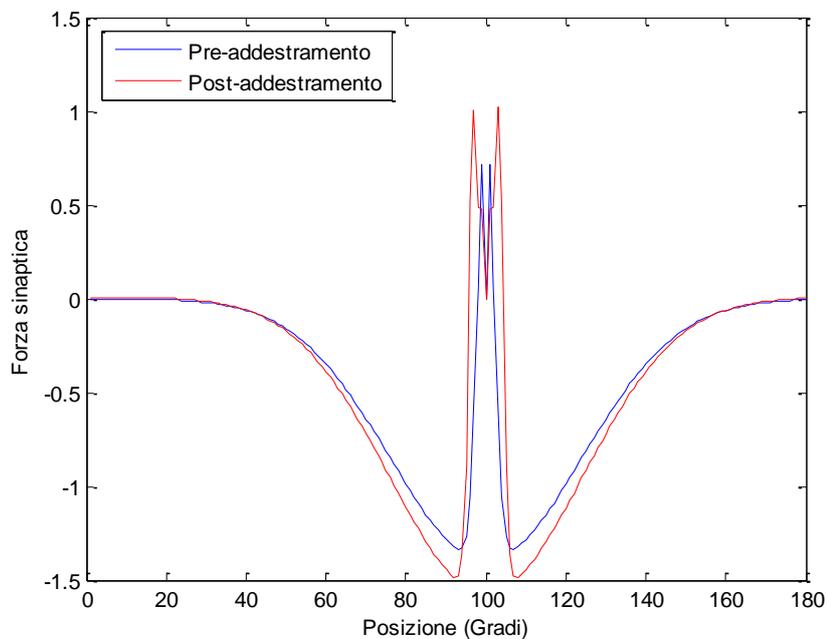


Figura 4.17. Modifiche delle sinapsi laterali visive.

4.2.2.2 Valutazione dell'aftereffect

Un modo per verificare che ci sia stata la generazione dell'aftereffect, è analizzare l'attivazione nell'area visiva dopo la fase di adattamento. In particolare, alla rete addestrata è stato presentato un input sonoro centrato alla posizione e alla frequenza

usate nel training ($i=20$ e $j=80$). La Figura 4.18 riporta l'attivazione nell'area acustica bidimensionale e il profilo di attività rispetto alla posizione dei neuroni uditivi, in seguito a tale stimolazione.

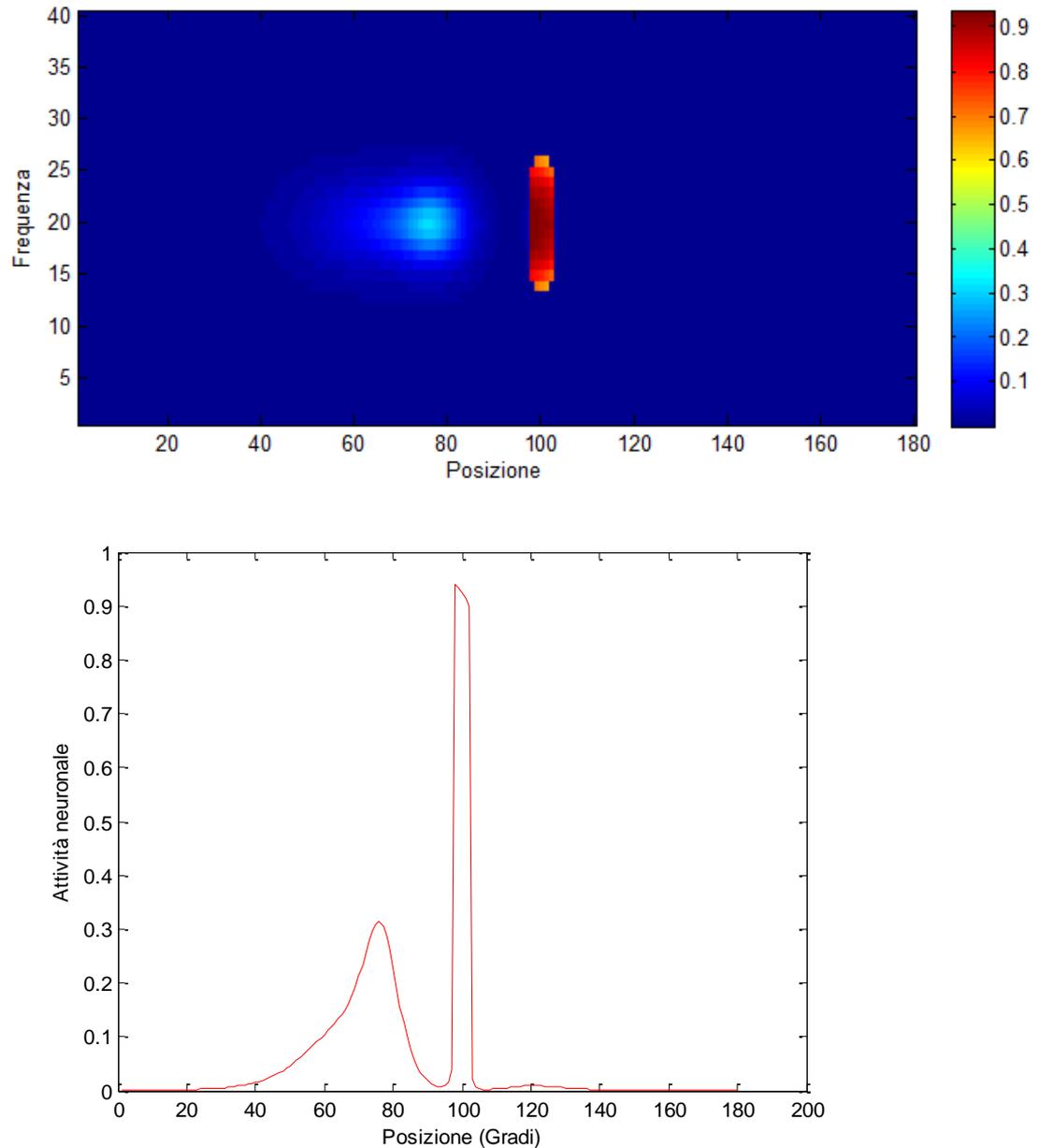


Figura 4.18. Attivazione nell'area uditiva (sopra) e profilo di attività dei neuroni acustici (sotto) dopo l'addestramento.

Si osserva che l'attività dei neuroni acustici continua a rimanere molto alta nella posizione in cui è stato applicato l'input visivo durante il training. Questo dovrebbe

già indicare che la percezione dello spazio acustico è stata alterata a causa dell'esposizione al conflitto cross-modale.

Comunque la grandezza che meglio rappresenta la posizione percepita dello stimolo è il baricentro, che, come affermato nel Capitolo 3, rappresenta il valore medio delle posizioni codificate da ciascun neurone, pesate dalle rispettive attività neuronali. Il valore del baricentro ottenuto dopo la presentazione di uno stimolo sonoro di posizione 80 e frequenza 20 è risultato pari a 85.26° . Questo significa che c'è stato uno spostamento a lungo termine nella percezione dello spazio acustico.

Una domanda che potrebbe sorgere è se anche la percezione dello spazio visivo subisce dei cambiamenti dopo l'adattamento alla situazione di conflitto cross-modale. Per rispondere a questa domanda, il modello addestrato è stato stimolato con un input unicamente visivo centrato alla posizione di training ($j=100$). In Figura 4.19 è illustrato il conseguente profilo di attività nell'area visiva. E' evidente che l'attivazione dei neuroni di quest'area non subisce variazioni dopo il training, infatti essa resta centrata alla posizione di applicazione dell'input visivo. Inoltre il relativo baricentro risulta pari a 100.01° . Si può quindi affermare che non avvengono ricalibramenti nella percezione dello spazio visivo.

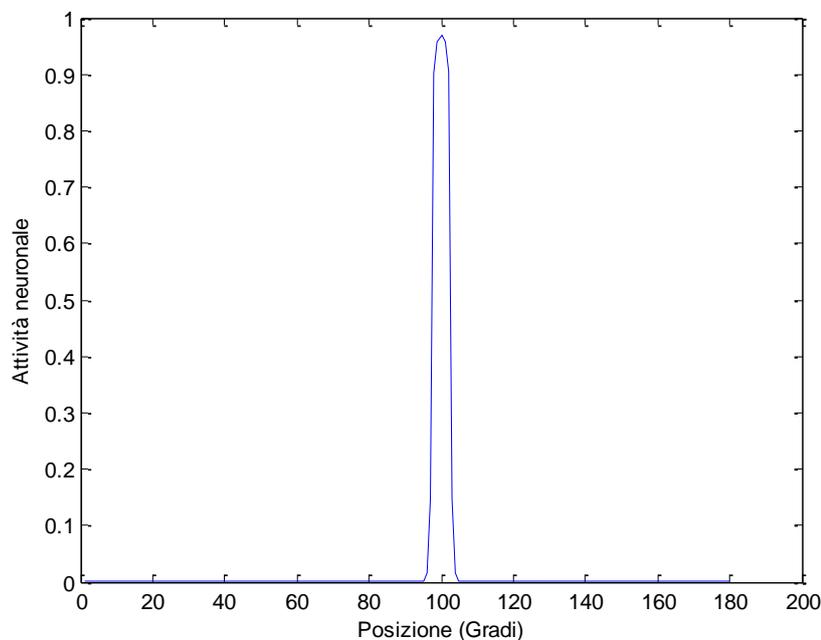


Figura 4.19. Profilo di attività nell'area visiva dopo il training a seguito della presentazione di un input visivo in posizione 100.

4.2.2.3 Verifica del trasferimento dell'aftereffect attraverso le frequenze

I risultati post-addestramento descritti finora sono stati ricavati stimolando la rete addestrata con un input sonoro alla stessa frequenza di training (1100 Hz). La verifica della generalizzazione in frequenza dell'aftereffect è stata effettuata applicando al modello input acustici centrati alla posizione 80, a frequenze diverse da quella di adattamento, in modo tale da coprire l'intero range di frequenza codificato dallo strato uditivo della rete (200 - 6000 Hz). In Figura 4.20 è riportato l'aftereffect per ogni frequenza di stimolazione.

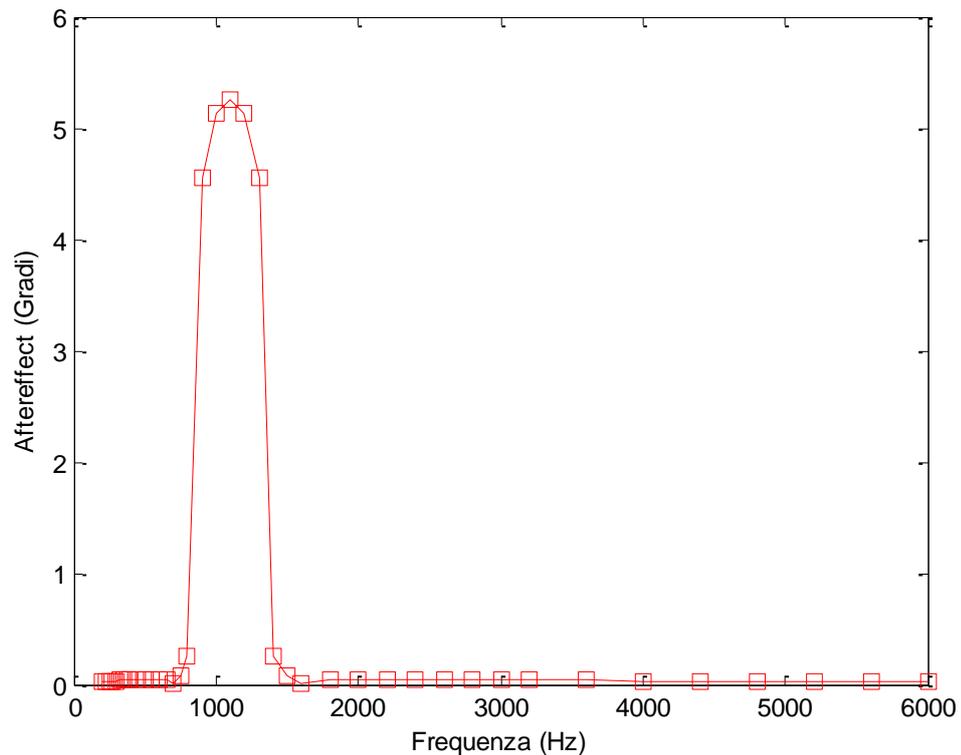


Figura 4.20. Mancato trasferimento dell'aftereffect attraverso le frequenze.

Si può notare che lo shift percettivo coinvolge solo i neuroni acustici che codificano per frequenze vicine a quella di training, interessando un intervallo di 400 Hz. Il mancato trasferimento dell'aftereffect attraverso le frequenze sonore è in accordo con gli esiti delle ricerche di Recanzone [12,18] e di Lewald [19].

Questi risultati possono essere spiegati analizzando le variazioni delle sinapsi laterali acustiche entranti in neuroni che codificano per frequenze distanti da quella di

adattamento. In Figura 4.21 sono mostrate le sinapsi laterali che arrivano in un neurone di posto $i=28$ (2200 Hz) e $j=100$.

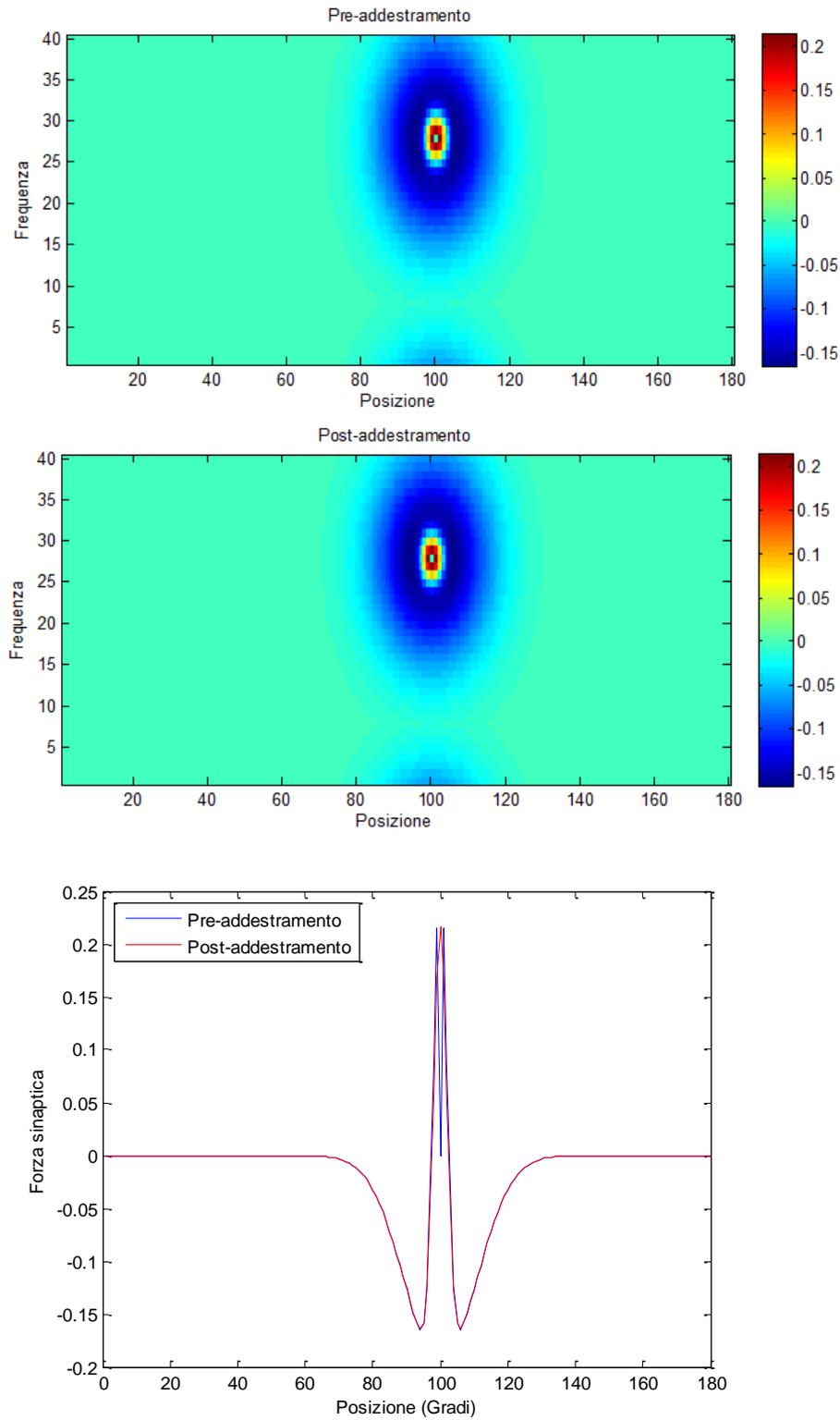


Figura 4.21. Modifiche delle sinapsi entranti in un neurone acustico che codifica per una frequenza diversa da quella di training.

Le immagini superiore e centrale mostrano le sinapsi che entrano nel neurone considerato nell'area acustica rispettivamente pre- e post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi, considerate però soltanto lungo la posizione per una frequenza fissa ($i = 28$). E' possibile osservare che tali sinapsi non subiscono variazioni rilevanti rispetto alla loro condizione di pre-addestramento; per questo motivo l'aftereffect rimane limitato a frequenze vicine a quella di training.

4.2.3 Post-addestramento – Secondo paradigma di training

I risultati di seguito riportati sono stati ottenuti dopo che la rete è stata addestrata mediante stimoli acustici e visivi simultanei e spazialmente separati, in posizioni dello spazio variabili. I due tipi di input sono infatti stati presentati in 9 diverse posizioni (20° , 40° , 60° , 80° , 100° , 120° , 140° , 160° e 180°), con una distanza spaziale l'uno dall'altro costante pari a $+20^\circ$ (ovvero lo stimolo visivo è stato sempre presentato a destra di quello acustico ad una distanza di 20°). La frequenza dell'input acustico è stata mantenuta al valore fisso di 1100 Hz, che è quello codificato dai neuroni di posto 20 rispetto alla frequenza nello strato uditivo. La procedura di allenamento della rete si è composta di dieci prove. In ciascuna delle prove gli stimoli sono stati dati in ognuna delle 9 posizioni. L'ordine con cui queste sono state fornite è stato scelto in modo casuale. La durata dell'addestramento rispetto a ogni posizione è stata di 100 ms. Durante ciascuna prova, le sinapsi laterali sia acustiche sia visive si modificano secondo le regole di apprendimento descritte nel Capitolo 3. Dopo tale procedura, per verificare la presenza dell'aftereffect, il modello è stato stimolato solamente con un input acustico, la cui posizione è stata fatta variare tra quelle utilizzate durante il periodo di adattamento. Per verificare se l'effetto off-line del ventriloquismo si generalizza per frequenze diverse da quella adoperata nella fase di training, è stata di volta in volta variata anche la frequenza dell'input, in modo tale da coprire tutto il range codificato dallo strato acustico della rete (200 - 6000 Hz).

4.2.3.1 Modifiche delle sinapsi

Un primo aspetto da considerare riguarda le modifiche che subiscono le sinapsi laterali dei due strati a causa dell'addestramento. I cambiamenti sinaptici relativi all'area acustica sono illustrati in Figura 4.22.

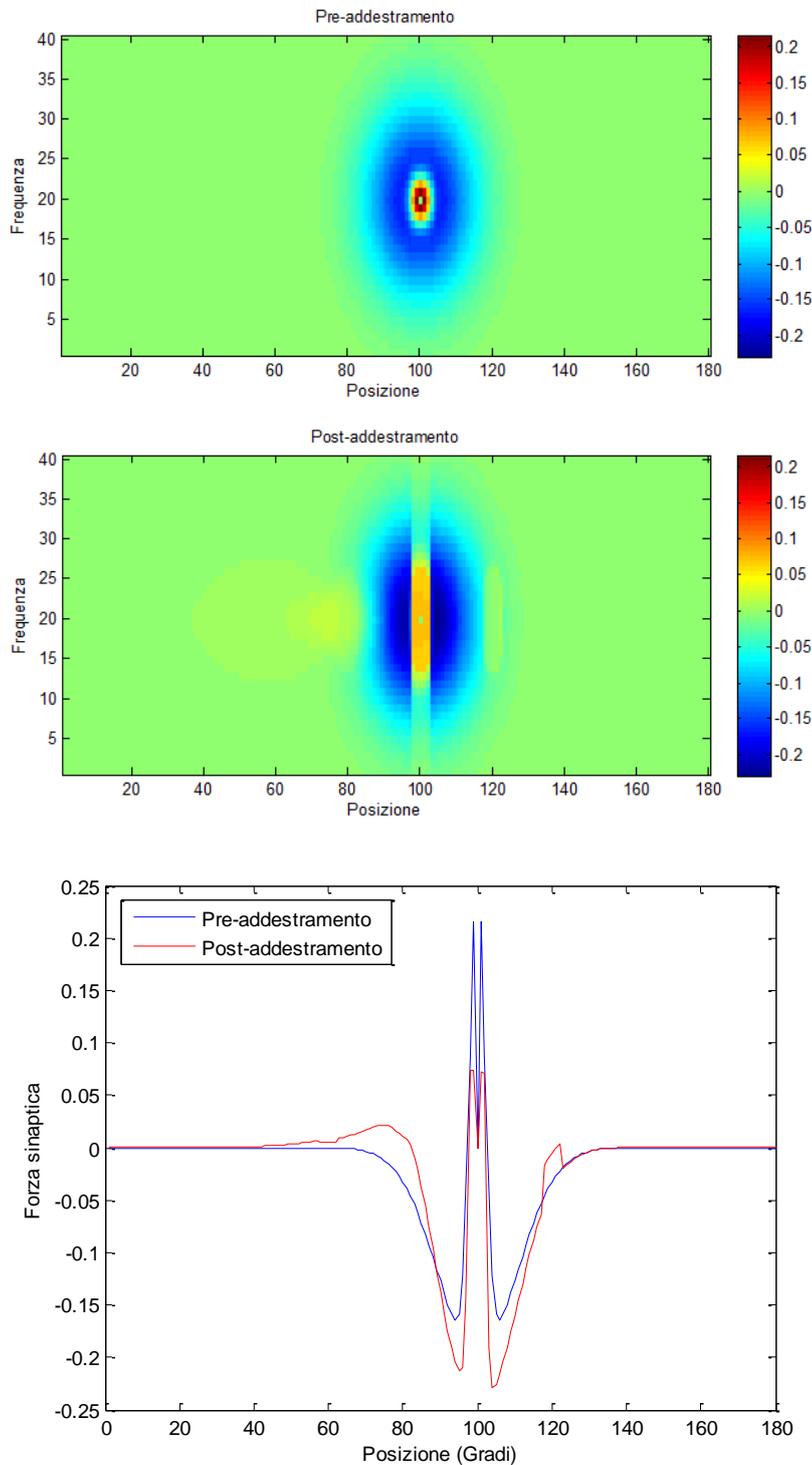


Figura 4.22. Modifiche delle sinapsi laterali acustiche.

Le immagini superiore e centrale mostrano le sinapsi che entrano nel neurone di posto $i=20$ e $j=100$ nell'area acustica rispettivamente pre- e post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi, considerate però soltanto lungo la posizione per una frequenza fissa ($i=20$). Si può osservare che rispetto alla condizione di pre-addestramento, dopo il training il cappello messicano non ha più una forma simmetrica: il neurone considerato riceve infatti più eccitazione dalle cellule alla sua sinistra e una maggiore inibizione dai neuroni alla sua destra. L'asimmetria delle sinapsi porterà, come descritto più avanti, a un'alterazione nella percezione acustica.

Anche le sinapsi laterali visive vanno incontro a delle modifiche, come mostrato in Figura 4.23, che illustra le connessioni entranti nel neurone visivo di posto $j=100$.

In questo caso, dopo la fase di addestramento, le sinapsi eccitatorie provenienti dai neuroni vicini a quello in esame risultano molto potenziate, mentre le sinapsi inibitorie che arrivano dalle cellule più lontane sono leggermente rinforzate. Tuttavia la forma del cappello messicano continua a rimanere simmetrica. Il mantenimento della simmetria consentirà di non alterare la percezione dello spazio visivo.

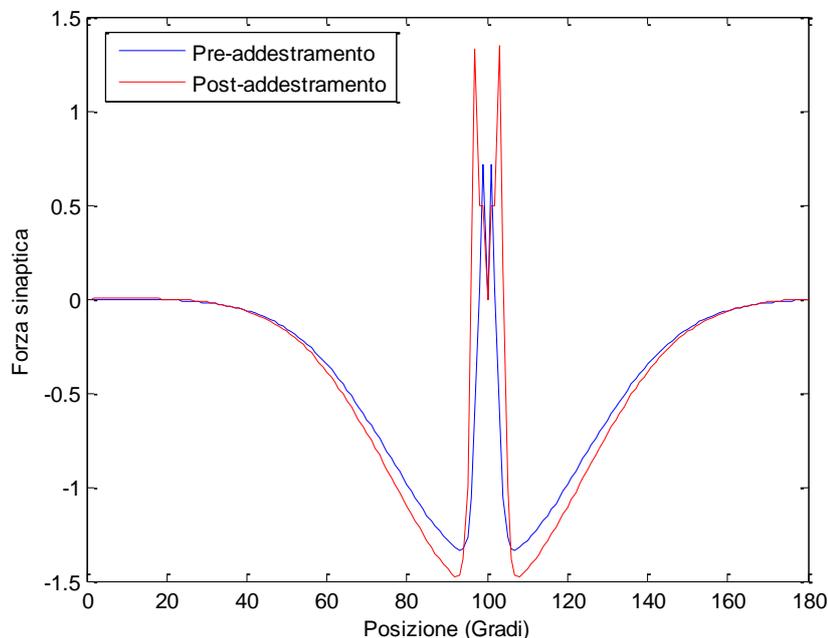


Figura 4.23. Modifiche delle sinapsi laterali visive.

4.2.3.2 Valutazione dell'aftereffect

Una caratteristica della rete rappresentativa della generazione dell'aftereffect del ventriloquismo è data dall'attivazione dei neuroni dopo il periodo di training. A tal proposito in Figura 4.24 viene riportata la bolla di attivazione nell'area uditiva (sopra) e il profilo di attività dei neuroni acustici (sotto), dopo aver presentato alla rete un input acustico esterno centrato nel punto $i=20$ e $j=80$, applicato quindi in una delle posizioni usate nel training e alla stessa frequenza dell'adattamento.

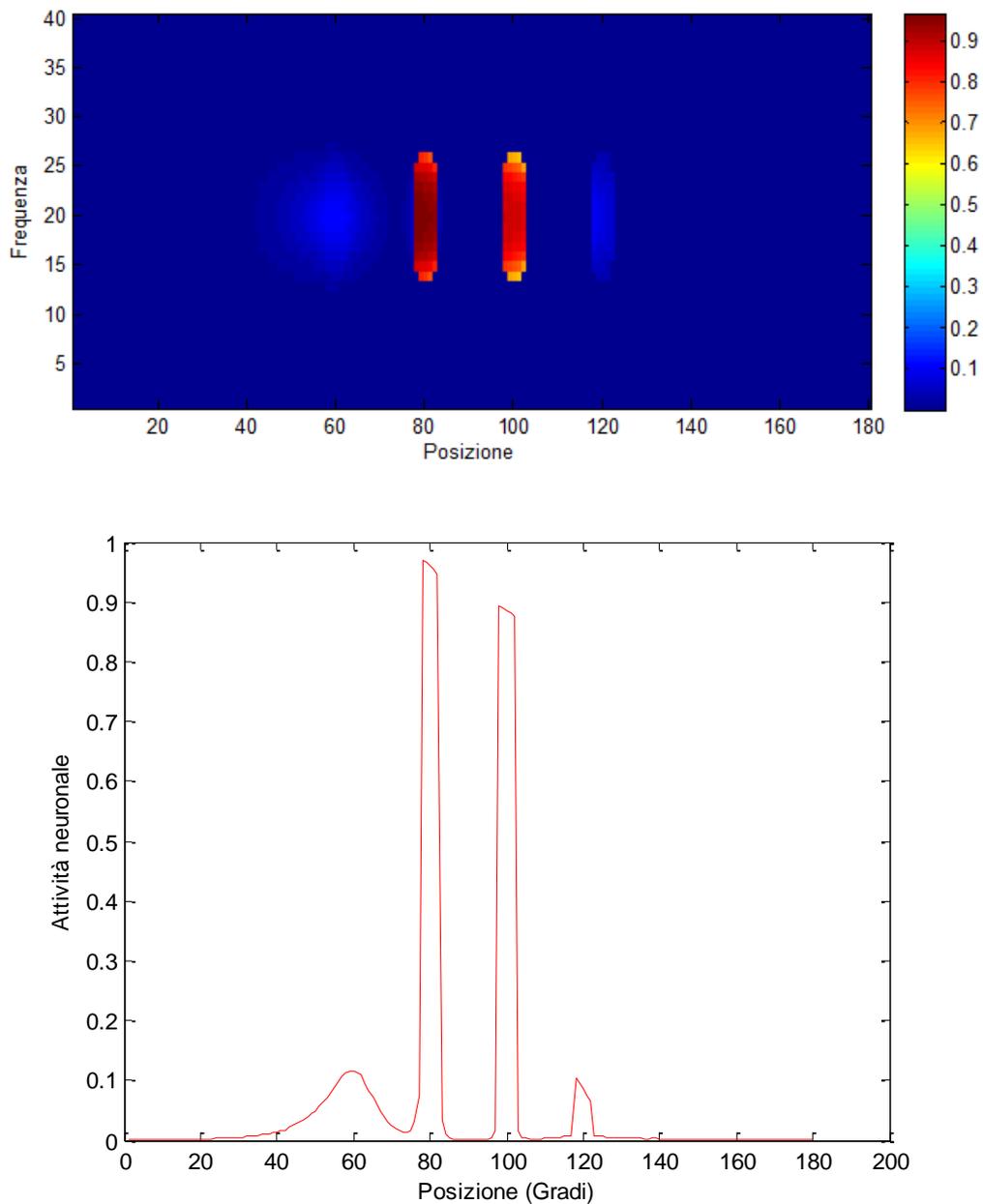


Figura 4.24. Attivazione nell'area uditiva (sopra) e profilo di attività dei neuroni acustici (sotto) dopo l'addestramento.

Si notano subito due picchi nell'attivazione dei neuroni acustici; essi sono probabilmente dovuti al fatto che nella fase di adattamento il modello è stato addestrato con degli input in diverse posizioni dello spazio e sono state così potenziate le attività di neuroni in molte posizioni. Per capire meglio quanto appena detto è possibile fare un semplice esempio. Si supponga che le prime posizioni addestrate siano 80 per l'input acustico e 100 per quello visivo; a seguito di questa esposizione al conflitto cross-modale, il profilo di attività dei neuroni acustici si sposta verso la posizione in cui è stato applicato lo stimolo visivo, come indicato dalla Figura 4.25.

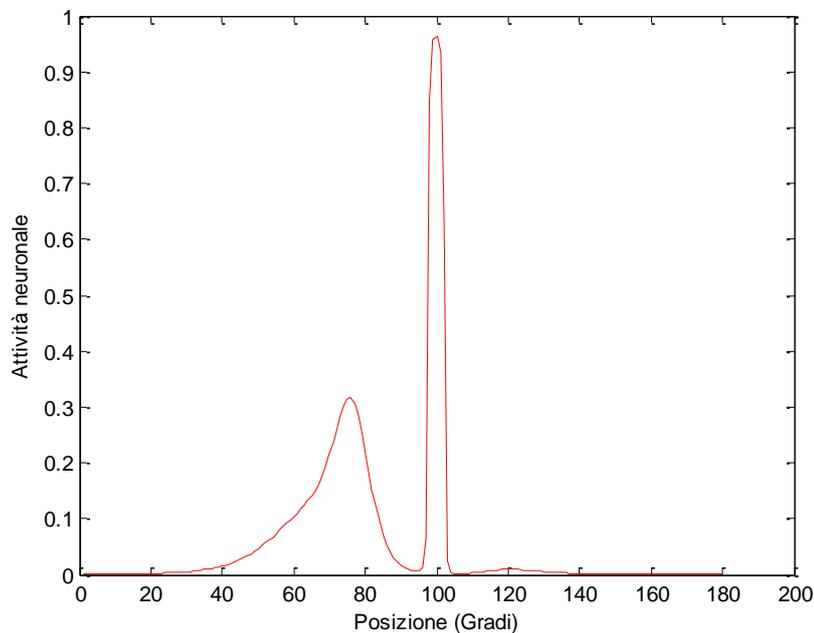


Figura 4.25. Profilo di attività dei neuroni acustici durante l'addestramento con un input uditivo in posizione 80 e uno visivo in posizione 100.

Durante il periodo di adattamento si arriverà in una fase in cui vengono addestrate le posizioni 60 per l'input acustico e 80 per quello visivo; tutto ciò produce un profilo di attivazione come quello riportato in Figura 4.26. Anche in questo caso l'attività delle cellule uditive è attratta dalla posizione dello stimolo visivo.

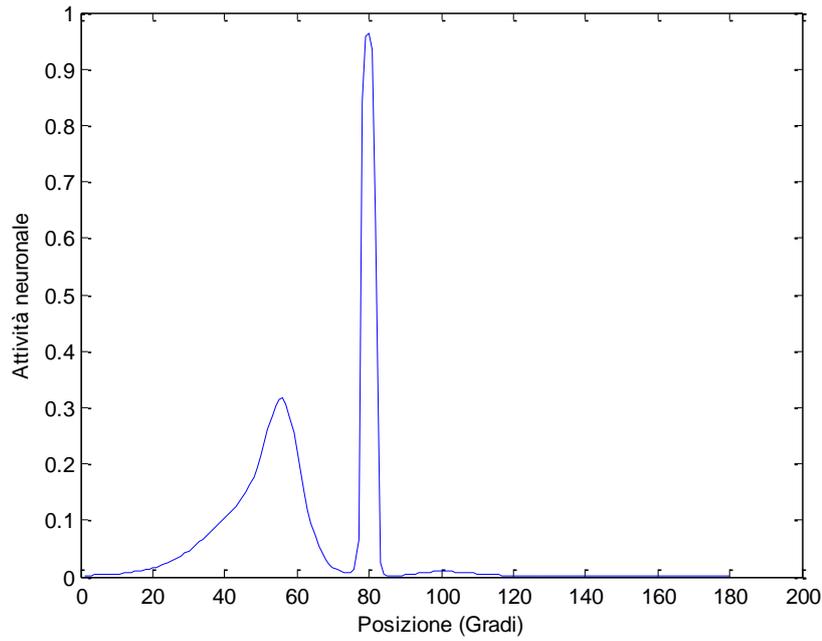


Figura 4.26. Profilo di attività dei neuroni acustici durante l’addestramento con un input uditivo in posizione 60 e uno visivo in posizione 80.

Si consideri a questo punto la Figura 4.27, nella quale vengono riportati contemporaneamente i due precedenti profili di attività.

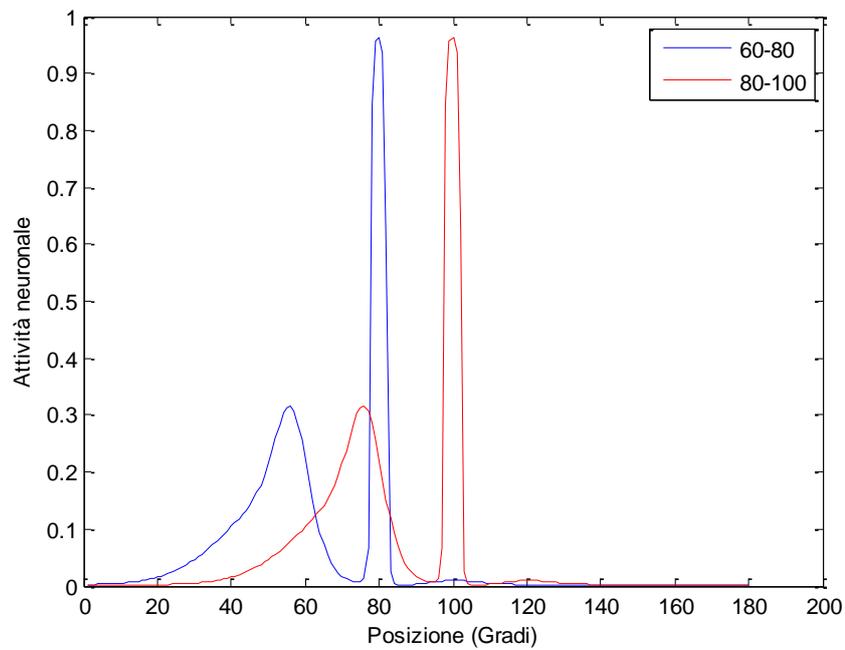


Figura 4.27. Sovrapposizione dei profili di attività dei neuroni acustici.

Si osserva che l'attivazione nell'intorno della posizione 80 generata dalla prima esposizione si somma a quella provocata dalla successiva esposizione al conflitto audio-visivo, perciò i neuroni che avevano un'attività bassa dopo la fase iniziale del training vengono fortemente eccitati nelle fasi successive, portando alla generazione di due picchi nel profilo di attività. L'esempio appena descritto può essere riproposto per tutte le altre posizioni addestrate.

La grandezza che comunque deve essere maggiormente considerata per valutare l'aftereffect è il baricentro. In particolare, per quantificare lo shift nella percezione, di volta in volta alla rete addestrata è stato presentato un input acustico esterno in una delle 9 posizioni usate durante l'addestramento ed è stato registrato il corrispondente valore del baricentro. In queste prove la frequenza dello stimolo sonoro è sempre stata mantenuta uguale a quella utilizzata durante la fase di addestramento, ovvero 1100 Hz. Il valore medio baricentrico delle 9 posizioni addestrate è risultato pari a 85.17° . C'è quindi uno spostamento sistematico nella posizione percepita. L'aftereffect è stato valutato anche per posizioni dello stimolo acustico non utilizzate in fase di training. Infatti in un'altra simulazione, lo stimolo è stato fatto variare in 19 posizioni, delle quali 10 non sono mai state usate durante l'addestramento. I risultati sono riportati in Figura 4.28, che mostra la posizione percepita dell'input sonoro in termini di baricentro, in funzione della posizione in cui è stato applicato lo stimolo durante il test di verifica dell'aftereffect, insieme alla retta di regressione per i dati ottenuti. In corrispondenza delle posizioni che non sono state adoperate nel training, l'aftereffect si presenta con minore entità (circa 1.66°). Tuttavia l'aftereffect complessivo risulta molto evidente, infatti la retta di regressione ha una pendenza di quasi 1 e un'intercetta sull'asse delle ordinate di 3.9° , che rappresenta circa il 20% della distanza spaziale audio-visiva usata nell'addestramento. Questa percentuale è in accordo con il range riportato in letteratura (tra il 20% e il 60% [19,20]).

Anche per questo secondo paradigma di training, è stato verificato che non ci fossero shift nella percezione visiva. Infatti sono state compiute ulteriori simulazioni in cui la rete addestrata è stata stimolata solo con input visivi in diverse posizioni. La Figura 4.29 riporta il profilo di attività nella area visiva dopo che è stato presentato un segnale visivo centrato alla posizione 100. Si nota subito che l'attivazione dei neuroni visivi non è stata per nulla influenzata dall'esposizione al conflitto, come

dimostra anche lo shift percettivo che ha un valore di 0.01° . Si può perciò concludere che lo strato visivo non subisce modifiche tali da alterarne la percezione.

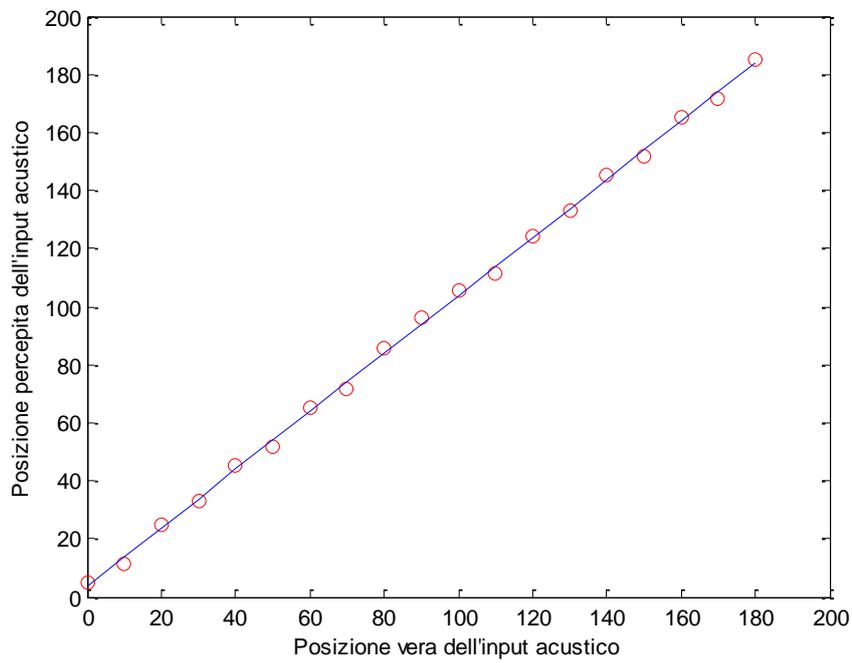


Figura 4.28. Verifica dell'aftereffect per i neuroni acustici.

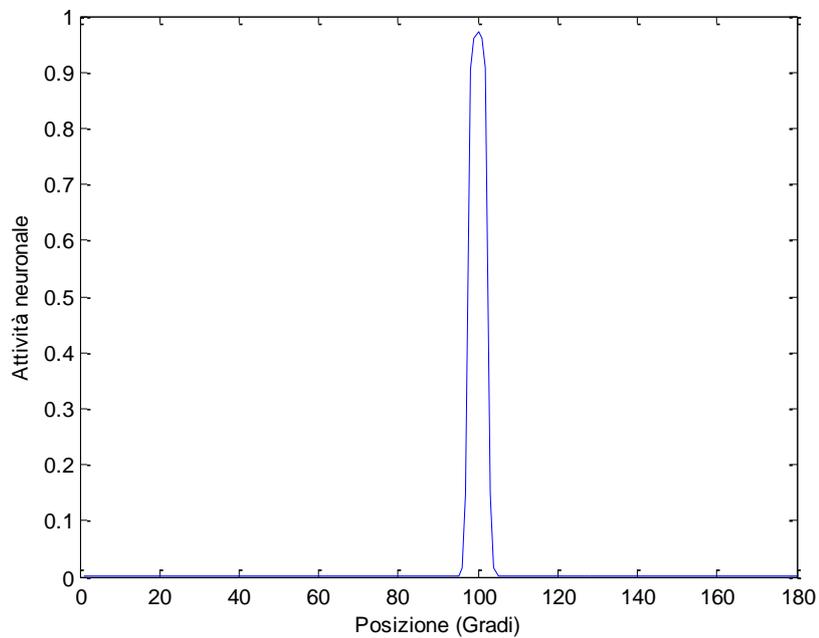


Figura 4.29. Profilo di attività nell'area visiva dopo il training a seguito della presentazione di un input visivo in posizione 100.

4.2.3.3 Verifica del trasferimento dell'aftereffect attraverso le frequenze

Le prove descritte finora per il post-addestramento, sono state ottenute applicando al modello addestrato input esterni in diverse posizioni dello spazio e mantenendo gli stimoli acustici alla frequenza utilizzata in fase di adattamento, cioè 1100 Hz (valore codificato dai neuroni di coordinata $i = 20$ nello strato acustico.) Perciò, per capire se l'aftereffect del ventriloquismo rimane confinato alla frequenza di training o si generalizza per un range di frequenze più ampio, sono state compiute simulazioni nelle quali oltre che la posizione, veniva fatta variare anche la frequenza dell'input sonoro applicato. In particolare, per valutare l'eventuale trasferimento in frequenza dell'aftereffect, alla rete addestrata sono stati presentati gli input in ciascuna delle 9 posizioni adottate in fase di training e in ognuna delle 40 frequenze codificate dai neuroni dello strato uditivo. I risultati di queste prove sono mostrati in Figura 4.30, in cui l'aftereffect è stato calcolato come media, per ogni frequenza, degli aftereffect delle posizioni dell'input utilizzate durante l'addestramento.

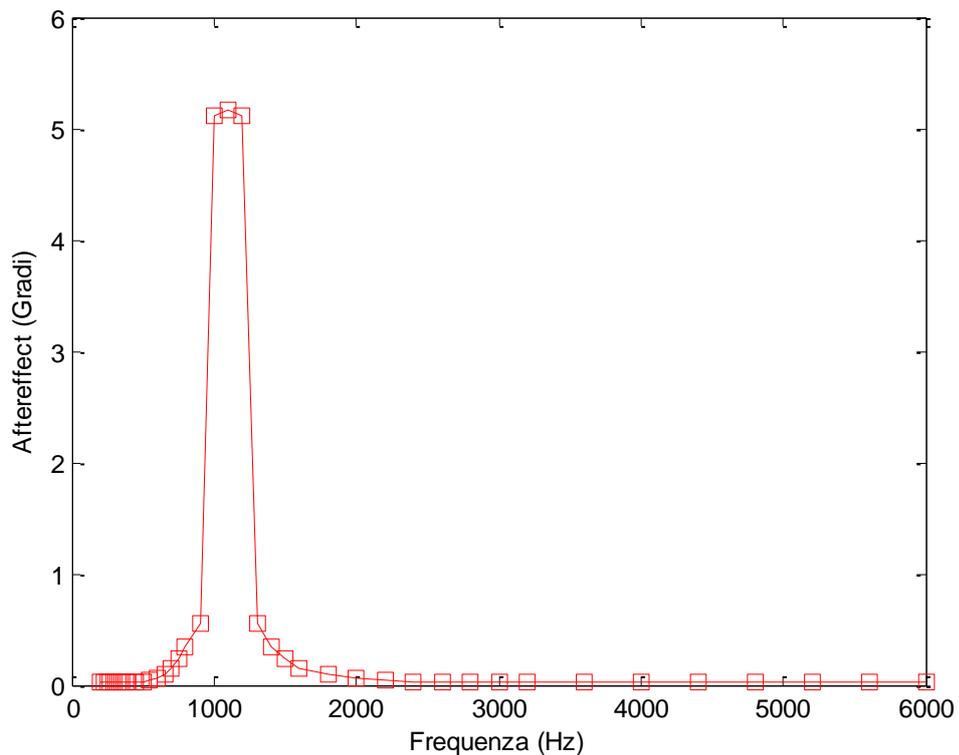


Figura 4.30. Mancato trasferimento dell'aftereffect attraverso le frequenze.

Dagli esiti di questa prova, si evince che l'effetto off-line del ventriloquismo non si generalizza per frequenze lontane da quella di addestramento, ma rimane limitato in un range di circa 200 Hz centrato al valore di adattamento. Questi risultati sono in accordo con quelli riportati da Recanzone [12,18] e da Lewald [19].

Il mancato trasferimento dell'aftereffect attraverso le frequenze sonore ricalca il tipo di plasticità cui vanno incontro le sinapsi laterali acustiche entranti nei neuroni che codificano per frequenze lontane da quella di training. La Figura 4.31 mostra le sinapsi laterali che arrivano in un neurone di posto $i=28$ (2200 Hz) e $j=100$. Le immagini superiore e centrale mostrano le sinapsi che entrano nel neurone considerato nell'area acustica rispettivamente pre- e post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi, considerate però soltanto lungo la posizione per una frequenza fissa ($i=28$). E' possibile osservare che tali connessioni non subiscono modifiche significative rispetto alla condizione di base e questo si traduce nell'assenza di aftereffect per frequenze diverse da quella di adattamento.

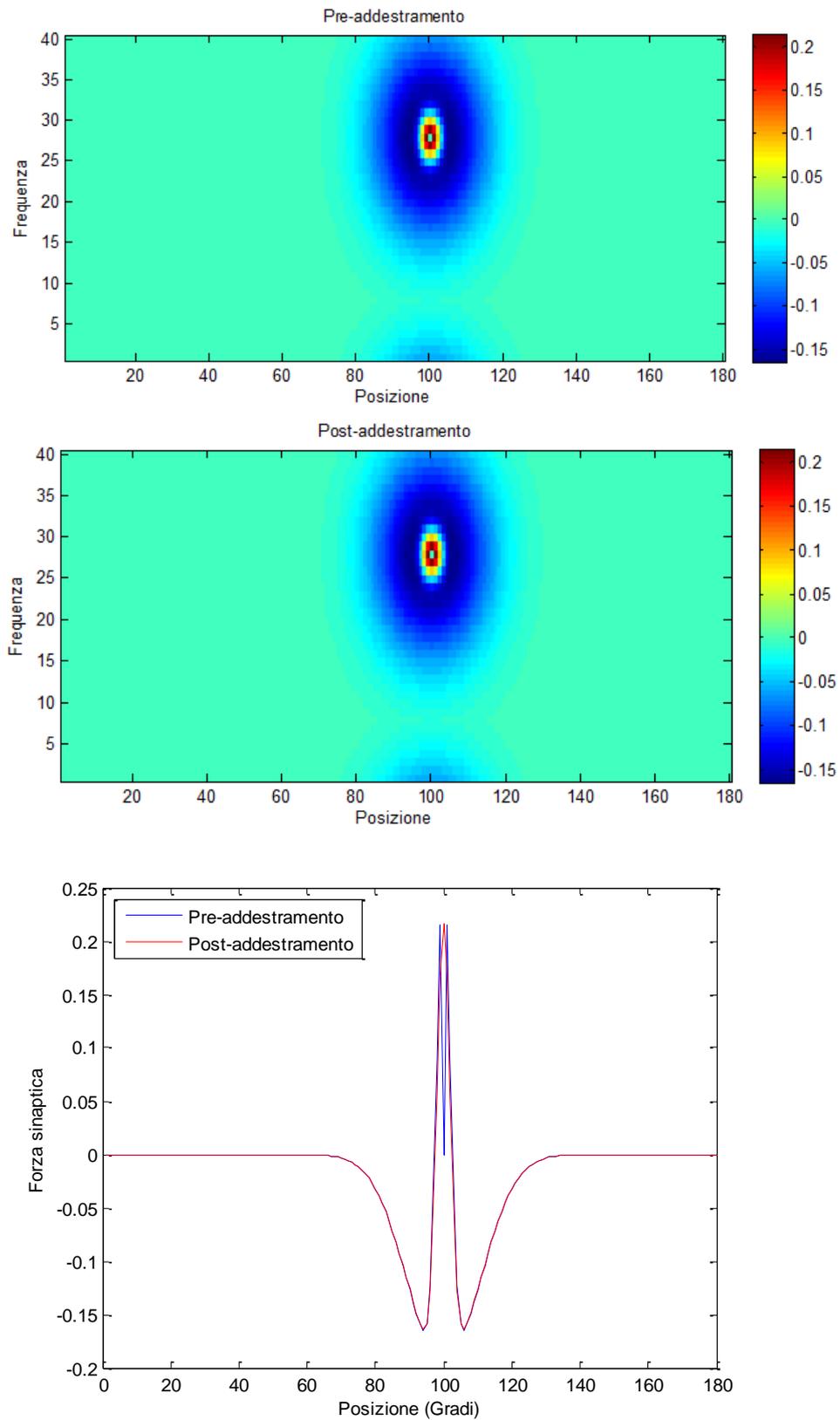


Figura 4.31. Modifiche delle sinapsi entranti in un neurone acustico che codifica per una frequenza diversa da quella di training.

4.3 Risultati della configurazione per la generalizzazione in frequenza dell'aftereffect

4.3.1 Pre-addestramento

4.3.1.1 Stimolazione unimodale visiva

In questa prova la rete è stata stimolata con un input esterno unicamente visivo centrato nella posizione 100, della stessa forma di quello descritto precedentemente per l'altro set di parametri. In seguito a questa stimolazione, il modello ha fornito un'attivazione nell'area acustica bidimensionale come quella riportata in Figura 4.32.

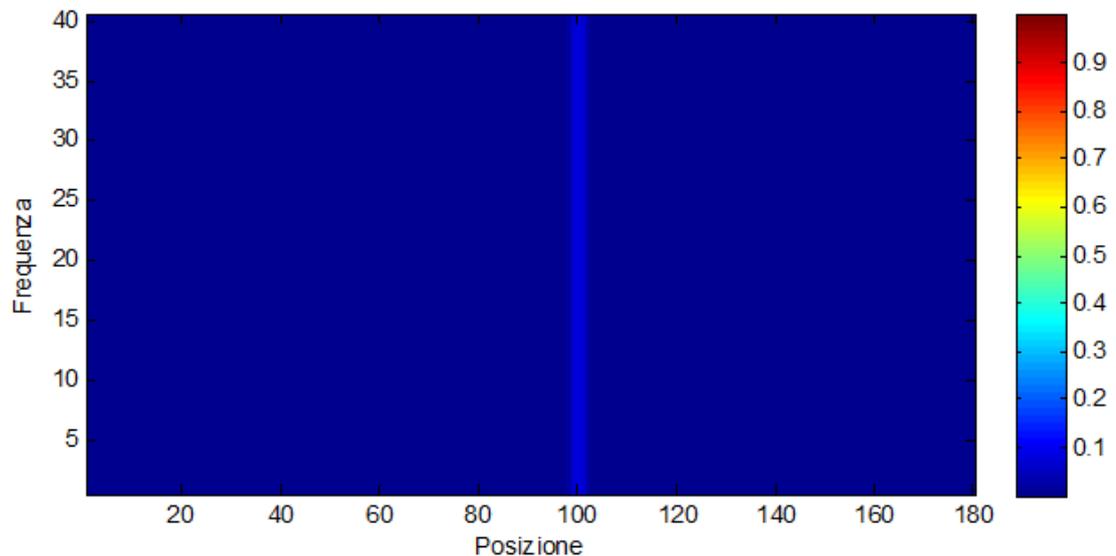


Figura 4.32. Attivazione nell'area acustica in seguito a stimolazione unimodale visiva.

Vengono inoltre riportati i profili di attività rispetto alla posizione (Figura 4.33). Come avveniva per la precedente configurazione, anche qui i neuroni acustici mostrano un'attivazione molto bassa in corrispondenza del punto in cui è centrato l'input esterno. Quindi nemmeno in questo caso c'è un effetto fantasma nella modalità non stimolata.

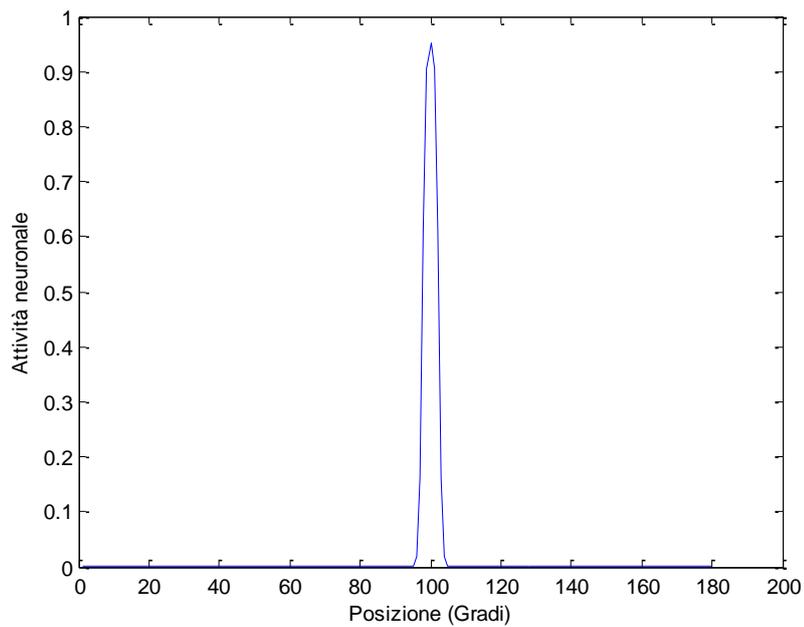
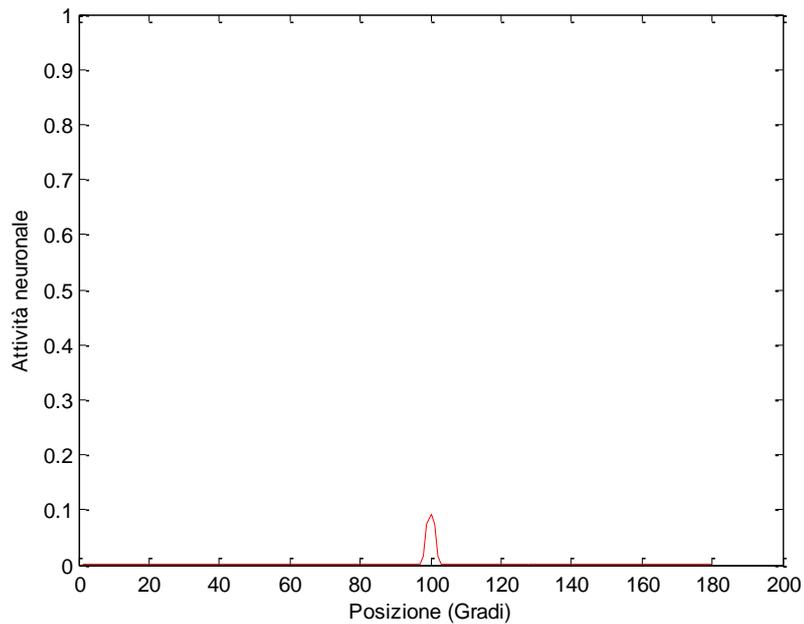


Figura 4.33. Profili di attività nelle aree acustica (sopra) e visiva (sotto) rispetto alla posizione in seguito a stimolazione unimodale visiva.

4.3.1.2 Stimolazione unimodale acustica

Al modello è stato applicato un input esterno acustico centrato nella posizione 80 e alla frequenza 20 (1100 Hz) e generato in MATLAB® secondo l'Eq. 3.6.

La forma dell'input acustico è mostrata in Figura 4.34.

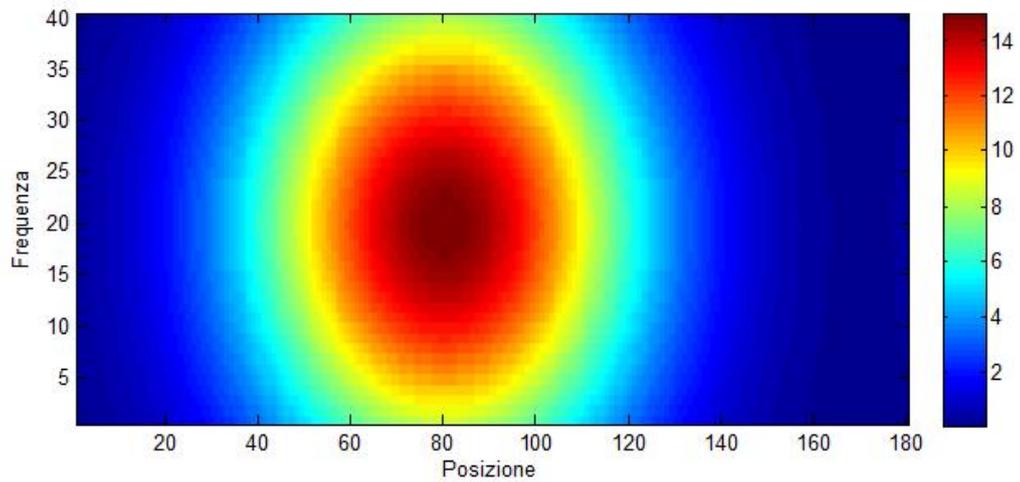


Figura 4.34. Input acustico.

La conseguente attivazione nell'area bidimensionale uditiva è mostrata in Figura 4.35.

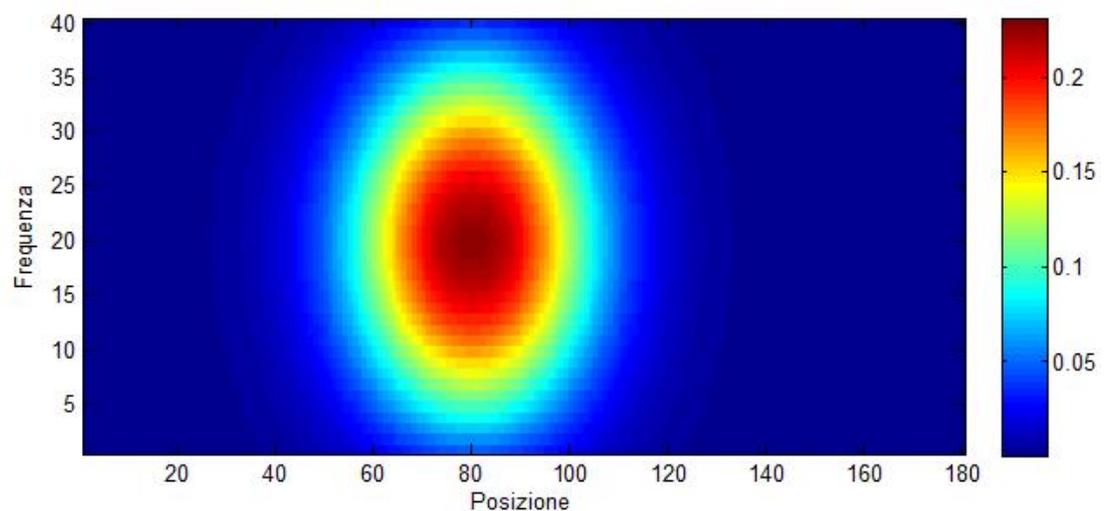


Figura 4.35. Attivazione nell'area uditiva in seguito a stimolazione unimodale acustica.

Per meglio comprendere la morfologia di tale attivazione, in Figura 4.36 viene riportato il profilo di attività acustica rispetto alla posizione, in corrispondenza della frequenza di 1100 Hz. Si nota che il profilo di attività dell'area acustica è più largo rispetto a quello dell'area visiva di Figura 4.33, un risultato ragionevole dato che la risoluzione spaziale uditiva è inferiore di quella visiva.

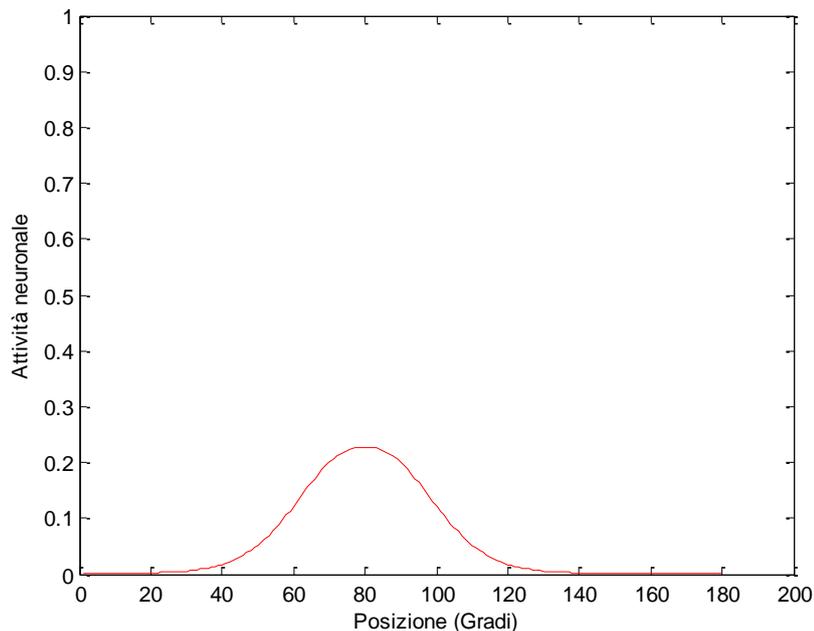


Figura 4.36. Profilo di attività nell'area acustica rispetto alla posizione in seguito a stimolazione unimodale acustica.

I neuroni uditivi codificano sia per la posizione che per la frequenza; così in Figura 4.37 viene mostrato il profilo di attività rispetto alla frequenza delle cellule uditive. Questa volta l'attivazione lungo la frequenza è molto larga e coinvolge neuroni anche lontani dalla frequenza di applicazione dell'input. Tutto ciò è coerente con la caratteristica della configurazione in esame per cui l'aftereffect del ventriloquismo dovrà trasferirsi attraverso le frequenze sonore.

Infine, in Figura 4.38 viene riportata la risposta dei neuroni visivi. Come nel caso di stimolazione unimodale visiva, anche qui non c'è effetto fantasma.

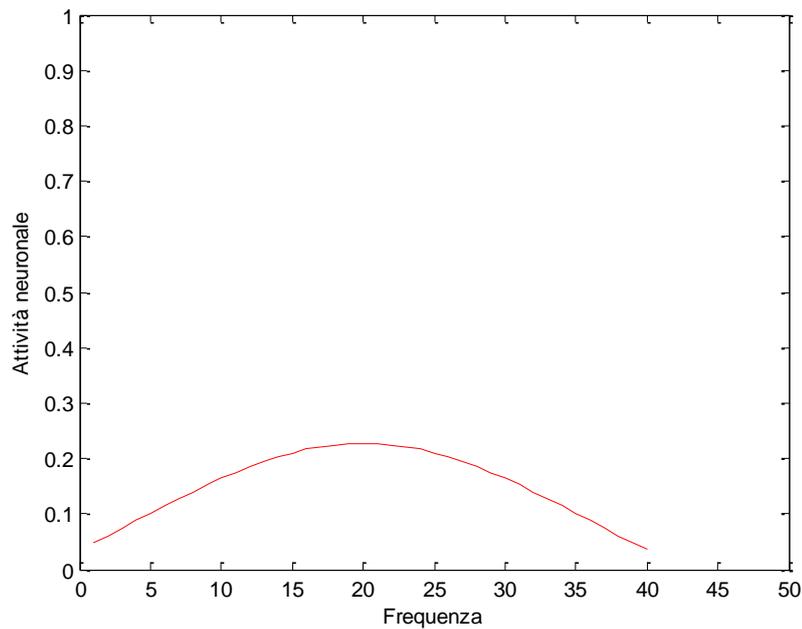


Figura 4.37. Profilo di attività rispetto alla frequenza nell'area uditiva in seguito a stimolazione unimodale acustica.

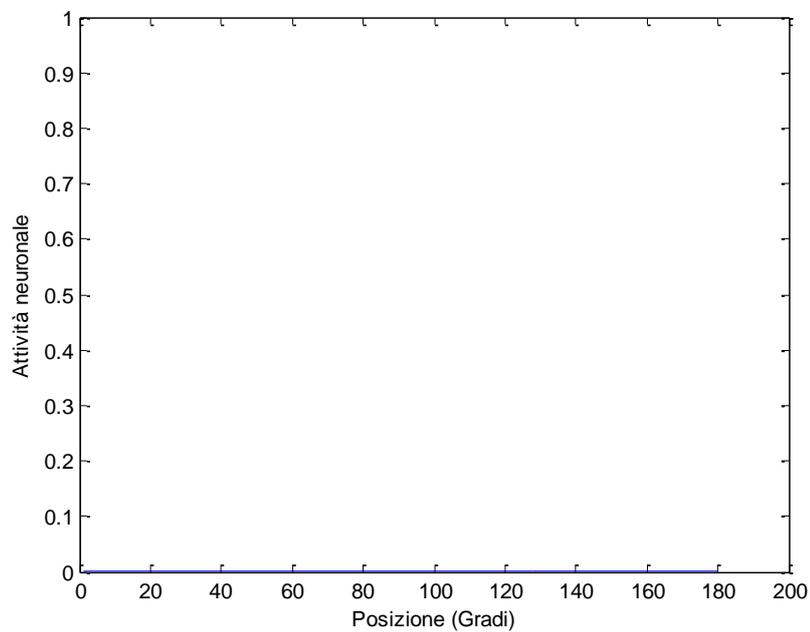


Figura 4.38. Profilo di attività nell'area visiva rispetto alla posizione in seguito a stimolazione unimodale visiva.

4.3.1.3 Stimoli spazialmente coincidenti

In questa simulazione, la rete ha ricevuto in ingresso un input visivo e uno sonoro simultaneamente e alla stessa posizione spaziale. I profili di attività nelle due aree sensoriali sono riportati in Figura 4.39. La Figura 4.40 mostra invece l'attivazione nell'intera area uditiva.

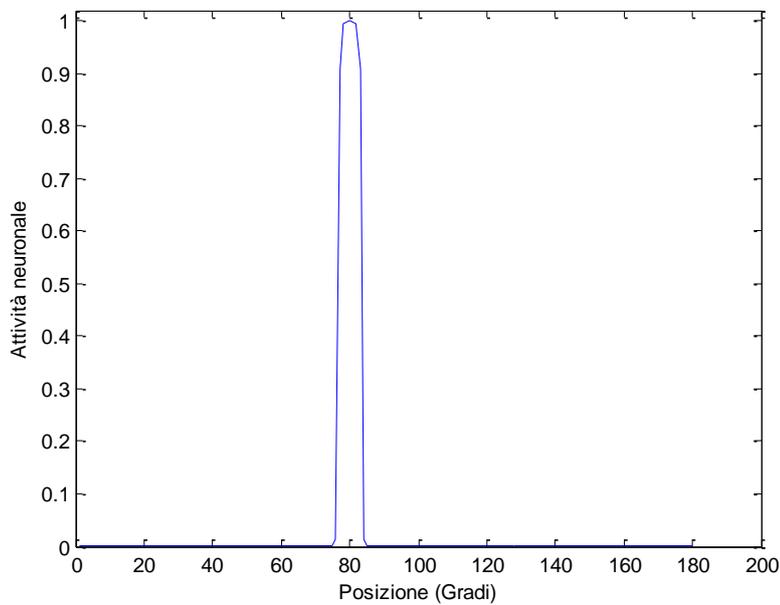
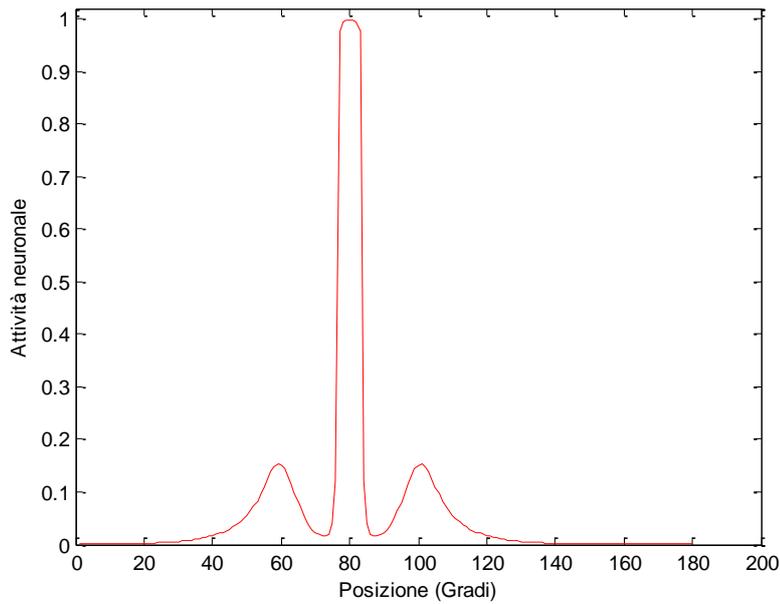


Figura 4.39. Profili di attività rispetto alla posizione nelle aree acustica (sopra) e visiva (sotto) dopo la presentazione di stimoli spazialmente coincidenti. Il profilo di attivazione acustica è mostrato in corrispondenza della frequenza di 1100 Hz.

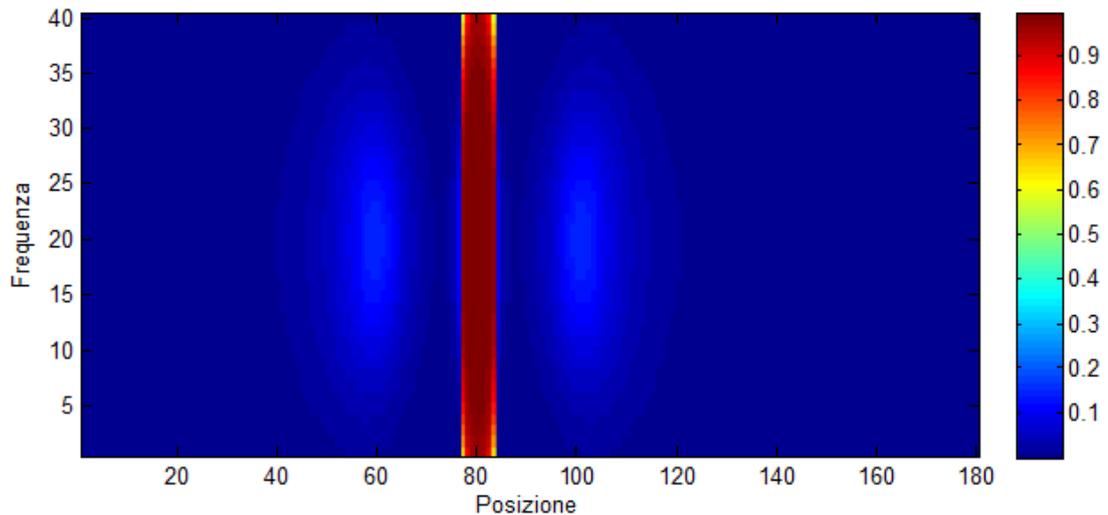


Figura 4.40. Attivazione nell'area uditiva dopo la presentazione di stimoli spazialmente coincidenti.

Si nota subito che le attività neuronali sono molto alte in corrispondenza della posizione in cui sono stati applicati i due stimoli esterni. L'aumento nell'attivazione nasce grazie alle sinapsi inter-area che connettono i due strati della rete; esse danno infatti luogo a un rafforzamento reciproco tra le attività nelle due modalità, coerentemente con le prove condotte in vivo [23], in base alle quali la capacità di localizzazione acustica viene migliorata da un concomitante segnale visivo nella stessa posizione di quello acustico.

4.3.1.4 Stimoli spazialmente separati

In questa prova, alla rete sono stati presentati un input acustico e uno visivo contemporanei e spazialmente separati, centrati rispettivamente in posizione 80 e 100. La frequenza centrale dello stimolo acustico è 1100 Hz ($i=20$). In Figura 4.41 viene illustrata l'attivazione nell'intera area acustica in risposta a tale stimolazione. A differenza della precedente configurazione, i neuroni coinvolti lungo la frequenza

sono praticamente tutti. L'attivazione quasi totale delle cellule acustiche è coerente con il fatto che la configurazione in esame è stata ideata per generalizzare l'aftereffect del ventriloquismo attraverso le frequenze sonore.

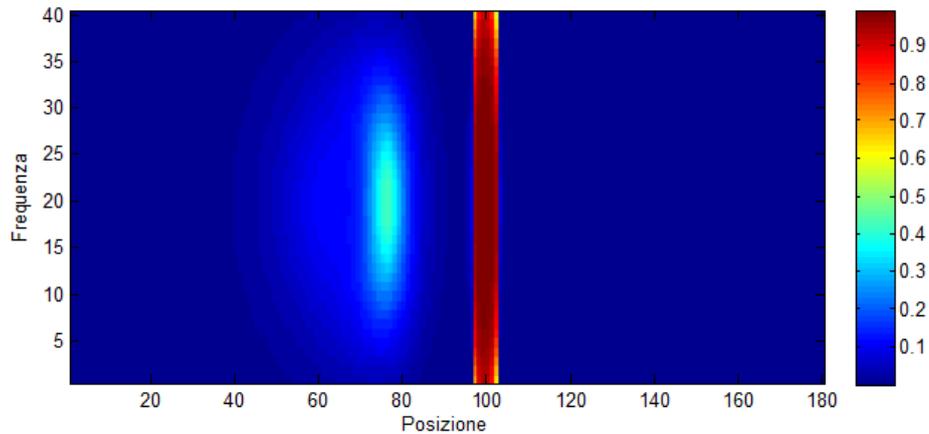


Figura 4.41. Attivazione nell'area acustica dopo la presentazione di due stimoli contemporanei e spazialmente separati.

Come per la configurazione precedente, è evidente che l'attivazione nello strato uditivo sia attratta dalla posizione di applicazione dello stimolo visivo. Per maggiore chiarezza, si riporta il profilo di attività acustica rispetto alla posizione (in corrispondenza della frequenza centrale di stimolazione pari a 1100 Hz), confrontato con il profilo di attività visiva (Figura 4.42).

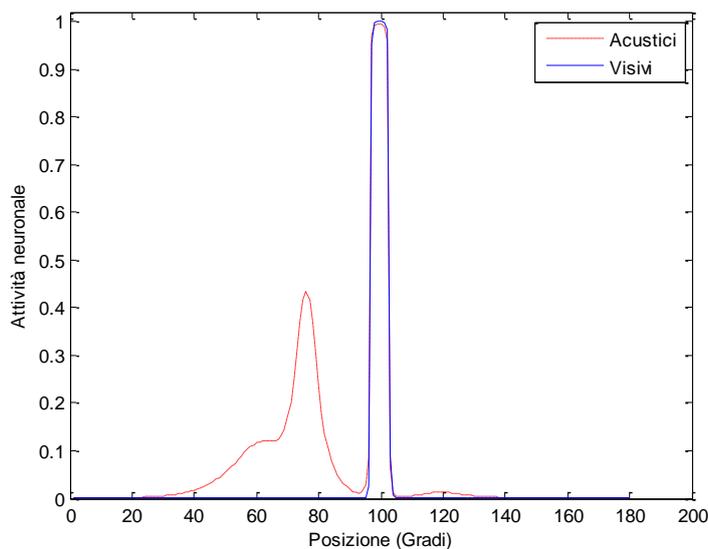


Figura 4.42. Profili di attività rispetto alla posizione nelle aree acustica e visiva dopo la presentazione di stimoli spazialmente separati.

La posizione percepita dell'input acustico in base alla metrica del baricentro risulta uguale a 87.07° , evidenziando perciò uno spostamento nella percezione di 7.07° rispetto al punto di applicazione dell'input sonoro; si tratta di un valore coerente con i risultati riportati in letteratura [15,16,18]. Non ci sono invece alterazioni rilevanti nella percezione visiva, come indicato dal baricentro pari a 99.55° .

Anche per la configurazione in esame sono state fatte delle simulazioni in cui la posizione spaziale dell'input visivo è stata mantenuta fissa ($p^v=100$), mentre quella dell'input acustico è stata fatta variare ($p^a=70, 75, 80, 85, 90, 95, 98$), al fine di valutare l'influenza che ha la distanza spaziale tra i due input rispetto all'effetto on-line del ventriloquismo. In Figura 4.43 è illustrato lo spostamento nella percezione acustica (in rosso) e in quella visiva (in blu), calcolato come posizione percepita meno posizione reale dello stimolo acustico, in funzione della differenza tra la posizione in cui è centrato l'input visivo (p^v) e quella in cui è applicato lo stimolo acustico (p^a). I risultati forniti dal modello si mantengono nel range dei dati ottenuti in esperimenti in vivo [15,16,17].

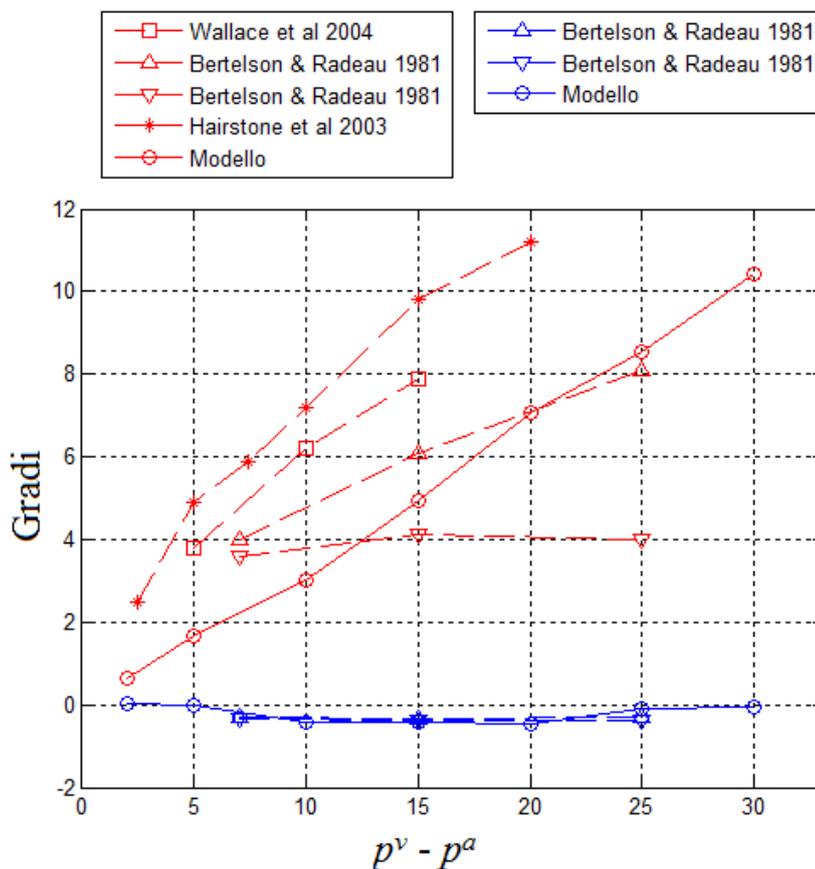


Figura 4.43. Confronto dei risultati del modello con quelli della letteratura.

4.3.2 Post-addestramento – Primo paradigma di training

Il primo paradigma di addestramento e le relative procedure di verifica dell'aftereffect utilizzate per la configurazione in esame sono gli stessi di quelli descritti nel sottoparagrafo 4.2.2.

4.3.2.1 Modifiche delle sinapsi

Un primo aspetto da considerare è rappresentato dal fenomeno di plasticità a cui vanno incontro le sinapsi laterali a causa del periodo di training. La Figura 4.44 mostra le modifiche delle sinapsi laterali relative all'area acustica. Le immagini superiore e centrale mostrano le sinapsi che entrano nel neurone di posto $i=20$ e $j=100$ nell'area acustica rispettivamente pre- e post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi, considerate però soltanto lungo la posizione per una frequenza fissa ($i=20$). È possibile notare che rispetto alla condizione di pre-addestramento, dopo il training viene persa la simmetria del cappello messicano; infatti il neurone considerato riceve più eccitazione dalle cellule alla sua sinistra e una maggiore inibizione dai neuroni alla sua destra. L'asimmetria delle sinapsi porterà, come descritto più avanti, a un'alterazione nella percezione acustica.

Anche le sinapsi laterali visive subiscono dei cambiamenti, come mostrato in Figura 4.45, che illustra le connessioni entranti nel neurone visivo di posto $j=100$. Si può notare infatti che, dopo l'addestramento, sia le connessioni eccitatorie provenienti dai neuroni vicini sia quelle inibitorie che partono da cellule più distanti risultano potenziate, tuttavia la forma del cappello messicano rimane simmetrica. Questo tipo di variazione sinaptica farà in modo che non ci saranno shift duraturi nella percezione dello spazio visivo.

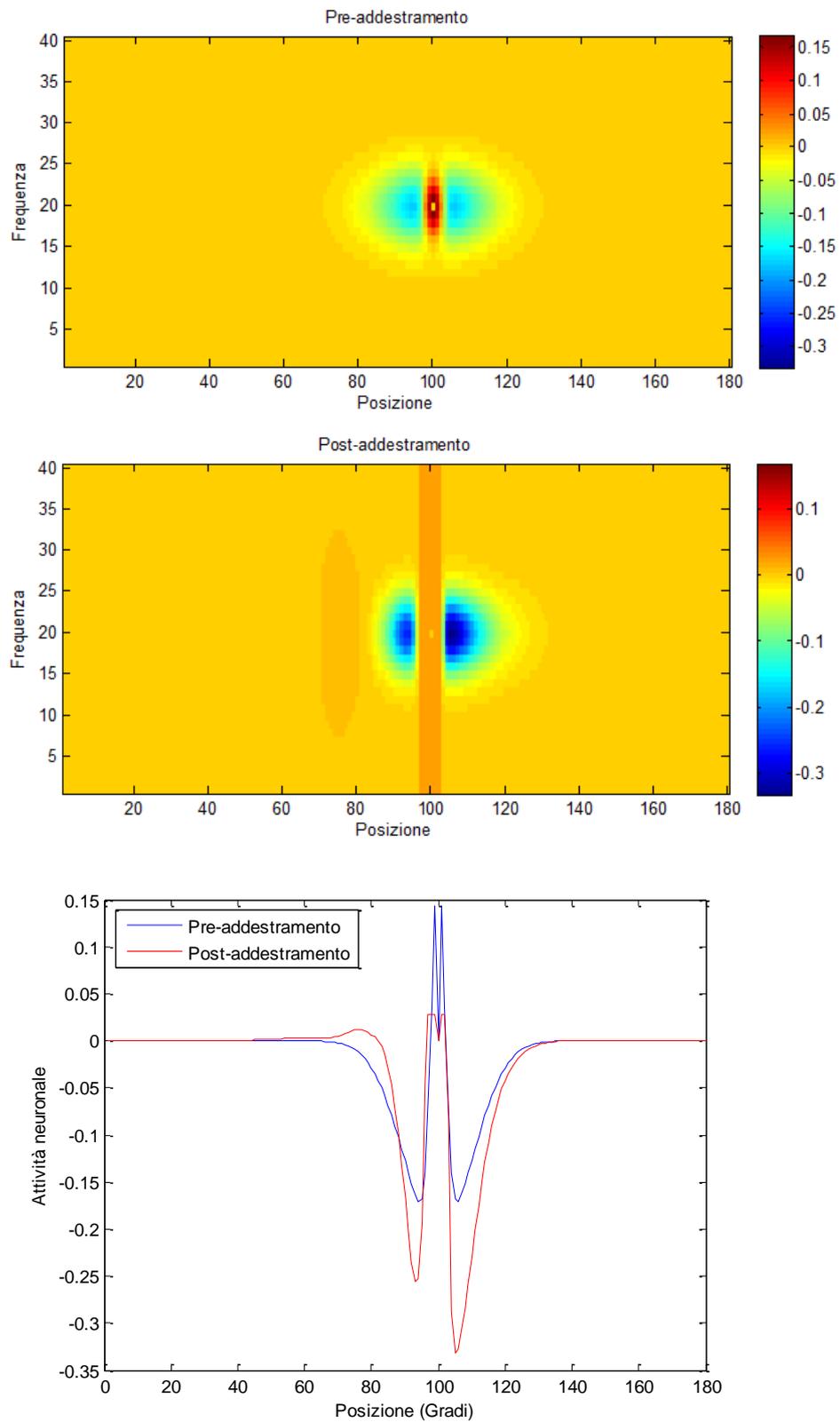


Figura 4.44. Modifiche delle sinapsi laterali acustiche.

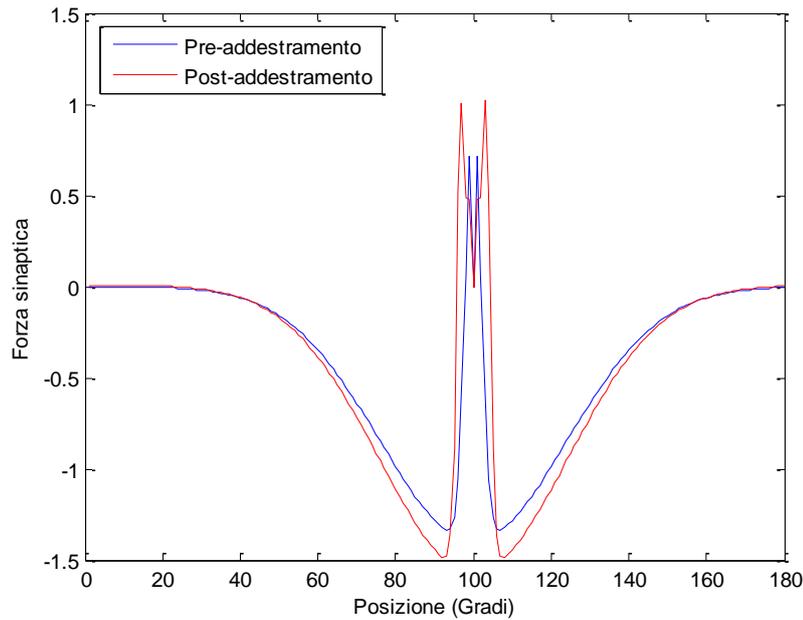


Figura 4.45. Modifiche delle sinapsi laterali visive.

4.3.2.2 Valutazione dell'aftereffect

La generazione dell'aftereffect può essere verificata valutando l'attivazione dei neuroni acustici dopo l'addestramento. Per fare questo la rete addestrata è stata stimolata con un input acustico centrato alla posizione e alla frequenza usate nel training ($i=20$ e $j=80$). In risposta a tale stimolazione, il modello produce un'attivazione nell'intera area uditiva e un profilo di attività dei neuroni acustici come quelli mostrati in Figura 4.46. Si nota subito che l'attività acustica è potentemente shiftata verso la posizione spaziale 100, cioè quella in cui era centrato l'input visivo durante il training. Questo shift è indicativo dell'alterazione che si è verificata a causa dell'adattamento.

Tuttavia, la grandezza che deve essere maggiormente valutata è il baricentro. Dopo la presentazione dell'input sonoro di coordinate $i=20$ e $j=80$, il valore del baricentro è risultato pari a 87.09° , evidenziando perciò l'insorgenza dell'aftereffect.

Come per la precedente configurazione, è stato verificato che non ci fosse un'alterazione nella percezione dello spazio visivo. Per fare questo, il modello addestrato ha ricevuto in ingresso uno stimolo solo visivo centrato alla posizione

100. Il corrispondente profilo di attività nell'area visiva è mostrato in Figura 4.47. Si nota subito che l'attivazione in quest'area resta uguale a quella della condizione di pre-addestramento. Inoltre il relativo valore del baricentro è uguale a 99.67° . Perciò si può concludere che non ci sono alterazioni significative nella percezione dello spazio visivo.

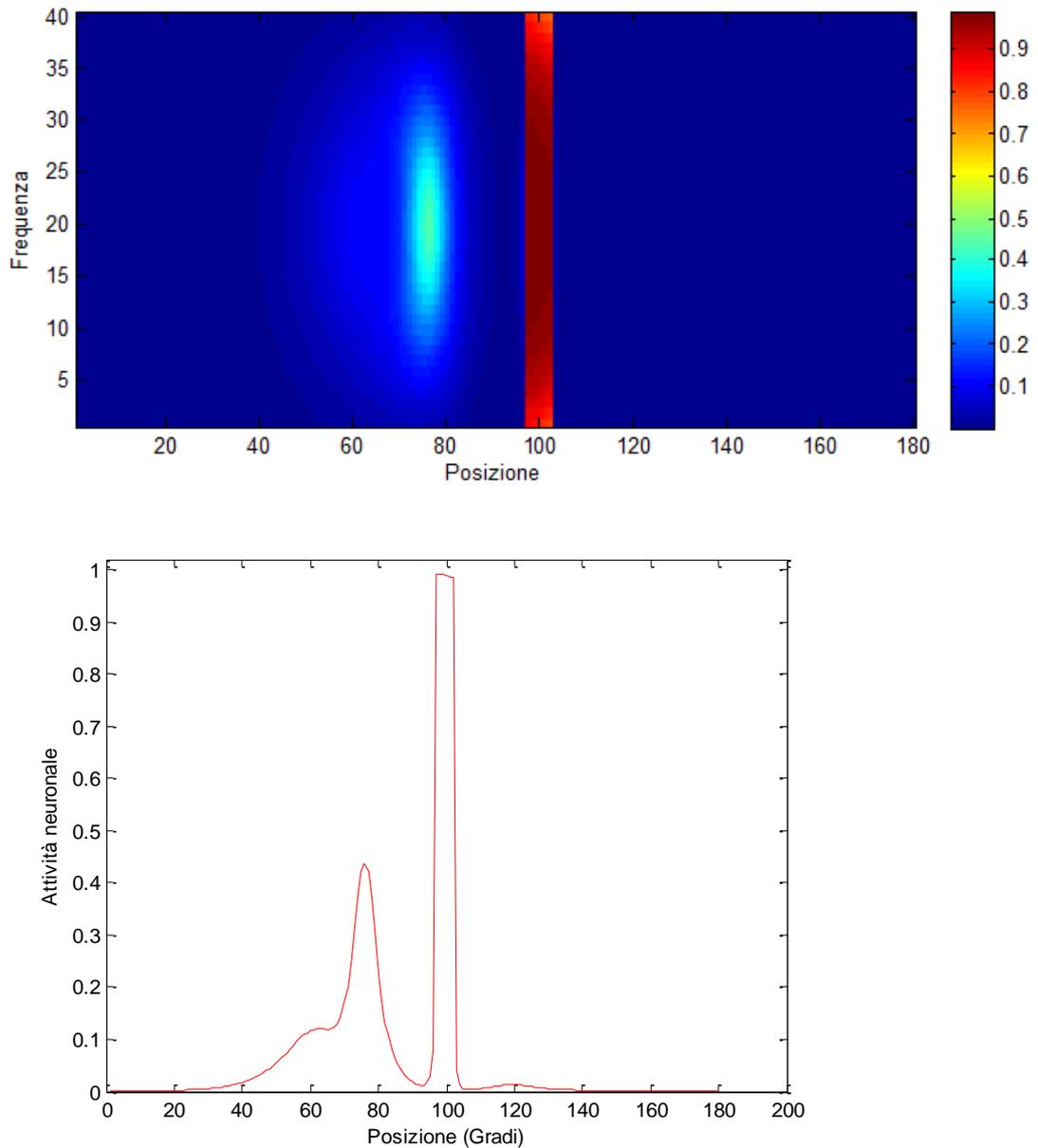


Figura 4.46. Attivazione nell'area uditiva (sopra) e profilo di attività dei neuroni acustici (sotto) dopo l'addestramento.

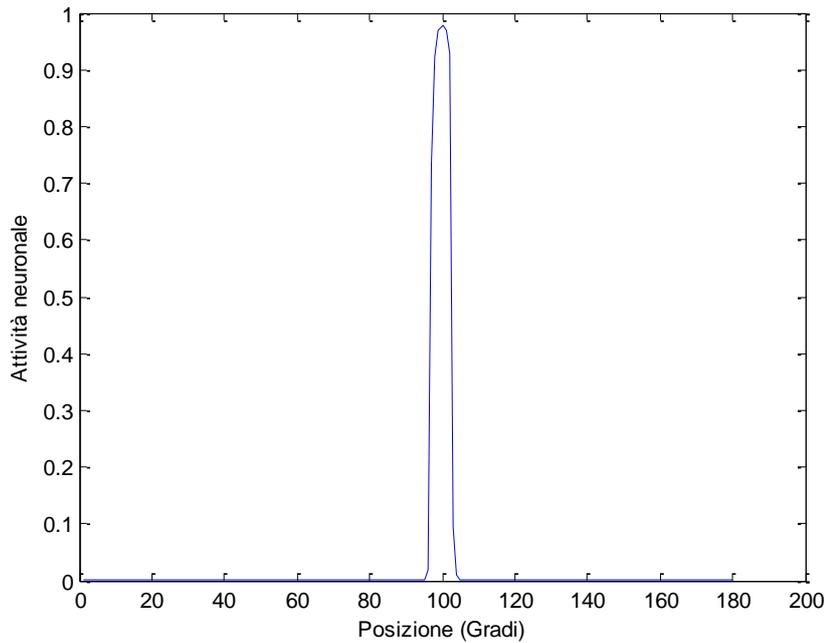


Figura 4.47. Profilo di attività nell'area visiva dopo il training a seguito della presentazione di un input visivo in posizione 100.

4.3.2.3 Verifica del trasferimento dell'aftereffect attraverso le frequenze

Per capire se l'effetto off-line del ventriloquismo si fosse generalizzato attraverso le frequenze sonore, alla rete addestrata è stato applicato uno stimolo acustico centrato in posizione 80, variandone di volta in volta la frequenza, fino a comprendere l'intero range di frequenza codificato dall'area uditiva del modello (200 - 6000 Hz). I risultati di tale simulazione sono mostrati in Figura 4.48, dove viene riportato l'aftereffect per ciascuna frequenza dello stimolo. Si osserva che lo shift percettivo è significativo per un intervallo di frequenze che va da 275 Hz a 4400 Hz. Dato che la frequenza di training è pari a 1100 Hz, è avvenuto un trasferimento dell'aftereffect per due ottave di frequenza. Questi risultati sono in accordo con quelli riportati da Frissen et al. [20].

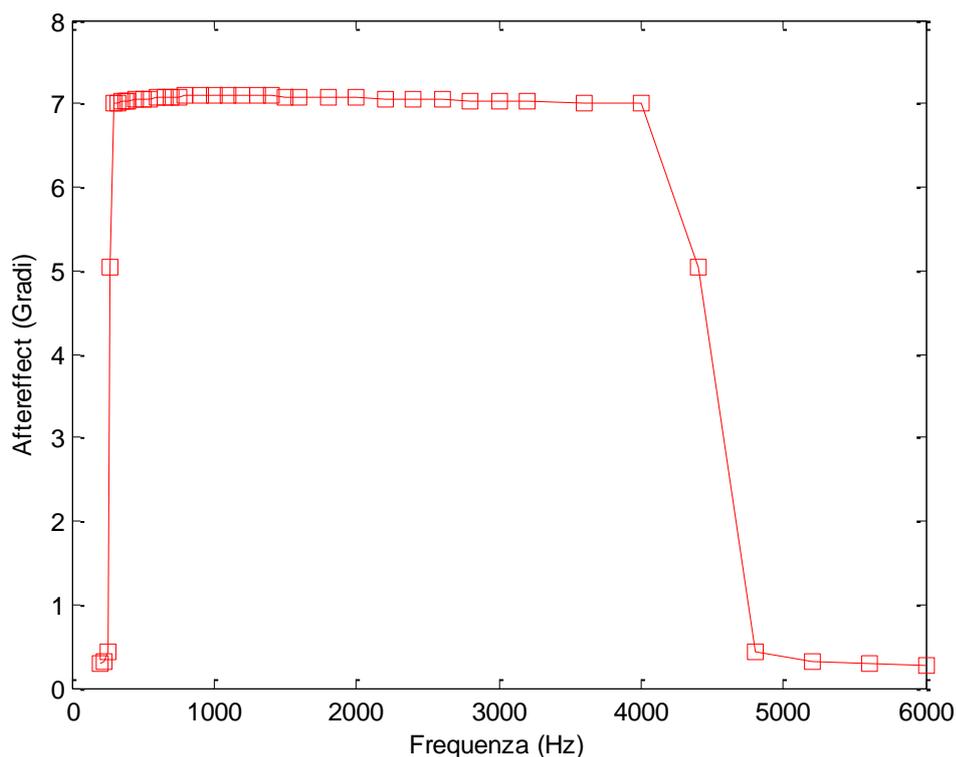


Figura 4.48 Trasferimento dell'aftereffect attraverso le frequenze.

La generalizzazione in frequenza dell'aftereffect rispecchia le variazioni che coinvolgono le sinapsi laterali entranti in neuroni acustici che codificano per frequenze lontane da quella di adattamento. La Figura 4.49 mostra le sinapsi laterali che arrivano in un neurone di posto $i=28$ (2200 Hz) e $j=100$. Le immagini superiore e centrale di tale figura mostrano le sinapsi che entrano nel neurone considerato nell'area acustica rispettivamente pre- e post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi considerate però soltanto lungo la posizione per una frequenza fissa ($i=28$). E' possibile osservare che le sinapsi post-addestramento entranti nel neurone considerato subiscono le stesse modifiche di quelle che arrivano sui neuroni che codificano per la frequenza di training (si veda la Figura 4.44). Questo tipo di plasticità sinaptica porta al trasferimento dell'aftereffect anche per frequenze diverse da quella di addestramento.

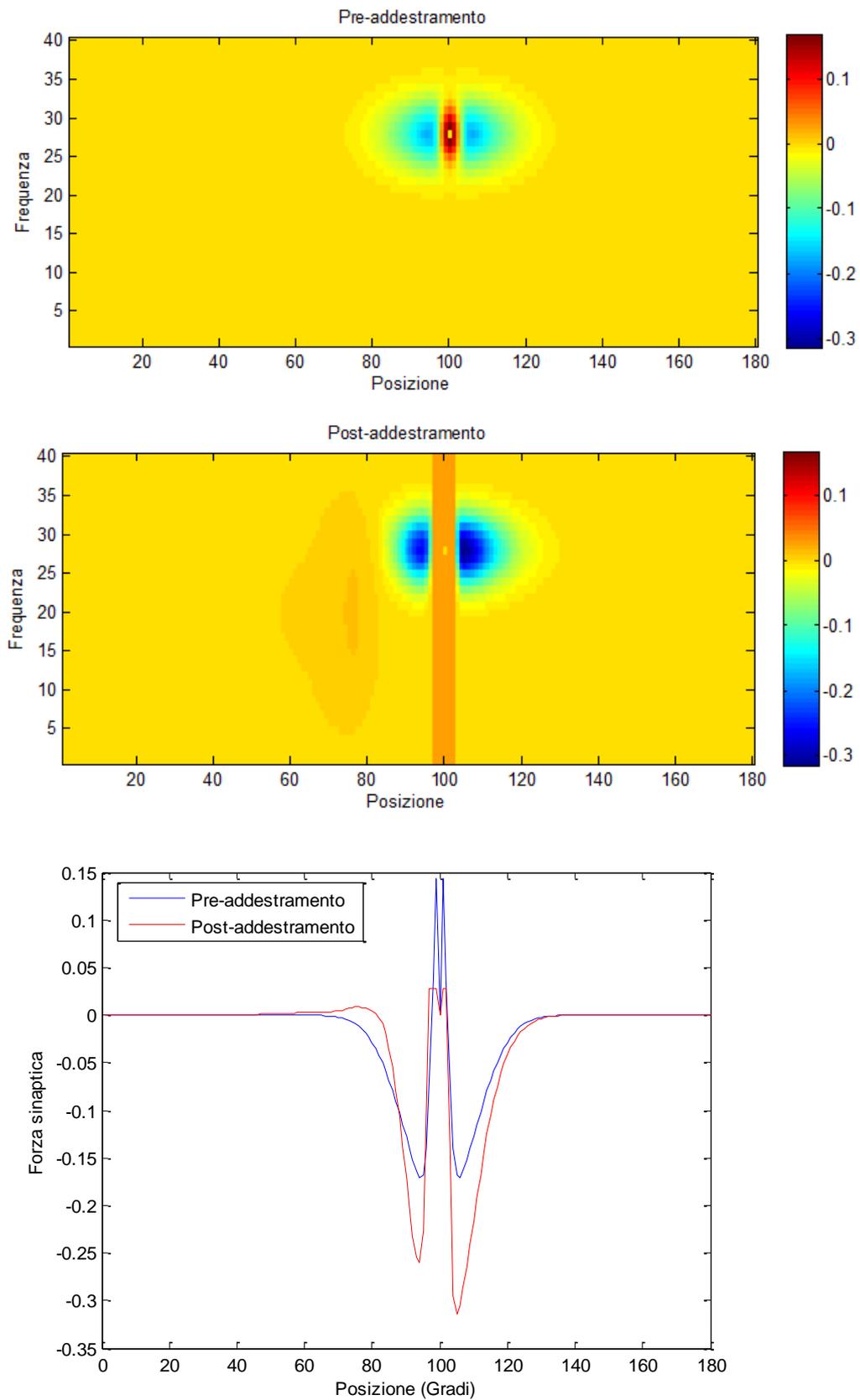


Figura 4.49. Modifiche delle sinapsi entranti in un neurone acustico che codifica per una frequenza diversa da quella di training.

4.3.3 Post-addestramento – Secondo paradigma di training

Il secondo paradigma di addestramento e le relative procedure di verifica dell'aftereffect utilizzate per la configurazione in esame sono gli stessi di quelli descritti nel sottoparagrafo 4.2.3.

4.3.3.1 Modifiche delle sinapsi

Anche qui sono state dapprima considerate le variazioni a carico delle sinapsi laterali provocate dalla fase di training. Nella Figura 4.50 sono riportati gli andamenti delle sinapsi laterali prima e dopo l'addestramento per il neurone acustico di posto $i=20$ e $j=100$. Le immagini superiore e centrale mostrano le sinapsi che entrano nel neurone di posto $i=20$ e $j=100$ nell'area acustica rispettivamente per il pre- e il post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi, considerate però soltanto lungo la posizione per una frequenza fissa ($i=20$). Dopo l'adattamento, il neurone considerato riceve una maggiore eccitazione da sinistra, e questo si traduce in una forma asimmetrica del cappello messicano, la quale provocherà uno sbilanciamento nella percezione dello spazio acustico.

Anche le sinapsi laterali visive vanno incontro a dei cambiamenti a causa del training, come indicato in Figura 4.51, che illustra le connessioni entranti nel neurone visivo di posto $j=100$. Dopo l'adattamento, si verifica un decremento nelle sinapsi eccitatorie dei neuroni vicini e un aumento nelle connessioni inibitorie delle cellule più distanti. Tali modifiche però non alterano la simmetria del cappello messicano; di conseguenza non ci saranno shift duraturi nella percezione dell'input visivo.

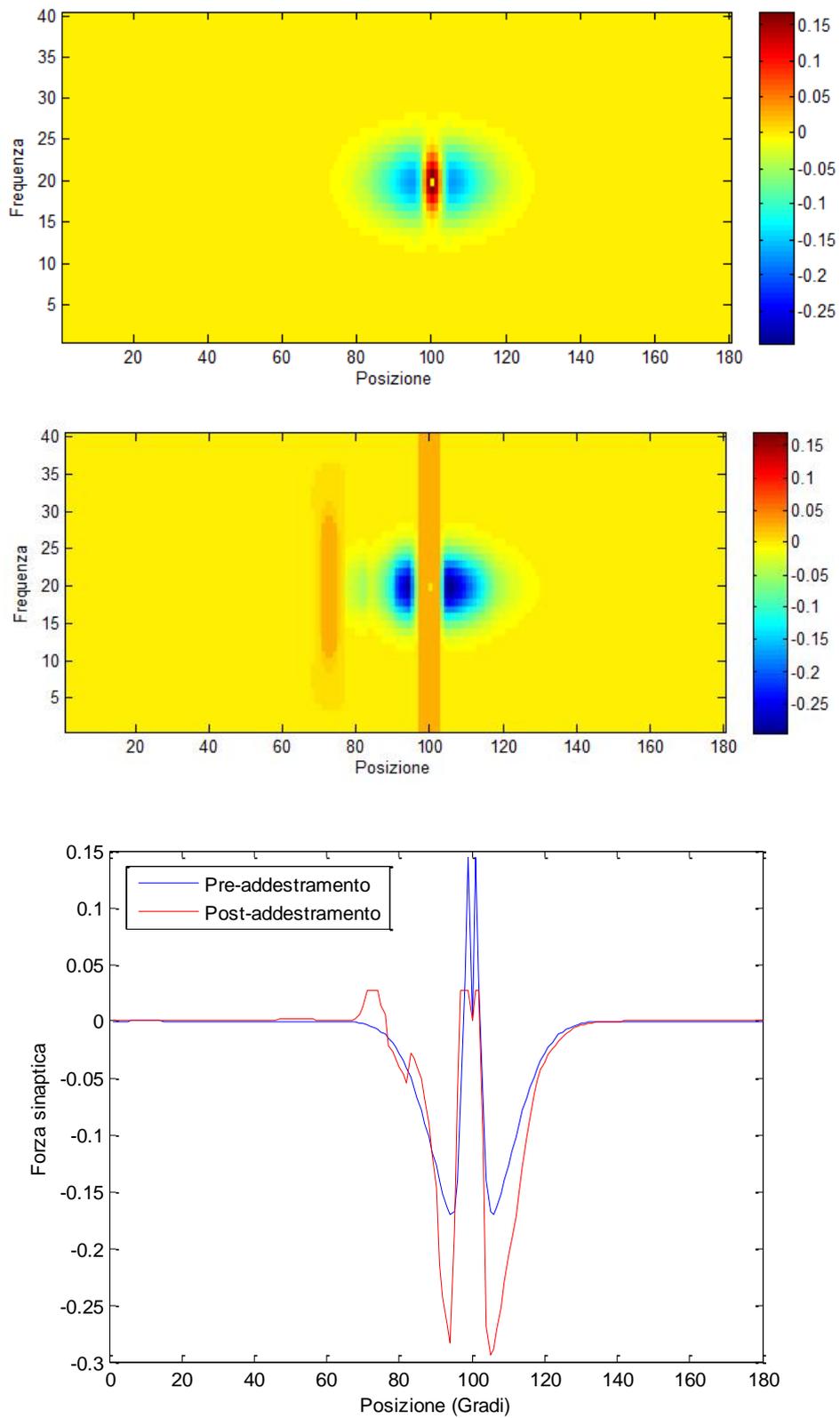


Figura 4.50. Modifiche delle sinapsi laterali acustiche.

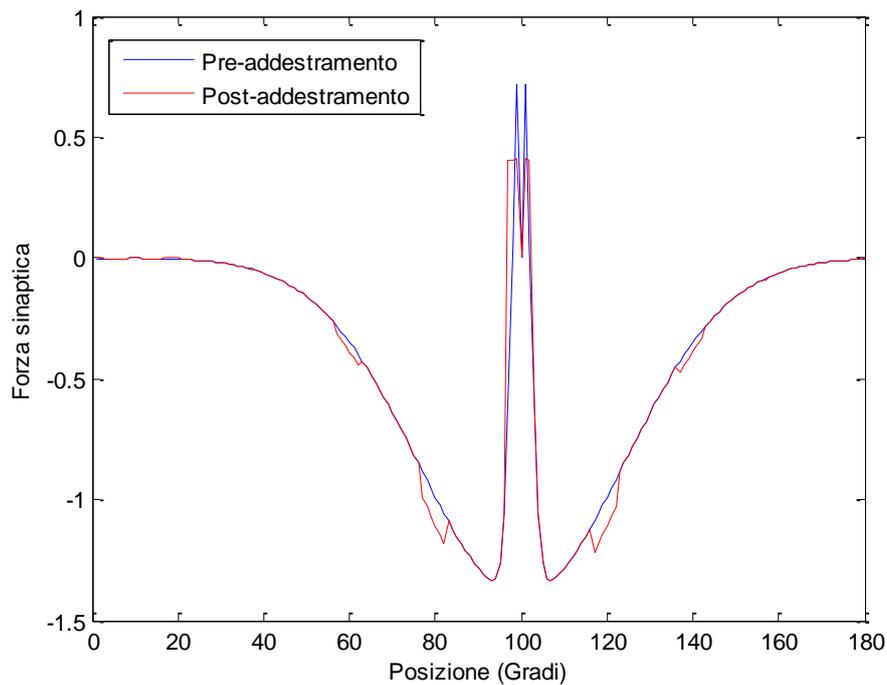


Figura 4.51. Modifiche delle sinapsi laterali visive.

4.3.3.2 Valutazione dell'aftereffect

Per verificare l'insorgenza dell'aftereffect del ventriloquismo, un aspetto che si può valutare è l'attivazione nell'area acustica in risposta alla presentazione di un input unicamente sonoro. La bolla di attivazione e il profilo di attività dei neuroni acustici dopo l'applicazione di uno stimolo uditivo centrato nel punto $i=20$ e $j=80$ sono mostrati in Figura 4.52. Come per la configurazione per la non generalizzazione dell'aftereffect, l'attivazione nello strato acustico è caratterizzata da due picchi. La spiegazione di questo risultato è la stessa di quella illustrata per il primo set di parametri.

Per valutare la presenza dell'aftereffect è comunque preferibile fare riferimento alla metrica del baricentro. Lo spostamento nella percezione è stato quantificato presentando al modello addestrato un input acustico esterno in una delle 9 posizioni usate durante il training ed è stato registrato il corrispondente valore del baricentro. Il valore medio baricentrico delle 9 posizioni addestrate è risultato pari a 7° .

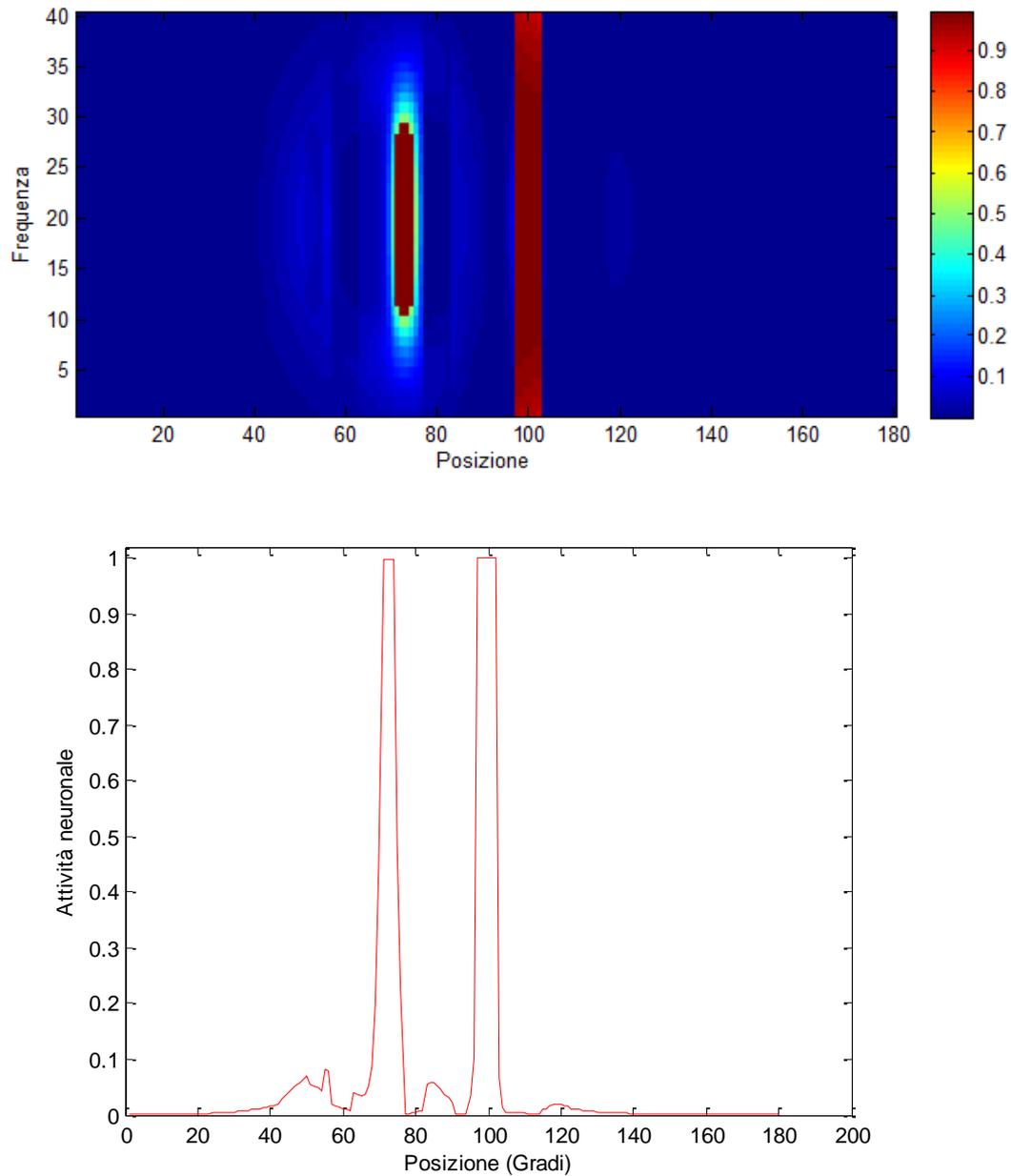


Figura 4.52. Attivazione nell'area uditiva (sopra) e profilo di attivazione dei neuroni acustici (sotto) dopo l'addestramento.

Lo shift percettivo è stato poi analizzato per posizioni spaziali non usate durante l'adattamento. Infatti in un'altra simulazione, lo stimolo è stato fatto variare in 19 posizioni, delle quali 10 non sono mai state usate nel training. La Figura 4.53 mostra gli esiti della prova: la posizione percepita dell'input sonoro in termini di baricentro è espressa in funzione della posizione in cui è stato applicato lo stimolo durante il test di verifica dell'aftereffect; è riportata anche la retta di regressione per i dati ottenuti.

Per le posizioni non addestrate l'effetto off-line del ventriloquismo si manifesta con minore entità (circa 2.34°). Tuttavia l'aftereffect complessivo risulta molto forte, infatti la retta di regressione ha una pendenza di quasi 1 e un'intercetta sull'asse delle ordinate di 4.57° , che rappresenta circa il 23% della distanza spaziale audio-visiva usata nell'addestramento, coerentemente con il range riportato in letteratura (tra il 20% e il 60% [19,20]). In queste simulazioni l'input acustico è sempre stato presentato alla stessa frequenza usata nel training (1100 Hz).

Anche per questa configurazione è stato verificato che non avvengono alterazioni a lungo termine nella percezione dello spazio visivo. Per fare questo, sono state compiute altre simulazioni in cui la rete addestrata è stata stimolata con input unicamente visivi in diverse posizioni. La Figura 4.54 riporta il profilo di attività nella area visiva dopo che è stato presentato un segnale visivo centrato alla posizione 100. L'attivazione neuronale appare non risentire affatto della fase di esposizione al conflitto cross-modale; inoltre il baricentro visivo risulta pari a 0.32° . Perciò nemmeno in questa configurazione avviene una ricalibrazione nella percezione visiva.

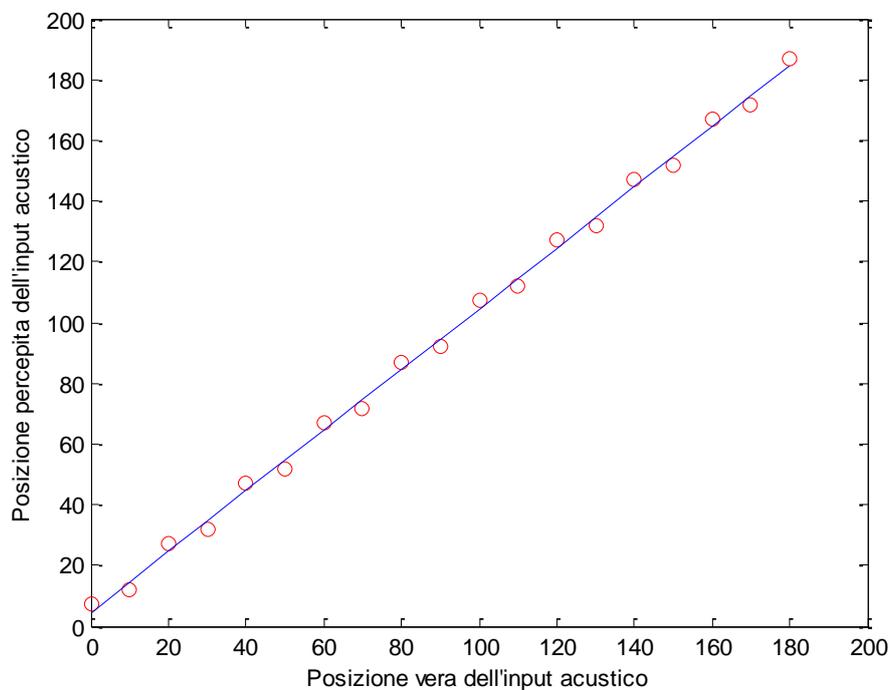


Figura 4.53. Verifica dell'aftereffect per i neuroni acustici.

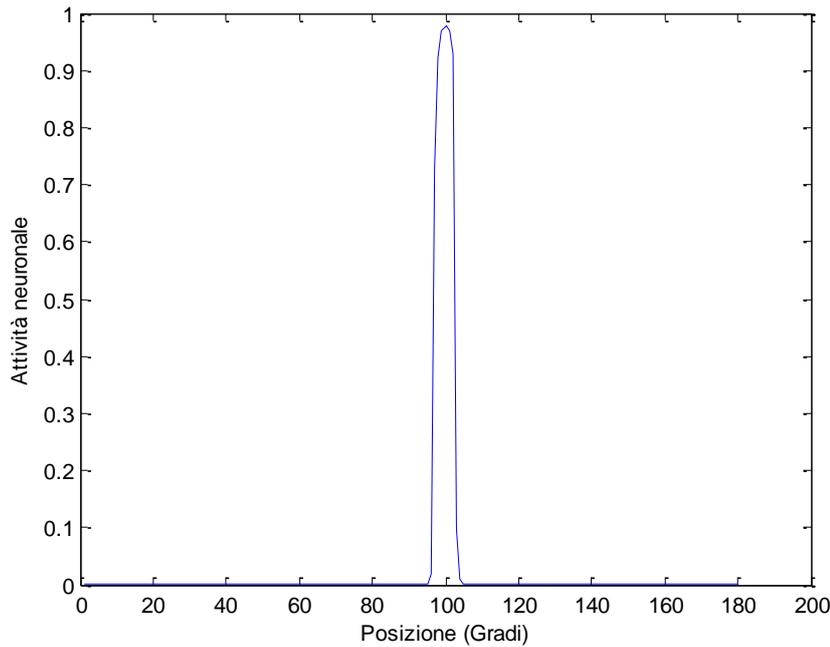


Figura 4.54. Profilo di attività nell'area visiva dopo il training a seguito della presentazione di un input visivo in posizione 100.

4.3.3.3 Verifica del trasferimento dell'aftereffect attraverso le frequenze

Vengono ora riportati i risultati che il modello addestrato ha fornito presentando input unicamente acustici a frequenze differenti rispetto a quella usata nel training (1100 Hz). La possibile generalizzazione in frequenza dell'aftereffect è stata valutata applicando input sonori in ciascuna delle 9 posizioni adottate in fase di training e in ognuna delle 40 frequenze codificate dai neuroni dell'area acustica. La Figura 4.55 riporta gli esiti di queste simulazioni; come per la configurazione precedente, l'aftereffect è stato calcolato come media, per ogni frequenza, degli aftereffect delle posizioni dell'input utilizzate durante l'addestramento. E' possibile osservare che lo spostamento nella percezione dello stimolo esterno rimane rilevante per un intervallo di frequenze che va da 250 Hz a 4800 Hz; considerando che la frequenza di training era di 1100 Hz, si può concludere che l'aftereffect si è trasferito per un intervallo di frequenza leggermente superiore a due ottave. Questi risultati sono in accordo con quelli ottenuti da Frissen et al. [20,21], come indicato in Figura 4.56.

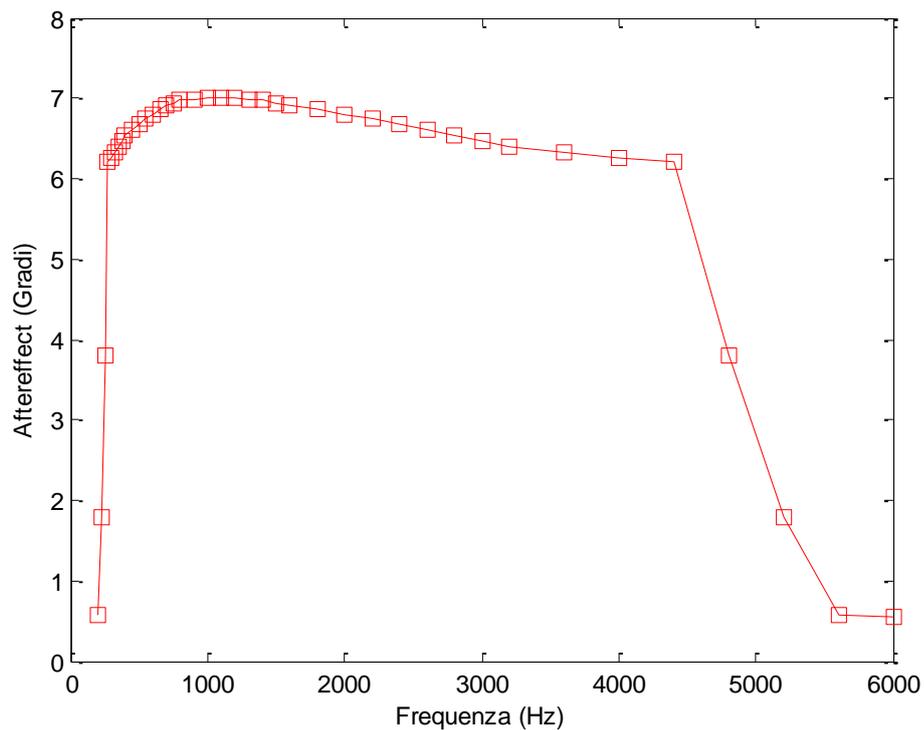


Figura 4.55. Trasferimento dell'aftereffect attraverso le frequenze.

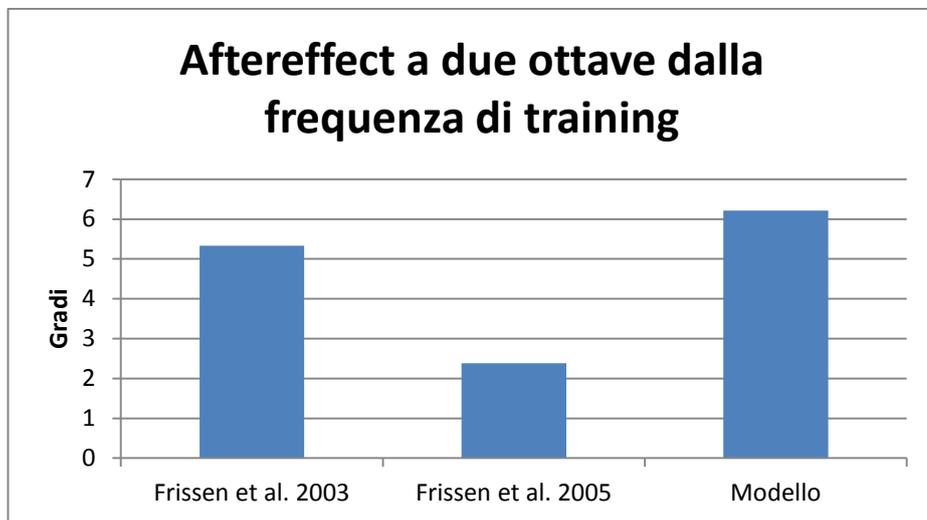


Figura 4.56. Confronto con i risultati della letteratura.

La generalizzazione in frequenza data da questo tipo di configurazione, rispecchia il fenomeno di plasticità a cui vanno incontro le sinapsi laterali che arrivano su neuroni

che codificano per frequenze lontane da quella di training. La Figura 4.57 mostra le sinapsi laterali che arrivano in un neurone di posto $i=28$ (2200 Hz) e $j=100$.

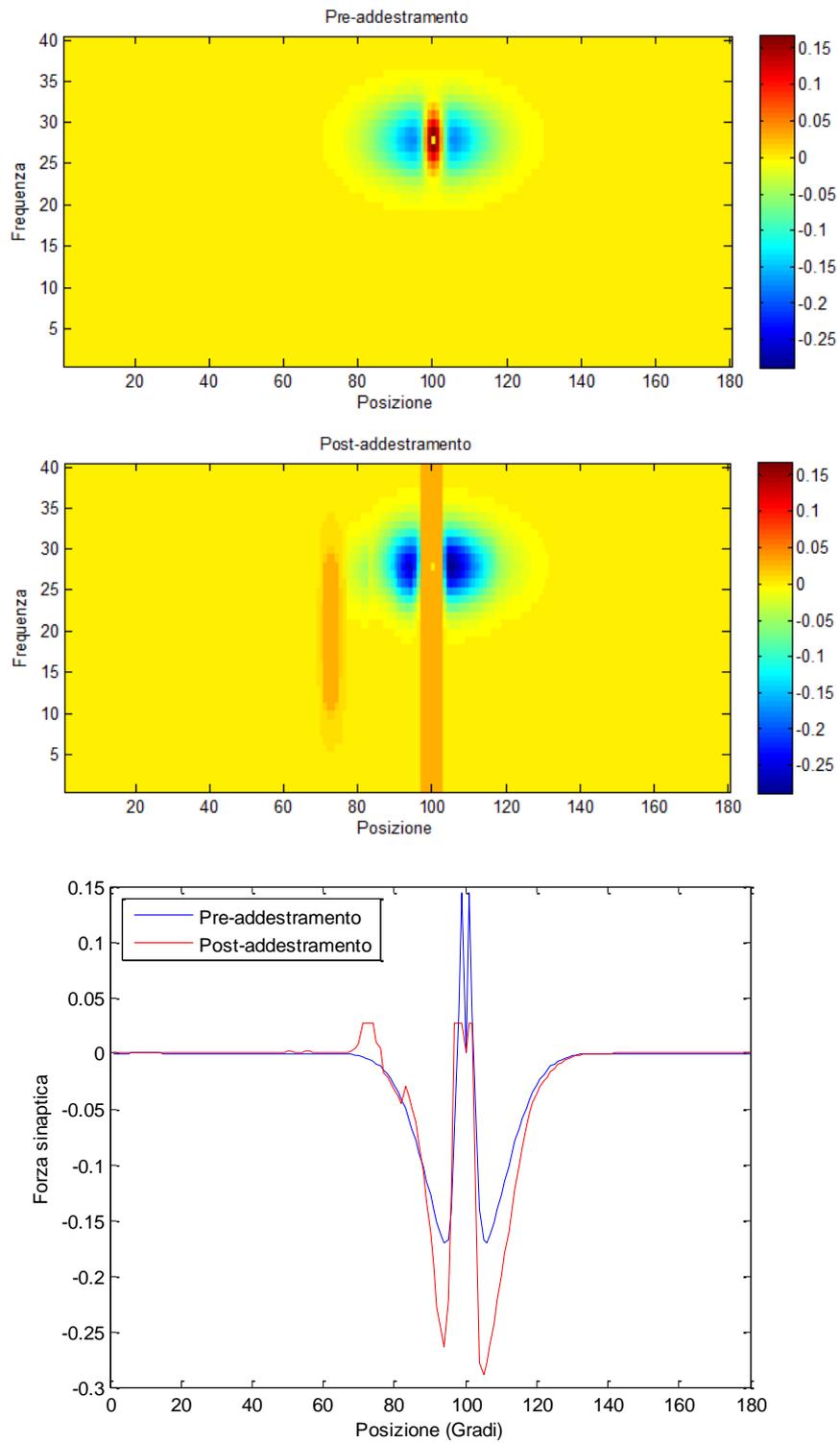


Figura 4.57. Modifiche delle sinapsi entranti in un neurone acustico che codifica per una frequenza diversa da quella di training.

Le immagini superiore e centrale mostrano le sinapsi che entrano nel neurone considerato nell'area acustica rispettivamente pre- e post-addestramento, mentre l'immagine inferiore rappresenta le stesse sinapsi, considerate però soltanto lungo la posizione per una frequenza fissa ($i=28$). E' possibile notare che le sinapsi laterali considerate si modificano in maniera praticamente identica a quelle entranti nei neuroni che codificano per la frequenza di training (si veda la Figura 4.50). Questo provoca evidentemente la generalizzazione in frequenza dell'aftereffect.

Capitolo 5

Risultati della configurazione per l'influenza dell'intensità dell'input acustico sull'aftereffect

5.1 Aspetti generali

In questo tipo di configurazione, l'attenzione maggiore è stata rivolta al peso che ha l'intensità dell'input acustico esterno sull'effetto off-line del ventriloquismo. In particolare, le simulazioni compiute per questo nuovo set di parametri sono servite a capire se variare la forza dello stimolo uditivo potesse avere delle conseguenze sul trasferimento dell'aftereffect attraverso le frequenze sonore. Perciò i risultati presentati di seguito sono stati ottenuti mantenendo tutti i parametri costanti tranne uno, ovvero l'intensità dell'input uditivo esterno (E_{0x}). L'intenzione di procedere con queste ulteriori valutazioni è nata dall'osservazione di alcuni risultati riportati in due lavori di Recanzone [24,25], che trattavano l'elaborazione spaziale e la risposta in frequenza e in intensità dei neuroni della corteccia uditiva dei macachi. Da questi studi sembrerebbe infatti che la risposta in frequenza dei neuroni acustici possa cambiare se varia l'intensità dell'input.

5.2 Pre-addestramento

5.2.1 Costruzione della FRA

Innanzitutto il modello è stato stimolato con un input esterno unicamente acustico in modo da costruire delle risposte in frequenza dei neuroni dell'area uditiva, che fossero confrontabili con quelle mostrate da Recanzone [24,25]. Nei due studi menzionati precedentemente, Recanzone, analizzando le proprietà delle cellule

neurali della corteccia uditiva dei macachi, ha riportato le risposte in frequenza in termini di *FRAs* (*Frequency Response Areas*): esse rappresentano i campi recettivi dei neuroni acustici rispetto alla frequenza e all'intensità dello stimolo sonoro applicato. La Figura 5.1 [24] illustra le *FRAs* di neuroni appartenenti a diverse aree corticali uditive, ciascuna con una propria frequenza caratteristica (*CF*).

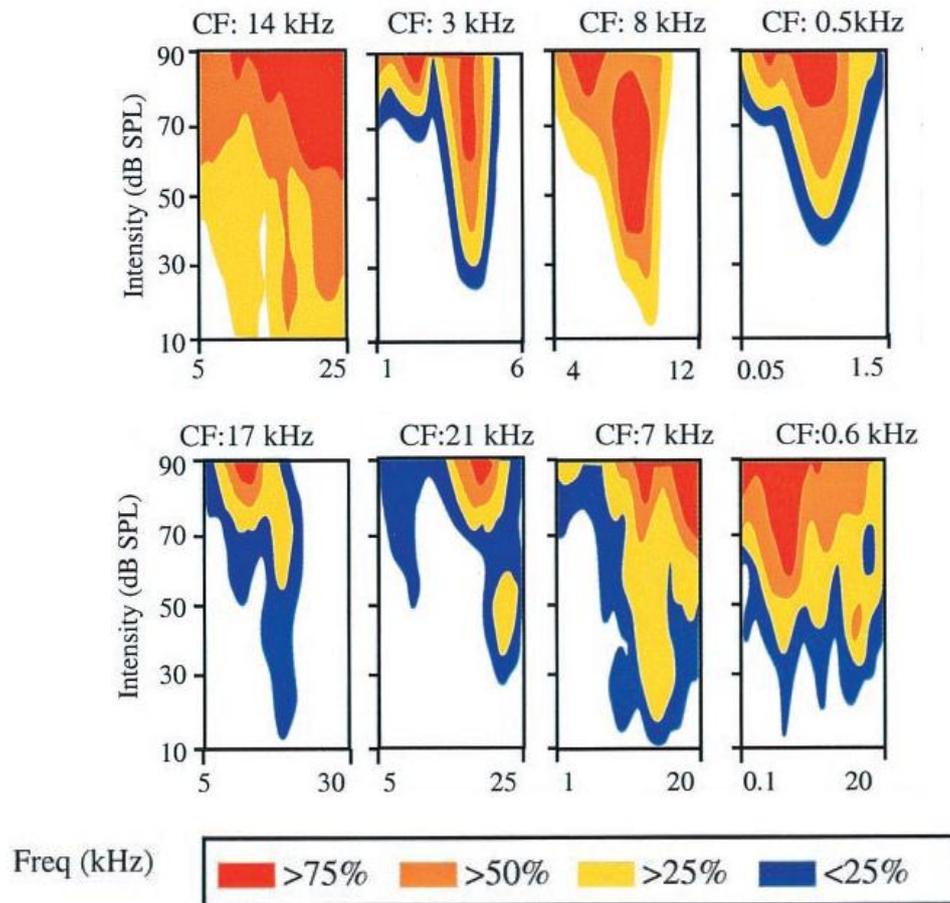


Figura 5.1. FRAs di neuroni corticali uditivi.

In questo caso, l'attivazione dei neuroni è stata calcolata come percentuale della risposta massima in seguito alla presentazione dell'input. Tale percentuale è rappresentata dai diversi colori presenti nell'immagine. E' possibile notare come variando la forza dello stimolo sonoro esterno, cambi non solo l'intensità della risposta, ma anche la sua ampiezza in frequenza. Infatti si osserva che, aumentando

l'intensità dell'input, il campo recettivo si allarga, nel senso che la risposta neurale diventa significativa per un range di frequenze molto ampio.

In analogia a quanto fatto da Recanzone [24,25], alla rete sono stati presentati stimoli uditivi variandone di volta in volta l'intensità. Gli input sono stati dati alla posizione $i=20$ e $j=80$ dell'area acustica e la loro forza è stata fatta variare attraverso 5 valori: 13, 15, 18, 20, 22. Così facendo sono stati registrati i profili di attività rispetto alla frequenza per ogni intensità testata, i quali sono stati normalizzati rispetto al valore massimo, ed è stata realizzata una FRA relativa allo strato acustico della rete neurale, come quello indicato in Figura 5.2.

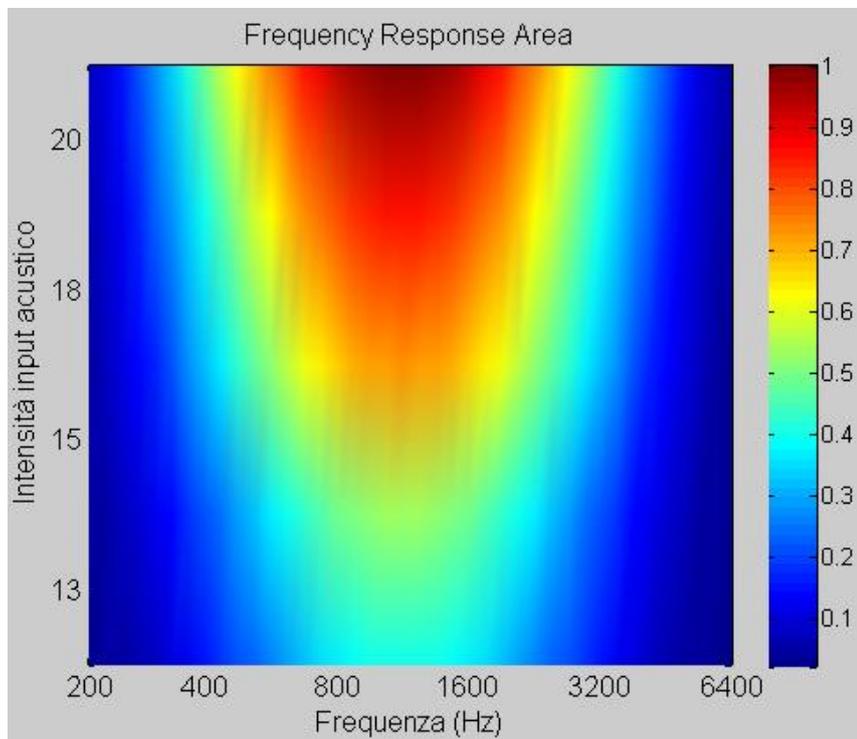


Figura 5.2. FRA della rete neurale.

Il modello riproduce il legame tra intensità dello stimolo esterno e ampiezza della risposta in frequenza dei neuroni acustici: un input più forte è associato a un campo recettivo in frequenza più ampio.

5.2.2 Stimolazione cross-modale e ventriloquismo

Per verificare che il modello con questo set di parametri riesce a riprodurre il fenomeno del ventriloquismo, alla rete sono stati dati in ingresso uno stimolo acustico e uno visivo simultanei ma spazialmente separati ($j=80$ per l'input acustico e $j=100$ per quello visivo). L'input acustico è stato centrato alla frequenza di 1100 Hz ($i=20$). Anche in questo tipo di prova, l'intensità dell'input sonoro è stata fatta variare tra i 5 valori indicati nel precedente paragrafo, mentre l'intensità dello stimolo visivo è stata mantenuta fissa ($E_{0y}=15$). Le simulazioni hanno portato ai profili di attività rispetto alla posizione nell'area uditiva mostrati in Figura 5.3. Si osserva immediatamente che l'attivazione dei neuroni acustici è fortemente spostata verso la posizione in cui è stato applicato l'input nell'altra modalità, con un andamento che cresce all'aumentare dell'intensità dello stimolo esterno.

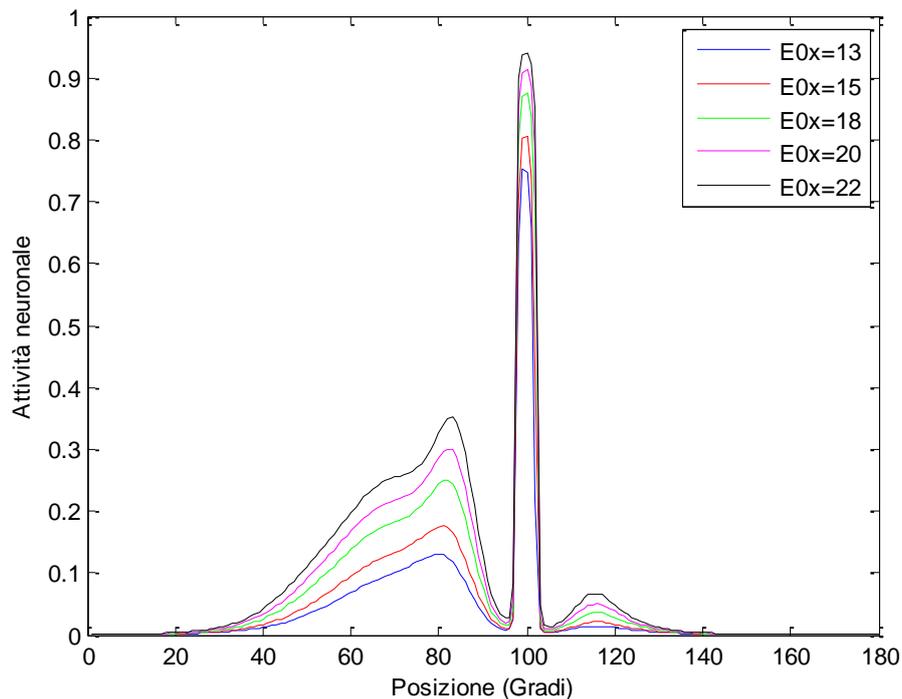


Figura 5.3. Profili di attività rispetto alla posizione corrispondente alla frequenza di 1100 Hz nell'area acustica per ogni intensità dell'input sonoro.

E' stato inoltre calcolato il baricentro per ciascuna forza dell'input e i valori corrispondenti sono riportati nella seguente tabella:

E_{0x}	Baricentro
13	86.51°
15	85.52°
18	84.31°
20	83.62°
22	83.04°

Osservando il valore del baricentro si nota come esso diminuisca al crescere dell'intensità dello stimolo applicato. Questo trend è da imputare probabilmente al fatto che, mantenendo costante la forza dell'input visivo, ogni volta che aumenta quella del segnale acustico, l'attrazione che l'attivazione dei neuroni visivi esercita sui neuroni uditivi diminuisce.

Dai risultati mostrati in questo paragrafo si può comunque concludere che il modello, anche con quest'ulteriore set di parametri, riproduce efficacemente l'effetto on-line del ventriloquismo.

Un aspetto interessante da considerare è la risposta in frequenza dei neuroni dell'area acustica, che si ottiene a seguito della stimolazione cross-modale sopra descritta.

A tal proposito, nella Figura 5.4 sono illustrati i profili di attività rispetto alla frequenza dei neuroni acustici, in corrispondenza della posizione 100, riportati per ciascuna intensità dell'ingresso. Si può notare che l'attivazione neuronale oltre che crescere in intensità, aumenta anche in ampiezza usando ingressi più forti. Se per esempio si confronta il profilo per $E_{0x}=13$ e quello per $E_{0x}=22$, risulta che i neuroni che presentano un'attivazione sopra 0.6 nel primo caso vanno da 450 Hz ($i=10$) a 2600 Hz ($i=30$), mentre nel secondo coprono un intervallo che va da 300 Hz ($i=5$) a 4000 Hz ($i=35$). Queste osservazioni inducono a pensare che probabilmente la forza del segnale in ingresso giocherà un ruolo importante nel trasferimento dell'aftereffect attraverso le frequenze.

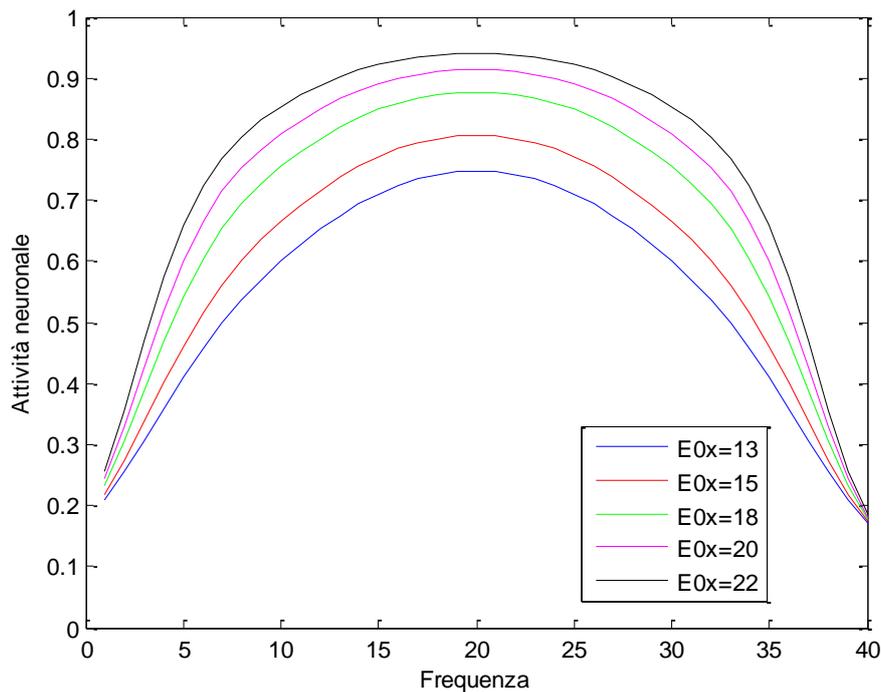


Figura 5.4. Profili di attività rispetto alla frequenza nell'area acustica per ogni intensità dell'input sonoro.

5.3 Post-addestramento

I risultati di post-addestramento sono stati ricavati dopo aver sottoposto il modello a un periodo di adattamento, in cui venivano presentati stimoli cross-modalità contemporanei ma spazialmente discrepanti. In particolare, il training si è composto di una prova della durata di 1000 ms. Durante la prova la posizione spaziale dell'input visivo e di quello acustico sono state tenute costanti ($j=100$ per il primo e $j=80$ per il secondo). Durante il training, le sinapsi laterali in entrambe le aree si sono addestrate secondo le regole descritte nel Capitolo 3. Anche la frequenza di applicazione del segnale uditivo è stata mantenuta allo stesso valore, pari a 1100 Hz (quindi $i=20$).

Come per i due precedenti set di parametri, l'effetto off-line del ventriloquismo è stato valutato presentando alla rete addestrata un input unicamente sonoro centrato

nella stessa posizione spaziale dell'addestramento, dapprima alla frequenza di training e dopo a tutte le frequenze codificate dai neuroni dello strato acustico (200 - 6000 Hz). L'intera procedura è stata eseguita per ogni intensità dell'input sonoro.

5.3.1 Valutazione dell'aftereffect

Un modo per capire se c'è uno shift di lunga durata nella percezione di segnali acustici è considerare l'attivazione dei neuroni uditivi rispetto alla posizione. Così, in Figura 5.5 sono illustrati i profili di attività rispetto alla posizione per le 5 diverse intensità dello stimolo acustico esterno. Dai risultati si evince che nei casi $E_{0x}=13$ ed $E_{0x}=15$ non si verifica uno spostamento dell'attività dei neuroni acustici verso quelli dell'altra modalità, nemmeno per la stessa frequenza di adattamento. Per gli altri tre valori di intensità (18, 20 e 22) c'è invece un chiaro spostamento dell'attivazione uditiva verso quella visiva. Questi profili di attività rispecchiano i valori del baricentro ottenuto per ciascuna forza dell'ingresso, come indicato nella seguente tabella:

E_{0x}	Baricentro
13	80.01°
15	80.09°
18	82.5°
20	82.71°
22	82.57°

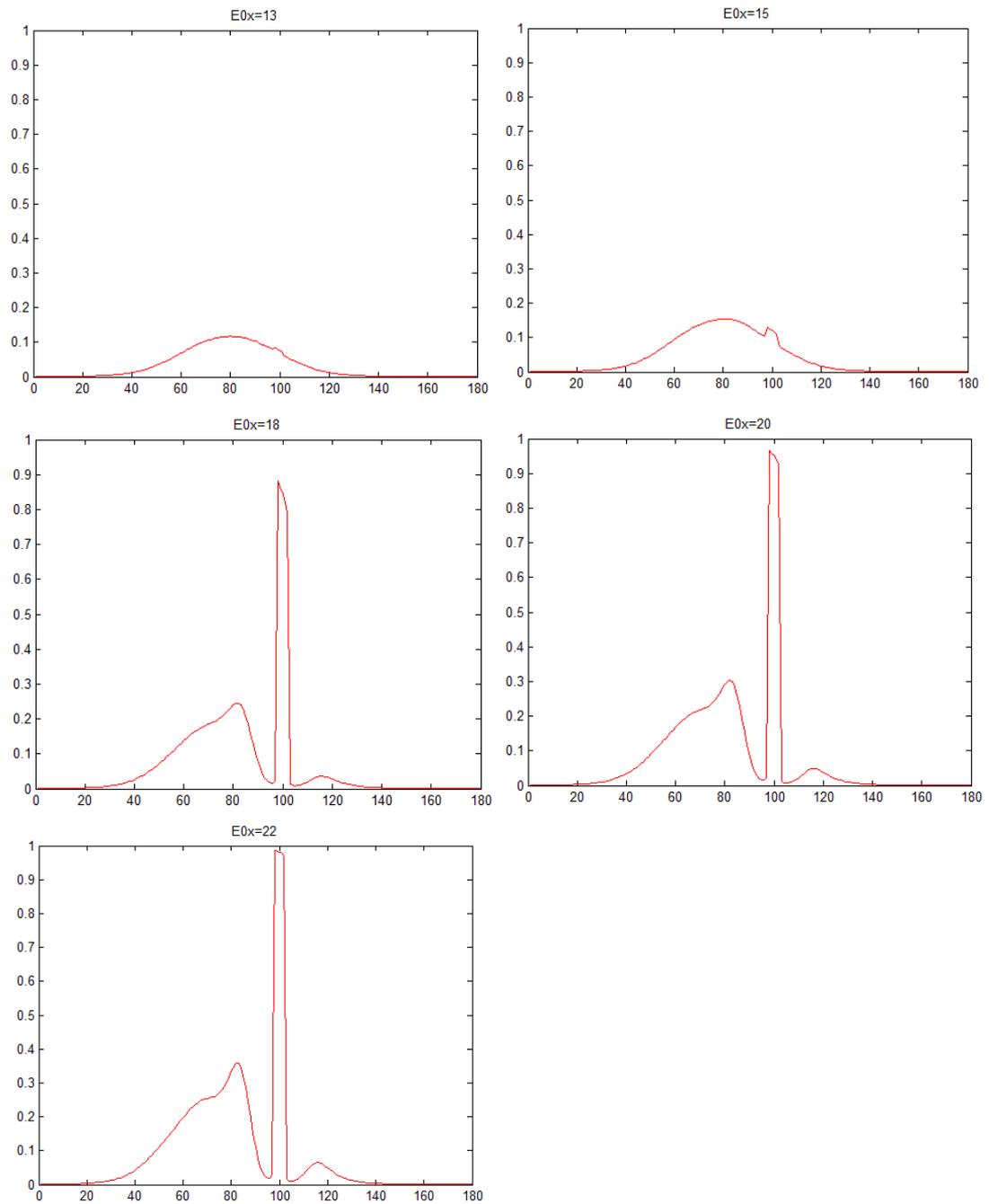


Figura 5.5. Profili di attività rispetto alla posizione nell'area acustica per ciascuna delle intensità dell'input. In ascissa è riportata la posizione in gradi e in ordinata l'entità dell'attivazione dei neuroni.

5.3.2 Verifica del trasferimento dell'aftereffect attraverso le frequenze

Vengono ora riportati i risultati ottenuti dopo che la rete addestrata è stata stimolata con input sonori a frequenze diverse da quella di training, al fine di capire se lo shift nella percezione uditiva fosse rimasto localizzato alla frequenza di addestramento oppure no. La Figura 5.5 illustra gli esiti di tali simulazioni per ogni intensità dell'input sonoro.

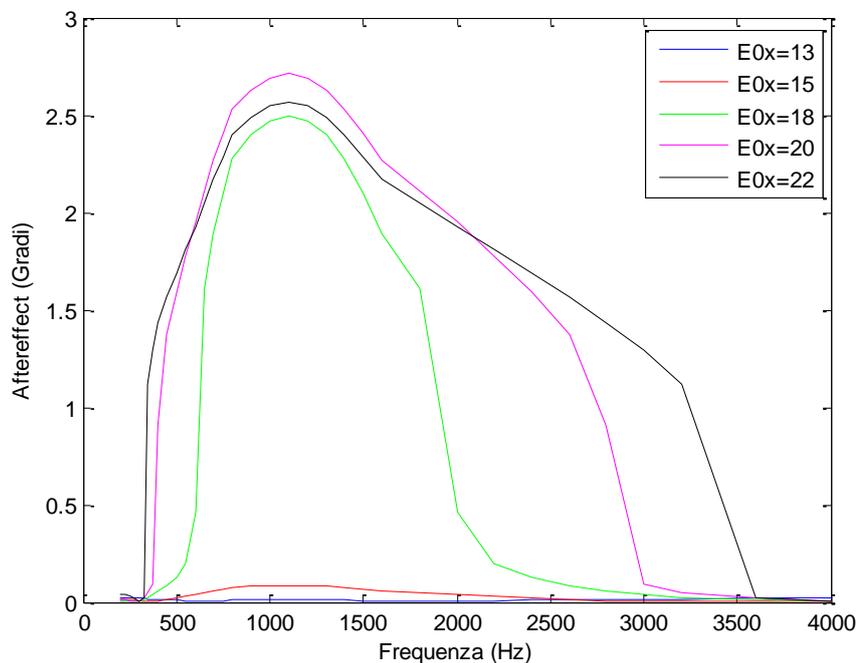


Figura 5.5. Generalizzazione dell'aftereffect.

Si nota subito che al crescere dell'intensità dell'input acustico esterno, l'aftereffect si trasferisce per un range di frequenze sempre più ampio; questo evidenzia l'esistenza di una dipendenza tra effetto off-line del ventriloquismo e intensità del segnale sonoro utilizzato per addestrare e testare la rete neurale.

Questa dipendenza potrebbe spiegare le differenze tra i risultati ottenuti da Recanzone [12,18] e quelli ricavati da Frissen et al. [20,21]. Come detto nei precedenti capitoli, le ricerche compiute da Recanzone [12,18] mostrano che

l'aftereffect del ventriloquismo resta limitato alla frequenza di training, mentre i risultati di Frissen et al. [20,21] evidenziano un trasferimento dell'aftereffect attraverso le frequenze sonore. Gli esperimenti di Recanzone sono stati condotti presentando ai soggetti segnali acustici di intensità pari a 45 dB SPL, in quelli di Frissen et al. invece sono stati usati stimoli sonori di 70 dB SPL. Si può quindi osservare che la generalizzazione in frequenza dell'aftereffect riportata da Frissen et al. è stata ottenuta con un'intensità dell'input uditivo maggiore rispetto a quella usata da Recanzone. Come suggerito dai risultati del modello, l'utilizzo di input acustici di diversa intensità può spiegare, almeno in parte, il diverso range di frequenza in cui l'aftereffect si trasferisce.

Capitolo 6

Conclusioni

Nel presente lavoro di tesi è stato sviluppato un modello di rete neurale per lo studio di un particolare aspetto dell'interazione spaziale visuo-acustica, ovvero il fenomeno del ventriloquismo, e dei meccanismi di plasticità sinaptica ad esso legati.

La rete neurale è costituita da due strati di neuroni, uno rappresentativo dell'area visiva e un altro dell'area acustica. Lo strato visivo è composto da una catena monodimensionale di neuroni: ogni neurone codifica per una posizione dello spazio. Lo strato acustico è invece di tipo bidimensionale: ciascun neurone codifica per una posizione dello spazio e per una frequenza sonora.

I neuroni appartenenti alla stessa area unimodale sono collegati da sinapsi laterali della forma a cappello messicano, ovvero ogni neurone eccita i suoi vicini e inibisce le cellule più lontane. I neuroni di strati differenti sono connessi da sinapsi inter-area: ciascun neurone acustico manda una sinapsi eccitatoria a un solo neurone visivo, e cioè quello che codifica per la sua stessa posizione spaziale; ogni cellula visiva invia sinapsi a tutti i neuroni uditivi che si trovano nella sua stessa locazione spaziale, ma che codificano per frequenze diverse.

Gli ingressi della rete neurale sono rappresentati da stimoli uditivi e visivi; la risposta della rete è costituita dall'attivazione delle due popolazioni di neuroni. Per poter confrontare i dati sperimentali (per lo più comportamentali) con le predizioni del modello, l'attivazione di ciascuna popolazione è stata decodificata in termini di posizione percepita dello stimolo, utilizzando la semplice metrica del baricentro

A differenza delle sinapsi inter-area, mantenute fisse, le sinapsi laterali possono modificarsi in seguito a una fase di addestramento mediante regole di Hebb.

L'addestramento si è composto di prove in cui alla rete venivano ripetutamente applicati input acustici e visivi concomitanti, ma in diverse posizioni dello spazio.

Il modello così realizzato è in grado di riprodurre l'effetto on-line del ventriloquismo: se la rete riceve in ingresso un input sonoro e uno visivo temporalmente coincidenti, ma spazialmente separati, in uscita viene fornita una

posizione dello stimolo acustico spostata verso quella in cui è stato applicato il contemporaneo segnale visivo.

Inoltre, il modello riesce a simulare l'aftereffect del ventriloquismo: se alla rete addestrata viene presentato un ingresso unicamente sonoro, la posizione percepita di tale stimolo continua a essere spostata nella direzione dell'input visivo usato in fase di adattamento.

Il modello è stato testato in tre diverse configurazioni, corrispondenti a tre set di parametri differenti, al fine di approfondire le caratteristiche in frequenza dell'aftereffect del ventriloquismo.

Una configurazione genera un aftereffect che rimane limitato alla frequenza di stimolazione usata durante l'addestramento. Un altro set di parametri invece riproduce la generalizzazione dell'aftereffect attraverso le frequenze sonore. Il set di parametri usato in questa seconda configurazione (che riproduce la generalizzazione in frequenza) si differenzia dal precedente per una maggiore ampiezza del campo recettivo in frequenza dei neuroni acustici e per un più elevato rapporto eccitazione/inibizione laterale all'interno dell'area uditiva. Infine la terza configurazione mostra, a parità di set di parametri, la dipendenza della generalizzazione dell'effetto off-line del ventriloquismo dall'intensità dell'input acustico.

Il mancato trasferimento in frequenza dell'aftereffect, prodotto dal modello testato nella prima configurazione, riflette il tipo di plasticità a cui vanno incontro le sinapsi laterali uditive a causa dell'addestramento: le sinapsi entranti in neuroni acustici che codificano per frequenze vicine a quella usata nell'adattamento sono le uniche a modificarsi, mentre le altre non subiscono variazioni significative rispetto alla loro condizione di base.

La generalizzazione in frequenza dell'aftereffect fornita dal secondo set di parametri si spiega considerando anche in questo caso le modifiche subite dalle sinapsi laterali acustiche durante l'adattamento: le connessioni che arrivano in neuroni che rispondono a frequenze distanti da quella di training subiscono modifiche simili a quelle entranti nel neurone che codifica per la frequenza di addestramento. In questa configurazione, il modello riesce così a trasferire l'aftereffect per due ottave di frequenza.

Nei due casi considerati, la generalizzazione o meno in frequenza dell'aftereffect è imposta a priori dal set di parametri utilizzati. Tuttavia i due casi illustrati sono utili in quanto mostrano come differenze nei parametri, ad esempio dovute ad una variabilità tra soggetti o tra gli stessi elementi neurali, possono spiegare risultati comportamentali contrastanti.

Un risultato degno di rilievo raggiunto in questa tesi è quello ricavato testando la rete neurale con il terzo set di parametri. In questo caso infatti, il modello fornisce in uscita uno shift percettivo a lungo termine che coinvolge un range di frequenze sempre più ampio all'aumentare dell'intensità dell'input sonoro usato per addestrare e testare la rete neurale. Tali risultati suggeriscono che il trasferimento o meno dell'aftereffect attraverso le frequenze sonore dipende anche dalla forza dello stimolo esterno. Questo legame tra generalizzazione dell'aftereffect ed intensità dell'input spiega parzialmente i risultati discordanti presenti in letteratura [12,18-21], riguardanti la generalizzazione in frequenza del fenomeno in esame. Sembrerebbe infatti che negli studi in cui tale generalizzazione è stata raggiunta [20,21], l'intensità dell'input sonoro fosse maggiore rispetto agli esperimenti in cui non c'è stato trasferimento [12,18,19]. L'ipotesi suggerita dal modello potrebbe essere testata in vivo, progettando uno studio sperimentale che analizza le proprietà di generalizzazione dell'aftereffect variando in modo sistematico l'intensità dell'input acustico.

Per quanto riguarda i limiti del modello, un primo aspetto da considerare concerne le frequenze sonore codificate dallo strato uditivo della rete. Il range rappresentato nel modello copre 40 valori di frequenza, che vanno da 200 Hz a 6000 Hz, con un incremento raddoppiato ogni 8 valori. L'intervallo di frequenze codificato è quindi abbastanza limitato e questa scelta è stata dettata da motivi computazionali. Sarebbe perciò auspicabile aumentare il numero di frequenze codificate dai neuroni uditivi della rete, al fine di riprodurre una mappa tonotopica più ampia e quindi maggiormente simile a quella presente negli esseri umani.

Un'altra caratteristica che potrebbe essere rivalutata è il tipo di addestramento sinaptico. Nel modello sviluppato, le uniche sinapsi che si addestrano sono quelle laterali, mentre le connessioni inter-area rimangono fisse. Si potrebbe pensare di modificare il paradigma di training in modo da consentire anche alle sinapsi che

collegano i due strati unimodali di modificarsi mediante regole di Hebb, anche se questo probabilmente porterebbe a un effetto fantasma nella modalità non stimolata. Infine sarebbe opportuno cercare di stabilire una corrispondenza tra le intensità dell'input acustico usate nel modello, che sono adimensionali, e quelle presenti nei lavori di letteratura, che sono invece espresse in dB SPL [12,18-21]. Questo consentirebbe di effettuare migliori confronti tra i risultati del modello e quelli ottenuti da esperimenti in vivo.

In conclusione, il modello di rete neurale descritto riproduce bene diversi risultati sperimentali, spiega parzialmente alcune discordanze presenti in letteratura in riferimento all'aftereffect del ventriloquismo, è stato realizzato sulla base di meccanismi biologici plausibili (connessioni sinaptiche, regole di apprendimento di Hebb), infine suggerisce nuovi esperimenti da realizzare in vivo per testare le predizioni da esso fornite.

Bibliografia

- [1] Graziano M. S. e Gross C. G., «The representation of extrapersonal space: a possible role for bimodal, visuo-tactile neurons», in *The Cognitive Neurosciences*, MIT press, 1995, pagg 1021–1034.
- [2] B. E. Stein e M. A. Meredith, *The Merging of the Senses*. Mit Press, 1993.
- [3] E. Macaluso e J. Driver, «Multisensory spatial interactions: a window onto functional integration in the human brain», *Trends Neurosci.*, vol 28, n° 5, pagg 264–271, Mag 2005.
- [4] F. Leo. *Il ruolo del collicolo superiore nell'orientamento spaziale*. Tesi di dottorato. Dottorato di ricerca in Psicologia Generale e Clinica. Ciclo XXI. 2009. Rel. Làdavas E.
- [5] C. Passamonti. *Effetti dell'integrazione visuo-acustica in pazienti con disturbo di campo visivo*. Tesi di dottorato. Dottorato di ricerca in Psicologia Generale e Clinica. Ciclo XXI. 2009. Rel. Làdavas E.
- [6] F. Bremmer, A. Schlack, N. J. Shah, O. Zafiris, M. Kubischik, K. Hoffmann, K. Zilles, e G. R. Fink, «Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys», *Neuron*, vol 29, n° 1, pagg 287–296, Gen 2001.
- [7] E. Macaluso e J. Driver, «Spatial attention and crossmodal interactions between vision and touch», *Neuropsychologia*, vol 39, n° 12, pagg 1304–1316, 2001.
- [8] L. Leinonen, J. Hyvärinen, e A. R. Sovijärvi, «Functional properties of neurons in the temporo-parietal association cortex of awake monkey», *Exp Brain Res*, vol 39, n° 2, pagg 203–215, 1980.
- [9] C. J. Robinson e H. Burton, «Organization of somatosensory receptive fields in cortical areas 7b, retroinsula, postauditory and granular insula of *M. fascicularis*», *J. Comp. Neurol.*, vol 192, n° 1, pagg 69–92, Lug 1980.

- [10] C. E. Schroeder e J. Foxe, «Multisensory contributions to low-level, ‘unisensory’ processing», *Curr. Opin. Neurobiol.*, vol 15, n° 4, pagg 454–458, Ago 2005.
- [11] J. P. Rauschecker, «Parallel processing in the auditory cortex of primates», *Audiol. Neurootol.*, vol 3, n° 2–3, pagg 86–103, Giu 1998.
- [12] G. H. Recanzone, «Interactions of auditory and visual stimuli in space and time», *Hear. Res.*, vol 258, n° 1–2, pagg 89–99, Dic 2009.
- [13] G. H. Recanzone e M. L. Sutter, «The biological basis of audition», *Annu Rev Psychol.*, vol 59, pagg 119–142, 2008.
- [14] M. L. Phan, K. L. Schendel, G. H. Recanzone, e L. C. Robertson, «Auditory and visual spatial localization deficits following bilateral parietal lobe lesions in a patient with Balint’s syndrome», *J Cogn Neurosci*, vol 12, n° 4, pagg 583–600, Lug 2000.
- [15] P. Bertelson e M. Radeau, «Cross-modal bias and perceptual fusion with auditory-visual spatial discordance», *Percept Psychophys*, vol 29, n° 6, pagg 578–584, Giu 1981.
- [16] M. T. Wallace, G. E. Roberson, W. D. Hairston, B. E. Stein, J. W. Vaughan, e J. A. Schirillo, «Unifying multisensory signals across time and space», *Exp Brain Res*, vol 158, n° 2, pagg 252–258, Set 2004.
- [17] W. D. Hairston, M. T. Wallace, J. W. Vaughan, B. E. Stein, J. L. Norris, e J. A. Schirillo, «Visual localization ability influences cross-modal bias», *J Cogn Neurosci*, vol 15, n° 1, pagg 20–29, Gen 2003.
- [18] G. H. Recanzone, «Rapidly induced auditory plasticity: the ventriloquism aftereffect», *Proc. Natl. Acad. Sci. U.S.A.*, vol 95, n° 3, pagg 869–875, Feb 1998.
- [19] J. Lewald, «Rapid adaptation to auditory-visual spatial disparity», *Learn. Mem.*, vol 9, n° 5, pagg 268–278, Ott 2002.

- [20] I. Frissen, J. Vroomen, B. de Gelder, e P. Bertelson, «The aftereffects of ventriloquism: are they sound-frequency specific?», *Acta Psychol (Amst)*, vol 113, n° 3, pagg 315–327, Lug 2003.
- [21] I. Frissen, J. Vroomen, B. de Gelder, e P. Bertelson, «The aftereffects of ventriloquism: generalization across sound-frequencies», *Acta Psychol (Amst)*, vol 118, n° 1–2, pagg 93–100, Feb 2005.
- [22] E. Magosso, C. Cuppini, e M. Ursino, «A neural network model of ventriloquism effect and aftereffect», *PLoS ONE*, vol 7, n° 8, pag e42503, 2012.
- [23] N. Bolognini, F. Leo, C. Passamonti, B. E. Stein, e E. Làdavas, «Multisensory-mediated auditory localization», *Perception*, vol 36, n° 10, pagg 1477–1485, 2007.
- [24] G. H. Recanzone, «Spatial processing in the auditory cortex of the macaque monkey», *Proc. Natl. Acad. Sci. U.S.A.*, vol 97, n° 22, pagg 11829–11835, Ott 2000.
- [25] G. H. Recanzone, D. C. Guard, e M. L. Phan, «Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey», *J. Neurophysiol.*, vol 83, n° 4, pagg 2315–2331, Apr 2000.